Integral Calculus: Mathematics 103

Notes by Leah Edelstein-Keshet:
University of British Columbia

January 2, 2010

Contents

List of Figures

List of Tables

Preface

Integral calculus arose originally to solve very practical problems that merchants, landowners, and ordinary people faced on a daily basis. Among such pressing problems were the following: How much should one pay for a piece of land? If that land has an irregular shape, i.e. is not a simple geometrical shape, how should its area (and therefore, its cost) be calculated? How much olive oil or wine, are you getting when you purchase a barrel-full? Barrels come is a variety of shapes and sizes. If the barrel is not close to cylindrical, what is its volume (and thus, a reasonable price to pay)? In most such transactions, the need to accurately measure an area or a volume went well beyond the available results of geometry. (It was known how to compute areas of rectangles, triangles, and polygons. Volumes of cylinders and cubes were also known, but these were at best crude approximations to actual shapes and objects encountered in commerce.) This led to motivation for the development of the topic we now call integral calculus.

Essentially, the approach is based on the idea of “divide and conquer”: that is, cut up the geometric shape into smaller pieces, and approximate those pieces by regular shapes that can be quantified using simple geometry. In computing the area of an irregular shape, add up the areas of the (approximately regular) little parts in your “dissection”, to arrive at an approximation of the desired area of the shape. Depending on how fine the dissection (i.e. how many little parts), this approximation could be quite crude, or fairly accurate. The idea of applying a limit to obtain the true dimensions of the object was a flash of inspiration that led to modern day calculus. Similar ideas apply to computing the volume of a 3D object by successive subdivisions.

It is the aim of a calculus course to develop the language to deal with such concepts, to make such concepts systematic, and to find convenient and relevant shortcuts that can be used to solve a variety of problems that have common features. More than that, it is the purpose of this course to show that ideas developed in the original context of geometry (finding areas or volumes of 2D or 3D shapes) can be generalized and extended to a variety of applications that have little to do with geometry.

One area of application is that of computing total change given some time-dependent rate of change. We encounter many cases where a process changes at a rate that varies over time: the rate of production of hormone changes over a day, the rate of flow of water in a river changes over the seasons, or the rate of motion of a vehicle (i.e. its velocity) changes over its path. Computing the total change over some time span turns out to be closely related to the same underlying concept of “divide and conquer”: namely, subdivide (the time interval) and add up approximate changes over each of the smaller subintervals. The same idea applies to quantities that are distributed not in time but rather over space.

We show the connection between material that is spatially distributed in a nonuniform way (e.g. a density that varies from point to point) and total amount of material (obtained by the same process of integration).

A theme that unites much of the approach is that integral calculus has both analytic (i.e. pencil and paper) calculations - but these apply to a limited set of cases, and analogous numerical (i.e. computer-enabled) calculations. The two go hand-in-hand, with concepts that are closely linked. A set of computer labs using a spreadsheet tool are an important part of this course. The importance of seeing calculus from these two distinct but related perspectives is stressed: on the one hand, analytic computations can be very powerful and helpful, but at the same time, many interesting problems are too challenging to be handled by integration techniques. Here is where the same ideas, used in the context of simple computer algorithms, comes in handy. For this reason, the importance of understanding the concepts (not just the technical results, or the “formulae” for integrals) is vital: Ideas used to develop the analytic techniques on which calculus is based can be adapted to develop good working methods for harnessing computer power to solve problems. This is particularly useful in cases where the analytic methods are not sufficient or too technically challenging.

This set of lecture notes grew out of many years of teaching of Mathematics 103. The material is organized as follows: In Chapter 1 we develop the basic formulae for areas and volumes of elementary shapes, and show how to set up summations that describe compound objects made up of many such shapes. An example to motivate these ideas is the volume and surface area of a branching structure. In Chapter 2, we turn attention to the classic problem of defining and computing the area of a two-dimensional region, leading to the notion of the definite integral. In Chapter 3, we discuss the linchpin of Integral Calculus, namely the Fundamental Theorem that connects derivatives and integrals. This allows us to find a great shortcut to the analytic computations described in Chapter 2. Applications of these ideas to calculating total change from rates of change, and to computing volumes and masses are discussed in Chapters 4 and 5.

To expand our reach to other cases, we discuss the techniques on integration in Chapter 6. Here, we find that the chain rule of calculus reappears (in the form of substitution integrals), and a variety of miscellaneous tricks are devised to simplify integrals. Among these, the most important is integration by parts, a technique that has independent applications in many areas of science.

We study the ideas of probability in Chapters 7 and 8. Here we rediscover the connection between discrete sums and continuous integration, and apply the techniques to computing expected values for random variables. The connection between the mean (in probability) and the center of mass (of a density distributed in space) is illustrated.

Many scientific problems are phrased in terms of rules about rates of change. Quite often such rules take the form of differential equations. In an earlier differential calculus course, the student will have made acquaintance with the topic of such equations and qualitative techniques associated with interpreting their solutions. With the methods of integral calculus in hand, we can solve some types of differential equations analytically. This is discussed in Chapter 9.

The course concludes with the development of some notions of infinite sums and convergence in Chapter 10. Of prime importance, the Taylor series is developed and discussed in this concluding chapter.

Chapter 1

Areas, volumes and simple sums

1.1 Introduction

This introductory chapter has several aims. First, we concentrate here a number of basic formulae for areas and volumes that are used later in developing the notions of integral calculus. Among these are areas of simple geometric shapes and formulae for sums of certain common sequences. An important idea is introduced, namely that we can use the sum of areas of elementary shapes to approximate the areas of more complicated objects, and that the approximation can be made more accurate by a process of refinement.

We show using examples how such ideas can be used in calculating the volumes or areas of more complex objects. In particular, we conclude with a detailed exploration of the structure of branched airways in the lung as an application of ideas in this chapter.

1.2 Areas of simple shapes

One of the main goals in this course will be calculating areas enclosed by curves in the plane and volumes of three dimensional shapes. We will find that the tools of calculus will provide important and powerful techniques for meeting this goal. Some shapes are simple enough that no elaborate techniques are needed to compute their areas (or volumes). We briefly survey some of these simple geometric shapes and list what we know or can easily determine about their area or volume.

The areas of simple geometrical objects, such as rectangles, parallelograms, triangles, and circles are given by elementary formulae. Indeed, our ability to compute areas and volumes of more elaborate geometrical objects will rest on some of these simple formulae, summarized below.

Rectangular areas

Most integration techniques discussed in this course are based on the idea of carving up irregular shapes into rectangular strips. Thus, areas of rectangles will play an important part in those methods.

- The area of a rectangle with base b and height h is

$$A = b \cdot h$$

- Any parallelogram with height h and base b also has area, $A = b{\cdot}h$. See Figure 1.1(a) and (b)

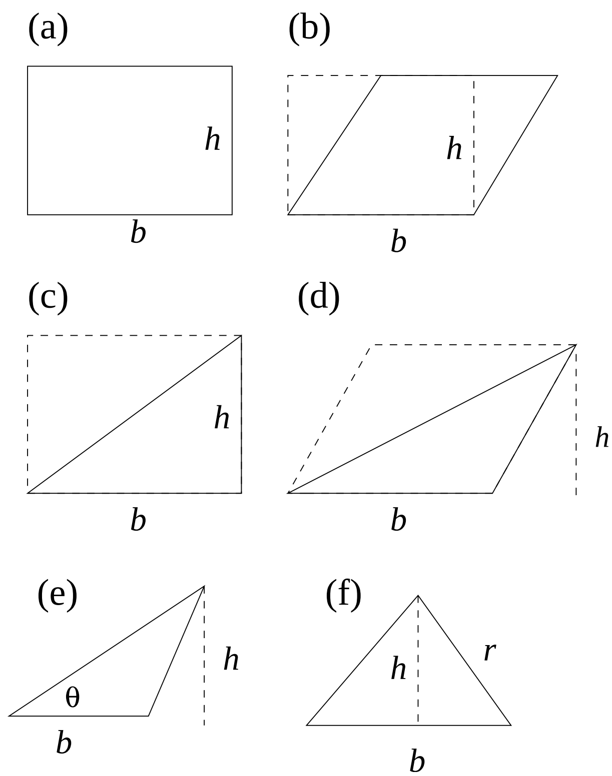

Figure 1.1. *Planar regions whose areas are given by elementary formulae.*

Areas of triangular shapes

A few illustrative examples in this chapter will be based on dissecting shapes (such as regular polygons) into triangles. The areas of triangles are easy to compute, and we summarize this review material below. However, triangles will play a less important role in subsequent integration methods.

- The area of a triangle can be obtained by slicing a rectangle or parallelogram in half, as shown in Figure 1.1(c) and (d). Thus, any triangle with base b and height h has area

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

- In some cases, the height of a triangle is not given, but can be determined from other information provided. For example, if the triangle has sides of length b and r with enclosed angle θ, as shown on Figure 1.1(e) then its height is simply $h = r\sin(\theta)$, and its area is
$$A = (1/2)br\sin(\theta)$$
- If the triangle is isosceles, with two sides of equal length, r, and base of length b, as in Figure 1.1(f) then its height can be obtained from Pythagoras's theorem, i.e. $h^2 = r^2 - (b/2)^2$ so that the area of the triangle is
$$A = (1/2)b\sqrt{r^2 - (b/2)^2}.$$

1.2.1 Example 1: Finding the area of a polygon using triangles: a "dissection" method

Using the simple ideas reviewed so far, we can determine the areas of more complex geometric shapes. For example, let us compute the area of a regular polygon with n equal sides, where the length of each side is $b = 1$. This example illustrates how a complex shape (the polygon) can be dissected into simpler shapes, namely triangles[1].

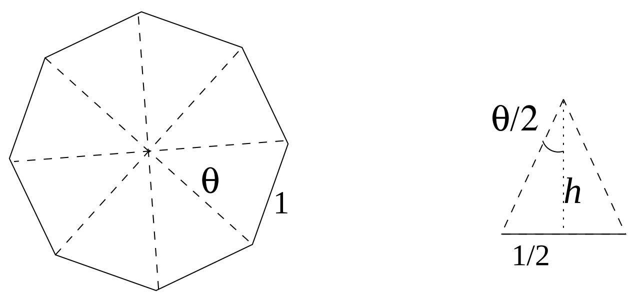

Figure 1.2. *An equilateral n-sided polygon with sides of unit length can be dissected into n triangles. One of these triangles is shown at right. Since it can be further divided into two Pythagorean triangles, trigonometric relations can be used to find the height h in terms of the length of the base $1/2$ and the angle $\theta/2$.*

Solution

The polygon has n sides, each of length $b = 1$. We dissect the polygon into n isosceles triangles, as shown in Figure 1.2. We do not know the heights of these triangles, but the angle θ can be found. It is
$$\theta = 2\pi/n$$
since together, n of these identical angles make up a total of 360° or 2π radians.

[1] This calculation will be used again to find the area of a circle in Section 1.2.2. However, note that in later chapters, our dissections of planar areas will focus mainly on rectangular pieces.

Let h stand for the height of one of the triangles in the dissected polygon. Then trigonometric relations relate the height to the base length as follows:

$$\frac{\text{opp}}{\text{adj}} = \frac{b/2}{h} = \tan(\theta/2)$$

Using the fact that $\theta = 2\pi/n$, and rearranging the above expression, we get

$$h = \frac{b}{2\tan(\pi/n)}$$

Thus, the area of each of the n triangles is

$$A = \frac{1}{2}bh = \frac{1}{2}b\left(\frac{b}{2\tan(\pi/n)}\right).$$

The statement of the problem specifies that $b = 1$, so

$$A = \frac{1}{2}\left(\frac{1}{2\tan(\pi/n)}\right).$$

The area of the entire polygon is then n times this, namely

$$A_{\text{n-gon}} = \frac{n}{4\tan(\pi/n)}.$$

For example, the area of a square (a polygon with 4 equal sides, $n = 4$) is

$$A_{\text{square}} = \frac{4}{4\tan(\pi/4)} = \frac{1}{\tan(\pi/4)} = 1,$$

where we have used the fact that $\tan(\pi/4) = 1$.

As a second example, the area of a hexagon (6 sided polygon, i.e. $n = 6$) is

$$A_{\text{hexagon}} = \frac{6}{4\tan(\pi/6)} = \frac{3}{2(1/\sqrt{3})} = \frac{3\sqrt{3}}{2}.$$

Here we used the fact that $\tan(\pi/6) = 1/\sqrt{3}$.

1.2.2 Example 2: How Archimedes discovered the area of a circle: dissect and "take a limit"

As we learn early in school the formula for the area of a circle of radius r, $A = \pi r^2$. But how did this convenient formula come about? and how could we relate it to what we know about simpler shapes whose areas we have discussed so far. Here we discuss how this formula for the area of a circle was determined long ago by Archimedes using a clever "dissection" and approximation trick. We have already seen part of this idea in dissecting a polygon into triangles, in Section 1.2.1. Here we see a terrifically important second step that formed the "leap of faith" on which most of calculus is based, namely taking a limit as the number of subdivisions increases [2].

First, we recall the definition of the constant π:

[2] This idea has important parallels with our later development of integration. Here it involves adding up the areas of triangles, and then taking a limit as the number of triangles gets larger. Later on, we do much the same, but using rectangles in the dissections.

Definition of π

In any circle, π is the ratio of the circumference to the diameter of the circle. (Comment: expressed in terms of the radius, this assertion states the obvious fact that the ratio of $2\pi r$ to $2r$ is π.)

Shown in Figure 1.3 is a sequence of regular polygons inscribed in the circle. As the number of sides of the polygon increases, its area gradually becomes a better and better approximation of the area inside the circle. Similar observations are central to integral calculus, and we will encounter this idea often. We can compute the area of any one of these polygons by dissecting into triangles. All triangles will be isosceles, since two sides are radii of the circle, whose length we'll call r.

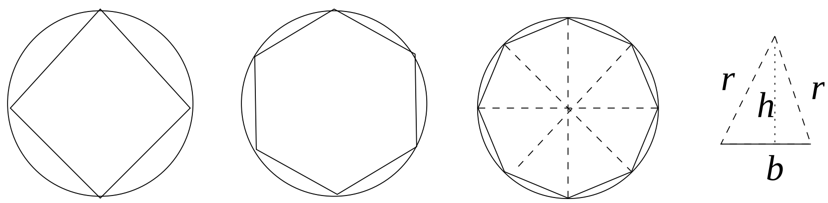

Figure 1.3. *Archimedes approximated the area of a circle by dissecting it into triangles.*

Let r denote the radius of the circle. Suppose that at one stage we have an n sided polygon. (If we knew the side length of that polygon, then we already have a formula for its area. However, this side length is not known to us. Rather, we know that the polygon should fit exactly inside a circle of radius r.) This polygon is made up of n triangles, each one an isosceles triangle with two equal sides of length r and base of undetermined length that we will denote by b. (See Figure 1.3.) The area of this triangle is

$$A_{\text{triangle}} = \frac{1}{2}bh.$$

The area of the whole polygon, A_n is then

$$A = n \cdot (\text{area of triangle}) = n\frac{1}{2}bh = \frac{1}{2}(nb)h.$$

We have grouped terms so that (nb) can be recognized as the perimeter of the polygon (i.e. the sum of the n equal sides of length b each). Now consider what happens when we increase the number of sides of the polygon, taking larger and larger n. Then the height of each triangle will get closer to the radius of the circle, and the perimeter of the polygon will get closer and closer to the perimeter of the circle, which is (by definition) $2\pi r$. i.e. as $n \to \infty$,

$$h \to r, \quad (nb) \to 2\pi r$$

so

$$A = \frac{1}{2}(nb)h \to \frac{1}{2}(2\pi r)r = \pi r^2$$

We have used the notation "$\to$" to mean that in the limit, as n gets large, the quantity of interest "approaches" the value shown. This argument proves that the area of a circle must be

$$A = \pi r^2.$$

One of the most important ideas contained in this little argument is that by approximating a shape by a larger and larger number of simple pieces (in this case, a large number of triangles), we get a better and better approximation of its area. This idea will appear again soon, but in most of our standard calculus computations, we will use a collection of rectangles, rather than triangles, to approximate areas of interesting regions in the plane.

Areas of other shapes

We concentrate here the area of a circle and of other shapes.

- The area of a circle of radius r is

$$A = \pi r^2.$$

- The surface area of a sphere of radius r is

$$S_{\text{ball}} = 4\pi r^2.$$

- The surface area of a right circular cylinder of height h and base radius r is

$$S_{\text{cyl}} = 2\pi r h.$$

Units

The units of area can be meters^2 (m^2), centimeters^2 (cm^2), square inches, etc.

1.3 Simple volumes

Later in this course, we will also be computing the volumes of 3D shapes. As in the case of areas, we collect below some basic formulae for volumes of elementary shapes. These will be useful in our later discussions.

1. The volume of a cube of side length s (Figure 1.4a), is

$$V = s^3.$$

2. The volume of a rectangular box of dimensions $h,\ w,\ l$ (Figure 1.4b) is

$$V = hwl.$$

3. The volume of a cylinder of base area A and height h, as in Figure 1.4(c), is

$$V = Ah.$$

   This applies for a cylinder with flat base of any shape, circular or not.

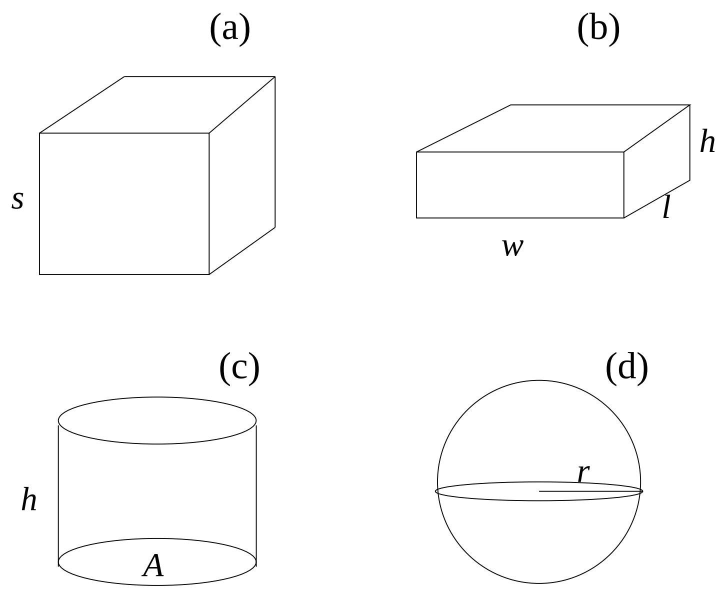

Figure 1.4. *3-dimensional shapes whose volumes are given by elementary formulae*

4. In particular, the volume of a cylinder with a circular base of radius r, (e.g. a disk) is

$$V = h(\pi r^2).$$

5. The volume of a sphere of radius r (Figure 1.4d), is

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

6. The volume of a spherical shell (hollow sphere with a shell of some small thickness, τ) is approximately

$$V \approx \tau \cdot (\text{surface area of sphere}) = 4\pi\tau r^2.$$

7. Similarly, a cylindrical shell of radius r, height h and small thickness, τ has volume given approximately by

$$V \approx \tau \cdot (\text{surface area of cylinder}) = 2\pi\tau r h.$$

Units

The units of volume are meters^3 (m^3), centimeters^3 (cm^3), cubic inches, etc.

1.3.1 Example 3: The Tower of Hanoi: a tower of disks

In this example, we consider how elementary shapes discussed above can be used to determine volumes of more complex objects. The Tower of Hanoi is a shape consisting of a number of stacked disks. It is a simple calculation to add up the volumes of these disks, but if the tower is large, and comprised of many disks, we would want some shortcut to avoid long sums[3].

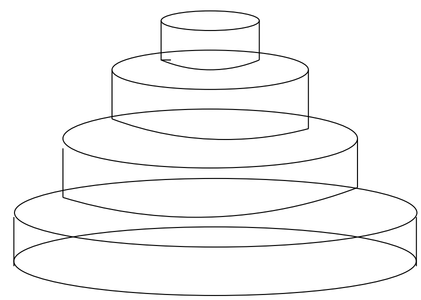

Figure 1.5. *Computing the volume of a set of disks. (This structure is sometimes called the tower of Hanoi after a mathematical puzzle by the same name.)*

(a) Compute the volume of a tower made up of four disks stacked up one on top of the other, as shown in Figure 1.5. Assume that the radii of the disks are 1, 2, 3, 4 units and that each disk has height 1.

(b) Compute the volume of a tower made up of 100 such stacked disks, with radii $r = 1, 2, \ldots, 99, 100$.

Solution

(a) The volume of the four-disk tower is calculated as follows:

$$V = V_1 + V_2 + V_3 + V_4,$$

where V_i is the volume of the i'th disk whose radius is $r = i,\ i = 1, 2 \ldots 4$. The height of each disk is $h = 1$, so

$$V = (\pi 1^2) + (\pi 2^2) + (\pi 3^2) + (\pi 4^2) = \pi(1 + 4 + 9 + 16) = 30\pi.$$

(b) The idea will be the same, but we have to calculate

$$V = \pi(1^2 + 2^2 + 3^2 + \ldots + 99^2 + 100^2).$$

It would be tedious to do this by adding up individual terms, and it is also cumbersome to write down the long list of terms that we will need to add up. This motivates inventing some helpful notation, and finding some clever way of performing such calculations.

[3]Note that the idea of computing a volume of a radially symmetric 3D shape by dissection into disks will form one of the main themes in Chapter 5. Here, the sums of the volumes of disks is exactly the same as the volume of the tower. Later on, the disks will only approximate the true 3D volume, and a limit will be needed to arrive at a "true volume".

1.4 Summations and the "Sigma" notation

We introduce the following notation for the operation of summing a list of numbers:

$$\boxed{S = a_1 + a_2 + a_3 + \ldots + a_N \equiv \sum_{k=1}^{N} a_k.}$$

The Greek symbol Σ ("Sigma") indicates summation. The symbol k used here is called the "index of summation" and it keeps track of where we are in the list of summands. The notation $k = 1$ that appears underneath Σ indicates where the sum begins (i.e. which term starts off the series), and the superscript N tells us where it ends. We will be interested in getting used to this notation, as well as in actually computing the value of the desired sum using a variety of shortcuts.

Example 4a: Summation notation

Suppose we want to form the sum of ten numbers, each equal to 1. We would write this as

$$S = 1 + 1 + 1 + \ldots 1 \equiv \sum_{k=1}^{10} 1.$$

The notation ... signifies that we have left out some of the terms (out of laziness, or in cases where there are too many to conveniently write down.) We could have just as well written the sum with another symbol (e.g. n) as the index, i.e. the same operation is implied by

$$\sum_{n=1}^{10} 1.$$

To compute the value of the sum we use the elementary fact that the sum of ten ones is just 10, so

$$S = \sum_{k=1}^{10} 1 = 10.$$

Example 4b: Sum of squares

Expand and sum the following:

$$S = \sum_{k=1}^{4} k^2.$$

Solution

$$S = \sum_{k=1}^{4} k^2 = 1 + 2^2 + 3^2 + 4^2 = 1 + 4 + 9 + 16 = 30.$$

(We have already seen this sum in part (a) of The Tower of Hanoi.)

Example 4c: Common factors

Add up the following list of 100 numbers (only a few of them are shown):

$$S = 3 + 3 + 3 + 3 + \ldots + 3.$$

Solution

There are 100 terms, all equal, so we can take out a common factor

$$S = 3 + 3 + 3 + 3 + \ldots + 3 = \sum_{k=1}^{100} 3 = 3 \sum_{k=1}^{100} 1 = 3(100) = 300.$$

Example 4d: Finding the pattern

Write the following terms in summation notation:

$$S = \frac{1}{3} + \frac{1}{9} + \frac{1}{27} + \frac{1}{81}.$$

Solution

We recognize that there is a pattern in the sequence of terms, namely, each one is $1/3$ raised to an increasing integer power, i.e.

$$S = \frac{1}{3} + \left(\frac{1}{3}\right)^2 + \left(\frac{1}{3}\right)^3 + \left(\frac{1}{3}\right)^4.$$

We can represent this with the "Sigma" notation as follows:

$$S = \sum_{n=1}^{4} \left(\frac{1}{3}\right)^n.$$

The "index" n starts at 1, and counts up through 2, 3, and 4, while each term has the form of $(1/3)^n$. This series is a **geometric series**, to be explored shortly. In most cases, a standard geometric series starts off with the value 1. We can easily modify our notation to include additional terms, for example:

$$S = \sum_{n=0}^{5} \left(\frac{1}{3}\right)^n = 1 + \frac{1}{3} + \left(\frac{1}{3}\right)^2 + \left(\frac{1}{3}\right)^3 + \left(\frac{1}{3}\right)^4 + \left(\frac{1}{3}\right)^5.$$

Learning how to compute the sum of such terms will be important to us, and will be described later on in this chapter.

1.4.1 Manipulations of sums

Since addition is commutative and distributive, sums of lists of numbers satisfy many convenient properties. We give a few examples below:

Example 5a: Simple operations

Simplify the following expression:

$$\sum_{k=1}^{10} 2^k - \sum_{k=3}^{10} 2^k.$$

Solution

$$\sum_{k=1}^{10} 2^k - \sum_{k=3}^{10} 2^k = (2 + 2^2 + 2^3 + \cdots + 2^{10}) - (2^3 + \cdots + 2^{10}) = 2 + 2^2.$$

We could have arrived at this conclusion directly from

$$\sum_{k=1}^{10} 2^k - \sum_{k=3}^{10} 2^k = \sum_{k=1}^{2} 2^k = 2 + 2^2 = 2 + 4 = 6.$$

The idea is that all but the first two terms in the first sum will cancel. The only remaining terms are those corresponding to $k = 1$ and $k = 2$.

Example 5b: Expanding

Expand the following expression:

$$\sum_{n=0}^{5} (1 + 3^n).$$

Solution

$$\sum_{n=0}^{5} (1 + 3^n) = \sum_{n=0}^{5} 1 + \sum_{n=0}^{5} 3^n.$$

1.5 Summation formulas

In this section we introduce a few examples of useful sums and give formulae that provide a shortcut to dreary calculations.

The sum of consecutive integers (Gauss' formula)

We first show that the sum of the first N integers is:

$$\boxed{S = 1 + 2 + 3 + \ldots + N = \sum_{k=1}^{N} k = \frac{N(N+1)}{2}. \qquad (1.1)}$$

The following trick is due to Gauss. By aligning two copies of the above sum, one written backwards, we can easily add them up one by one vertically. We see that:

$$\begin{array}{rcccccccccc} S = & 1 & + & 2 & + & \ldots & + & (N-1) & + & N \\ + & & & & & & & & & \\ S = & N & + & (N-1) & + & \ldots & + & 2 & + & 1 \end{array}$$

$$2S = \quad (1+N) \quad + \quad (1+N) \quad + \quad \ldots \quad + \quad (1+N) \quad + \quad (1+N)$$

Thus, there are N times the value $(N+1)$ above, so that

$$2S = N(1+N), \quad \text{so} \quad S = \frac{N(1+N)}{2}.$$

Thus, Gauss' formula is confirmed.

Example: Adding up the first 1000 integers

Suppose we want to add up the first 1000 integers. This formula is very useful in what would otherwise be a huge calculation. We find that

$$S = 1+2+3+\ldots+1000 = \sum_{k=1}^{1000} k = \frac{1000(1+1000)}{2} = 500(1001) = 500500.$$

Two other useful formulae are those for the sums of consecutive squares and of consecutive cubes:

The sum of the first N consecutive square integers

$$S_2 = 1^2+2^2+3^2+\ldots+N^2 = \sum_{k=1}^{N} k^2 = \frac{N(N+1)(2N+1)}{6}. \tag{1.2}$$

The sum of the first N consecutive cube integers

$$S_3 = 1^3+2^3+3^3+\ldots+N^3 = \sum_{k=1}^{N} k^3 = \left(\frac{N(N+1)}{2}\right)^2. \tag{1.3}$$

In the Appendix, we show how the formula for the sum of square integers can be proved by a technique called *mathematical induction*.

1.5.1 Example 3, revisited: Volume of a Tower of Hanoi

Armed with the formula for the sum of squares, we can now return to the problem of computing the volume of a tower of 100 stacked disks of heights 1 and radii $r = 1,\ 2, \ldots, 99,\ 100$. We have

$$V = \pi(1^2+2^2+3^2+\ldots+99^2+100^2) = \pi\sum_{k=1}^{100} k^2 = \pi\frac{100(101)(201)}{6} = 338,350\pi \text{ cubic units.}$$

Examples: Evaluating the sums

Compute the following two sums:

$$\text{(a) } S_a = \sum_{k=1}^{20}(2-3k+2k^2), \qquad \text{(b) } S_b = \sum_{k=10}^{50} k.$$

Solutions

(a) We can separate this into three individual sums, each of which can be handled by algebraic simplification and/or use of the summation formulae developed so far.

$$S_a = \sum_{k=1}^{20}(2-3k+2k^2) = 2\sum_{k=1}^{20} 1 - 3\sum_{k=1}^{20} k + 2\sum_{k=1}^{20} k^2.$$

Thus, we get

$$S_a = 2(20) - 3\left(\frac{20(21)}{2}\right) + 2\left(\frac{(20)(21)(41)}{6}\right) = 5150.$$

(b) We can express the second sum as a difference of two sums:

$$S_b = \sum_{k=10}^{50} k = \left(\sum_{k=1}^{50} k\right) - \left(\sum_{k=1}^{9} k\right).$$

Thus

$$S_b = \left(\frac{50(51)}{2} - \frac{9(10)}{2}\right) = 1275 - 45 = 1230.$$

1.6 Summing the geometric series

Consider a sum of terms that all have the form r^k, where r is some real number and k is an integer power. We refer to a series of this type as a **geometric series**. We have already seen one example of this type in a previous section. Below we will show that the sum of such a series is given by:

$$\boxed{S_N = 1 + r + r^2 + r^3 + \ldots + r^N = \sum_{k=0}^{N} r^k = \frac{1-r^{N+1}}{1-r}} \tag{1.4}$$

where $r \neq 1$. We call this sum a (finite) geometric series. We would like to find an expression for terms of this form in the general case of any real number r, and finite number of terms N. First we note that there are $N+1$ terms in this sum, so that if $r = 1$ then

$$S_N = 1 + 1 + 1 + \ldots 1 = N + 1$$

(a total of $N+1$ ones added.) If $r \neq 1$ we have the following trick:

$$\begin{array}{rccccccccc} S = & 1 & + & r & + & r^2 & + & \ldots & + & r^N \\ - & & & & & & & & & \\ rS = & & & r & + & r^2 & + & \ldots & + & r^{N+1} \end{array}$$

Subtracting leads to

$$S - rS = (1 + r + r^2 + \ldots + r^N) - (r + r^2 + \ldots + r^N + r^{N+1})$$

Most of the terms on the right hand side cancel, leaving

$$S(1 - r) = 1 - r^{N+1}.$$

Now dividing both sides by $1 - r$ leads to

$$S = \frac{1 - r^{N+1}}{1 - r},$$

which was the formula to be established.

Example: Geometric series

Compute the following sum:

$$S_c = \sum_{k=0}^{10} 2^k.$$

Solution

This is a geometric series

$$S_c = \sum_{k=0}^{10} 2^k = \frac{1 - 2^{10+1}}{1 - 2} = \frac{1 - 2048}{-1} = 2047.$$

1.7 Prelude to infinite series

So far, we have looked at several examples of finite series, i.e. series in which there are only a finite number of terms, N (where N is some integer). We would like to investigate how the sum of a series behaves when more and more terms of the series are included. It is evident that in many cases, such as Gauss's series (1.1), or sums of squared or cubed integers (e.g., Eqs. (1.2) and (1.3)), the series simply gets larger and larger as more terms are included. We say that such series *diverge* as $N \to \infty$. Here we will look specifically for series that *converge*, i.e. have a finite sum, even as more and more terms are included[4].

Let us focus again on the geometric series and determine its behaviour when the number of terms is increased. Our goal is to find a way of attaching a meaning to the expression

$$S_n = \sum_{k=0}^{\infty} r^k,$$

when the series becomes an *infinite series*. We will use the following definition:

[4] Convergence and divergence of series is discussed in fuller depth in Chapter 10 in the context of Taylor Series. However, these concepts are so important that it was felt necessary to introduce some preliminary ideas early in the term.

1.7.1 The infinite geometric series

Definition

An infinite series that has a finite sum is said to be *convergent*. Otherwise it is *divergent*.

Definition

Suppose that S is an (infinite) series whose terms are a_k. Then the *partial sums*, S_n, of this series are

$$S_n = \sum_{k=0}^{n} a_k.$$

We say that the sum of the infinite series is S, and write

$$S = \sum_{k=0}^{\infty} a_k,$$

provided that

$$S = \lim_{n\to\infty} \sum_{k=0}^{n} a_k.$$

That is, we consider the infinite series as the limit of the partial sums as the number of terms n is increased. In this case we also say that the infinite series converges to S.

We will see that only under certain circumstances will infinite series have a finite sum, and we will be interested in exploring two questions:

1. Under what circumstances does an infinite series have a finite sum.
2. What value does the partial sum approach as more and more terms are included.

In the case of a geometric series, the sum of the series, (1.4) depends on the number of terms in the series, n via r^{n+1}. Whenever $r > 1$, or $r < -1$, this term will get bigger in magnitude as n increases, whereas, for $0 < r < 1$, this term decreases in magnitude with n. We can say that

$$\lim_{n\to\infty} r^{n+1} = 0 \text{ provided } |r| < 1.$$

These observations are illustrated by two specific examples below. This leads to the following conclusion:

The sum of an infinite geometric series,

$$S = 1 + r + r^2 + \ldots + r^k + \ldots = \sum_{k=0}^{\infty} r^k,$$

exists provided $|r| < 1$ and is

$$S = \frac{1}{1-r}. \tag{1.5}$$

Examples of convergent and divergent geometric series are discussed below.

1.7.2 Example: A geometric series that converges.

Consider the geometric series with $r = \frac{1}{2}$, i.e.

$$S_n = 1 + \frac{1}{2} + \left(\frac{1}{2}\right)^2 + \left(\frac{1}{2}\right)^3 + \ldots + \left(\frac{1}{2}\right)^n = \sum_{k=0}^{n} \left(\frac{1}{2}\right)^k.$$

Then

$$S_n = \frac{1 - (1/2)^{n+1}}{1 - (1/2)}.$$

We observe that as n increases, i.e. as we retain more and more terms, we obtain

$$\lim_{n\to\infty} S_n = \lim_{n\to\infty} \frac{1 - (1/2)^{n+1}}{1 - (1/2)} = \frac{1}{1 - (1/2)} = 2.$$

In this case, we write

$$\sum_{n=0}^{\infty} \left(\frac{1}{2}\right)^n = 1 + \frac{1}{2} + (\frac{1}{2})^2 + \ldots = 2$$

and we say that "the (infinite) series *converges* to 2".

1.7.3 Example: A geometric series that diverges

In contrast, we now investigate the case that $r = 2$: then the series consists of terms

$$S_n = 1 + 2 + 2^2 + 2^3 + \ldots + 2^n = \sum_{k=0}^{n} 2^k = \frac{1 - 2^{n+1}}{1 - 2} = 2^{n+1} - 1$$

We observe that as n grows larger, the sum continues to grow indefinitely. In this case, we say that the sum *does not converge*, or, equivalently, that the sum *diverges*.

It is important to remember that an infinite series, i.e. a sum with infinitely many terms added up, can exhibit either one of these two very different behaviours. It may converge in some cases, as the first example shows, or *diverge* (fail to converge) in other cases. We will see examples of each of these trends again. It is essential to be able to distinguish the two. Divergent series (or series that diverge under certain conditions) must be handled with particular care, for otherwise, we may find contradictions or seemingly reasonable calculations that have meaningless results.

1.8 Application of geometric series to the branching structure of the lungs

In this section, we will compute the volume and surface area of the branched airways of lungs[5]. We use the summation formulae to arrive at the results, and we also illustrate how the same calculation could be handled using a simple spreadsheet.

[5] This section provides an example of how to set up a biologically relevant calculation based on geometric series. It is further studied in the homework problems. A similar example is given as an exercise for the student in Lab 1 of this calculus course.

Our lungs pack an amazingly large surface area into a confined volume. Most of the oxygen exchange takes place in tiny sacs called *alveoli* at the terminal branches of the airways passages. The bronchial tubes conduct air, and distribute it to the many smaller and smaller tubes that eventually lead to those alveoli. The principle of this efficient organ for oxygen exchange is that these very many small structures present a very large surface area. Oxygen from the air can diffuse across this area into the bloodstream very efficiently.

The lungs, and many other biological "distribution systems" are composed of a branched structure. The initial segment is quite large. It bifurcates into smaller segments, which then bifurcate further, and so on, resulting in a geometric expansion in the number of branches, their collective volume, length, etc. In this section, we apply geometric series to explore this branched structure of the lung. We will construct a simple mathematical model and explore its consequences. The model will consist in some well-formulated assumptions about the way that "daughter branches" are related to their "parent branch". Based on these assumptions, and on tools developed in this chapter, we will then predict properties of the structure as a whole. We will be particularly interested in the volume V and the surface area S of the airway passages in the lungs[6].

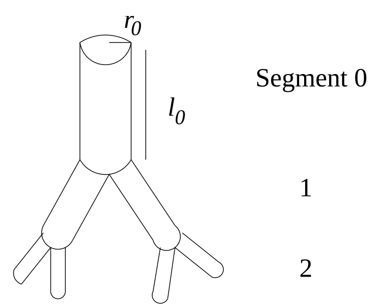

Figure 1.6. *Air passages in the lungs consist of a branched structure. The index n refers to the branch generation, starting from the initial segment, labeled 0. All segments are assumed to be cylindrical, with radius r_n and length ℓ_n in the n'th generation.*

1.8.1 Assumptions

- The airway passages consist of many "generations" of branched segments. We label the largest segment with index "0", and its daughter segments with index "1", their successive daughters "2", and so on down the structure from large to small branch segments. We assume that there are M "generations", i.e. the initial segment has undergone M subdivisions. Figure 1.6 shows only generations 0, 1, and 2. (Typically, for human lungs there can be up to 25-30 generations of branching.)
- At each generation, every segment is approximated as a cylinder of radius r_n and length ℓ_n.

[6]The surface area of the bronchial tubes does not actually absorb much oxygen, in humans. However, as an example of summation, we will compute this area and compare how it grows to the growth of the volume from one branching layer to the next.

radius of first segment	r_0	0.5 cm
length of first segment	ℓ_0	5.6 cm
ratio of daughter to parent length	α	0.9
ratio of daughter to parent radius	β	0.86
number of branch generations	M	30
average number daughters per parent	b	1.7

Table 1.1. *Typical structure of branched airway passages in lungs.*

- The number of branches grows along the "tree". On average, each parent branch produces b daughter branches. In Figure 1.6, we have illustrated this idea for $b = 2$. A branched structure in which each branch produces two daughter branches is described as a **bifurcating** tree structure (whereas **trifurcating** implies $b = 3$). In real lungs, the branching is slightly irregular. Not every level of the structure bifurcates, but in general, averaging over the many branches in the structure b is smaller than 2. In fact, the rule that links the number of branches in generation n, here denoted x_n with the number (of smaller branches) in the next generation, x_{n+1} is

$$x_{n+1} = bx_n. \tag{1.6}$$

  We will assume, for simplicity, that b is a constant. Since the number of branches is growing down the length of the structure, it must be true that $b > 1$. For human lungs, on average, $1 < b < 2$. Here we will take b to be constant, i.e. $b = 1.7$. In actual fact, this simplification cannot be precise, because we have just one segment initially ($x_0 = 1$), and at level 1, the number of branches x_1 should be some small *integer*, not a number like "1.7". However, as in many mathematical models, some accuracy is sacrificed to get intuition. Later on, details that were missed and are considered important can be corrected and refined.

- The ratios of radii and lengths of daughters to parents are approximated by "proportional scaling". This means that the relationship of the radii and lengths satisfy simple rules: The lengths are related by

$$\ell_{n+1} = \alpha\ell_n, \tag{1.7}$$

  and the radii are related by

$$r_{n+1} = \beta r_n, \tag{1.8}$$

  with α and β positive constants. For example, it could be the case that the radius of daughter branches is 1/2 or 2/3 that of the parent branch. Since the branches decrease in size (while their number grows), we expect that $0 < \alpha < 1$ and $0 < \beta < 1$.

Rules such as those given by equations (1.7) and (1.8) are often called *self-similar growth* laws. Such concepts are closely linked to the idea of fractals, i.e. theoretical structures produced by iterating such growth laws indefinitely. In a real biological structure, the number of generations is finite. (However, in some cases, a finite geometric series is well-approximated by an infinite sum.)

Actual lungs are not fully symmetric branching structures, but the above approximations are used here for simplicity. According to physiological measurements, the scale factors for sizes of daughter to parent size are in the range $0.65 \le \alpha, \beta \le 0.9$. (K. G. Horsfield, G. Dart, D. E. Olson, and G. Cumming, (1971) J. Appl. Phys. 31, 207217.) For the purposes of this example, we will use the values of constants given in Table 1.1.

1.8.2 A simple geometric rule

The three equations that govern the rules for successive branching, i.e. equations (1.6), (1.7), and (1.8), are examples of a very generic "geometric progression" recipe. Before returning to the problem at hand, let us examine the implications of this recursive rule, when it is applied to generating the whole structure. Essentially, we will see that the rule linking two generations implies an exponential growth. To see this, let us write out a few first terms in the progression of the sequence $\{x_n\}$:

$$\begin{aligned}
\text{initial value: } & x_0 \\
\text{first iteration: } & x_1 = bx_0 \\
\text{second iteration: } & x_2 = bx_1 = b(bx_0) = b^2 x_0 \\
\text{third iteration: } & x_3 = bx_2 = b(b^2 x_0) = b^3 x_0 \\
& \vdots
\end{aligned}$$

By the same pattern, at the n'th generation, the number of segments will be

$$n\text{'th iteration:} \quad x_n = bx_{n-1} = b(bx_{n-2}) = b(b(bx_{n-3})) = \ldots = \underbrace{(b \cdot b \cdots b)}_{n \text{ factors}} x_0 = b^n x_0.$$

We have arrived at a simple, but important result, namely:

> The rule linking two generations,
> $$x_n = bx_{n-1} \tag{1.9}$$
> implies that the n'th generation will have grown by a factor b^n, i.e.,
> $$x_n = b^n x_0. \tag{1.10}$$

This connection between the rule linking two generations and the resulting number of members at each generation is useful in other circumstances. Equation (1.9) is sometimes called a *recursion relation*, and its solution is given by equation (1.10). We will use the same idea to find the connection between the volumes, and surface areas of successive segments in the branching structure.

1.8.3 Total number of segments

We used the result of Section 1.8.2 and the fact that there is one segment in the 0'th generation, i.e. $x_0 = 1$, to conclude that at the n'th generation, the number of segments is

$$x_n = x_0 b^n = 1 \cdot b^n = b^n.$$

For example, if $b = 2$, the number of segments grows by powers of 2, so that the tree bifurcates with the pattern 1, 2, 4, 8, etc.

To determine how many branch segments there are in total, we add up over all generations, $0, 1, \ldots M$. This is a geometric series, whose sum we can compute. Using equation (1.4), we find

$$N = \sum_{n=0}^{M} b^n = \left(\frac{1 - b^{M+1}}{1 - b} \right).$$

Given b and M, we can then predict the exact number of segments in the structure. The calculation is summarized further on for values of the branching parameter, b, and the number of branch generations, M, given in Table 1.1.

1.8.4 Total volume of airways in the lung

Since each lung segment is assumed to be cylindrical, its volume is

$$v_n = \pi r_n^2 \ell_n.$$

Here we mean just a single segment in the n'th generation of branches. (There are b^n such identical segments in the n'th generation, and we will refer to the volume of all of them together as V_n below.)

The length and radius of segments also follow a geometric progression. In fact, the same idea developed above can be used to relate the length and radius of a segment in the n'th, generation segment to the length and radius of the original 0'th generation segment, namely,

$$\ell_n = \alpha \ell_{n-1} \Rightarrow \quad \ell_n = \alpha^n \ell_0,$$

and

$$r_n = \beta r_{n-1} \Rightarrow \quad r_n = \beta^n r_0.$$

Thus the volume of one segment in generation n is

$$v_n = \pi r_n^2 \ell_n = \pi (\beta^n r_0)^2 (\alpha^n \ell_0) = (\alpha \beta^2)^n \underbrace{(\pi r_0^2 \ell_0)}_{v_0}.$$

This is just a product of the initial segment volume $v_0 = \pi r_0^2 \ell_0$, with the n'th power of a certain factor(α, β). (That factor takes into account that both the radius and the length are being scaled down at every successive generation of branching.) Thus

$$v_n = (\alpha \beta^2)^n v_0.$$

The total volume of all (b^n) segments in the n'th layer is

$$V_n = b^n v_n = b^n(\alpha\beta^2)^n v_0 = (\underbrace{b\alpha\beta^2}_{a})^n\, v_0.$$

Here we have grouped terms together to reveal the simple structure of the relationship: one part of the expression is just the initial segment volume, while the other is now a "scale factor" that includes not only changes in length and radius, but also in the number of branches. Letting the constant a stand for that scale factor, $a = (b\alpha\beta^2)$ leads to the result that the volume of all segments in the n'th layer is

$$V_n = a^n v_0.$$

The total volume of the structure is obtained by summing the volumes obtained at each layer. Since this is a geometric series, we can use the summation formula. i.e., Equation (1.4). Accordingly, total airways volume is

$$V = \sum_{n=0}^{30} V_n = v_0 \sum_{n=0}^{30} a^n = v_0 \left(\frac{1 - a^{M+1}}{1 - a} \right).$$

The similarity of treatment with the previous calculation of number of branches is apparent. We compute the value of the constant a in Table 1.2, and find the total volume in Section 1.8.6.

1.8.5 Total surface area of the lung branches

The surface area of a single segment at generation n, based on its cylindrical shape, is

$$s_n = 2\pi r_n \ell_n = 2\pi(\beta^n r_0)(\alpha^n \ell_0) = (\alpha\beta)^n\, \underbrace{(2\pi r_0 \ell_0)}_{s_0},$$

where s_0 is the surface area of the initial segment. Since there are b^n branches at generation n, the total surface area of all the n'th generation branches is thus

$$S_n = b^n(\alpha\beta)^n s_0 = (\underbrace{b\alpha\beta}_{c})^n s_0,$$

where we have let c stand for the scale factor $c = (b\alpha\beta)$. Thus,

$$S_n = c^n s_0.$$

This reveals the similar nature of the problem. To find the total surface area of the airways, we sum up,

$$S = s_0 \sum_{n=0}^{M} c^n = s_0 \left(\frac{1 - c^{M+1}}{1 - c} \right).$$

We compute the values of s_0 and c in Table 1.2, and summarize final calculations of the total airways surface area in section 1.8.6.

volume of first segment	$v_0 = \pi r_0^2 \ell_0$	4.4 cm^3
surface area of first segment	$s_0 = 2\pi r_0 \ell_0$	17.6 cm^2
ratio of daughter to parent segment volume	$(\alpha\beta^2)$	0.66564
ratio of daughter to parent segment surface area	$(\alpha\beta)$	0.774
ratio of net volumes in successive generations	$a = b\alpha\beta^2$	1.131588
ratio of net surface areas in successive generations	$c = b\alpha\beta$	1.3158

Table 1.2. *Volume, surface area, scale factors, and other derived quantities. Because a and c are bases that will be raised to large powers, it is important to that their values are fairly accurate, so we keep more significant figures.*

1.8.6 Summary of predictions for specific parameter values

By setting up the model in the above way, we have revealed that each quantity in the structure obeys a simple geometric series, but with distinct "bases" b, a and c and coefficients $1, v_0$, and s_0. This approach has shown that the formula for geometric series applies in each case. Now it remains to merely "plug in" the appropriate quantities. In this section, we collect our results, use the sample values for a model "human lung" given in Table 1.1, or the resulting derived scale factors and quantities in Table 1.2 to finish the task at hand.

Total number of segments

$$N = \sum_{n=0}^{M} b^n = \left(\frac{1-b^{M+1}}{1-b}\right) = \left(\frac{1-(1.7)^{31}}{1-1.7}\right) = 1.9898 \cdot 10^7 \approx 2 \cdot 10^7.$$

According to this calculation, there are a total of about 20 million branch segments overall (including all layers, form top to bottom) in the entire structure!

Total volume of airways

Using the values for a and v_0 computed in Table 1.2, we find that the total volume of all segments in the n'th generation is

$$V = v_0 \sum_{n=0}^{30} a^n = v_0 \left(\frac{1-a^{M+1}}{1-a}\right) = 4.4 \frac{(1-1.131588^{31})}{(1-1.131588)} = 1510.3\ \mathrm{cm}^3.$$

Recall that 1 litre = 1000 cm^3. Then we have found that the lung airways contain about 1.5 litres.

Total surface area of airways

Using the values of s_0 and c in Table 1.2, the total surface area of the tubes that make up the airways is

$$S = s_0 \sum_{n=0}^{M} c^n = s_0 \left(\frac{1 - c^{M+1}}{1 - c} \right) = 17.6 \frac{(1 - 1.3158^{31})}{(1 - 1.3158)} = 2.76 \cdot 10^5 \text{ cm}^2.$$

There are 100 cm per meter, and $(100)^2 = 10^4$ cm^2 per m^2. Thus, the area we have computed is equivalent to about 28 square meters!

1.8.7 Exploring the problem numerically

Up to now, all calculations were done using the formulae developed for geometric series. However, sometimes it is more convenient to devise a computer algorithm to implement “rules” and perform repetitive calculations in a problem such as discussed here. The advantage of that approach is that it eliminates tedious calculations by hand, and, in cases where summation formulae are not know to us, reduces the need for analytical computations. It can also provide a shortcut to visual summary of the results. The disadvantage is that it can be less obvious how each of the values of parameters assigned to the problem affects the final answers.

A spreadsheet is an ideal tool for exploring iterated rules such as those given in the lung branching problem[7]. In Figure 1.7 we show the volumes and surface areas associated with the lung airways for parameter values discussed above. Both layer by layer values and cumulative sums leading to total volume and surface area are shown in each of (a) and (c). In (b) and (d), we compare these results to similar graphs in the case that one parameter, the branching number, b is adjusted from 1.7 (original value) to 2. The contrast between the graphs shows how such a small change in this parameter can significantly affect the results.

1.8.8 For further independent study

The following problems can be used for further independent exploration of these ideas.

1. In our model, we have assumed that, on average, a parent branch has only “1.7” daughter branches, i.e. that $b = 1.7$. Suppose we had assumed that $b = 2$. What would the total volume V be in that case, keeping all other parameters the same? Explain why this is biologically impossible in the case $M = 30$ generations. For what value of M would $b = 2$ lead to a reasonable result?

2. Suppose that the first 5 generations of branching produce 2 daughters each, but then from generation 6 on, the branching number is $b = 1.7$. How would you set up this variant of the model? How would this affect the calculated volume?

3. In the problem we explored, the net volume and surface area keep growing by larger and larger increments at each “generation” of branching. We would describe this as “unbounded growth”. Explain why this is the case, paying particular attention to the scale factors a and c.

[7] See Lab 1 for a similar problem that is also investigated using a spreadsheet.

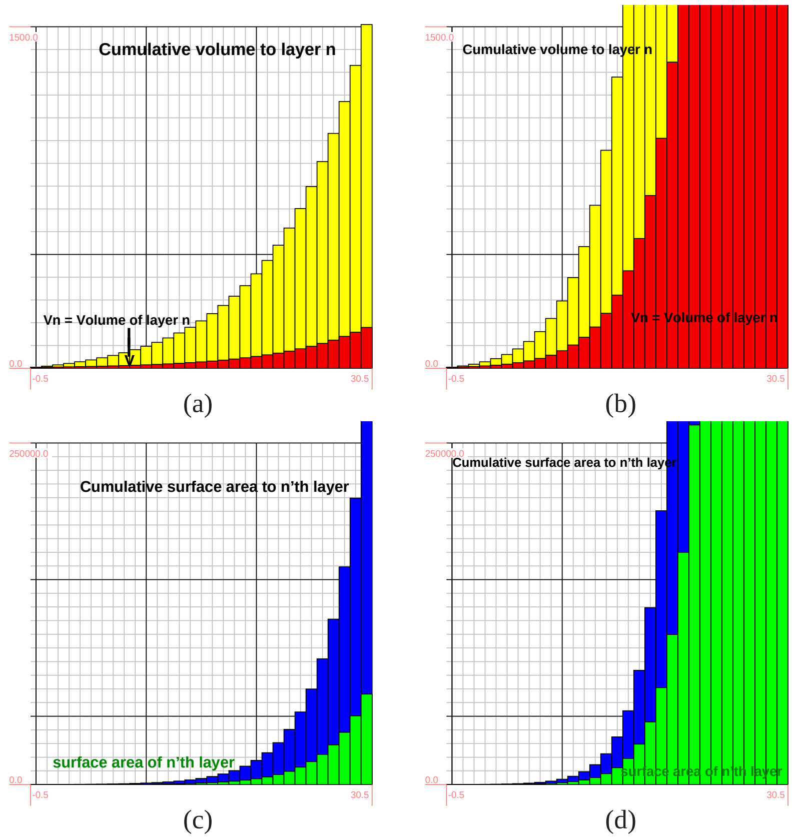

Figure 1.7. *(a) V_n, the volume of layer n (red bars), and the cumulative volume down to layer n (yellow bars) are shown for parameters given in Table 1.1. (b) Same as (a) but assuming that parent segments always produce two daughter branches (i.e. $b = 2$). The graphs in (a) and (b) are shown on the same scale to accentuate the much more dramatic growth in (b). (c) and (d): same idea showing the surface area of n'th layer (green) and the cumulative surface area to layer n (blue) for original parameters (in c), as well as for the value $b = 2$ (in d).*

4. Suppose we want a set of tubes with a large surface area but small total volume. Which *single* factor or parameter should we change (and how should we change it) to correct this feature of the model, i.e. to predict that the total volume of the branching tubes remains roughly constant while the surface area increases as branching layers are added.

5. Determine how the branching properties of real human lungs differs from our assumed model, and use similar ideas to refine and correct our estimates. You may want to investigate what is known about the actual branching parameter b, the number of generations of branches, M, and the ratios of lengths and radii that we have assumed. Alternately, you may wish to find parameters for other species and do a

comparative study of lungs in a variety of animal sizes.

6. Branching structures are ubiquitous in biology. Many species of plants are based on a regular geometric sequence of branching. Consider a tree that trifurcates (i.e. produces 3 new daughter branches per parent branch, $b = 3$). Explain (a) What biological problem is to be solved in creating such a structure (b) What sorts of constraints must be satisfied by the branching parameters to lead to a viable structure. This is an open-ended problem.

1.9 Summary

In this chapter, we collected useful formulae for areas and volumes of simple 2D and 3D shapes. A summary of the most important ones is given below. Table 1.3 lists the areas of simple shapes, Table 1.4 the volumes and Table 1.5 the surface areas of 3D shapes.

We used areas of triangles to compute areas of more complicated shapes, including regular polygons. We used a polygon with N sides to approximate the area of a circle, and then, by letting N go to infinity, we were able to prove that the area of a circle of radius r is $A = \pi r^2$. This idea, and others related to it, will form a deep underlying theme in the next two chapters and later on in this course.

We introduced some notation for series and collected useful formulae for summation of such series. These are summarized in Table 1.6. We will use these extensively in our next chapter.

Finally, we investigated geometric series and studied a biological application, namely the branching structure of lungs.

Object	dimensions	area, A
triangle	base b, height h	$\frac{1}{2}bh$
rectangle	base b, height h	bh
circle	radius r	πr^2

Table 1.3. Areas of planar regions

Object	dimensions	volume, V
box	base b, height h, width w	hwb
circular cylinder	radius r, height h	$\pi r^2 h$
sphere	radius r	$\frac{4}{3}\pi r^3$
cylindrical shell*	radius r, height h, thickness τ	$2\pi r h \tau$
spherical shell*	radius r, thickness τ	$4\pi r^2 \tau$

Table 1.4. Volumes of 3D shapes. * *Assumes a thin shell, i.e. small τ.*

Object	dimensions	surface area, S
box	base b, height h, width w	$2(bh + bw + hw)$
circular cylinder	radius r, height h	$2\pi r h$
sphere	radius r	$4\pi r^2$

Table 1.5. Surface areas of 3D shapes

Sum	Notation	Formula	Comment
$1 + 2 + 3 + \ldots + N$	$\sum_{k=1}^{N} k$	$\frac{N(1+N)}{2}$	Gauss' formula
$1^2 + 2^2 + 3^2 + \ldots + N^2$	$\sum_{k=1}^{N} k^2$	$\frac{N(N+1)(2N+1)}{6}$	Sum of squares
$1^3 + 2^3 + 3^3 + \ldots + N^3$	$\sum_{k=1}^{N} k^3$	$\left(\frac{N(N+1)}{2}\right)^2$	Sum of cubes
$1 + r + r^2 + r^3 \ldots r^N$	$\sum_{k=0}^{N} r^k$	$\frac{1-r^{N+1}}{1-r}$	Geometric sum

Table 1.6. Useful summation formulae.

Chapter 2

Areas

2.1 Areas in the plane

A long-standing problem of integral calculus is how to compute the area of a region in the plane. This type of geometric problem formed part of the original motivation for the development of calculus techniques, and we will discuss it in many contexts in this course. We have already seen examples of the computation of areas of especially simple geometric shapes in Chapter 1. For triangles, rectangles, polygons, and circles, no advanced methods (beyond simple geometry) are needed. However, beyond these elementary shapes, such methods fail, and a new idea is needed. We will discuss such ideas in this chapter, and in Chapter 3.

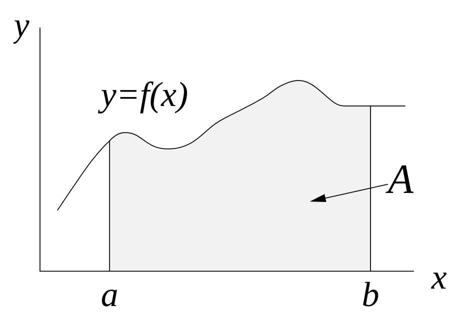

Figure 2.1. *We consider the problem of determining areas of regions such bounded by the x axis, the lines $x = a$ and $x = b$ and the graph of some function, $y = f(x)$.*

We now consider the problem of determining the area of a region in the plane that has the following special properties: The region is formed by straight lines on three sides, and by a smooth curve on one of its edges, as shown in Figure 2.1. You might imagine that the shaded portion of this figure is a plot of land bounded by fences on three sides, and by a river on the fourth side. A farmer wishing to purchase this land would want to know exactly how large an area is being acquired. Here we set up the calculation of that area.

More specifically, we use a cartesian coordinate system to describe the region: we require that it falls between the x-axis, the lines $x = a$ and $x = b$, and the graph of a function $y = f(x)$. This is required for the process described below to work[8]. We will first restrict attention to the case that $f(x) > 0$ for all points in the interval $a \leq x \leq b$ as we concentrate on "real areas". Later, we generalize our results and lift this restriction.

We will approximate the area of the region shown in Figure 2.1 by dissecting it into smaller regions (rectangular strips) whose areas are easy to determine. We will refer to this type of procedure as a **Riemann sum**. In Figure 2.2, we illustrate the basic idea using a region bounded by the function $y = f(x) = x^2$ on $0 \leq x \leq 1$. It can be seen that the

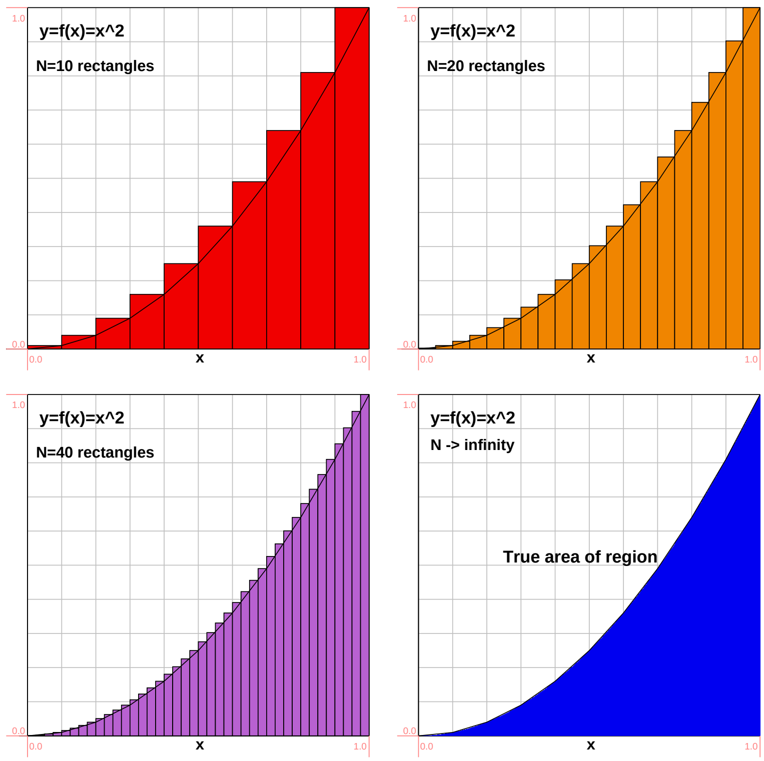

Figure 2.2. *The function $y = x^2$ for $0 \leq x \leq 1$ is shown, with rectangles that approximate the area under its curve. As we increase the number of rectangular strips, the total area of the strips becomes a better and better approximation of the desired "true" area. Shown are the intermediate steps $N = 10$, $N = 20$, $N = 40$ and the true area for $N \to \infty$*

approximation is fairly coarse when the number of rectangles is small[9]. However, if the number of rectangles is increased, (as shown in subsequent panels of this same figure), we

[8]Not all planar areas have this property. Later examples indicate how to deal with some that do not.

[9]That is, the area of the rectangles is very different from the area of the region of interest.

obtain a better and better approximation of the true area. In the limit as N, the number of rectangles, approaches infinity, the area of the desired region is obtained. This idea will form the core of this chapter. The reader will note a similarity with the idea we already encountered in obtaining the area of a circle, though in that context, we had used a dissection of the circle into approximating triangles.

With this idea in mind, in Section 2.2, we compute the area of the region shown in Figure 2.2 in two ways. First, we use a simple spreadsheet to do the computations for us. This is meant to illustrate the "numerical approach".

Then, as the alternate analytic approach , we set up the Riemann sum corresponding to the function shown in Figure 2.2. We will find that carefully setting up the calculation of areas of the approximating rectangles will be important. Making a cameo appearance in this calculation will be the formula for the sums of square integers developed in the previous chapter. A new feature will be the limit $N \to \infty$ that introduces the final step of arriving at the smooth region shown in the final panel of Figure 2.2.

2.2 Computing the area under a curve by rectangular strips

2.2.1 First approach: Numerical integration using a spreadsheet

The same tool that produces Figure 2.2 can be used to calculate the areas of the steps for each of the panels in the figure. To do this, we fix N for a given panel, (e.g. $N = 10, 20$, or 40), find the corresponding value of Δx, and set up a calculation which adds up the areas of steps, i.e. $\sum x^2 \Delta x$ in a given panel. The ideas are analogous to those described in Section 2.2.2, but a spreadsheet does the number crunching for us.

Using a spreadsheet, for example, we find the following results at each stage: For $N = 10$ strips, the area is 0.3850 units^2, for $N = 20$ strips it is 0.3588, for $N = 40$ strips, the area is 0.3459. If we increase N greatly, e.g. set $N = 1000$ strips, which begins to approximate the limit of $N \to \infty$, then the area obtained is 0.3338 units^2[10].

This example illustrates that areas can be computed "numerically" - indeed many of the laboratory exercises that accompany this course will be based on precisely this idea. The advantage of this approach is that it requires only elementary "programming" - i.e. the assembly of a simple **algorithm**, i.e. a set of instructions. Once assembled, we can use essentially the same algorithm to explore various functions, intervals, number of rectangles, etc. Lab 2 in this course will motivate the student to explore this numerical integration approach, and later labs will expand and generalize the idea to a variety of settings.

In our second approach, we set up the problem analytically. We will find that results are similar. However, we will get deeper insight by understanding what happens in the limit as the number of strips N gets very large.

[10]Note that all these values are approximations, correct to 4 decimal places. Compare with the exact calculations in Section 2.2.2

2.2.2 Second approach: Analytic computation using Riemann sums

In this section we consider the detailed steps involved in analytically computing the area of the region bounded by the function

$$y = f(x) = x^2, \quad 0 \le x \le 1.$$

By this we mean that we use "pen-and-paper" calculations, rather than computational aids to determine that area.

We set up the rectangles (as shown in Figure 2.2, with detailed labeling in Figures 2.3), determine the heights and areas of these rectangle, sum their total area, and then determine how this value behaves as the rectangles get more numerous (and thinner).

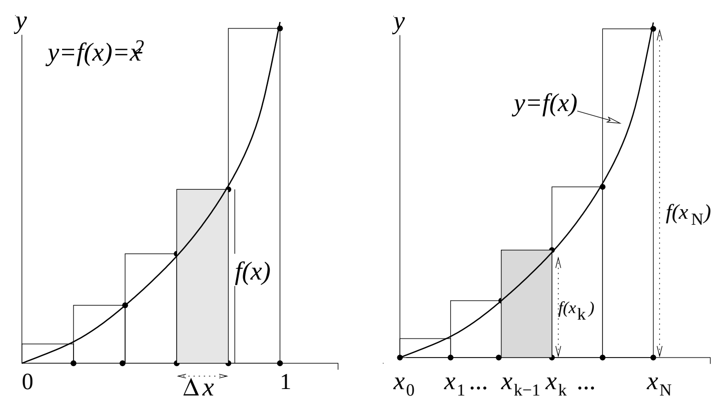

Figure 2.3. *The region under the graph of $y = f(x)$ for $0 \le x \le 1$ will be approximated by a set of N rectangles. A rectangle (shaded) has base width Δx and height $f(x)$. Since $0 \le x \le 1$, and the all rectangles have the same base width, it follows that $\Delta x = 1/N$. In the panel on the right, the coordinates of base corners and two typical heights of the rectangles have been labeled. Here $x_0 = 0, x_N = 1$ and $x_k = k\Delta x$.*

The interval of interest in this problem is $0 \le x \le 1$. Let us subdivide this interval into N equal subintervals. Then each has width $1/N$. (We will refer to this width as Δx, as shown in Figure 2.3, as it forms a difference of successive x coordinates.) The coordinates of the endpoints of these subintervals will be labeled $x_0, x_1, \ldots, x_k, \ldots, x_N$, where the value $x_0 = 0$ and $x_N = 1$ are the endpoints of the original interval. Since the points are equally spaced, starting at $x_0 = 0$, the coordinate x_k is just k steps of size $1/N$ along the x axis, i.e. $x_k = k(1/N) = k/N$. In the right panel of Figure 2.3, some of these coordinates have been labeled. For clarity, we show only the first few points, together with a representative pair x_{k-1} and x_k inside the region.

Let us look more carefully at one of the rectangles. Suppose we look at the rectangle labeled k. Such a representative k-th rectangle is shown shaded in Figures 2.3. The height of this rectangle is determined by the value of the function, since one corner of the rectangle is "glued" to the curve. The choice shown in Figure 2.3 is to affix the right corner of each

rectangle (k)	right x coord (x_k)	height $f(x_k)$	area a_k
1	$(1/N)$	$(1/N)^2$	$(1/N)^2\Delta x$
2	$(2/N)$	$(2/N)^2$	$(2/N)^2\Delta x$
3	$(3/N)$	$(3/N)^2$	$(3/N)^2\Delta x$
$\vdots$			
k	(k/N)	$(k/N)^2$	$(k/N)^2\Delta x$
$\vdots$			
N	$(N/N) = 1$	$(N/N)^2 = 1$	$(1)\Delta x$

Table 2.1. *The label, position, height, and area a_k of each rectangular strip is shown above. Each rectangle has the same base width, $\Delta x = 1/N$. We approximate the area under the curve $y = f(x) = x^2$ by the sum of the values in the last column, i.e. the total area of the rectangles.*

rectangle on the curve. This implies that the height of the k-th rectangle is obtained from substituting x_k into the function, i.e. height = $f(x_k)$. The base of every rectangle is the same, i.e. base = $\Delta x = 1/N$. This means that the area of the k-th rectangle, shown shaded, is

$$a_k = \text{height} \times \text{base} = f(x_k)\Delta x$$

We now use three facts:

$$f(x_k) = x_k^2, \quad \Delta x = \frac{1}{N}, \quad x_k = \frac{k}{N}.$$

Then the area of the k'th rectangle is

$$a_k = \text{height} \times \text{base} = f(x_k)\Delta x = \underbrace{\left(\frac{k}{N}\right)^2}_{f(x_k)}\underbrace{\left(\frac{1}{N}\right)}_{\Delta x}.$$

A list of rectangles, and their properties are shown in Table 2.1. This may help the reader to see the pattern that emerges in the summation. (In general this table is not needed in our work, and it is presented for this example only, to help visualize how heights of rectangles behave.) The total area of all rectangular strips (a sum of the values in the right column of Table 2.1) is

$$A_{N \text{ strips}} = \sum_{k=1}^{N} a_k = \sum_{k=1}^{N} f(x_k)\Delta x = \sum_{k=1}^{N} \left(\frac{k}{N}\right)^2 \left(\frac{1}{N}\right). \tag{2.1}$$

The expressions shown in Eqn. (2.1) is a Riemann sum. A recurring theme underlying integral calculus is the relationship between Riemann sums and definite integrals, a concept introduced later on in this chapter.

We now rewrite this sum in a more convenient form so that summation formulae developed in Chapter 1 can be used. In this sum, only the quantity k changes from term to term. All other quantities are common factors, so that

$$A_{N \text{ strips}} = \left(\frac{1}{N^3}\right)\sum_{k=1}^{N} k^2.$$

The formula (1.2) for the sum of square integers can be applied to the summation, resulting in

$$A_{N \text{ strips}} = \left(\frac{1}{N^3}\right)\frac{N(N+1)(2N+1)}{6} = \frac{(N+1)(2N+1)}{6N^2}. \tag{2.2}$$

In the box below, we use Eqn. (2.2) to compute that approximate area for values of N shown in the first three panels of Fig 2.2. Note that these are comparable to the values we obtained "numerically" in Section 2.2.1. (We plug in the value of N into (2.2) and use a calculator to obtain the results below.)

If $N = 10$ **strips** (Figure 2.2a), the width of each strip is 0.1 unit. According to equation 2.2, the area of the 10 strips (shown in red) is

$$A_{10 \text{ strips}} = \frac{(10+1)(2 \cdot 10+1)}{6 \cdot 10^2} = 0.385.$$

If $N = 20$ **strips** (Figure 2.2b), $\Delta x = 1/20 = 0.05$, and

$$A_{20 \text{ strips}} = \frac{(20+1)(2 \cdot 20+1)}{6 \cdot 20^2} = 0.35875.$$

If $N = 40$ **strips** (Figure 2.2c), $\Delta x = 1/40 = 0.025$ and

$$A_{40 \text{ strips}} = \frac{(40+1)(2 \cdot 40+1)}{6 \cdot 40^2} = 0.3459375.$$

We will define **the true area** under the graph of the function $y = f(x)$ over the given interval to be:

$$A = \lim_{N\to\infty} A_{N \text{ strips}}.$$

This means that the true area is obtained by letting the number of rectangular strips, N, get very large, (while the width of each one, $\Delta x = 1/N$ gets very small.)

In the example discussed in this section, the true area is found by taking the limit as N gets large in equation (2.2), i.e.,

$$A = \lim_{N\to\infty}\left(\frac{1}{N^2}\right)\frac{(N+1)(2N+1)}{6} = \frac{1}{6}\lim_{N\to\infty}\frac{(N+1)(2N+1)}{N^2}.$$

To evaluate this limit, note that when N gets very large, we can use the approximations, $(N+1) \approx N$ and $(2N+1) \approx 2N$ so that (simplifying and cancelling common factors)

$$\lim_{N\to\infty}\frac{(N+1)(2N+1)}{N^2} = \lim_{N\to\infty}\frac{(N)}{N}\frac{(2N)}{N} = 2.$$

The result is:

$$A = \frac{1}{6}(2) = \frac{1}{3} \approx 0.333. \tag{2.3}$$

Thus, the true area of the region (Figure 2.2d) is is 1/3 units^2.

2.2.3 Comments

Many student who have had calculus before in highschool, ask "why do we bother with such tedious calculations, when we could just use integration?". Indeed, our development of Riemann sums foreshadows and anticipates the idea of a definite integral, and in short order, some powerful techniques will help to shortcut such technical calculations. There are two reasons why we linger on Riemann sums. First, in order to understand integration adequately, we must understand the underlying "technology" and concepts; this proves vital in understanding how to use the methods, and when things can go wrong. It also helps to understand what integrals represent in applications that occur later on. Second, even though we will shortly have better tools for analytical calculations, the ideas of setting up area approximations using rectangular strips is very similar to the way that the spreadsheet computations are designed. (However, the summation is handled automatically using the spreadsheet, and no "formulae" are needed.) In Section 2.2.1, we gave only few details of the steps involved. The student will find that understanding the ideas of Section 2.2.2 will go hand-in-hand with understanding the numerical approach of Section 2.2.1.

The ideas outlined above can be applied to more complicated situations. In the next section we consider a practical problem in which a similar calculation is carried out.

2.3 The area of a leaf

Leaves act as solar energy collectors for plants. Hence, their surface area is an important property. In this section we use our techniques to determine the area of a rhododendron leaf, shown in Figure 2.4. For simplicity of treatment, we will first consider a function designed to mimic the shape of the leaf in a simple system of units: we will scale distances by the length of the leaf, so that its profile is contained in the interval $0 \leq x \leq 1$. We later ask how to modify this treatment to describe similarly curved leaves of arbitrary length and width, and leaves that are less symmetric. As shown in Figure 2.4, a simple parabola, of the form

$$y = f(x) = x(1 - x),$$

provides a convenient approximation to the top edge of the leaf. To check that this is the case, we observe that at $x = 0$ and $x = 1$, the curve intersects the x axis. At $0 < x < 1$, the curve is above the axis. Thus, the area between this curve and the x axis, is one half of the leaf area.

We set up the computation of approximating rectangular strips as before, by subdividing the interval of interest into N rectangular strips. We can set up the calculation systematically, as follows:

$$\text{length of interval} = 1 - 0 = 1$$

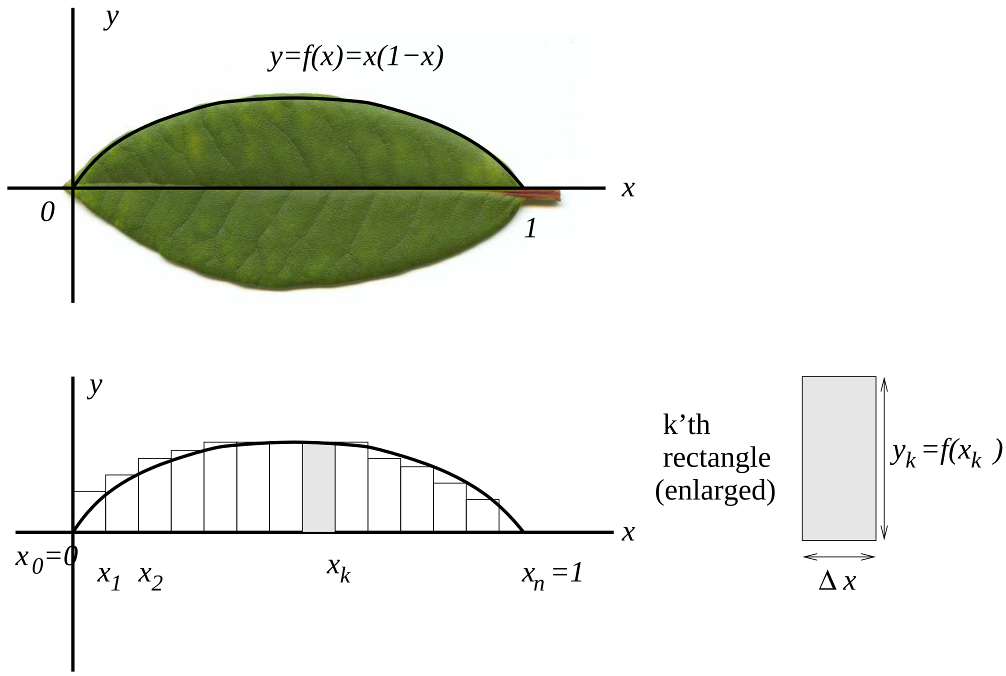

Figure 2.4. *In this figure we show how the area of a leaf can be approximated by rectangular strips.*

number of segments, N

width of rectangular strips, $\Delta x = \dfrac{1}{N}$

the k'th x value, $x_k = k\dfrac{1}{N} = \dfrac{k}{N}$

height of k'th rectangular strip, $f(x_k) = x_k(1 - x_k)$

The representative k'th rectangle is shown shaded in Figure 2.4: Its area is

$$a_k = \text{base} \times \text{height} = \Delta x \cdot f(x_k) = \underbrace{\left(\frac{1}{N}\right)}_{\Delta x} \cdot \underbrace{\left(\frac{k}{N}(1 - \frac{k}{N})\right)}_{f(x_k)}.$$

The total area of these rectangular strips is:

$$A_{N \text{ strips}} = \sum_{k=1}^{N} a_k = \sum_{k=1}^{N} \Delta x \cdot f(x_k) = \sum_{k=1}^{N} \left(\frac{1}{N}\right) \cdot \left(\frac{k}{N}(1 - \frac{k}{N})\right).$$

Simplifying the result (so we can use summation formulae) leads to:

$$A_{N \text{ strips}} = \left(\frac{1}{N}\right)\sum_{k=1}^{N}\left(\frac{k}{N}(1-\frac{k}{N})\right) = \left(\frac{1}{N^2}\right)\sum_{k=1}^{N}k - \left(\frac{1}{N^3}\right)\sum_{k=1}^{N}k^2.$$

Using the summation formulae (1.1) and(1.2) from Chapter 1 results in:

$$A_{N \text{ strips}} = \left(\frac{1}{N^2}\right)\left(\frac{N(N+1)}{2}\right) - \left(\frac{1}{N^3}\right)\left(\frac{(2N+1)N(N+1)}{6}\right).$$

Simplifying, and regrouping terms, we get

$$A_{N \text{ strips}} = \frac{1}{2}\left(\frac{(N+1)}{N}\right) - \frac{1}{6}\left(\frac{(2N+1)(N+1)}{N^2}\right).$$

This is the area for a finite number, N, of rectangular strips. As before, the **true area** is obtained as the limit as N goes to infinity, i.e. $A = \lim_{N\to\infty} A_{N \text{ strips}}$. We obtain:

$$A = \lim_{N\to\infty}\frac{1}{2}\left(\frac{(N+1)}{N}\right) - \lim_{N\to\infty}\frac{1}{6}\left(\frac{(2N+1)(N+1)}{N^2}\right) = \frac{1}{2} - \frac{1}{6}\cdot 2 = \frac{1}{6}.$$

Taking the limit leads to

$$A = \frac{1}{2} - \frac{1}{6}\cdot 2 = \frac{1}{2} - \frac{1}{3} = \frac{1}{6}.$$

Thus the area of the entire leaf (twice this area) is 1/3.

Remark:

The function in this example can be written as $y = x - x^2$. For part of this expression, we have seen a similar calculation in Section 2.2. This example illustrates an important property of sums, namely the fact that we can rearrange the terms into simpler expressions that can be summed individually.

In the homework problems accompanying this chapter, we investigate how to describe leaves with arbitrary lengths and widths, as well as leaves with shapes that are tapered, broad, or less symmetric than the current example.

2.4 Area under an exponential curve

In the precious examples, we considered areas under curves described by a simple quadratic functions. Each of these led to calculations in which sums of integers or square integers appeared. Here we demonstrate an example in which a geometric sum will be used. Recall that we derived Eqn. (1.4) in Chapter 1, for a finite geometric sum.

We will find the area under the graph of the function $y = f(x) = e^{2x}$ over the interval between $x = 0$ and $x = 2$. In evaluating a limit in this example, we will also use the fact that the exponential function has a linear approximation as follows:

$$e^z \approx 1 + z$$

(See Linear Approximations in an earlier calculus course.)

As before, we subdivide the interval into N pieces, each of width $2/N$. Proceeding systematically as before, we write

length of interval $= 2 - 0 = 2$

number of segments $= N$

width of rectangular strips, $\Delta x = \dfrac{2}{N}$

the k'th x value, $x_k = k\dfrac{2}{N} = \dfrac{2k}{N}$

height of k'th rectangular strip, $f(x_k) = e^{x_k} = e^{2(2k/N)} = e^{4k/N}$

We observe that the length of the interval (here 2) has affected the details of the calculation. As before, the area of the k'th rectangle is

$$a_k = \text{base} \times \text{height} = \Delta x \times f(x_k) = \left(\frac{2}{N}\right) e^{4k/N},$$

and the total area of all the rectangles is

$$A_{N \text{ strips}} = \left(\frac{2}{N}\right) \sum_{k=1}^{N} e^{4k/N} = \left(\frac{2}{N}\right) \sum_{k=1}^{N} r^k = \left(\frac{2}{N}\right) \left(\sum_{k=0}^{N} r^k - r^0\right),$$

where $r = e^{4/N}$. This is a finite geometric series. Because the series starts with $k = 1$ and not with $k = 0$, the sum is

$$A_{N \text{ strips}} = \left(\frac{2}{N}\right) \left[\frac{(1 - r^{N+1})}{(1 - r)} - 1\right].$$

After some simplification and using $r = e^{4/N}$, we find that

$$A_{N \text{ strips}} = \left(\frac{2}{N}\right) e^{4/N} \frac{1 - e^4}{1 - e^{4/N}} = 2\frac{1 - e^4}{N(e^{-4/N} - 1)}.$$

We need to determine what happens when N gets very large. We can use the linear approximation

$$e^{-4/N} \approx 1 - 4/N$$

to evaluate the limit of the term in the denominator, and we find that

$$A = \lim_{N\to\infty} 2\frac{1 - e^4}{N(e^{-4/N} - 1)} = \lim_{N\to\infty} 2\frac{1 - e^4}{-N(1 + 4/N - 1)} = 2\frac{e^4 - 1}{4} \approx 26.799.$$

2.5 Extensions and other examples

More general interval

To calculate the area under the curve $y = f(x) = x^2$ over the interval $2 \leq x \leq 5$ using N rectangles, the width of each one would be $\Delta x = (5 - 2)/N = 3/N$, (i.e., length of interval divided by N). Since the interval starts at $x_0 = 2$, and increments in units of $(3/N)$, the k'th coordinate is $x_k = 2 + k(3/N) = 2 + (3k/N)$. The area of the k'th rectangle is then $A_K = f(x_k) \times \Delta x = [(2 + (3k/N))^2](3/N)$, and this is to be summed over k. A similar algebraic simplification, summation formulae, and limit is needed to calculate the true area.

Other examples

In the Appendix 11.2 we discuss a number of other examples with several modifications: First, in Appendix 11.2.1, we show how to set up a Riemann sum for a more complicated quadratic function on a general interval, $a \le x \le b$.

Second, we show how Riemann sums can be set up for left, rather than right endpoint approximations. The results are entirely analogous.

2.6 The definite integral

We now introduce a central concept that will form an important theme in this course, that of the definite integral. We begin by defining a new piece of notation relevant to the topic in this chapter, namely the area associated with the graph of a function. For a function $y =$

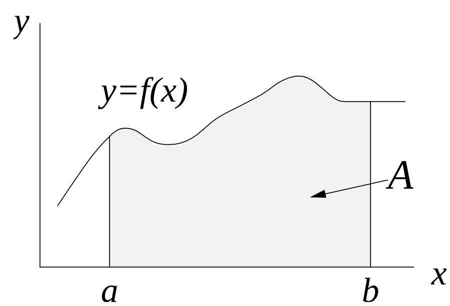

Figure 2.5. *The shaded area A corresponds to the definite integral I of the function $f(x)$ over the interval $a \le x \le b$.*

$f(x) > 0$ that is bounded and continuous[11] on an interval $[a, b]$ (also written $a \le x \le b$), we define the *definite integral*,

$$I = \int_a^b f(x)\,dx \tag{2.4}$$

to be the area A of the region under the graph of the function between the endpoints a and b. See Figure 2.5.

2.6.1 Remarks

1. The definite integral is a number.

[11] A function is said to be bounded if its graph stays between some pair of horizontal lines. It is continuous if there are no “breaks” in its graph.

2. The value of the definite integral depends on the function, and on the two end points of the interval.

3. From previous remarks, we have a procedure to calculate the value of the definite integral by dissecting the region into rectangular strips, summing up the total area of the strips, and taking a limit as N, the number of strips gets large. (The calculation may be non-trivial, and might involve sums that we have not discussed in our simple examples so far, but in principle the procedure is well-defined.)

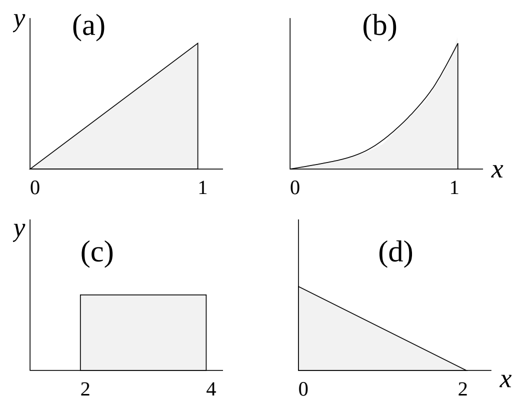

Figure 2.6. *Examples (1-4) relate areas shown above to* definite integrals*.*

2.6.2 Examples

We have calculated the areas of regions bounded by particularly simple functions. To practice notation, we write down the corresponding definite integral in each case. Note that in many of the examples below, we need no elaborate calculations, but merely use previously known or recently derived results, to familiarize the reader with the new notation just defined.

Example (1)

The area under the function $y = f(x) = x$ over the interval $0 \leq x \leq 1$ is triangular, with base and height 1. The area of this triangle is thus $A = (1/2)$base$\times$ height= 0.5 (Figure 2.6a). Hence,

$$\int_0^1 x\,dx = 0.5.$$

Example (2)

In Section 2.2, we also computed the area under the function $y = f(x) = x^2$ on the interval $0 \le x \le 1$ and found its area to be 1/3 (See Eqn. (2.3) and Fig. 2.6(b)). Thus

$$\int_0^1 x^2\, dx = 1/3 \simeq 0.333.$$

Example (3)

A constant function of the form $y = 1$ over an interval $2\ lex \le 4$ would produce a rectangular region in the plane, with base (4-2)=2 and height 1 (Figure 2.6(c)). Thus

$$\int_2^4 1\, dx = 2.$$

Example (4)

The function $y = f(x) = 1 - x/2$ (Figure 2.6(d)) forms a triangular region with base 2 and height 1, thus

$$\int_0^2 (1 - x/2)\, dx = 1.$$

2.7 The area as a function

In Chapter 3, we will elaborate on the idea of the definite integral and arrive at some very important connection between differential and integral calculus. Before doing so, we have to extend the idea of the definite integral somewhat, and thereby define a new function, $A(x)$.

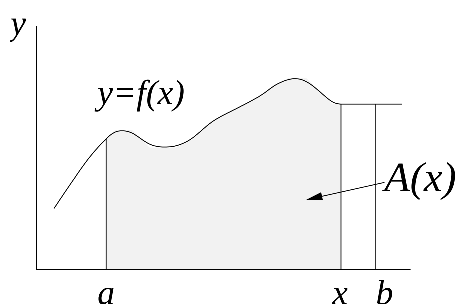

Figure 2.7. *We define a new function $A(x)$ to be the area associated with the graph of some function $y = f(x)$ from the fixed endpoint a up to the endpoint x, where $a \le x \le b$.*

We will investigate how the area under the graph of a function changes as one of the endpoints of the interval moves. We can think of this as a function that gradually changes (i.e. the area accumulates) as we sweep across the interval (a, b) from left to right in Figure 2.1. The function $A(x)$ represents the area of the region shown in Figure 2.7.

Extending our definition of the definite integral, we might be tempted to use the notation

$$A(x) = \int_a^x f(x)\, dx.$$

However, there is a slight problem with this notation: the symbol x is used in slightly confusing ways, both as the argument of the function and as the variable endpoint of the interval. To avoid possible confusion, we will prefer the notation

$$A(x) = \int_a^x f(s)\, ds.$$

(or some symbol other than s used as a placeholder instead of x.)

An analogue already seen is the sum

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

where N denotes the "end" of the sum, and k keeps track of where we are in the process of summation. The symbol s, sometimes called a "dummy variable" is analogous to the summation symbol k.

In the upcoming Chapter 3, we will investigate properties of this new "area function" $A(x)$ defined above. This will lead us to the *Fundamental Theorem of Calculus*, and will provide new and powerful tools to replace the dreary summations that we had to perform in much of Chapter 2. Indeed, we are about to discover the amazing connection between a function, the area $A(x)$ under its curve, and the derivative of $A(x)$.

2.8 Summary

In this chapter, we showed how to calculate the area of a region in the plane that is bounded by the x axis, two lines of the form $x = a$ and $x = b$, and the graph of a positive function $y = f(x)$. We also introduced the terminology "definite integral" (Section 2.6) and the notation (2.4) to represent that area.

One of our main efforts here focused on how to actually compute that area by the following set of steps:

- Subdivide the interval $[a, b]$ into smaller intervals (width Δx).
- Construct rectangles whose heights approximate the height of the function above the given interval.
- Add up the areas of these approximating rectangles. (Here we often used summation formulae from Chapter 1.) The resulting expression, such as Eqn. (2.1), for example, was denoted a Riemann sum.
- Find out what happens to this total area in the limit when the width Δx goes to zero (or, in other words, when the number of rectangles N goes to infinity).

We showed both the analytic approach, using Riemann sums and summation formulae to find areas, as well as numerical approximations using a spreadsheet tool to arrive at similar results. We then used a variety of examples to illustrate the concepts and arrive at computed areas.

As a final important point, we noted that the area “under the graph of a function” can itself be considered a function. This idea will emerge as particularly important and will lead us to the key concept linking the geometric concept of areas with the analytic properties of antiderivatives. We shall see this link in the Fundamental Theorem of Calculus, in Chapter 3.

Chapter 3

The Fundamental Theorem of Calculus

In this chapter we will formulate one of the most important results of calculus, the Fundamental Theorem. This result will link together the notions of an integral and a derivative. Using this result will allow us to replace the technical calculations of Chapter 2 by much simpler procedures involving antiderivatives of a function.

3.1 The definite integral

In Chapter 2, we defined the definite integral, I, of a function $f(x) > 0$ on an interval $[a, b]$ as the area under the graph of the function over the given interval $a \leq x \leq b$. We used the notation

$$I = \int_a^b f(x)dx$$

to represent that quantity. We also set up a technique for computing areas: the procedure for calculating the value of I is to write down a sum of areas of rectangular strips and to compute a limit as the number of strips increases:

$$I = \int_a^b f(x)dx = \lim_{N\to\infty} \sum_{k=1}^{N} f(x_k)\Delta x, \tag{3.1}$$

where N is the number of strips used to approximate the region, k is an index associated with the k'th strip, and $\Delta x = x_{k+1} - x_k$ is the width of the rectangle. As the number of strips increases ($N \to \infty$), and their width decreases ($\Delta x \to 0$), the sum becomes a better and better approximation of the true area, and hence, of the definite integral, I. Example of such calculations (tedious as they were) formed the main theme of Chapter 2 .

We can generalize the definite integral to include functions that are not strictly positive, as shown in Figure 3.1. To do so, note what happens as we incorporate strips corresponding to regions of the graph below the x axis: These are associated with negative values of the function, so that the quantity $f(x_k)\Delta x$ in the above sum would be negative for each rectangle in the "negative" portions of the function. This means that regions of the graph below the x axis will contribute negatively to the net value of I.

If we refer to A_1 as the area corresponding to regions of the graph of $f(x)$ above the x axis, and A_2 as the total area of regions of the graph under the x axis, then we will find that the value of the definite integral I shown above will be

$$I = A_1 - A_2.$$

Thus the notion of "area under the graph of a function" must be interpreted a little carefully when the function dips below the axis.

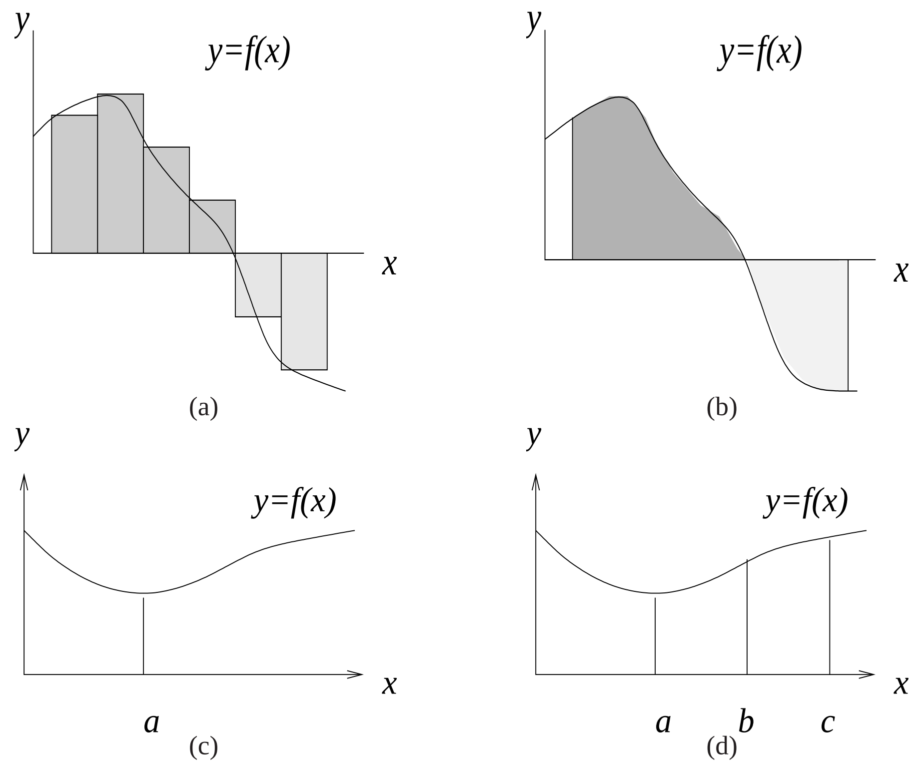

Figure 3.1. *(a) If $f(x)$ is negative in some regions, there are terms in the sum (3.1) that carry negative signs: this happens for all rectangles in parts of the graph that dip below the x axis. (b) This means that the definite integral $I = \int_a^b f(x)dx$ will correspond to the difference of two areas, $A_1 - A_2$ where A_1 is the total area (dark) of positive regions minus the total area (light) of negative portions of the graph. Properties of the definite integral: (c) illustrates Property 1. (d) illustrates Property 2.*

3.2 Properties of the definite integral

The following properties of a definite integral stem from its definition, and the procedure for calculating it discussed so far. For example, the fact that summation satisfies the distributive property means that an integral will satisfy the same the same property. We illustrate some of these in Fig 3.1.

1. $\int_a^a f(x)dx = 0$,
2. $\int_a^c f(x)dx = \int_a^b f(x)dx + \int_b^c f(x)dx$,
3. $\int_a^b Cf(x)dx = C\int_a^b f(x)dx$,
4. $\int_a^b (f(x)+g(x))dx = \int_a^b f(x) + \int_a^b g(x)dx$,
5. $\int_a^b f(x)dx = -\int_b^a f(x)dx$.

Property 1 states that the "area" of a region with no width is zero. Property 2 shows how a region can be broken up into two pieces whose total area is just the sum of the individual areas. Properties 3 and 4 reflect the fact that the integral is actually just a sum, and so satisfies properties of simple addition. Property 5 is obtained by noting that if we perform the summation "in the opposite direction", then we must replace the previous "rectangle width" given by $\Delta x = x_{k+1} - x_k$ by the new "width" which is of opposite sign: $x_k - x_{k+1}$. This accounts for the sign change shown in Property 5.

3.3 The area as a function

In Chapter 2, we investigated how the area under the graph of a function changes as one of the endpoints of the interval moves. We defined a function that represents the area under the graph of a function f, from some fixed starting point, a to an endpoint x.

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

This endpoint is considered as a variable[12], i.e. we will be interested in the way that this area changes as the endpoint varies (Figure 3.2(a)). We will now investigate the interesting connection between $A(x)$ and the original function, $f(x)$.

We would like to study how $A(x)$ changes as x is increased ever so slightly. Let $\Delta x = h$ represent some (very small) increment in x. (*Caution: do not confuse h with height here. It is actually a step size along the x axis*.) Then, according to our definition,

$$A(x+h) = \int_a^{x+h} f(t)\,dt.$$

[12]Recall that the "dummy variable" t inside the integral is just a "place holder", and is used to avoid confusion with the endpoint of the integral (x in this case). Also note that the value of $A(x)$ does not depend in any way on t, so any letter or symbol in its place would do just as well.

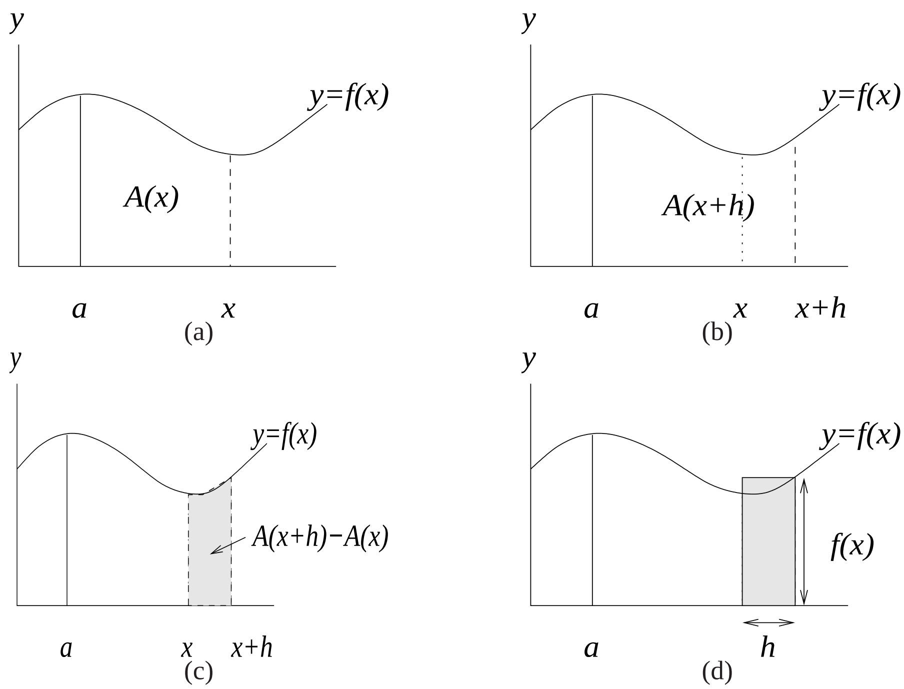

Figure 3.2. *When the right endpoint of the interval moves by a distance h, the area of the region increases from $A(x)$ to $A(x+h)$. This leads to the important Fundamental Theorem of Calculus, given in Eqn. (3.2).*

In Figure 3.2(a)(b), we illustrate the areas represented by $A(x)$ and by $A(x+h)$, respectively. The difference between the two areas is a thin sliver (shown in Figure 3.2(c)) that looks much like a rectangular strip (Figure 3.2(d)). (Indeed, if h is small, then the approximation of this sliver by a rectangle will be good.) The height of this sliver is specified by the function f evaluated at the point x, i.e. by $f(x)$, so that the area of the sliver is approximately $f(x) \cdot h$. Thus,

$$A(x+h) - A(x) \approx f(x)h$$

or

$$\frac{A(x+h) - A(x)}{h} \approx f(x).$$

As h gets small, i.e. $h \to 0$, we get a better and better approximation, so that, in the limit,

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

The ratio above should be recognizable. It is simply the derivative of the area function, i.e.

$$f(x) = \frac{dA}{dx} = \lim_{h \to 0} \frac{A(x+h) - A(x)}{h}. \tag{3.2}$$

We have just given a simple argument in support of an important result, called the *Fundamental Theorem of Calculus*, which is restated below..

3.4 The Fundamental Theorem of Calculus

3.4.1 Fundamental theorem of calculus: Part I

Let $f(x)$ be a bounded and continuous function on an interval $[a, b]$. Let

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

Then for $a < x < b$,

$$\frac{dA}{dx} = f(x).$$

In other words, this result says that $A(x)$ is an "antiderivative" of the original function, $f(x)$[13].

Proof

See above argument. and Figure 3.2.

3.4.2 Example: an antiderivative

Recall the connection between functions and their derivatives. Consider the following two functions:

$$g_1(x) = \frac{x^2}{2}, \quad g_2 = \frac{x^2}{2} + 1.$$

Clearly, both functions have the same derivative:

$$g_1'(x) = g_2'(x) = x.$$

We would say that $x^2/2$ is an "antiderivative" of x and that $(x^2/2) + 1$ is also an "antiderivative" of x. In fact, *any* function of the form

$$g(x) = \frac{x^2}{2} + C \quad \text{where } C \text{ is any constant}$$

is also an "antiderivative" of x.

This example illustrates that adding a constant to a given function will not affect the value of its derivative, or, stated another way, antiderivatives of a given function are defined only up to some constant. We will use this fact shortly: if $A(x)$ and $F(x)$ are both antiderivatives of some function $f(x)$, then $A(x) = F(x) + C$.

[13] We often write "antiderivative", with no hyphen.

3.4.3 Fundamental theorem of calculus: Part II

Let $f(x)$ be a continuous function on $[a, b]$. Suppose $F(x)$ is *any* antiderivative of $f(x)$. Then for $a \le x \le b$,

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

Proof

From comments above, we know that a function $f(x)$ could have many different antiderivatives that differ from one another by some additive constant. We are told that $F(x)$ is an antiderivative of $f(x)$. But from Part I of the Fundamental Theorem, we know that $A(x)$ is also an antiderivative of $f(x)$. It follows that

$$A(x) = \int_a^x f(t)\, dt = F(x) + C, \quad \text{where } C \text{ is some constant.} \tag{3.3}$$

However, by property 1 of definite integrals,

$$A(a) = \int_a^a f(t) = F(a) + C = 0.$$

Thus,

$$C = -F(a).$$

Replacing C by $-F(a)$ in equation 3.3 leads to the desired result. Thus

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

Remark 1: Implications

This theorem has tremendous implications, because it allows us to use a powerful new tool in determining areas under curves. Instead of the drudgery of summations in order to compute areas, we will be able to use a shortcut: find an antiderivative, evaluate it at the two endpoints a, b of the interval of interest, and subtract the results to get the area. In the case of elementary functions, this will be very easy and convenient.

Remark 2: Notation

We will often use the notation

$$F(t)|_a^x = F(x) - F(a)$$

to denote the difference in the values of a function at two endpoints.

3.5 Review of derivatives (and antiderivatives)

By remarks above, we see that integration is related to “anti-differentiation”. This motivates a review of derivatives of common functions. Table 3.1 lists functions $f(x)$ and their derivatives $f'(x)$ (in the first two columns) and functions $f(x)$ and their antiderivatives $F(x)$ in the subsequent two columns. These will prove very helpful in our calculations of basic integrals.

function	derivative		function	antiderivative
$f(x)$	$f'(x)$		$f(x)$	$F(x)$
Cx	C		C	Cx
x^n	nx^{n-1}		x^m	$\frac{x^{m+1}}{m+1}$
$\sin(ax)$	$a\cos(ax)$		$\cos(bx)$	$(1/b)\sin(bx)$
$\cos(ax)$	$-a\sin(ax)$		$\sin(bx)$	$-(1/b)\cos(bx)$
$\tan(ax)$	$a\sec^2(ax)$		$\sec^2(bx)$	$(1/b)\tan(bx)$
e^{kx}	ke^{kx}		e^{kx}	e^{kx}/k
$\ln(x)$	$\dfrac{1}{x}$		$\dfrac{1}{x}$	$\ln(x)$
$\arctan(x)$	$\dfrac{1}{1+x^2}$		$\dfrac{1}{1+x^2}$	$\arctan(x)$
$\arcsin(x)$	$\dfrac{1}{\sqrt{1-x^2}}$		$\dfrac{1}{\sqrt{1-x^2}}$	$\arcsin(x)$

Table 3.1. *Common functions and their derivatives (on the left two columns) also result in corresponding relationships between functions and their antiderivatives (right two columns). In this table, we assume that* $m \neq -1, b \neq 0, k \neq 0$. *Also, when using* $\ln(x)$ *as antiderivative for* $1/x$, *we assume that* $x > 0$.

As an example, consider the polynomial

$$p(x) = a_0 + a_1x + a_2x^2 + a_3x^3 + \ldots$$

This polynomial could have many other terms (or even an infinite number of such terms, as we discuss much later, in Chapter 10). Its antiderivative can be found easily using the “power rule” together with the properties of addition of terms. Indeed, the antiderivative is

$$F(x) = C + a_0x + \frac{a_1}{2}x^2 + \frac{a_2}{3}x^3 + \frac{a_3}{4}x^4 + \ldots$$

This can be checked easily by differentiation[14].

3.6 Examples: Computing areas with the Fundamental Theorem of Calculus

3.6.1 Example 1: The area under a polynomial

Consider the polynomial

$$p(x) = 1 + x + x^2 + x^3.$$

(Here we have taken the first few terms from the example of the last section with coefficients all set to 1.) Then, computing

$$I = \int_0^1 p(x)\, dx$$

leads to

$$I = \int_0^1 (1 + x + x^2 + x^3)\, dx = (x + \frac{1}{2}x^2 + \frac{1}{3}x^3 + \frac{1}{4}x^4)\Big|_0^1 = 1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} \approx 2.083.$$

3.6.2 Example 2: Simple areas

Determine the values of the following definite integrals by finding antiderivatives and using the Fundamental Theorem of Calculus:

1. $I = \displaystyle\int_0^1 x^2\, dx$,
2. $I = \displaystyle\int_{-1}^1 (1 - x^2)\, dx$,
3. $I = \displaystyle\int_{-1}^1 e^{-2x}\, dx$,
4. $I = \displaystyle\int_0^{\pi} \sin\left(\frac{x}{2}\right)\, dx$,

Solutions

1. An antiderivative of $f(x) = x^2$ is $F(x) = (x^3/3)$, thus

$$I = \int_0^1 x^2 dx = F(x)\Big|_0^1 = (1/3)(x^3)\Big|_0^1 = \frac{1}{3}(1^3 - 0) = \frac{1}{3}.$$

[14] In fact, it is very good practice to perform such checks.

2. An antiderivative of $f(x) = (1 - x^2)$ is $F(x) = x - (x^3/3)$, thus
$$I = \int_{-1}^{1} (1-x^2)\, dx = F(x)\Big|_{-1}^{1} = \big(x - (x^3/3)\big)\Big|_{-1}^{1} = \big(1 - (1^3/3)\big) - \big((-1) - ((-1)^3/3)\big) = 4/3$$
See comment below for a simpler way to compute this integral.
3. An antiderivative of e^{-2x} is $F(x) = (-1/2)e^{-2x}$. Thus,
$$I = \int_{-1}^{1} e^{-2x}\, dx = F(x)\Big|_{-1}^{1} = (-1/2)(e^{-2x})\Big|_{-1}^{1} = (-1/2)(e^{-2} - e^{2}).$$
4. An antiderivative of $\sin(x/2)$ is $F(x) = -\cos(x/2)/(1/2) = -2\cos(x/2)$. Thus
$$I = \int_{0}^{\pi} \sin\left(\frac{x}{2}\right)\, dx = -2\cos(x/2)\Big|_{0}^{\pi} - 2(\cos(\pi/2) - \cos(0)) = -2(0-1) = 2.$$

Comment: The evaluation of Integral 2. in the examples above is tricky only in that signs can easily get garbled when we plug in the endpoint at -1. However, we can simplify our work by noting the symmetry of the function $f(x) = 1 - x^2$ on the given interval. As shown in Fig 3.3, the areas to the right and to the left of $x = 0$ are the same for the interval $-1 \le x \le 1$. This stems directly from the fact that the function considered is **even**[15]. Thus, we can immediately write
$$I = \int_{-1}^{1} (1 - x^2)\, dx = 2\int_{0}^{1} (1 - x^2)\, dx = 2\left(x - (x^3/3)\right)\Big|_{0}^{1} = 2\left(1 - (1^3/3)\right) = 4/3$$
Note that this calculation is simpler since the endpoint at $x = 0$ is trivial to plug in.

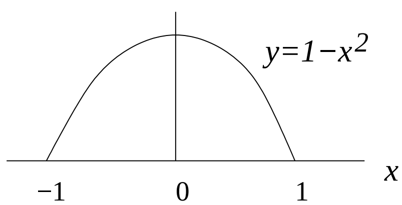

Figure 3.3. *We can exploit the symmetry of the function $f(x) = 1 - x^2$ in the second integral of Examples 3.6.2. We can integrate over $0 \le x \le 1$ and double the result.*

We state the general result we have obtained, which holds true for any function with even symmetry integrated on a symmetric interval about $x = 0$:

If $f(x)$ is an **even** function, then
$$\int_{-a}^{a} f(x)\, dx = 2\int_{-a}^{a} f(x)\, dx \qquad (3.4)$$

[15]Recall that a function $f(x)$ is **even** if $f(x) = f(-x)$ for all x. A function is **odd** if $f(x) = -f(-x)$.

3.6.3 Example 3: The area between two curves

The definite integral is an area of a somewhat special type of region, i.e., an axis, two vertical lines ($x = a$ and $x = b$) and the graph of a function. However, using additive (or subtractive) properties of areas, we can generalize to computing areas of other regions, including those bounded by the graphs of two functions.

(a) Find the area enclosed between the graphs of the functions $y = x^3$ and $y = x^{1/3}$ in the first quadrant.

(b) Find the area enclosed between the graphs of the functions $y = x^3$ and $y = x$ in the first quadrant.

(c) What is the relationship of these two areas? What is the relationship of the functions $y = x^3$ and $y = x^{1/3}$ that leads to this relationship between the two areas?

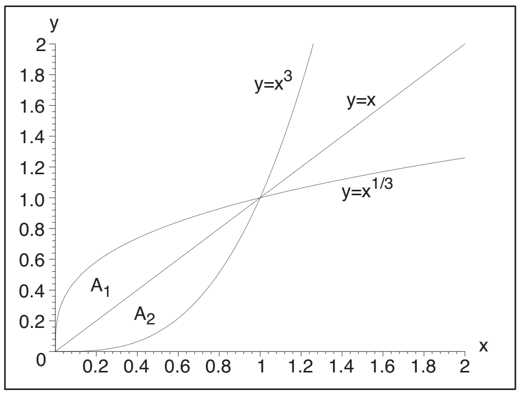

Figure 3.4. *In Example 3, we compute the areas A_1 and A_2 shown above.*

Solution

(a) The two curves, $y = x^3$ and $y = x^{1/3}$, intersect at $x = 0$ and at $x = 1$ in the first quadrant. Thus the interval that we will be concerned with is $0 < x < 1$. On this interval, $x^{1/3} > x^3$, so that the area we want to find can be expressed as:

$$A_1 = \int_0^1 \left(x^{1/3} - x^3\right) dx.$$

Thus,

$$A_1 = \left.\frac{x^{4/3}}{4/3}\right|_0^1 - \left.\frac{x^4}{4}\right|_0^1 = \frac{3}{4} - \frac{1}{4} = \frac{1}{2}.$$

(b) The two curves $y = x^3$ and $y = x$ also intersect at $x = 0$ and at $x = 1$ in the first quadrant, and on the interval $0 < x < 1$ we have $x > x^3$. The area can be represented as

$$A_2 = \int_0^1 \left(x - x^3\right) dx.$$

$$A_2 = \left.\frac{x^2}{2}\right|_0^1 - \left.\frac{x^4}{4}\right|_0^1 = \frac{1}{2} - \frac{1}{4} = \frac{1}{4}.$$

(c) The area calculated in (a) is twice the area calculated in (b). The reason for this is that $x^{1/3}$ is the inverse of the function x^3, which means geometrically that the graph of $x^{1/3}$ is the mirror image of the graph of x^3 reflected about the line $y = x$. Therefore, the area A_1 between $y = x^{1/3}$ and $y = x^3$ is twice as large as the area A_2 between $y = x$ and $y = x^3$ calculated in part (b): $A_1 = 2A_2$ (see Figure 3.4).

3.6.4 Example 4: Area of land

Find the exact area of the piece of land which is bounded by the y axis on the west, the x axis in the south, the lake described by the function $y = f(x) = 100 + (x/100)^2$ in the north and the line $x = 1000$ in the east.

Solution

The area is

$$A = \int_0^{1000} \left(100 + \left(\frac{x}{100}\right)^2\right) dx. = \int_0^{1000} \left(100 + \left(\frac{1}{10000}\right) x^2\right) dx.$$

Note that the multiplicative constant (1/10000) is not affected by integration. The result is

$$A = 100x\Big|_0^{1000} + \left.\frac{x^3}{3}\right|_0^{1000} \cdot \left(\frac{1}{10000}\right) = \frac{4}{3}10^5.$$

3.7 Qualitative ideas

In some cases, we are given a sketch of the graph of a function, $f(x)$, from which we would like to construct a sketch of the associated function $A(x)$. This sketching skill is illustrated in the figures shown in this section.

Suppose we are given a function as shown in the top left hand panel of Figure 3.5. We would like to assemble a sketch of

$$A(x) = \int_a^x f(t)dt$$

which corresponds to the area associated with the graph of the function f. As x moves from left to right, we show how the "area" accumulated along the graph gradually changes. (See $A(x)$ in bottom panels of Figure 3.5): We start with no area, at the point $x = a$ (since, by definition $A(a) = 0$) and gradually build up to some net positive amount, but then we encounter a portion of the graph of f below the x axis, and this subtracts from the amount accrued. (Hence the graph of $A(x)$ has a little peak that corresponds to the point at which $f = 0$.) Every time the function $f(x)$ crosses the x axis, we see that $A(x)$ has either a maximum or minimum value. This fits well with our idea of $A(x)$ as the antiderivative of $f(x)$: Places where $A(x)$ has a critical point coincide with places where $dA/dx = f(x) = 0$.

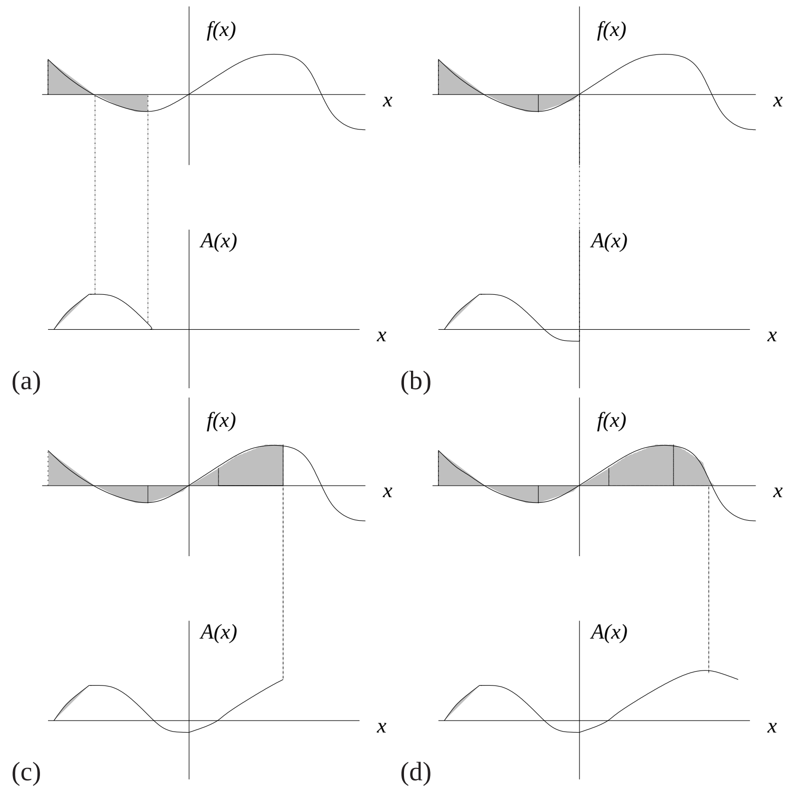

Figure 3.5. *Given a function $f(x)$, we here show how to sketch the corresponding "area function" $A(x)$. (The relationship is that $f(x)$ is the derivative of $A(x)$*

Sketching the function $A(x)$ is thus analogous to sketching a function $g(x)$ when we are given a sketch of its derivative $g'(x)$. Recall that this was one of the skills we built up in learning the connection between functions and their derivatives in a first semester calculus course.

Remarks

The following remarks may be helpful in gaining confidence with sketching the "area" function $A(x) = \int_a^x f(t)\, dt$, from the original function $f(x)$:

1. The endpoint of the interval, a on the x axis indicates the place at which $A(x) = 0$. This follows from Property 1 of the definite integral, i.e. from the fact that $A(a) = \int_a^a f(t)\, dt = 0$.

2. Whenever $f(x)$ is positive, $A(x)$ is an increasing function - this follows from the fact that the area continues to accumulate as we "sweep across" positive regions of $f(x)$.

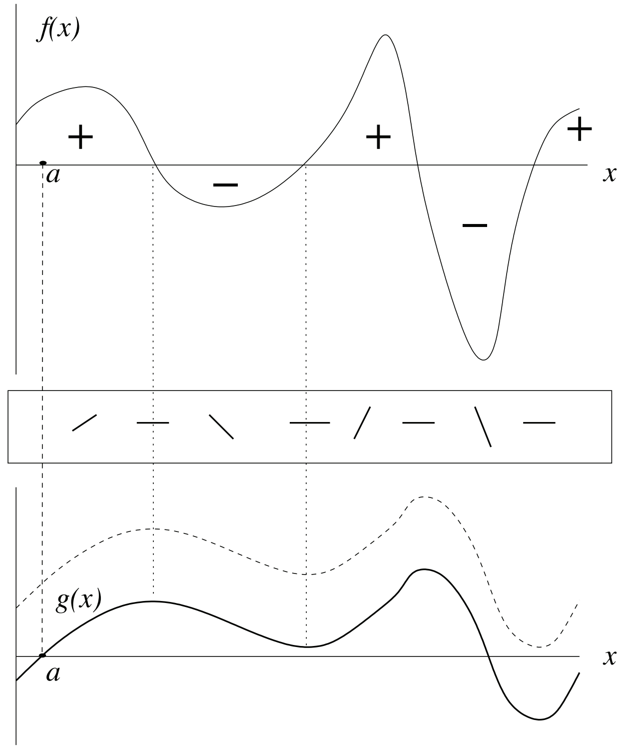

Figure 3.6. *Given a function $f(x)$ (top, solid line), we assemble a plot of the corresponding function $g(x) = \int_a^x f(t)dt$ (bottom, solid line). $g(x)$ is an antiderivative of $f(x)$. Whether $f(x)$ is positive (+) or negative (-) in portions of its graph, determines whether $g(x)$ is increasing or decreasing over the given intervals. Places where $f(x)$ changes sign correspond to maxima and minima of the function $g(x)$ (Two such places are indicated by dotted vertical lines). The box in the middle of the sketch shows configurations of tangent lines to $g(x)$ based on the sign of $f(x)$. Where $f(x) = 0$, those tangent lines are horizontal. The function $g(x)$ is drawn as a smooth curve whose direction is parallel to the tangent lines shown in the box. While the function $f(x)$ has many antiderivatives (e.g., dashed curve parallel to $g(x)$), only one of these satisfies $g(a) = 0$ as required by Property 1 of the definite integral. (See dashed vertical line at $x = a$). This determines the height of the desired function $g(x)$.*

3. Wherever $f(x)$, changes sign, the function $A(x)$ has a local minimum or maximum. This means that either the area stops increasing (if the transition is from positive to negative values of f), or else the area starts to increase (if f crosses from negative to positive values).

4. Since $dA/dx = f(x)$ by the Fundamental Theorem of Calculus, it follows that (tak-

ing a derivative of both sides) $d^2A/dx^2 = f'(x)$. Thus, when $f(x)$ has a local maximum or minimum, (i.e. $f'(x) = 0$), it follows that $A''(x) = 0$. This means that at such points, the function $A(x)$ would have an inflection point.

Given a function $f(x)$, Figure 3.6 shows in detail how to sketch the corresponding function

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

3.7.1 Example: sketching $A(x)$

Consider the $f(x)$ whose graph is shown in the top part of Figure 3.7. Sketch the corresponding function $g(x) = \int_a^x f(x)dx$.

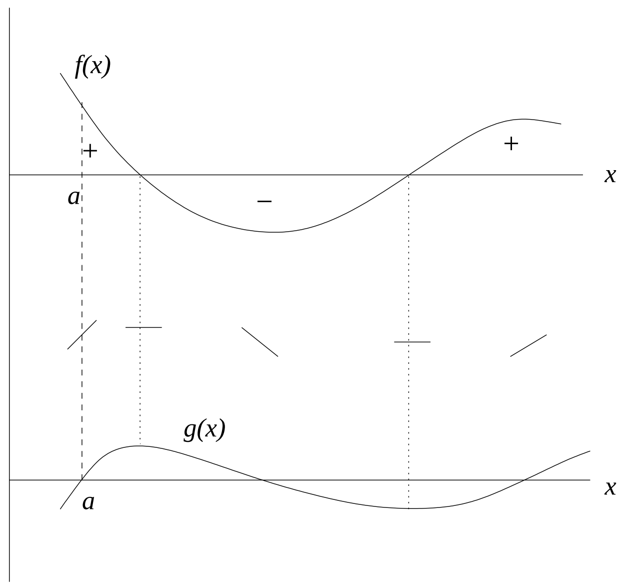

Figure 3.7. *The original functions, $f(x)$ is shown above. The corresponding functions $g(x)$ is drawn below.*

Solution

See Figure 3.7

3.8 Some fine print

The Fundamental Theorem has a number of restrictions that must be satisfied before its results can be applied. In this section we look at some examples in which care must be used.

3.8.1 Function unbounded I

Consider the definite integral

$$\int_0^2 \frac{1}{x}\, dx.$$

The function $f(x) = \frac{1}{x}$ is undefined at $x = 0$, and unbounded on any interval that contains the point $x = 0$. Hence, we cannot evaluate this integral using the Fundamental theorem, and indeed, we say that "*this integral does not exist*".

3.8.2 Function unbounded II

Consider the definite integral

$$\int_{-1}^1 \frac{1}{x^2} dx.$$

This function is also undefined (and hence not continuous) at $x = 0$. The Fundamental Theorem of Calculus cannot be applied. Technically, although one can "go through the motions" of computing an antiderivative, evaluating it at both endpoints, and getting a numerical answer, the result so obtained would be simply wrong. We say that his integral does not exist.

3.8.3 Example: Function discontinuous or with distinct parts

Suppose we are given the integral

$$I = \int_{-1}^2 |x|\, dx.$$

This function is actually made up of two distinct parts, namely

$$f(x) = \begin{cases} x & \text{if } x > 0 \\ -x & \text{if } x < 0. \end{cases}$$

The integral I must therefore be split up into two parts, namely

$$I = \int_{-1}^2 |x|\, dx = \int_{-1}^0 (-x)\, dx + \int_0^2 x\, dx.$$

We find that

$$I = -\left.\frac{x^2}{2}\right|_{-1}^0 + \left.\frac{x^2}{2}\right|_0^2 = -\left[0 - \frac{1}{2}\right] + \left[\frac{4}{2} - 0\right] = 2.5$$

3.8.4 Function undefined

Now let us examine the integral

$$\int_{-1}^1 x^{1/2}\, dx.$$

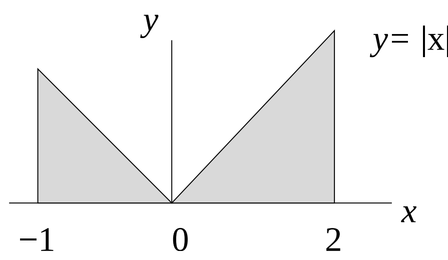

Figure 3.8. *In this example, to compute the integral over the interval* $-1 \leq x \leq 2$, *we must split up the region into two distinct parts.*

We see that there is a problem here. Recall that $x^{1/2} = \sqrt{x}$. Hence, the function is not defined for $x < 0$ and the interval of integration is inappropriate. Hence, this integral does not make sense.

3.8.5 Infinite domain ("improper integral")

Consider the integral

$$I = \int_0^b e^{-rx}\, dx, \quad \text{where } r > 0, \text{ and } b > 0 \text{ are constants.}$$

Simple integration using the antiderivative in Table 3.1 (for $k = -r$) leads to the result

$$I = \left. \frac{e^{-rx}}{-r} \right|_0^b = -\frac{1}{r}\left(e^{-rb} - e^0\right) = \frac{1}{r}\left(1 - e^{-rb}\right).$$

This is the area under the exponential curve between $x = 0$ and $x = b$. Now consider what happens when b, the upper endpoint of the integral increases, so that $b \to \infty$. Then the value of the integral becomes

$$I = \lim_{b\to\infty} \int_0^b e^{-rx}\, dx = \lim_{b\to\infty} \frac{1}{r}\left(1 - e^{-rb}\right) = \frac{1}{r}(1-0) = \frac{1}{r}.$$

(We used the fact that $e^{-rb} \to 0$ as $b \to \infty$.) We have, in essence, found that

$$I = \int_0^\infty e^{-rx}\, dx = \frac{1}{r}. \tag{3.5}$$

An integral of the form (3.5) is called an **improper integral**. Even though the domain of integration of this integral is infinite, $(0, \infty)$, observe that the value we computed is finite, so long as $r \neq 0$. Not all such integrals have a bounded finite value. Learning to distinguish between those that do and those that do not will form an important theme in Chapter 10.

Regions that need special treatment

So far, we have learned how to compute areas of regions in the plane that are bounded by one or more curves. In all our examples so far, the basis for these calculations rests on imagining rectangles whose heights are specified by one or another function. Up to now, all the rectangular strips we considered had bases (of width Δx) on the x axis. In Figure 3.9 we observe an example in which it would not be possible to use this technique. We are

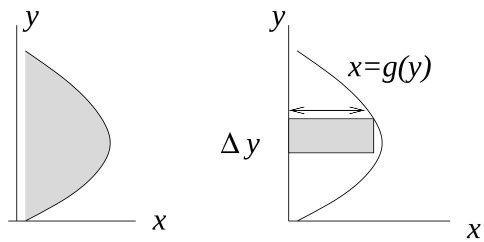

Figure 3.9. *The area in the region shown here is best computed by integrating in the y direction. If we do so, we can use the curved boundary as a single function that defines the region. (Note that the curve cannot be expressed in the form of a function in the usual sense, $y = f(x)$, but it can be expressed in the form of a function $x = f(y)$.)*

asked to find the area between the curve $y^2 - y + x = 0$ and the y axis. However, one and the same curve, $y^2 - y + x = 0$ forms the boundary from both the top and the bottom of the region. We are unable to set up a series of rectangles with bases along the x axis whose heights are described by this curve. This means that our definite integral (which is really just a convenient way of carrying out the process of area computation) has to be handled with care.

Let us consider this problem from a "new angle", i.e. with rectangles based on the y axis, we can achieve the desired result. To do so, let us express our curve in the form

$$x = g(y) = y - y^2.$$

Then, placing our rectangles along the interval $0 < y < 1$ on the y axis (each having base of width Δy) leads to the integral

$$I = \int_0^1 g(y)\, dy = \int_0^1 (y - y^2) dy = \left(\frac{y^2}{2} - \frac{y^3}{3} \right) \Big|_0^1 = \frac{1}{2} - \frac{1}{3} = \frac{1}{6}.$$

3.9 Summary

In this chapter we first recapped the definition of the definite integral in Section 3.1, recalled its connection to an area in the plane under the graph of some function $f(x)$, and examined its basic properties.

If one of the endpoints, x of the integral is allowed to vary, the area it represents, $A(x)$, becomes a function of x. Our construction in Figure 3.2 showed that there is a connection between the derivative $A'(x)$ of the area and the function $f(x)$. Indeed, we showed that $A'(x) = f(x)$ and argued that this makes $A(x)$ an antiderivative of the function $f(x)$.

This important connection between integrals and antiderivatives is the crux of Integral Calculus, forming the Fundamental Theorem of Calculus. Its significance is that finding areas need not be as tedious and labored as the calculation of Riemann sums that formed the bulk of Chapter 2. Rather, we can take a shortcut using antidifferentiation.

Motivated by this very important result, we reviewed some common functions and derivatives, and used this to relate functions and their antiderivatives in Table 3.1. We used these antiderivatives to calculate areas in several examples. Finally, we extended the treatment to include qualitative sketches of functions and their antiderivatives.

As we will see in upcoming chapters, the ideas presented here have a much wider range of applicability than simple area calculations. Indeed, we will shortly show that the same concepts can be used to calculate net changes in continually varying processes, to compute volumes of various shapes, to determine displacement from velocity, mass from densities, as well as a host of other quantities that involve a process of accumulation. These ideas will be investigated in Chapters 4, and 5.

Chapter 4

Applications of the definite integral to velocities and rates

4.1 Introduction

In this chapter, we encounter a number of applications of the definite integral to practical problems. We will discuss the connection between acceleration, velocity and displacement of a moving object, a topic we visited in an earlier, Differential Calculus Course. Here we will show that the notion of antiderivatives and integrals allows us to deduce details of the motion of an object from underlying Laws of Motion. We will consider both uniform and accelerated motion, and recall how air resistance can be described, and what effect it induces.

An important connection is made in this chapter between a rate of change (e.g. rate of growth) and the total change (i.e. the net change resulting from all the accumulation and loss over a time span). We show that such examples also involve the concept of integration, which, fundamentally, is a cumulative summation of infinitesimal changes. This allows us to extend the utility of the mathematical tools to a variety of novel situations. We will see examples of this type in Sections 4.3 and 4.4.

Several other important ideas are introduced in this chapter. We encounter for the first time the idea of spatial density, and see that integration can also be used to "add up" the total amount of material distributed over space. In Section 5.2.2, this idea is applied to the density of cars along a highway. We also consider mass distributions and the notion of a center of mass.

Finally, we also show that the definite integral is useful for determining the average value of a function, as discussed in Section 4.6. In all these examples, the important step is to properly set up the definite integral that corresponds to the desired net change. Computations at this stage are relatively straightforward to emphasize the process of setting up the appropriate integrals and understanding what they represent.

4.2 Displacement, velocity and acceleration

Recall from our study of derivatives that for $x(t)$ the position of some particle at time t, $v(t)$ its velocity, and $a(t)$ the acceleration, the following relationships hold:

$$\frac{dx}{dt} = v, \quad \frac{dv}{dt} = a.$$

(**Velocity** is the derivative of position and **acceleration** is the derivative of velocity.) This means that position is an anti-derivative of velocity and velocity is an anti-derivative of acceleration.

Since position, $x(t)$, is an anti-derivative of velocity, $v(t)$, by the Fundamental Theorem of Calculus, it follows that over the time interval $T_1 \leq t \leq T_2$,

$$\int_{T_1}^{T_2} v(t)\,dt = x(t)\Big|_{T_1}^{T_2} = x(T_2) - x(T_1). \tag{4.1}$$

The quantity on the right hand side of Eqn. (4.1) is a **displacement**,, i.e., the difference between the position at time T_1 and the position at time T_2. In the case that $T_1 = 0, T_2 = T$, we have

$$\int_0^T v(t)\,dt = x(T) - x(0),$$

as the displacement over the time interval $0 \leq t \leq T$.

Similarly, since velocity is an anti-derivative of acceleration, the Fundamental Theorem of Calculus says that

$$\int_{T_1}^{T_2} a(t)\,dt = v(t)\Big|_{T_1}^{T_2} = v(T_2) - v(T_1). \tag{4.2}$$

as above, we also have that

$$\int_0^T a(t)\,dt = v(t)\Big|_0^T = v(T) - v(0)$$

is the net change in velocity between time 0 and time T, (though this quantity does not have a special name).

4.2.1 Geometric interpretations

Suppose we are given a graph of the velocity $v(t)$, as shown on the left of Figure 4.1. Then by the definition of the definite integral, we can interpret $\int_{T_1}^{T_2} v(t)\,dt$ as the "area" associated with the curve (counting positive and negative contributions) between the endpoints T_1 and T_2. Then according to the above observations, this area represents the displacement of the particle between the two times T_1 and T_2.

Similarly, by previous remarks, the area under the curve $a(t)$ is a geometric quantity that represents the net change in the velocity, as shown on the right of Figure 4.1.

Next, we consider two examples where either the acceleration or the velocity is constant. We use the results above to compute the displacements in each case.

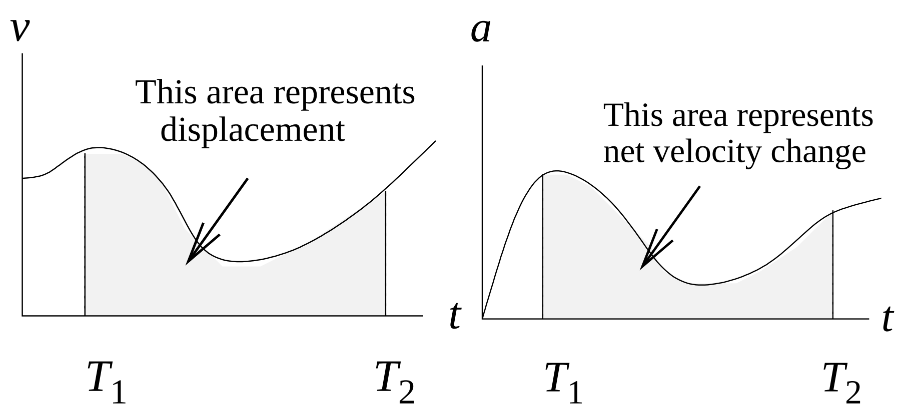

Figure 4.1. *The total area under the velocity graph represents net displacement, and the total area under the graph of acceleration represents the net change in velocity over the interval* $T_1 \leq t \leq T_2$.

4.2.2 Displacement for uniform motion

We first examine the simplest case that the velocity is constant, i.e. $v(t) = v =$ constant. Then clearly, the acceleration is zero since $a = dv/dt = 0$ when v is constant. Thus, by direct antidifferentiation,

$$\int_0^T v\, dt = vt\Big|_0^T = v(T-0) = vT.$$

However, applying result (4.1) over the time interval $0 \leq t \leq T$ also leads to

$$\int_0^T v\, dt = x(T) - x(0).$$

Therefore, it must be true that the two expressions obtained above must be equal, i.e.

$$x(T) - x(0) = vT.$$

Thus, for uniform motion, the displacement is proportional to the velocity and to the time elapsed. The final position is

$$x(T) = x(0) + vT.$$

This is true for all time T, so we can rewrite the results in terms of the more familiar (lower case) notation for time, t, i.e.

$$x(t) = x(0) + vt. \tag{4.3}$$

4.2.3 Uniformly accelerated motion

In this case, the acceleration a is a constant. Thus, by direct antidifferentiation,

$$\int_0^T a\, dt = at\Big|_0^T = a(T-0) = aT.$$

However, using Equation (4.2) for $0 \le t \le T$ leads to

$$\int_0^T a\, dt = v(T) - v(0).$$

Since these two results must match, $v(T) - v(0) = aT$ so that

$$v(T) = v(0) + aT.$$

Let us refer to the initial velocity $V(0)$ as v_0. The above connection between velocity and acceleration holds for any final time T, i.e., it is true for all t that:

$$v(t) = v_0 + at. \tag{4.4}$$

This just means that velocity at time t is the initial velocity incremented by an increase (over the given time interval) due to the acceleration. From this we can find the displacement and position of the particle as follows: Let us call the initial position $x(0) = x_0$. Then

$$\int_0^T v(t)\, dt = x(T) - x_0. \tag{4.5}$$

But

$$I = \int_0^T v(t)\, dt = \int_0^T (v_0 + at)\, dt = \left(v_0 t + a\frac{t^2}{2} \right)\Big|_0^T = \left(v_0 T + a\frac{T^2}{2} \right). \tag{4.6}$$

So, setting Equations (4.5) and (4.6) equal means that

$$x(T) - x_0 = v_0 T + a\frac{T^2}{2}.$$

But this is true for *all* final times, T, i.e. this holds for any time t so that

$$x(t) = x_0 + v_0 t + a\frac{t^2}{2}. \tag{4.7}$$

This expression represents the position of a particle at time t given that it experienced a constant acceleration. The initial velocity v_0, initial position x_0 and acceleration a allowed us to predict the position of the object $x(t)$ at any later time t. That is the meaning of Eqn. (4.7)[16].

4.2.4 Non-constant acceleration and terminal velocity

In general, the acceleration of a falling body is not actually uniform, because frictional forces impede that motion. A better approximation to the rate of change of velocity is given by the **differential equation**

$$\frac{dv}{dt} = g - kv. \tag{4.8}$$

[16]Of course, Eqn. (4.7) only holds so long as the object is accelerating. Once the a falling object hits the ground, for example, this equation no longer holds.

We will assume that initially the velocity is zero, i.e. $v(0) = 0$.

This equation is a mathematical statement that relates changes in velocity $v(t)$ to the constant acceleration due to gravity, g, and drag forces due to friction with the atmosphere. A good approximation for such drag forces is the term kv, proportional to the velocity, with k, a positive constant, representing a frictional coefficient. Because $v(t)$ appears both in the derivative and in the expression kv, we cannot apply the methods developed in the previous section directly. That is, we do not have an expression that depends on time whose antiderivative we would calculate. The derivative of $v(t)$ (on the left) is connected to the unknown $v(t)$ on the right.

Finding the velocity and then the displacement for this type of motion requires special techniques. In Chapter 9, we will develop a systematic approach, called Separation of Variables to find analytic solutions to equations such as (4.8).

Here, we use a special procedure that allows us to determine the velocity in this case. We first recall the following result from first term calculus material:

The differential equation and initial condition

$$\frac{dy}{dt} = -ky, \quad y(0) = y_0 \tag{4.9}$$

has a solution

$$y(t) = y_0 e^{-kt}. \tag{4.10}$$

Equation (4.8) implies that

$$a(t) = g - kv(t),$$

where $a(t)$ is the acceleration at time t. Taking a derivative of both sides of this equation leads to

$$\frac{da}{dt} = -k\frac{dv}{dt} = -ka.$$

We observe that this equation has the same form as equation (4.9) (with a replacing y), which implies (according to 4.10) that $a(t)$ is given by

$$a(t) = C\, e^{-kt} = a_0\, e^{-kt}.$$

Initially, at time $t = 0$, the acceleration is $a(0) = g$ (since $a(t) = g - kv(t)$, and $v(0) = 0$). Therefore,

$$a(t) = g\, e^{-kt}.$$

Since we now have an explicit formula for acceleration vs time, we can apply direct integration as we did in the examples in Sections 4.2.2 and 4.2.3. The result is:

$$\int_0^T a(t)\, dt = \int_0^T g\, e^{-kt}\, dt = g \int_0^T e^{-kt}\, dt = g \left[\frac{e^{-kt}}{-k}\right]\Bigg|_0^T = g\,\frac{(e^{-kT} - 1)}{-k} = \frac{g}{k}\left(1 - e^{-kT}\right).$$

In the calculation, we have used the fact that the antiderivative of e^{-kt} is e^{-kt}/k. (This can be verified by simple differentiation.)

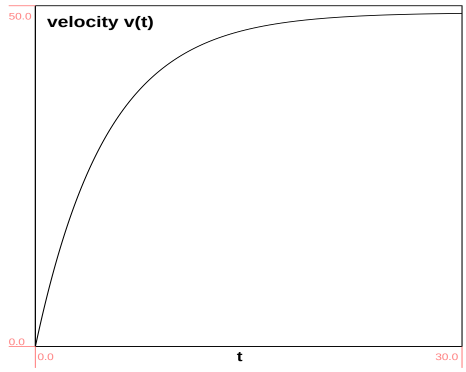

Figure 4.2. *Terminal velocity (m/s) for acceleration due to gravity g=9.8 m/s^2, and $k = 0.2$/s. The velocity reaches a near constant 49 m/s by about 20 s.*

As before, based on equation (4.2) this integral of the acceleration over $0 \leq t \leq T$ must equal $v(T) - v(0)$. But $v(0) = 0$ by assumption, and the result is true for *any* final time T, so, in particular, setting $T = t$, and combining both results leads to an expression for the velocity at any time:

$$v(t) = \frac{g}{k}\ \left(1 - e^{-kt}\right). \tag{4.11}$$

We will study the differential equation (4.8) again in Section 9.3.2, in the context of a more detailed discussion of differential equations

From our result here, we can also determine how the velocity behaves in the long term: observe that for $t \to \infty$, the exponential term $e^{-kt} \to 0$, so that

$$v(t) \to \frac{g}{k}(1 - \text{ very small quantity}) \approx \frac{g}{k}.$$

Thus, when drag forces are in effect, the falling object does not continue to accelerate indefinitely: it eventually attains a **terminal velocity**. We have seen that this limiting velocity is $v = g/k$. The object continues to fall at this (approximately constant) speed as shown in Figure 4.2. The terminal velocity is also a steady state value of Eqn. (4.8), i.e. a value of the velocity at which no further change occurs.

4.3 From rates of change to total change

In this section, we examine several examples in which the rate of change of some process is specified. We use this information to obtain the total change[17] that occurs over some time period.

[17]We will use the terminology "total change" and "net change" interchangeably in this section.

Changing temperature

We must carefully distinguish between information about the time dependence of some function, from information about the rate of change of some function. Here is an example of these two different cases, and how we would handle them

(a) The temperature of a cup of juice is observed to be

$$T(t) = 25(1 - e^{-0.1t})^{\circ}\text{Celcius}$$

where t is time in minutes. Find the change in the temperature of the juice between the times $t = 1$ and $t = 5$.

(b) The **rate of change** of temperature of a cup of coffee is observed to be

$$f(t) = 8e^{-0.2t^{\circ}}\text{Celcius per minute}$$

where t is time in minutes. What is the **total change** in the temperature between $t = 1$ and $t = 5$ minutes ?

Solutions

(a) In this case, we are given the temperature as a function of time. To determine what **net change** occurred between times $t = 1$ and $t = 5$, we find the temperatures at each time point and subtract: That is, the change in temperature between times $t = 1$ and $t = 5$ is simply

$$T(5) - T(1) = 25(1 - e^{-0.5}) - 25(1 - e^{-0.1}) = 25(0.94 - 0.606) = 7.47^{\circ}\text{Celcius}.$$

(b) Here, we do not know the temperature at any time, but we are given information about **the rate of change**. (Carefully note the subtle difference in the wording.) To get the total change, we would sum up all the small changes, $f(t)\Delta t$ (over N subintervals of duration $\Delta t = (5 - 1)/N = 4/N$) for t starting at 1 and ending at 5 min. We obtain a sum of the form $\sum f(t_k)\Delta t$ where t_k is the k'th time point. Finally, we take a limit as the number of subintervals increases ($N \to \infty$). By now, we recognize that this amounts to a process of integration. Based on this variation of the same concept we can take the usual shortcut of integrating the rate of change, $f(t)$, from $t = 1$ to $t = 5$. To do so, we apply the Fundamental Theorem as before, reducing the amount of computation to finding antiderivatives. We compute:

$$I = \int_1^5 f(t)\,dt = \int_1^5 8e^{-0.2t}dt = -40e^{-0.2t}\Big|_1^5 = -40e^{-1} + 40e^{-0.2},$$

$$I = 40(e^{-0.2} - e^{-1}) = 40(0.8187 - 0.3678) = 18.$$

Only in the second case did we need to use a definite integral to find a net change, since we were given the way that the *rate of change* depended on time. Recognizing the subtleties of the wording in such examples will be an important skill that the reader should gain.

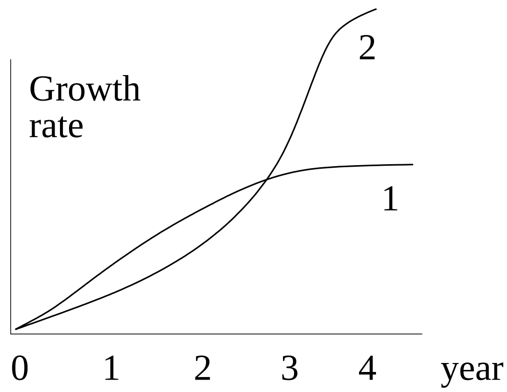

Figure 4.3. *Growth rates of two trees over a four year period. Tree 1 initially has a higher growth rate, but tree 2 catches up and grows faster after year 3.*

4.3.1 Tree growth rates

The rate of growth in height for two species of trees (in feet per year) is shown in Figure 4.3. If the trees start at the same height, which tree is taller after 1 year? After 4 years?

Solution

In this problem we are provided with a sketch, rather than a formula for the growth rate of the trees. Our solution will thus be *qualitative* (i.e. descriptive), rather than *quantitative*. (This means we do not have to calculate anything; rather, we have to make some important observations about the behaviour shown in Fig 4.3.)

We recognize that the net change in height of each tree is of the form

$$H_i(T) - H_i(0) = \int_0^T g_i(t)dt, \quad i = 1, 2.$$

where $i = 1$ for tree 1, $i = 2$ for tree 2, $g_i(t)$ is the growth rate as a function of time (shown for each tree in Figure 4.3) and $H_i(t)$ is the height of tree "i" at time t. But, by the Fundamental Theorem of Calculus, this definite integral corresponds to the area under the curve $g_i(t)$ from $t = 0$ to $t = T$. Thus we must interpret the net change in height for each tree as the area under its growth curve. We see from Figure 4.3 that at $t = 1$ year, the area under the curve for tree 1 is greater, so it has grown more. At $t = 4$ years the area under the second curve is greatest so tree 2 has grown the most by that time.

4.3.2 Radius of a tree trunk

The trunk of a tree, assumed to have the shape of a cylinder, grows incrementally, so that its cross-section consists of "rings". In years of plentiful rain and adequate nutrients, the tree grows faster than in years of drought or poor soil conditions. Suppose the rainfall pattern

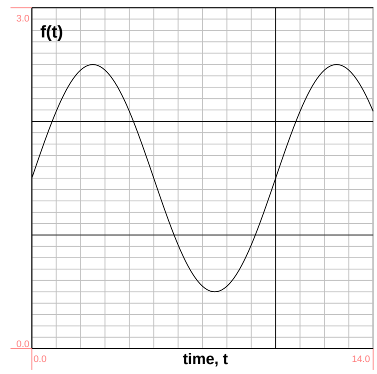

Figure 4.4. *Rate of change of radius, $f(t)$ for a growing tree over a period of 14 years.*

has been cyclic, so that, over a period of 14 years, the growth rate of the radius of the tree trunk (in cm/year) is given by the function

$$f(t) = 1.5 + \sin(\pi t/5),$$

as shown in Figure 4.4. Let the height of the tree trunk be approximately constant over this ten year period, and assume that the density of the trunk is approximately 1 gm/cm^3.

(a) If the radius was initially r_0 at time $t = 0$, what will the radius of the trunk be at time t later?

(b) What is the ratio of the mass of the tree trunk at $t = 10$ years and $t = 0$ years? (i.e. find the ratio mass(10)/mass(0).)

Solution

(a) Let $R(t)$ denote the trunk's radius at time t. The rate of change of the radius of the tree is given by the function $f(t)$, and we are told that at $t = 0$, $R(0) = r_0$. A graph of this growth rate over the first fifteen years is shown in Figure 4.4. The net change in the radius is

$$R(t) - R(0) = \int_0^t f(s)\, ds = \int_0^t (1.5 + \sin(\pi s/5))\, ds.$$

Integrating the above, we get

$$R(t) - R(0) = \left(1.5t - \frac{\cos(\pi s/5)}{\pi/5}\right)\Bigg|_0^t.$$

Here we have used the fact that the antiderivative of $\sin(ax)$ is $-(\cos(ax)/a)$.

Thus, using the initial value, $R(0) = r_0$ (which is a constant), and evaluating the integral, leads to

$$R(t) = r_0 + 1.5t - \frac{5\cos(\pi t/5)}{\pi} + \frac{5}{\pi}.$$

(The constant at the end of the expression stems from the fact that $\cos(0) = 1$.) A graph of the radius of the tree over time (using $r_0 = 1$) is shown in Figure 4.5. This function is equivalent to the area associated with the function shown in Figure 4.4. Notice that Figure 4.5 confirms that the radius keeps growing over the entire period, but that its growth rate (slope of the curve) alternates between higher and lower values.

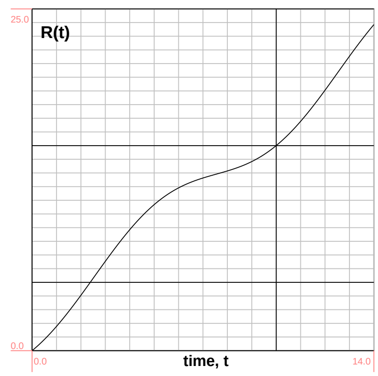

Figure 4.5. *The radius of the tree, $R(t)$, as a function of time, obtained by integrating the rate of change of radius shown in Fig. 4.4.*

After ten years we have

$$R(10) = r_0 + 15 - \frac{5}{\pi}\cos(2\pi) + \frac{5}{\pi}.$$

But $\cos(2\pi) = 1$, so

$$R(10) = r_0 + 15.$$

(b) The mass of the tree is density times volume, and since the density in this example is constant, 1 gm/cm^3, we need only obtain the volume at $t = 10$. Taking the trunk to be cylindrical means that the volume at any given time is

$$V(t) = \pi[R(t)]^2 h.$$

The ratio we want is then

$$\frac{V(10)}{V(0)} = \frac{\pi[R(10)]^2 h}{\pi r_0^2 h} = \frac{[R(10)]^2}{r_0^2} = \left(\frac{r_0 + 15}{r_0}\right)^2.$$

In this problem we used simple anti-differentiation to compute the desired total change. We also related the graph of the radial growth rate in Fig. 4.4 to that of the resulting radius at time t, in Fig. 4.5.

4.3.3 Birth rates and total births

After World War II, the birth rate in western countries increased dramatically. Suppose that the number of babies born (in millions per year) was given by

$$b(t) = 5 + 2t, \;\; 0 \le t \le 10,$$

where t is time in years after the end of the war.

(a) How many babies in total were born during this time period (i.e in the first 10 years after the war)?

(b) Find the time T_0 such that the total number of babies born from the end of the war up to the time T_0 was precisely 14 million.

Solution

(a) To find the number of births, we would integrate the birth rate, $b(t)$ over the given time period. The *net change* in the population due to births (neglecting deaths) is

$$P(10)-P(0) = \int_0^{10} b(t)\, dt = \int_0^{10} (5+2t)\, dt = (5t+t^2)|_0^{10} = 50+100 = 150\text{[million babies]}.$$

(b) Denote by T the time at which the total number of babies born was 14 million. Then, [in units of million]

$$I = \int_0^T b(t)\, dt = 14 = \int_0^T (5+2t)\, dt = 5T + T^2$$

equating $I = 14$ leads to the quadratic equation, $T^2 + 5T - 14 = 0$, which can be written in the factored form, $(T - 2)(T + 7) = 0$. This has two solutions, but we reject $T = -7$ since we are looking for time after the War. Thus we find that $T = 2$ years, i.e it took two years for 14 million babies to have been born.

While this problem involves simple integration, we had to solve for a quantity (T) based on information about behaviour of that integral. Many problems in real application involve such slight twists on the ideas of integration.

4.4 Production and removal

The process of integration can be used to convert rates of production and removal into net amounts present at a given time. The example in this section is of this type. We investigate a process in which a substance accumulates as it is being produced, but disappears through some removal process. We would like to determine when the quantity of material increases, and when it decreases.

Circadean rhythm in hormone levels

Consider a hormone whose level in the blood at time t will be denoted by $H(t)$. We will assume that the level of hormone is regulated by two separate processes: one might be the secretion rate of specialized cells that produce the hormone. (The production rate of hormone might depend on the time of day, in some cyclic pattern that repeats every 24 hours or so.) This type of cyclic pattern is called *circadean* rhythm. A competing process might be the removal of hormone (or its deactivation by some enzymes secreted by other cells.) In this example, we will assume that both the production rate, $p(t)$, and the removal rate, $r(t)$, of the hormone are time-dependent, periodic functions with somewhat different phases.

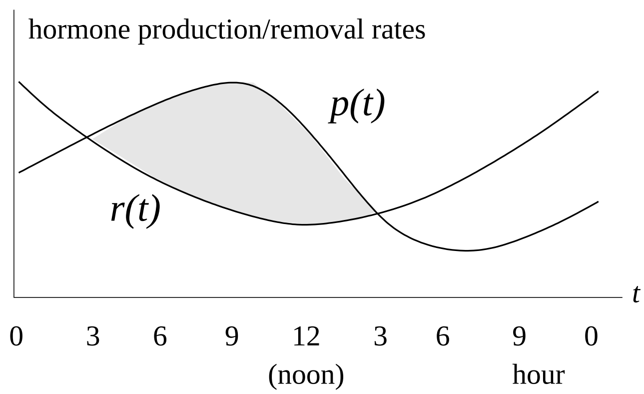

Figure 4.6. *The rate of hormone production $p(t)$ and the rate of removel $r(t)$ are shown here. We want to use these graphs to deduce when the level of hormone is highest and lowest.*

A typical example of two such functions are shown in Figure 4.6. This figure shows the production and removal rates over a period of 24 hours, starting at midnight. Our first task will be to use properties of the graph in Figure 4.6 to answer the following questions:

1. When is the production rate, $p(t)$, maximal?
2. When is the removal rate $r(t)$ minimal?
3. At what time is the hormone level in the blood highest?
4. At what time is the hormone level in the blood lowest?
5. Find the maximal level of hormone in the blood over the period shown, assuming that its basal (lowest) level is $H = 0$.

Solutions

1. We see directly from Fig. 4.6 that production rate is maximal at about 9:00 am.

2. Similarly, removal rate is minimal at noon.

3. To answer this question we note that the total amount of hormone produced over a time period $a \leq t \leq b$ is
$$P_{\text{total}} = \int_a^b p(t)dt.$$
The total amount removed over time interval $a \leq t \leq b$ is
$$R_{\text{total}} = \int_a^b r(t)dt.$$
This means that the net change in hormone level over the given time interval (amount produced minus amount removed) is
$$H(b) - H(a) = P_{\text{total}} - R_{\text{total}} = \int_a^b (p(t) - r(t))dt.$$
We interpret this integral as the *area between the curves* $p(t)$ and $r(t)$. But we must use caution here: For any time interval over which $p(t) > r(t)$, this integral will be positive, and the hormone level will have increased. Otherwise, if $r(t) < p(t)$, the integral yields a negative result, so that the hormone level has decreased. This makes simple intuitive sense: If production is greater than removal, the level of the substance is accumulating, whereas in the opposite situation, the substance is decreasing. With these remarks, we find that the hormone level in the blood will be *greatest* at 3:00 pm, when the greatest (positive) area between the two curves has been obtained.

4. Similarly, the least hormone level occurs after a period in which the removal rate has been larger than production for the longest stretch. This occurs at 3:00 am, just as the curves cross each other.

5. Here we will practice integration by actually fitting some cyclic functions to the graphs shown in Figure 4.6. Our first observation is that if the length of the cycle (also called the *period*) is 24 hours, then the *frequency* of the oscillation is $\omega = (2\pi)/24 = \pi/12$. We further observe that the functions shown in the Figure 4.7 have the form
$$p(t) = A(1 + \sin(\omega t)), \quad r(t) = A(1 + \cos(\omega t)).$$
Intersection points occur when
$$p(t) = r(t)$$
$$A(1 + \sin(\omega t)) = A(1 + \cos(\omega t)),$$
$$\sin(\omega t) = \cos(\omega t)),$$
$$\Rightarrow \tan(\omega t) = 1.$$
This last equality leads to $\omega t = \pi/4, 5\pi/4$. But then, the fact that $\omega = \pi/12$ implies that $t = 3, 15$. Thus, over the time period $3 \leq t \leq 15$ hrs, the hormone level is

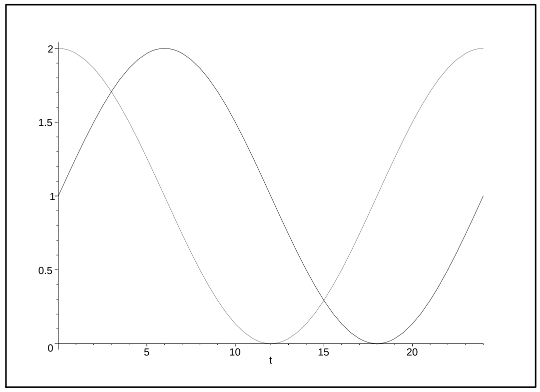

Figure 4.7. *The functions shown above are trigonometric approximations to the rates of hormone production and removal from Figure 4.6*

increasing. For simplicity, we will take the amplitude $A = 1$. (In general, this would just be a multiplicative constant in whatever solution we compute.) Then the net increase in hormone over this period is calculated from the integral

$$H_{\text{total}} = \int_3^{15} [p(t) - r(t)]\, dt = \int_3^{15} [(1 + \sin(\omega t)) - (1 + \cos(\omega t))]\, dt$$

In the problem set, the reader is asked to compute this integral and to show that the amount of hormone that accumulated over the time interval $3 \leq t \leq 15$, i.e. between 3:00 am and 3:00 pm is $24\sqrt{2}/\pi$.

4.5 Present value of a continuous income stream

Here we discuss the value of an annuity, which is a kind of savings account that guarantees a continuous stream of income. You would like to pay P dollars to purchase an annuity that will pay you an income $f(t)$ every year from now on, for $t > 0$. In some cases, we might want a constant income every year, in which case $f(t)$ would be constant. More generally, we can consider the case that at each future year t, we ask for income $f(t)$ that could vary from year to year. If the bank interest rate is r, how much should you pay now?

Solution

If we invest P dollars (the "principal" i.e., the amount deposited) in the bank with interest r (compounded continually) then the amount $A(t)$ in the account at time t (in years), will grow as follows:

$$A(t) = P\left(1 + \frac{r}{n}\right)^{nt},$$

where r is the yearly interest rate (e.g. 5%) and n is the number of times per year that interest is compound (e.g. $n = 2$ means interest compounded twice per year, $n = 12$ means monthly compounded interest, etc.) Define

$$h = \frac{r}{n}. \quad \text{Then} \quad n = \frac{r}{h}.$$

Then at time t, we have that

$$\begin{aligned} A(t) &= P(1+h)^{\frac{1}{h}rt} \\ &= P\left[(1+h)^{\frac{1}{h}}\right]^{rt} \\ &\approx Pe^{rt} \quad \text{for large } n \text{ or small } h. \end{aligned}$$

Here we have used the fact that when h is small (i.e. frequent intervals of compounding) the expression in square brackets above can be approximated by e, the base of the natural logarithms. Recall that

$$e = \lim_{h\to 0}\left[(1+h)^{\frac{1}{h}}\right].$$

(This result was obtained in a first semester calculus course by selecting the base of exponentials such that the derivative of e^x is just e^x itself.) Thus, we have found that the amount in the bank at time t will grow as

$$\boxed{A(t) = Pe^{rt}, \quad \text{(assuming continually compounded interest).} \qquad (4.12)}$$

Having established the exponential growth of an investment, we return to the question of how to set up an annuity for a continuous stream of income in the future. Rewriting Eqn. (4.12), the principle amount that we should invest in order to have $A(t)$ to spend at time t is

$$P = A(t)e^{-rt}.$$

Suppose we want to have $f(t)$ spending money for each year t. We refer to the *present value* of year t as the quantity

$$P = f(t)e^{-rt}.$$

(i.e. We must pay P now, in the present, to get $f(t)$ in a future year t.) Summing over all the years, we find that the present value of the continuous income stream is

$$P = \sum_{t=1}^{L} f(t)e^{-rt} \cdot \underbrace{1}_{\text{“}\Delta t''} \approx \int_0^L f(t)e^{-rt}\,dt,$$

where L is the expected number of years left in the lifespan of the individual to whom this annuity will be paid, and where we have approximated a sum of payments by an integral (of a continuous income stream). One problem is that we do not know in advance how long the lifespan L will be. As a crude approximation, we could assume that this income stream continues forever, i.e. that $L \approx \infty$. In such an approximation, we have to compute the integral:

$$P = \int_0^{\infty} f(t)e^{-rt}\, dt. \tag{4.13}$$

The integral in Eqn. (4.13) is an **improper integral** (i.e. integral over an unbounded domain), as we have already encountered in Section 3.8.5. We shall have more to say about the properties of such integrals, and about their technical definition, existence, and properties in Chapter 10. We refer to the quantity

$$P = \int_0^{\infty} f(t)e^{-rt}\, dt, \tag{4.14}$$

as the *present value of a continuous income stream* $f(t)$.

Example: Setting up an annuity

Suppose we want an annuity that provides us with an annual payment of $10,000$ from the bank, i.e. in this case $f(t) = \$10,000$ is a function that has a constant value for every year. Then from Eqn (4.14),

$$P = \int_0^{\infty} 10000e^{-rt}\, dt = 10000 \int_0^{\infty} e^{-rt}\, dt.$$

By a previous calculation in Section 3.8.5, we find that

$$P = 10000 \cdot \frac{1}{r},$$

e.g. if interest rate is 5% (and assumed constant over future years), then

$$P = \frac{10000}{0.05} = \$200,000.$$

Therefore, we need to pay \$200,000 today to get $10,000$ annually for every future year.

4.6 Average value of a function

In this final example, we apply the definite integral to computing the average height of a function over some interval. First, we define what is meant by average value in this context.[18]

Given a function

$$y = f(x)$$

over some interval $a \le x \le b$, we will define average value of the function as follows:

[18]In Chapters 5 and 8, we will encounter a different type of average (also called mean) that will represent an average horizontal position or center of mass. It is important to avoid confusing these distinct notions.

Definition

The average value of $f(x)$ over the interval $a \le x \le b$ is

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

Example 1

Find the average value of the function $y = f(x) = x^2$ over the interval $2 < x < 4$.

Solution

$$\bar{f} = \frac{1}{4-2}\int_2^4 x^2\,dx = \frac{1}{2}\frac{x^3}{3}\bigg|_2^4 = \frac{1}{6}\left(4^3 - 2^3\right) = \frac{28}{3}$$

Example 2: Day length over the year

Suppose we want to know the average length of the day during summer and spring. We will assume that day length follows a simple periodic behaviour, with a cycle length of 1 year (365 days). Let us measure time t in days, with $t = 0$ at the spring equinox, i.e. the date in spring when night and day lengths are equal, each being 12 hrs. We will refer to the number of daylight hours on day t by the function $f(t)$. (We will also call $f(t)$ the day-length on day t.

A simple function that would describe the cyclic changes of day length over the seasons is

$$f(t) = 12 + 4\sin\left(\frac{2\pi t}{365}\right).$$

This is a periodic function with period 365 days as shown in Figure 4.8. Its maximal value is $16h$ and its minimal value is $8h$. The average day length over spring and summer, i.e. over the first (365/2) $\approx$ 182 days is:

$$\begin{aligned}
\bar{f} &= \frac{1}{182}\int_0^{182} f(t)dt \\
&= \frac{1}{182}\int_0^{182}\left(12 + 4\sin(\frac{\pi t}{182})\right)\,dt \\
&= \frac{1}{182}\left(12t - \frac{4\cdot 182}{\pi}\cos(\frac{\pi t}{182})\right)\bigg|_0^{182} \\
&= \frac{1}{182}\left(12\cdot 182 - \frac{4\cdot 182}{\pi}[\cos(\pi) - \cos(0)]\right) \\
&= 12 + \frac{8}{\pi} \approx 14.546 \text{ hours}
\end{aligned} \tag{4.15}$$

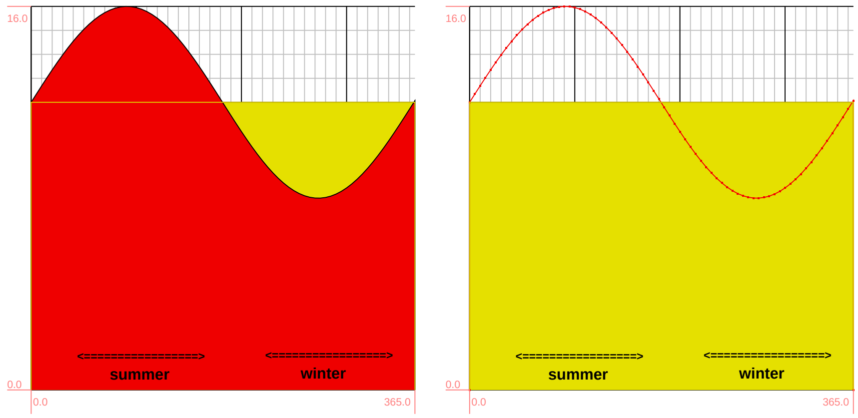

Figure 4.8. *We show the variations in day length (cyclic curve) as well as the average day length (height of rectangle) in this figure.*

Thus, on average, the day is 14.546 hrs long during the spring and summer.

In Figure 4.8, we show the entire day length cycle over one year. It is left as an exercise for the reader to show that the average value of f over the entire year is 12 hrs. (This makes intuitive sense, since overall, the short days in winter will average out with the longer days in summer.)

Figure 4.8 also shows geometrically what the average value of the function represents. Suppose we determine the area associated with the graph of $f(x)$ over the interval of interest. (This area is painted red (dark) in Figure 4.8, where the interval is $0 \leq t \leq 365$, i.e. the whole year.) Now let us draw in a rectangle over the same interval ($0 \leq t \leq 365$) having the same total area. (See the big rectangle in Figure 4.8, and note that its area matches with the darker, red region.) The height of the rectangle represents the average value of $f(t)$ over the interval of interest.

4.7 Summary

In this chapter, we arrived at a number of practical applications of the definite integral.

1. In Section 4.2, we found that for motion at constant acceleration a, the displacement of a moving object can be obtained by integrating twice: the definite integral of acceleration is the velocity $v(t)$, and the definite integral of the velocity is the displacement.

$$v(t) = v_0 + \int_0^t a\, ds. \quad x(t) = x(0) + \int_0^t v(s)\, ds.$$

   (Here we use the "dummy variable" s inside the integral, but the meaning is, of course, the same as in the previous presentation of the formulae.) We showed that at any time t, the position of an object initially at x_0 with velocity v_0 is

$$x(t) = x_0 + v_0 t + a\frac{t^2}{2}.$$

2. We extended our analysis of a moving object to the case of non-constant acceleration (Section 4.2.4), when air resistance tends to produce a drag force to slow the motion of a falling object. We found that in that case, the acceleration gradually decreases, $a(t) = ge^{-kt}$. (The decaying exponential means that $a \to 0$ as t increases.) Again, using the definite integral, we could compute a velocity,

$$v(t) = \int_0^t a(s)\, ds = \frac{g}{k}(1 - e^{-kt}).$$

3. We illustrated the connection between rates of change (over time) and total change (between on time point and another). In general, we saw that if $r(t)$ represents a rate of change of some process, then

$$\int_a^b r(s)\, ds = \text{Total change over the time interval } a \leq t \leq b.$$

   This idea was discussed in Section 4.3.

4. In the concluding Section 4.6, we discussed the average value of a function $f(x)$ over some interval $a \leq x \leq b$,

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

   In the next few chapters we encounters a vast assortment of further examples and practical applications of the definite integral, to such topics as mass, volumes, length, etc. In some of these we will be called on to "dissect" a geometric shape into pieces that are not simple rectangles. The essential idea of an integral as a sum of many (infinitesimally) small pieces will, nevertheless be the same.

Chapter 5

Applications of the definite integral to calculating volume, mass, and length

5.1 Introduction

In this chapter, we consider applications of the definite integral to calculating geometric quantities such as volumes of geometric solids, masses, centers of mass, and lengths of curves.

The idea behind the procedure described in this chapter is closely related to those we have become familiar with in the context of computing areas. That is, we first imagine an approximation using a finite number of pieces to represent a desired result. Then, a limiting process of refinement leads to the desired result. The technology of the definite integral, developed in Chapters 2 and 3 applies directly. This means that we need not re-derive the link between Riemann Sums and the definite integral, we can use these as we did in Chapter 4.

In the first parts of this chapter, we will calculate the total mass of a continuous density distribution. In this context, we will also define the concept of a center of mass. We will first motivate this concept for a discrete distribution made up of a number of finite masses. Then we show how the same concept can be applied in the continuous case. The connection between the discrete and continuous representation will form an important link with our study of analogous concepts in probability, in Chapters 7 and 8.

In the second part of this chapter, we will consider how to dissect certain three dimensional solids into a set of simpler parts whose volumes are easy to compute. We will use familiar formulae for the volumes of disks and cylindrical shells, and carefully construct a summation to represent the desired volume. The volume of the entire object will then be obtained by summing up volumes of a stack of disks or a set of embedded shells, and considering the limit as the thickness of the dissection cuts gets thinner. There are some important differences between material in this chapter and in previous chapters. Calculating volumes will stretch our imagination, requiring us to visualize 3-dimensional (3D) objects, and how they can be subdivided into shells or slices. Most of our effort will be aimed at understanding how to set up the needed integral. We provide a number of examples of this procedure, but first we review the basics of elementary volumes that will play the dominant role in our calculations.

5.2 Mass distributions in one dimension

We start our discussion with a number of example of mass distributed along a single dimension. First, we consider a discrete collection of masses and then generalize to a continuous density. We will be interested in computing the total mass (by summation or integration) as well as other properties of the distribution.

In considering the example of mass distributions, it becomes an easy step to develop the analogous concepts for continuous distributions. This allows us to recapitulate the link between finite sums and definite integrals that we developed in earlier chapters. Examples in this chapter also further reinforce the idea of density (in the context of mass density). Later, we will find that these ideas are equally useful in the context of probability, explored in Chapters 7 and 8.

5.2.1 A discrete distribution: total mass of beads on a wire

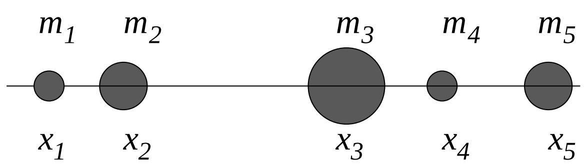

Figure 5.1. *A discrete distribution of masses along a (one dimensional) wire.*

In Figure 5.1 we see a number of beads distributed along a thin wire. We will label each bead with an index, $i = 1 \dots n$ (there are five beads so that $n = 5$). Each bead has a certain position (that we will think of as the value of x_i) and a mass that we will call m_i. We will think of this arrangement as a *discrete mass distribution*: both the masses of the beads, and their positions are of interest to us. We would like to describe some properties of this distribution.

The total mass of the beads, M, is just the sum of the individual masses, so that

$$M = \sum_{i=1}^{n} m_i. \tag{5.1}$$

5.2.2 A continuous distribution: mass density and total mass

We now consider a continuous mass distribution where the mass per unit length ("density") changes gradually from one point to another. For example, the bar in Figure 5.2 has a density that varies along its length.

The portion at the left is made of lighter material, or has a lower density than the portions further to the right. We will denote that density by $\rho(x)$ and this carries units of mass per unit length. (The density of the material along the length of the bar is shown in the graph.) How would we find the total mass?

Suppose the bar has length L and let x $(0 \leq x \leq L)$ denote position along that bar. Let us imagine dividing up the bar into small pieces of length Δx as shown in Figure 5.2.

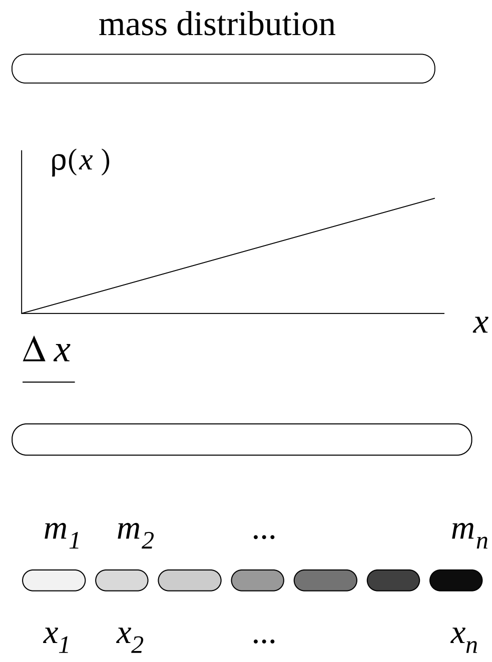

Figure 5.2. *Top: A continuous mass distribution along a one dimensional bar, discussed in Example 5.3.3. The density of the bar (mass per unit length), $\rho(x)$ is shown on the graph. Bottom: the discretized approximation of this same distribution. Here we have subdivided the bar into n smaller pieces, each of length Δx. The mass of each piece is approximately $m_k = \rho(x_k)\Delta x$ where $x_k = k\Delta x$. The total mass of the bar ("sum of all the pieces") will be represented by an integral (5.2) as we let the size, Δx, of the pieces become infinitesimal.*

The coordinates of the endpoints of those pieces are then

$$x_0 = 0, \ldots, x_k = k\Delta x, \quad \ldots, \quad x_N = L$$

and the corresponding masses of each of the pieces are approximately

$$m_k = \rho(x_k)\Delta x.$$

(Observe that units are correct, that is mass$_k$=(mass/length)· length. Note that Δx has units of length.) The total mass is then a sum of masses of all the pieces, and, as we have seen in an earlier chapter, this sum will approach the integral

$$M = \int_0^L \rho(x)dx \tag{5.2}$$

as we make the size of the pieces smaller.

We can also define a **cumulative function** for the mass distribution as

$$M(x) = \int_0^x \rho(s)ds. \tag{5.3}$$

Then $M(x)$ is the total mass in the part of the interval between the left end (assumed at 0) and the position x. The idea of a cumulative function becomes useful in discussions of probability in Chapter 8.

5.2.3 Example: Actin density inside a cell

Biologists often describe the density of protein, receptors, or other molecules in cells. One example is shown in Fig. 5.3. Here we show a keratocyte, which is a cell from the scale of a fish. A band of actin filaments (protein responsible for structure and motion of the

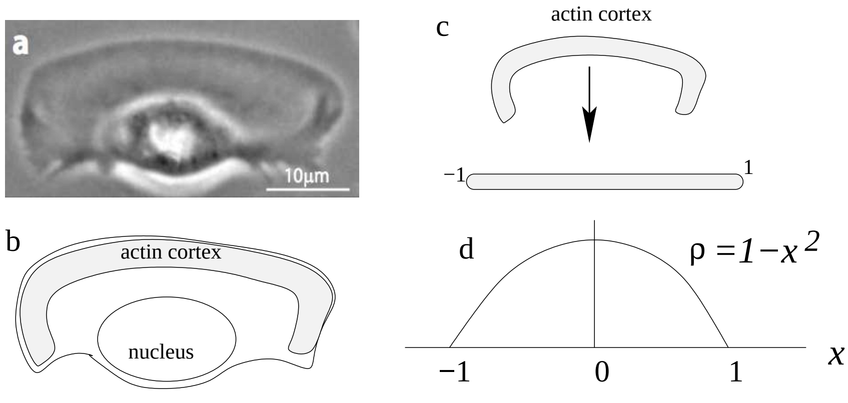

Figure 5.3. *A cell (keratocyte) shown in (a) has a dense distribution of actin in a band called the actin cortex. In (b) we show a schematic sketch of the actin cortex (shaded). In (c) that band of actin is scaled and straightened out so that it occupies a length corresponding to the interval $-1 \le x \le 1$. We are interested in the distribution of actin filaments across that band. That distribution is shown in (d). Note that actin is densest in the middle of the band. (a) Credit to Alex Mogilner.*

cell) are found at the edge of the cell in a band called the **actin cortex**. It has been found experimentally that the density of actin is greatest in the middle of the band, i.e. the position corresponding to the midpoint of the edge of the cell shown in Fig. 5.3a. According to Alex Mogilner[19], the density of actin across the cortex in filaments per edge μm is well approximated by a distribution of the form

$$\rho(x) = \alpha(1 - x^2) \quad -1 \le x \le 1,$$

where x is the fraction of distance[20] from midpoint to the end of the band (Fig. 5.3c and d). Here $\rho(x)$ is an actin filament density in units of filaments per μm. That is, ρ is the number

[19] Alex Mogilner is a professor of mathematics who specializes in cell biology and the actin cytoskeleton

[20] Note that 1μm (read " 1 micro-meter" or "micron") is 10^{-6}meters, and is appropriate for measuring lengths of small objects such as cells.

of actin fibers per unit length.

We can find the total number of actin filaments, N in the band by integration, i.e.

$$N = \int_{-1}^{1} \alpha(1-x^2)\,dx = \alpha \int_{-1}^{1} (1-x^2)\,dx.$$

The integral above has already been computed (Integral 2.) in the Examples 3.6.2 of Chapter 3 and was found to be 4/3. Thus, we have that there are $N = 4\alpha/3$ actin filaments in the band.

5.3 Mass distribution and the center of mass

It is useful to describe several other properties of mass distributions. We first define the "center of mass", $\bar{x}$ which is akin to an average x coordinate.

5.3.1 Center of mass of a discrete distribution

The center of mass $\bar{x}$ of a mass distribution is given by:

$$\boxed{\bar{x} = \frac{1}{M}\sum_{i=1}^{n} x_i m_i\ .}$$

This can also be written in the form

$$\boxed{\bar{x} = \frac{\sum_{i=1}^{n} x_i m_i}{\sum_{i=1}^{n} m_i}.}$$

5.3.2 Center of mass of a continuous distribution

We can generalize the concept of the center of mass for a continuous mass density. Our usual approach of subdividing the interval $0 \le x \le L$ and computing a Riemann sum leads to

$$\bar{x} = \frac{1}{M}\sum_{i=1}^{n} x_i \rho(x_i)\Delta x.$$

As $\Delta x \to dx$, this becomes an integral. Based on this, it makes sense to define the *center of mass* of the continuous mass distribution as follows:

$$\boxed{\bar{x} = \frac{1}{M}\int_0^L x\rho(x)dx\ .}$$

We can also write this in the form

$$\bar{x} = \frac{\int_0^L x\rho(x)dx}{\int_0^L \rho(x)dx}.$$

5.3.3 Example: Center of mass vs average mass density

Here we distinguish between two (potentially confusing) quantities in the context of an example.

A long thin bar of length L is made of material whose density varies along the length of the bar. Let x be distance from one end of the bar. Suppose that the mass density is given by

$$\rho(x) = ax, \quad 0 \le x \le L.$$

This type of mass density is shown in a panel in Fig. 5.2.

(a) Find the total mass of the bar.

(b) Find the average mass density along the bar.

(c) Find the center of mass of the bar.

(d) Where along the length of the bar should you cut to get two pieces of equal mass?

Solution

(a) From our previous discussion, the total mass of the bar is

$$M = \int_0^L ax\,dx = \left.\frac{ax^2}{2}\right|_0^L = \frac{aL^2}{2}.$$

(b) The average mass density along the bar is computed just as one would compute the average value of a function: integrate the function over an interval and divide by the length of the interval. An example of this type appeared in Section 4.6. Thus

$$\bar{\rho} = \frac{1}{L}\int_0^L \rho(x)\,dx = \frac{1}{L}\left(\frac{aL^2}{2}\right) = \frac{aL}{2}$$

A bar having a uniform density $\bar{\rho} = aL/2$ would have the same total mass as the bar in this example. (This is the physical interpretation of average mass density.)

(c) The center of mass of the bar is

$$\bar{x} = \frac{\int_0^L xp(x)\,dx}{M} = \frac{1}{M}\int_0^L ax^2\,dx = \left.\frac{a}{M}\frac{x^3}{3}\right|_0^L = \frac{2a}{aL^2}\frac{L^3}{3} = \frac{2}{3}L.$$

Observe that the center of mass is an “average x coordinate”, which is not the same as the average mass density.

(d) We can use the cumulative function defined in Eqn. (5.3) to figure out where half of the mass is concentrated. Suppose we cut the bar at some position $x = s$. Then the mass of this part of the bar is

$$M_1 = \int_0^s \rho(x)\,dx = \frac{as^2}{2},$$

We ask for what values of s is it true that M_1 is exactly half the total mass? Using the result of part (a), we find that for this to be true, we must have

$$M_1 = \frac{M}{2}, \quad \Rightarrow \quad \frac{as^2}{2} = \frac{1}{2}\frac{aL^2}{2}$$

Solving for s leads to

$$s = \frac{1}{\sqrt{2}}L = \frac{\sqrt{2}}{2}L.$$

Thus, cutting the bar at a distance $(\sqrt{2}/2)L$ from $x = 0$ results in two equal masses.

Remark: the position that subdivides the mass into two equal pieces is analogous to the idea of a median. This concept will appear again in the context of probability in Chapter 8.

5.3.4 Physical interpretation of the center of mass

The center of mass has a physical interpretation: it is the point at which the mass would "balance". In the Appendix 11.3 we discuss this in detail.

5.4 Miscellaneous examples and related problems

The idea of mass density can be extended to related problems of various kinds. We give some examples in this section.

Up to now, we have seen examples of mass distributed in one dimension: beads on a wire, actin density along the edge of a cell, (in Chapter 4), or a bar of varying density. For the continuous distributions, we determined the total mass by integration. Underlying the integral we computed was the idea that the interval could be "dissected" into small parts (of width Δx), and a sum of pieces transformed into an integral. In the next examples, we consider similar ideas, but instead of dissecting the region into 1-dimensional intervals, we have slightly more interesting geometries.

5.4.1 Example: A glucose density gradient

A cylindrical test-tube of radius r, and height h, contains a solution of glucose which has been prepared so that the concentration of glucose is greatest at the bottom and decreases gradually towards the top of the tube. (This is called a *density gradient*). Suppose that the concentration c as a function of the depth x is $c(x) = 0.1 + 0.5x$ grams per centimeter^3. ($x = 0$ at the top of the tube, and $x = h$ at the bottom of the tube.) In Figure 5.4 we show a schematic version of what this gradient might look like. (In reality, the transition between high and low concentration would be smoother than shown in this figure.) Determine the total amount of glucose in the tube (in gm). Neglect the "rounded" lower portion of the tube: i.e. assume that it is a simple cylinder.

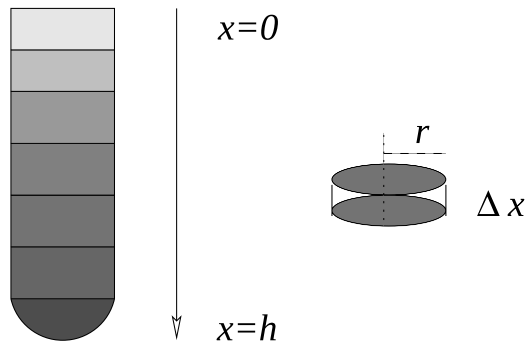

Figure 5.4. *A test-tube of radius r containing a gradient of glucose. A disk-shaped slice of the tube with small thickness Δx has approximately constant density.*

Solution

We assume a simple cylindrical tube and consider imaginary "slices" of this tube along its vertical axis, here labeled as the "x" axis. Suppose that the thickness of a slice is Δx. Then the volume of each of these (disk shaped) slices is $\pi r^2 \Delta x$. The amount of glucose in the slice is approximately equal to the concentration $c(x)$ multiplied by the volume of the slice, i.e. the small slice contains an amount $\pi r^2 \Delta x c(x)$ of glucose. In order to sum up the total amount over all slices, we use a definite integral. (As before, we imagine $\Delta x \to dx$ becoming "infinitesimal" as the number of slices increases.) The integral we want is

$$G = \pi r^2 \int_0^h c(x)\, dx.$$

Even though the geometry of the test-tube, at first glance, seems more complicated than the one-dimensional highway described in Chapter 4, we observe here that the integral that computes the total amount is still a sum over a single spatial variable, x. (Note the resemblance between the integrals

$$I = \int_0^L C(x)\, dx \quad \text{and} \quad G = \pi r^2 \int_0^h c(x)\, dx,$$

here and in the previous example.) We have neglected the complication of the rounded bottom portion of the test-tube, so that integration over its length (which is actually summation of disks shown in Figure 5.4) is a one-dimensional problem.

In this case the total amount of glucose in the tube is

$$G = \pi r^2 \int_0^h (0.1 + 0.5x) dx = \pi r^2 \left(0.1x + \frac{0.5x^2}{2} \right) \Bigg|_0^h = \pi r^2 \left(0.1h + \frac{0.5h^2}{2} \right).$$

Suppose that the height of the test-tube is $h = 10$ cm and its radius is $r = 1$ cm. Then the total mass of glucose is

$$G = \pi \left(0.1(10) + \frac{0.5(100)}{2} \right) = \pi\,(1 + 25) = 26\pi \text{ gm}.$$

In the next example, we consider a circular geometry, but the concept of dissecting and summing is the same. Our task is to determine how to set up the problem in terms of an integral, and, again, we must imagine which type of subdivision would lead to the summation (integration) needed to compute total amount.

5.4.2 Example: A circular colony of bacteria

A circular colony of bacteria has radius of $1\ cm$. At distance r from the center of the colony, the density of the bacteria, in units of one million cells per square centimeter, is observed to be $b(r) = 1 - r^2$ (Note: r is distance from the center in cm, so that $0 \le r \le 1$). What is the total number of bacteria in the colony?

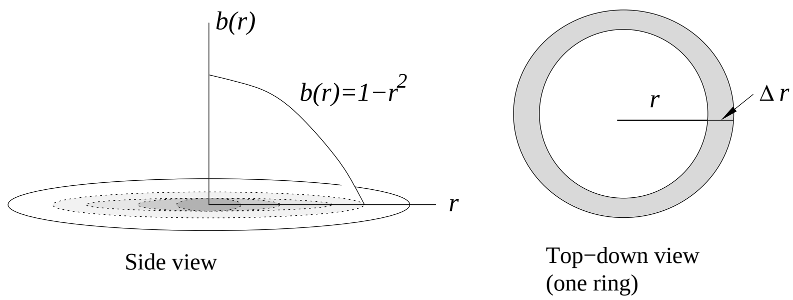

Figure 5.5. *A colony of bacteria with circular symmetry. A ring of small thickness Δr has roughly constant density. The superimposed curve on the left is the bacterial density $b(r)$ as a function of the radius r.*

Solution

Figure 5.5 shows a rough sketch of a flat surface with a colony of bacteria growing on it. We assume that this distribution is radially symmetric. The density as a function of distance from the center is given by $b(r)$, as shown in Figure 5.5. Note that the function describing density, $b(r)$ is smooth, but to accentuate the strategy of dissecting the region, we have shown a top-down view of a ring of nearly constant density on the right in Figure 5.5. We see that this ring occupies the region between two circles, e.g. between a circle of radius r and a slightly bigger circle of radius $r + \Delta r$. The area of that "ring"[21] would then be the area of the larger circle minus that of the smaller circle, namely

$$A_{\text{ring}} = \pi(r + \Delta r)^2 - \pi r^2 = \pi(2r\Delta r + (\Delta r)^2).$$

However, if we make the thickness of that ring really small ($\Delta r \to 0$), then the quadratic term is very very small so that

$$A_{\text{ring}} \approx 2\pi r \Delta r.$$

[21]Students commonly make the error of writing $A_{\text{ring}} = \pi(r + \Delta r - r)^2 = \pi(\Delta r)^2$. This trap should be avoided! It is clear that the correct expression has additional terms, since we really are computing a difference of two circular areas.

Consider all the bacteria that are found inside a "ring" of radius r and thickness Δr (see Figure 5.5.) The total number within such a ring is the product of the density, $b(r)$ and the area of the ring, i.e.

$$b(r) \cdot (2\pi r \Delta r) = 2\pi r(1 - r^2)\Delta r.$$

To get the total number in the colony we sum up over all the rings from $r = 0$ to $r = 1$ and let the thickness, $\Delta r \to dr$ become very small. But, as with other examples, this is equivalent to calculating a definite integral, namely:

$$B_{\text{total}} = \int_0^1 (1 - r)(2\pi r)\, dr = 2\pi \int_0^1 (1 - r^2) r dr = 2\pi \int_0^1 (r - r^3) dr.$$

We calculate the result as follows:

$$B_{\text{total}} = 2\pi \left(\frac{r^2}{2} - \frac{r^4}{4} \right) \Big|_0^1 = (\pi r^2 - \pi \frac{r^4}{2}) \Big|_0^1 = \pi - \frac{\pi}{2} = \frac{\pi}{2}.$$

Thus the total number of bacteria in the entire colony is $\pi/2$ million which is approximately 1.57 million cells.

5.5 Volumes of solids of revolution

We now turn to the problem of calculating volumes of 3D solids. Here we restrict attention to symmetric objects denoted *solids of revolution*. The outer surface of these objects is generated by revolving some curve around a coordinate axis. In Figure 5.7 we show one such curve, and the surface it forms when it is revolved about the y axis.

5.5.1 Volumes of cylinders and shells

Before starting the calculation, let us recall the volumes of some of the geometric shapes that are to be used as elementary pieces into which our shapes will be carved. See Figure 5.6.

1. The **volume of a cylinder** of height h having circular base of radius r, is

$$V_{\text{cylinder}} = \pi r^2 h.$$

2. The **volume of a circular disk** of thickness τ, and radius r (shown on the left in Figure 5.6), is a special case of the above,

$$V_{\text{disk}} = \pi r^2 \tau.$$

3. The **volume of a cylindrical shell** of height h, with circular radius r and small thickness τ (shown on the right in Figure 5.6) is

$$V_{\text{shell}} = 2\pi r h \tau.$$

(This approximation holds for $\tau << r$.)

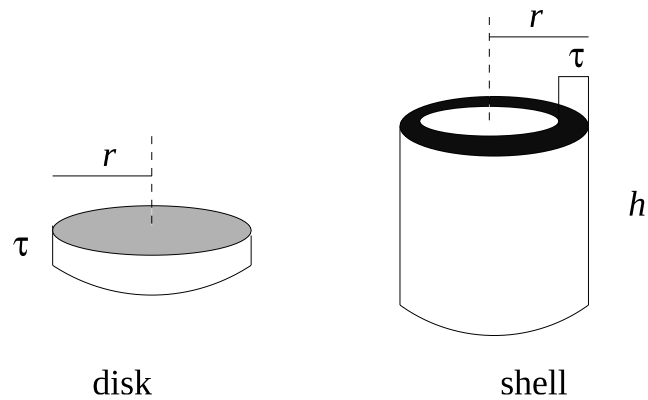

Figure 5.6. *The volumes of these simple 3D shapes are given by simple formulae. We use them as basic elements in computing more complicated volumes. Here we will present examples based on disks. In Appendix 11.4 we give an example based on shells.*

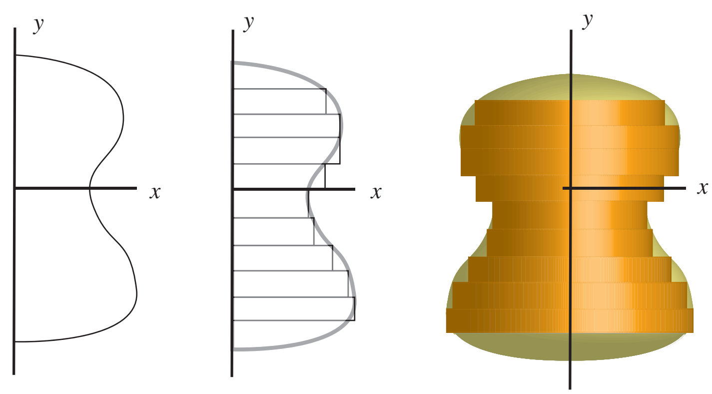

Figure 5.7. *A solid of revolution is formed by revolving a region in the xy-plane about the y-axis. We show how the region is approximated by rectangles of some given width, and how these form a set of approximating disks for the 3D solid of revolution.*

5.5.2 Computing the Volumes

Consider the curve in Figure 5.7 and the surface it forms when it is revolved about the y axis. In the same figure, we also show how a set of approximating rectangular strips associated with the planar region (grey rectangles) lead to a set of stacked disks (orange shapes) that approximate the volume of the solid (greenish object in Fig. 5.7). The total volume of the disks is not the same as the volume of the object but if we make the thickness of these disks very small, the approximation of the true volume is good. In the limit, as the thickness of the disks becomes infinitesimal, we arrive at the true volume of the solid of revolution. The reader should recognize a familiar theme. We used the same concept in computing areas using Riemann sums based on rectangular strips in Chapter 2.

Fig. 5.8 similarly shows a volume of revolution obtained by revolving the graph of the function $y = f(x)$ about the x axis. We note that if this surface is cut into slices, the radius of the cross-sections depend on the position of the cut. Let us imagine a stack of disks approximating this volume. One such disk has been pulled out and labeled for our inspection. We note that its radius (in the y direction) is given by the height of the graph of the function, so that $r = f(x)$. The thickness of the disk (in the x direction) is Δx. The volume of this single disk is then $v = \pi[f(x)]^2\Delta x$. Considering this disk to be based at the k'th coordinate point in the stack, i.e. at x_k, means that its volume is

$$v_k = \pi[f(x_k)]^2\Delta x.$$

Summing up the volumes of all such disks in the stack leads to the total volume of disks

$$V_{\text{disks}} = \sum_k \pi[f(x_k)]^2\Delta x.$$

When we increase the number of disks, making each one thinner so that $\Delta x \to 0$, we

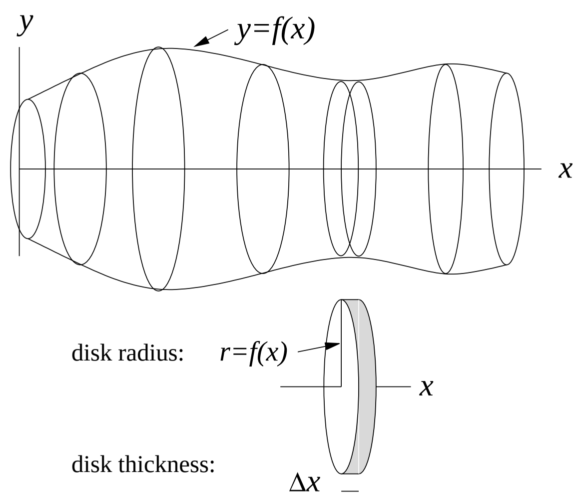

Figure 5.8. *Here the solid of revolution is formed by revolving the curve $y = f(x)$ about the y axis. A typical disk used to approximate the volume is shown. The radius of the disk (parallel to the y axis) is $r = y = f(x)$. The thickness of the disk (parallel to the x axis) is Δx. The volume of this disk is hence $v = \pi[f(x)]^2\Delta x$*

arrive at a definite integral,

$$V = \int_a^b \pi[f(x)]^2 dx.$$

In most of the examples discussed in this chapter, the key step is to make careful observation of the way that the radius of a given disk depends on the function that *generates* the surface.

(By this we mean the function that specifies the curve that forms the surface of revolution.) We also pay attention to the dimension that forms the disk thickness, Δx.

Some of our examples will involve surfaces revolved about the x axis, and others will be revolved about the y axis. In setting up these examples, a diagram is usually quite helpful.

Example 1: Volume of a sphere

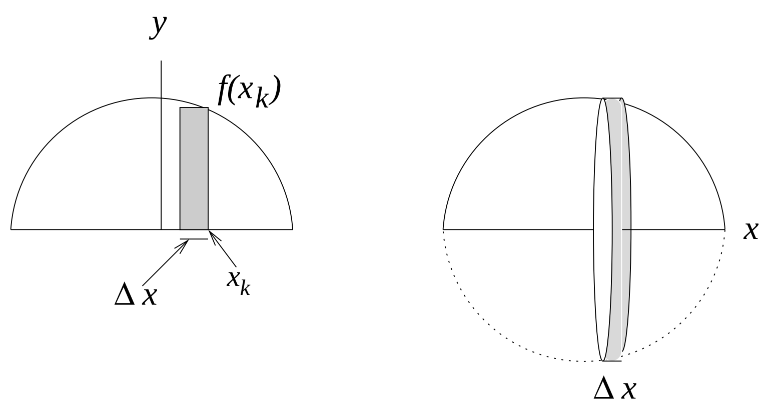

Figure 5.9. *When the semicircle (on the left) is rotated about the x axis, it generates a sphere. On the right, we show one disk generated by the revolution of the shaded rectangle.*

We can think of a sphere of radius R as a solid whose outer surface is formed by rotating a semi-circle about its long axis. A function that describe a semi-circle (i.e. the top half of the circle, $y^2 + x^2 = R^2$) is

$$y = f(x) = \sqrt{R^2 - x^2}.$$

In Figure 5.9, we show the sphere dissected into a set of disks, each of width Δx. The disks are lined up along the x axis with coordinates x_k, where $-R \leq x_k \leq R$. These are just integer multiples of the slice thickness Δx, so for example,

$$x_0 = -R, \quad x_1 = -R + \Delta x, \quad \ldots, \quad x_k = -R + k\Delta x \; .$$

The radius of the disk depends on its position[22]. Indeed, the radius of a disk through the x axis at a point x_k is specified by the function $r_k = f(x_k)$. The volume of the k'th disk is

$$V_k = \pi r_k^2 \Delta x.$$

By the above remarks, using the fact that the function $f(x)$ determines the radius, we have

$$V_k = \pi [f(x_k)]^2 \Delta x,$$

[22] Note that the radius is oriented along the y axis, so sometimes we may write this as $r_k = y_k = f(x_k)$

$$V_k = \pi \left[\sqrt{R^2 - x_k^2} \right]^2 \Delta x = \pi (R^2 - x_k^2) \Delta x.$$

The total volume of *all the disks* is

$$V = \sum_k V_k = \sum_k \pi [f(x_k)]^2 \Delta x = \pi \sum_k (R^2 - x_k^2) \Delta x.$$

as $\Delta x \to 0$, this sum becomes a definite integral, and represents the true volume. We start the summation at $x = -R$ and end at $x_N = R$ since the semi-circle extends from $x = -R$ to $x = R$. Thus

$$V_{\text{sphere}} = \int_{\text{-R}}^{R} \pi [f(x_k)]^2 \, dx = \pi \int_{\text{-R}}^{R} (R^2 - x^2) \, dx.$$

We compute this integral using the Fundamental Theorem of calculus, obtaining

$$V_{\text{sphere}} = \pi \left(R^2 x - \frac{x^3}{3} \right) \Big|_{-R}^{R}.$$

Observe that this is twice the volume obtained for the interval $0 < x < R$,

$$V_{\text{sphere}} = 2\pi \left(R^2 x - \frac{x^3}{3} \right) \Big|_{0}^{R} = 2\pi \left(R^3 - \frac{R^3}{3} \right).$$

We often use such symmetry properties to simplify computations. After simplification, we arrive at the familiar formula

$$V_{\text{sphere}} = \frac{4}{3} \pi R^3.$$

Example 2: Volume of a paraboloid

Consider the curve

$$y = f(x) = 1 - x^2.$$

If we rotate this curve about the y axis, we will get a paraboloid, as shown in Figure 5.10. In this section we show how to compute the volume by dissecting into disks stacked up along the y axis.

Solution

The object has the y axis as its axis of symmetry. Hence disks are stacked up along the y axis to approximate this volume. This means that the width of each disk is Δy. This accounts for the dy in the integral below. The volume of each disk is

$$V_{\text{disk}} = \pi r^2 \Delta y,$$

where the radius, r is now in the direction parallel to the x axis. Thus we must express radius as

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

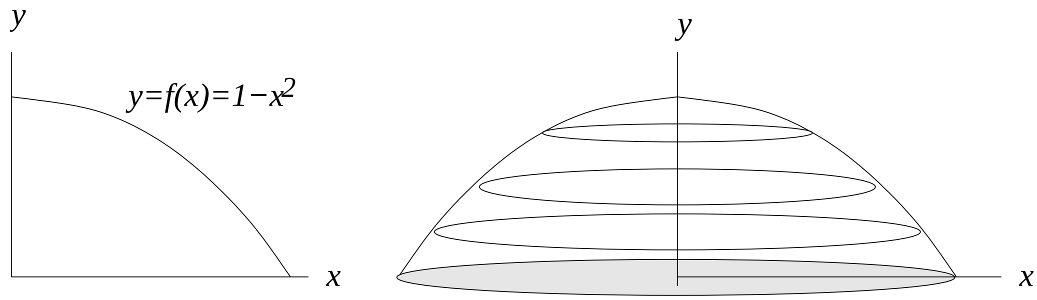

Figure 5.10. *The curve that generates the shape of a paraboloid (left) and the shape of the paraboloid (right).*

i.e, we invert the relationship to obtain x as a function of y. From $y = 1 - x^2$ we have $x^2 = 1 - y$ so $x = \sqrt{1-y}$. The radius of a disk at height y is therefore $r = x = \sqrt{1-y}$. The shape extends from a smallest value of $y = 0$ up to $y = 1$. Thus the volume is

$$V = \pi \int_0^1 [f(y)]^2 \, dy = \pi \int_0^1 [\sqrt{1-y}]^2 \, dy.$$

It is helpful to note that once we have identified the thickness of the disks (Δy), we are guided to write an integral in terms of the variable y, i.e. to reformulate the equation describing the curve. We compute

$$V = \pi \int_0^1 (1-y) \, dy = \pi \left(y - \frac{y^2}{2} \right) \bigg|_0^1 = \pi \left(1 - \frac{1}{2} \right) = \frac{\pi}{2}.$$

The above example was set up using disks. However, there are other options. In Appendix 11.4 we show yet another method, comprised of *cylindrical shells* to compute the volume of a cone. In some cases, one method is preferable to another, but here either method works equally well.

Example 3

Find the volume of the surface formed by rotating the curve

$$y = f(x) = \sqrt{x}, \quad 0 \le x \le 1$$

(a) about the x axis. (b) about the y axis.

Solution

(a) If we rotate this curve about the x axis, we obtain a bowl shape; dissecting this surface leads to disks stacked along the x axis, with thickness $\Delta x \to dx$, with radii in the y direction, i.e. $r = y = f(x)$, and with x in the range $0 \le x \le 1$. The volume will thus be

$$V = \pi \int_0^1 [f(x)]^2 \, dx = \pi \int_0^1 [\sqrt{x}]^2 \, dx = \pi \int_0^1 x \, dx = \pi \frac{x^2}{2}\bigg|_0^1 = \frac{\pi}{2}.$$

(b) When the curve is rotated about the y axis, it forms a surface with a sharp point at the origin. The disks are stacked along the y axis, with thickness $\Delta y \to dy$, and radii in the x direction. We must rewrite the function in the form

$$x = g(y) = y^2.$$

We now use the interval along the y axis, i.e. $0 < y < 1$ The volume is then

$$V = \pi \int_0^1 [f(y)]^2 \, dy = \pi \int_0^1 [y^2]^2 \, dy = \pi \int_0^1 y^4 \, dy = \pi \frac{y^5}{5}\bigg|_0^1 = \frac{\pi}{5}.$$

5.6 Length of a curve: Arc length

Analytic geometry provides a simple way to compute the length of a straight line segment, based on the distance formula[23]. Recall that, given points $P_1 = (x_1, y_1)$ and $P_2 = (x_2, y_2)$, the length of the line joining those points is

$$d = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}.$$

Things are more complicated for "curves" that are not straight lines, but in many cases, we are interested in calculating the length of such curves. In this section we describe how this can be done using the definite integral "technology".

The idea of dissection also applies to the problem of determining the length of a curve. In Figure 5.11, we see the general idea of subdividing a curve into many small "arcs". Before we look in detail at this construction, we consider a simple example, shown in Figure 5.12. In the triangle shown, by the Pythagorean theorem we have the length of the sloped side related as follows to the side lengths $\Delta x, \Delta y$:

$$\Delta \ell^2 = \Delta x^2 + \Delta y^2,$$

$$\Delta \ell = \sqrt{\Delta x^2 + \Delta y^2} = \left(\sqrt{1 + \frac{\Delta y^2}{\Delta x^2}} \right) \Delta x = \sqrt{1 + \left(\frac{\Delta y}{\Delta x}\right)^2} \; \Delta x.$$

We now consider a curve given by some function

$$y = f(x) \quad a < x < b,$$

as shown in Figure 5.11(a). We will approximate this curve by a set of line segments, as shown in Figure 5.11(b). To obtain these, we have selected some step size Δx along the x axis, and placed points on the curve at each of these x values. We connect the points with straight line segments, and determine the lengths of those segments. (The total length

[23]The reader should recall that this formula is a simple application of Pythagorean theorem.

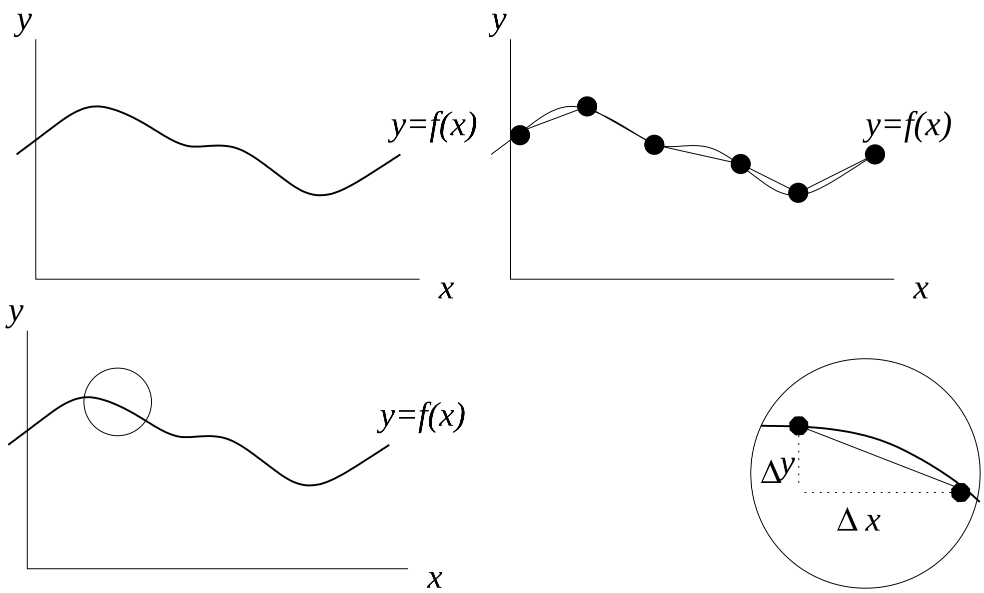

Figure 5.11. *Top: Given the graph of a function, $y = f(x)$ (at left), we draw secant lines connecting points on its graph at values of x that are multiples of Δx (right). Bottom: a small part of this graph is shown, and then enlarged, to illustrate the relationship between the arc length and the length of the secant line segment.*

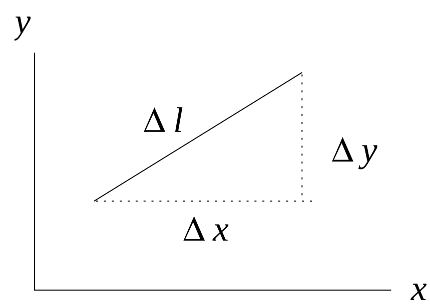

Figure 5.12. *The basic idea of arclength is to add up lengths Δl of small line segments that approximate the curve.*

of the segments is only an approximation of the length of the curve, but as the subdivision gets finer and finer, we will arrive at the true total length of the curve.)

We show one such segment enlarged in the circular inset in Figure 5.11. Its slope, shown at right is given by $\Delta y/\Delta x$. According to our remarks, above, the length of this segment is given by

$$\Delta\ell = \sqrt{1+\left(\frac{\Delta y}{\Delta x}\right)^2}\ \Delta x.$$

As the step size is made smaller and smaller $\Delta x \to dx$, $\Delta y \to dy$ and

$$\Delta\ell \to \sqrt{1+\left(\frac{dy}{dx}\right)^2}\ dx.$$

We recognize the ratio inside the square root as as the derivative, dy/dx. If our curve is given by a function $y = f(x)$ then we can rewrite this as

$$d\ell = \sqrt{1+(f'(x))^2}\ dx.$$

Thus, the length of the entire curve is obtained from summing (i.e. adding up) these small pieces, i.e.

$$L = \int_a^b \sqrt{1+(f'(x))^2}\ dx. \tag{5.4}$$

Example 1

Find the length of a line whose slope is -2 given that the line extends from $x = 1$ to $x = 5$.

Solution

We could find the equation of the line and use the distance formula. But for the purpose of this example, we apply the method of Equation (5.4): we are given that the slope $f'(x)$ is -2. The integral in question is

$$L = \int_1^5 \sqrt{1+(f'(x))^2}\, dx = \int_1^5 \sqrt{1+(-2)^2}\, dx = \int_5^1 \sqrt{5}\, dx.$$

We get

$$L = \sqrt{5}\int_1^5 dx = \sqrt{5}x\Big|_1^5 = \sqrt{5}[5-1] = 4\sqrt{5}.$$

Example 2

Find an integral that represents the length of the curve that forms the graph of the function

$$y = f(x) = x^3, \quad 1 < x < 2.$$

Solution

We find that

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

Thus, the integral is

$$L = \int_1^2 \sqrt{1 + (3x^2)^2}\, dx = \int_1^2 \sqrt{1 + 9x^4}\, dx.$$

At this point, we will not attempt to find the actual length, as we must first develop techniques for finding the anti-derivative for functions such as $\sqrt{1 + 9x^4}$.

Using the spreadsheet to calculate arclength

Most integrals for arclength contain square roots and functions that are not easy to integrate, simply because their antiderivatives are difficult to determine. However, now that we know the idea behind determining the length of a curve, we can apply the ideas developed have to approximate the length of a curve "numerically". The spreadsheet is a simple tool for doing the necessary summations.

As an example, we show here how to calculate the length of the curve

$$y = f(x) = 1 - x^2 \text{ for } 0 \le x \le 1$$

using a simple numerical procedure implemented on the spreadsheet.

Let us choose a step size of $\Delta x = 0.1$ along the x axis, for the interval $0 \le x \le 1$. We calculate the function, the slopes of the little segments (change in y divided by change in x), and from this, compute the length of each segment

$$\Delta \ell = \sqrt{1 + (\Delta y/\Delta x)^2}\ \Delta x$$

and the accumulated length along the curve from left to right, L which is just a sum of such values. The Table 5.6 shows steps in the calculation of the ratio $\Delta y/\Delta x$, the value of $\Delta \ell$, the cumulative sum, and, finally the total length L. The final value of $L = 1.4782$ represents the total length of the curve over the entire interval $0 < x < 1$.

In Figure 5.13(a) we show the actual curve $y = 1 - x^2$. with points placed on it at each multiple of Δx. In Figure 5.13(b), we show (in blue) how the lengths of the little straight-line segments connecting these points changes across the interval. (The segments on the left along the original curve are nearly flat, so their length is very close to Δx. The segments on the right part of the curve are much more sloped, and their lengths are thus bigger.) We also show (in red) how the total accumulated length L depends on the position x across the interval. This function represents the total arc-length of the curve $y = 1 - x^2$, from $x = 0$ up to a given x value. At $x = 1$ this function returns the value $y = L$, as it has added up the full length of the curve for $0 \le x \le 1$.

5.6.1 How the alligator gets its smile

The American alligator, *Alligator mississippiensis* has a set of teeth best viewed at some distance. The regular arrangement of these teeth, i.e. their spacing along the jaw is important in giving the reptile its famous bite. We will concern ourselves here with how that pattern of teeth is formed as the alligator develops from its embryonic stage to that of an

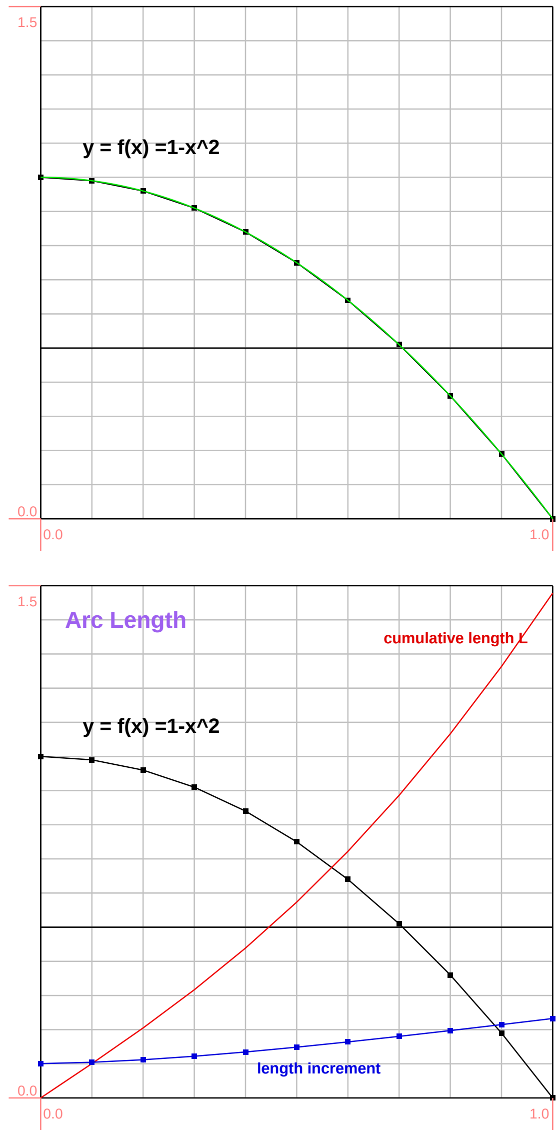

Figure 5.13. *The spreadsheet can be used to compute approximate values of integrals, and hence to calculate arclength. Shown here is the graph of the function* $y = f(x) = 1 - x^2$ *for* $0 \leq x \leq 1$*, together with the length increment and the cumulative arclength along that curve.*

x	$y = f(x)$	$\Delta y/\Delta x$	$\Delta\ell$	$L = \sum \Delta\ell$
0. 0	1.0000	-0.1	0.1005	0.0000
0.1	0.9900	-0.3	0.1044	0.1005
0.2	0.9600	-0.5	0.1118	0.2049
0.3	0.9100	-0.7	0.1221	0.3167
0.4	0.8400	-0.9	0.1345	0.4388
0.5	0.7500	-1.1	0.1487	0.5733
0.6	0.6400	-1.3	0.1640	0.7220
0.7	0.5100	-1.5	0.1803	0.8860
0.8	0.3600	-1.7	0.1972	1.0663
0.9	0.1900	-1.9	0.2147	1.2635
1. 0	0.0000	-2.1	0.2326	1.4782

Table 5.1. *For the function $y = f(x) = 1 - x^2$, and $0 \le x \le 1$, we show how to calculate an approximation to the arc-length using the spreadsheet.*

adult. As is the case in humans, the teeth on an alligator do not form or sprout simultaneously. In the development of the baby alligator, there is a sequence of initiation of teeth, one after the other, at well-defined positions along the jaw.

Paul Kulesa, a former student of James D Muray, set out to understand the pattern of development of these teeth, based on data in the literature about what happens at distinct stages of embryonic growth. Of interest in his research were several questions, including what determines the positions and timing of initiation of individual teeth, and what mechanisms lead to this pattern of initiation. One theory proposed by this group was that chemical signals that diffuse along the jaw at an early stage of development give rise to instructions that are interpreted by jaw cells: where the signal is at a high level, a tooth will start to initiate.

While we will not address the details of the mechanism of development here, we will find a simple application of the ideas of arclength in the developmental sequence of teething. Shown in Figure 5.14 is a smiling baby alligator (no doubt thinking of some future tasty meal). A close up of its smile (at an earlier stage of development) reveals the shape of the jaw, together with the sites at which teeth are becoming evident. (One of these sites, called primordia, is shown enlarged in an inset in this figure.)

Paul Kulesa found that the shape of the alligator's jaw can be described remarkably well by a parabola. A proper choice of coordinate system, and some experimentation leads to the equation of the best fit parabola

$$y = f(x) = -ax^2 + b$$

where $a = 0.256$, and $b = 7.28$ (in units not specified).

We show this curve in Figure 5.15(a). Also shown in this curve is a set of points at which teeth are found, labelled by order of appearance. In Figure 5.15(b) we see the same curve, but we have here superimposed the function $L(x)$ given by the arc length along the curve from the front of the jaw (i.e. the top of the parabola), i.e.

$$L(x) = \int_0^x \sqrt{1 + [f'(s)]^2}\, ds.$$

This curve measures distance along the jaw, from front to back. The distances of the teeth from one another, or along the curve of the jaw can be determined using this curve if we know the x coordinates of their positions.

The table below gives the original data, courtesy of Dr. Kulesa, showing the order of the teeth, their (x, y) coordinates, and the value of $L(x)$ obtained from the arclength formula. We see from this table that the teeth do not appear randomly, nor do they fill in the jaw in one sweep. Rather, they appear in several stages.

In Figure 5.15(c), we show the pattern of appearance: Plotting the distance along the jaw of successive teeth reveals that the teeth appear in waves of nearly equally-spaced sites. (By equally spaced, we refer to distance along the parabolic jaw.) The first wave (teeth 1, 2, 3) are followed by a second wave (4, 5, 6, 7), and so on. Each wave forms a linear pattern of distance from the front, and each successive wave fills in the gaps in a similar, equally spaced pattern.

The true situation is a bit more complicated: the jaw grows as the teeth appear as shown in 5.15(c). This has not been taken into account in our simple treatment here, where we illustrate only the essential idea of arc length application.

Tooth number	position		distance along jaw
	x	y	$L(x)$
1	1.95	6.35	2.1486
2	3.45	4.40	4.7000
3	4.54	2.05	7.1189
4	1.35	6.95	1.4000
5	2.60	5.50	3.2052
6	3.80	3.40	5.4884
7	5.00	1.00	8.4241
8	3.15	4.80	4.1500
9	4.25	2.20	6.3923
10	4.60	1.65	7.3705
11	0.60	7.15	0.6072
12	3.45	4.05	4.6572
13	5.30	0.45	9.2644

Table 5.2. *Data for the appearance of teeth, in the order in which they appear as the alligator develops. We can use arc-length computations to determine the distances between successive teeth.*

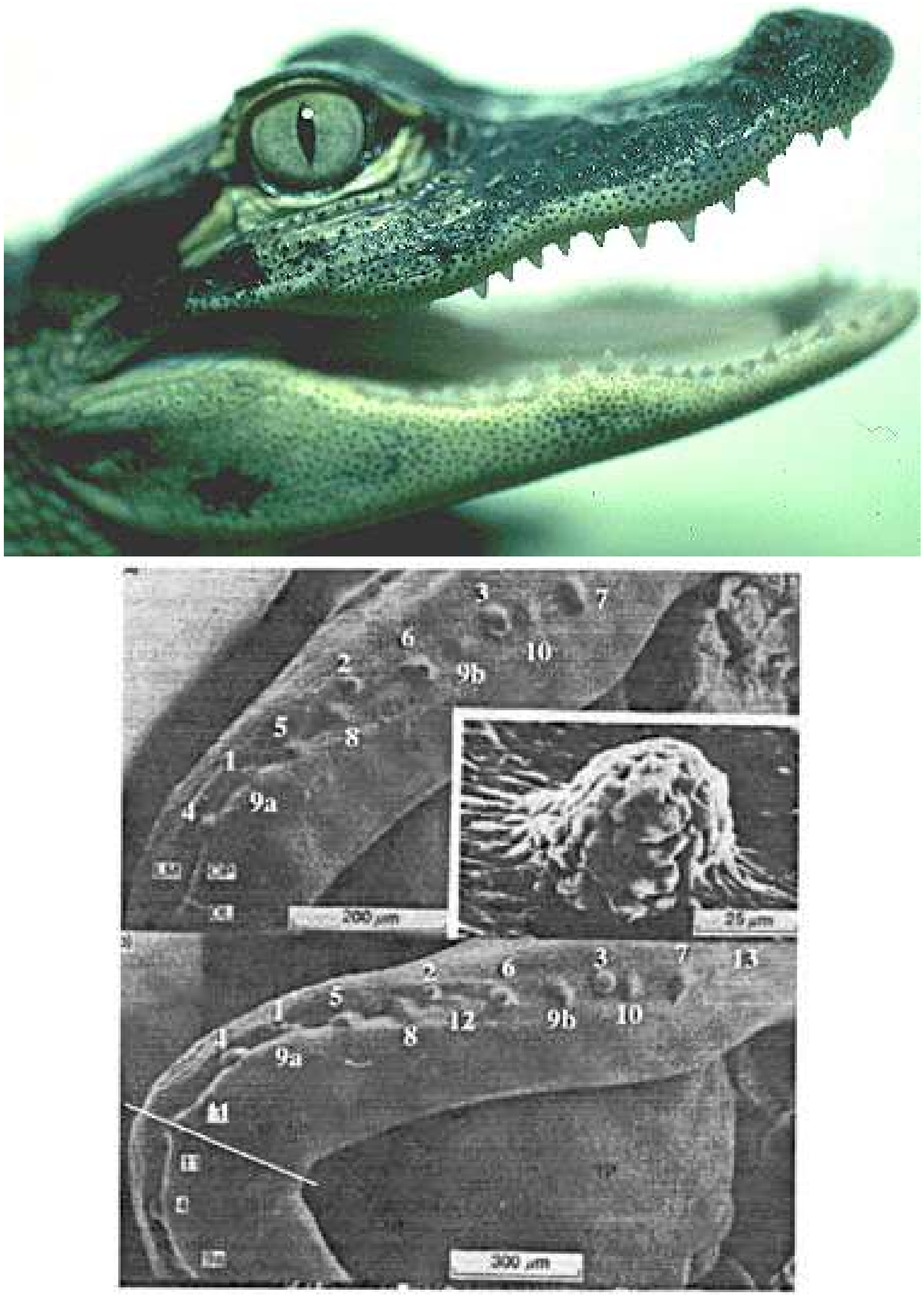

Figure 5.14. Alligator mississippiensis *and its teeth*

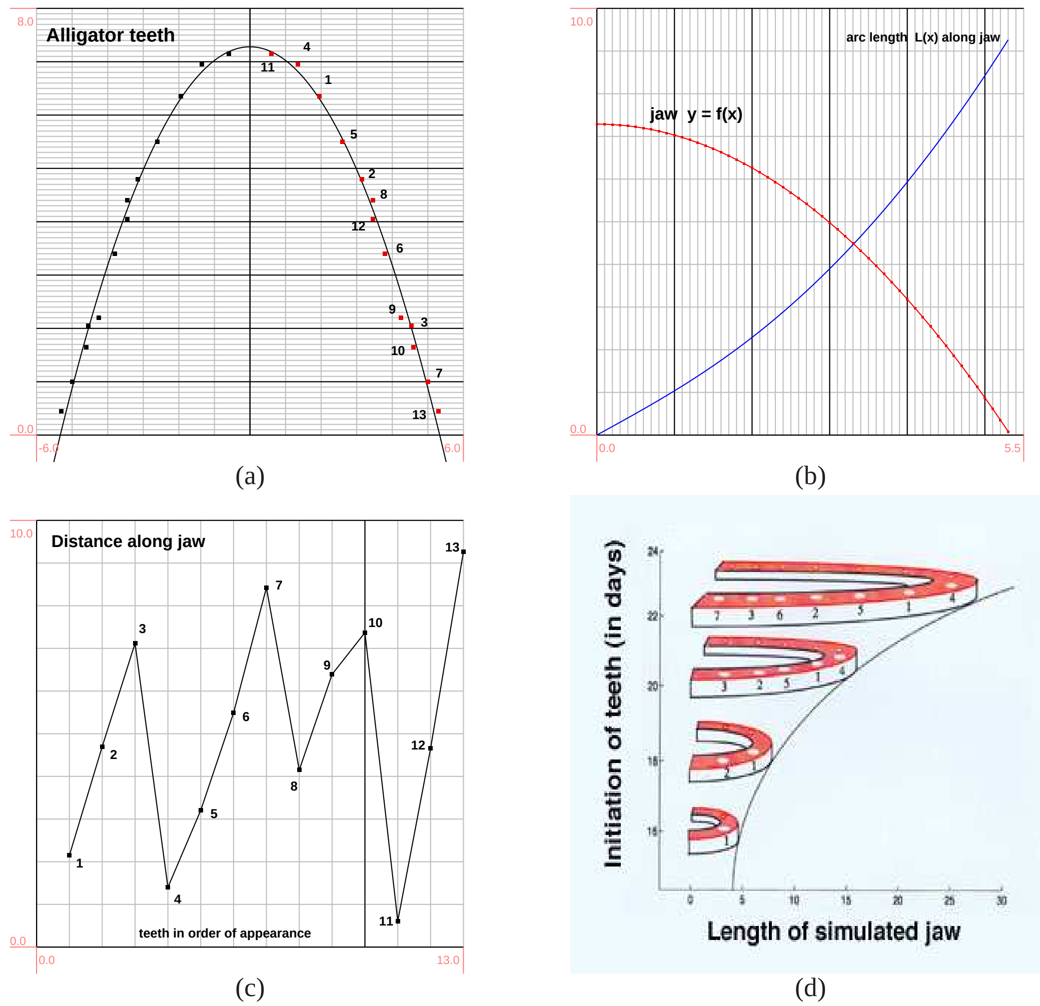

Figure 5.15. *(a) The parabolic shape of the jaw, showing positions of teeth and numerical order of emergence. (b) Arc length along the jaw from front to back. (c) Distance of successive teeth along the jaw. (d) Growth of the jaw.*

5.6.2 References

1. P.M. Kulesa and J.D. Murray (1995). Modelling the Wave-like Initiation of Teeth Primordia in the Alligator. FORMA. Cover Article. Vol. 10, No. 3, 259-280.

2. J.D. Murray and P.M. Kulesa (1996). On A Dynamic Reaction-Diffusion Mechanism for the Spatial Patterning of Teeth Primordia in the Alligator. Journal of Chemical Physics. J. Chem. Soc., Faraday Trans., 92 (16), 2927-2932.

3. P.M. Kulesa, G.C. Cruywagen, S.R. Lubkin, M.W.J. Ferguson and J.D. Murray (1996). Modelling the Spatial Patterning of Teeth Primordia in the Alligator. Acta Biotheoretica, 44, 153-164.

5.7 Summary

Here are the main points of the chapter:

1. We introduced the idea of a spatially distributed mass density $\rho(x)$ in Section 5.2.2. Here the definite integral represents
$$\int_a^b \rho(x)\,dx = \text{Total mass in the interval } a \le x \le b.$$

2. In this chapter, we defined the center of mass of a (discrete) distribution of n masses by
$$\bar{X} = \frac{1}{M}\sum_{i=0}^{n} x_i\, m_i. \tag{5.5}$$
We developed the analogue of this for a continuous mass distribution (distributed in the interval $0 \le x \le L$). We defined the center of mass of a continuous distribution by the definite integral
$$\bar{x} = \frac{1}{M}\int_0^L x\rho(x)dx\ . \tag{5.6}$$
Importantly, the quantities m_i in the sum (5.5) carry units of mass, whereas the analogous quantities in (5.6) are $\rho(x)dx$. [Recall that $\rho(x)$ is a mass per unit length in the case of mass distributed along a bar or straight line.]

3. We defined a cumulative function. In the discrete case, this was defined as In the continuous case, it is
$$M(x) = \int_0^x \rho(s)ds.$$

4. The mean is an average x coordinate, whereas the median is the x coordinate that splits the distribution into two equal masses (Geometrically, the median subdivides the graph of the distribution into two regions of equal areas). The mean and median are the same only in symmetric distributions. They differ for any distribution that is asymmetric. The mean (but not the median) is influenced more strongly by distant portions of the distribution.

5. In the later parts of this chapter, we showed how to compute volumes of various objects that have radial symmetry ("solids of revolution"). We showed that if the surface is generated by rotating the graph of a function $y = f(x)$ about the x axis (for $a \leq x \leq b$), then its volume can be described by an integral of the form

$$V = \int_a^b \pi [f(x)]^2 dx.$$

We used this idea to show that the volume of a sphere of radius R is $V_{sphere} = (4/3)\pi R^3$

In the Chapters 7 and 8, we find applications of the ideas of density and center of mass to the context of a probability distribution and its mean.

Chapter 6

Techniques of Integration

In this chapter, we expand our repertoire for antiderivatives beyond the "elementary" functions discussed so far. A review of the table of elementary antiderivatives (found in Chapter 3) will be useful. Here we will discuss a number of methods for finding antiderivatives. We refer to these collected "tricks" as methods of integration. As will be shown, in some cases, these methods are systematic (i.e. with clear steps), whereas in other cases, guesswork and trial and error is an important part of the process.

A primary method of integration to be described is **substitution**. A close relationship exists between the chain rule of differential calculus and the substitution method. A second very important method is **integration by parts**. Aside from its usefulness in integration per se, this method has numerous applications in physics, mathematics, and other sciences. Many other techniques of integration used to form a core of methods taught in such courses in integral calculus. Many of these are quite technical. Nowadays, with sophisticated mathematical software packages (including Maple and Mathematica), integration can be carried our automatically via computation called "symbolic manipulation", reducing our need to dwell on these technical methods.

6.1 Differential notation

We begin by familiarizing the reader with notation that appears frequently in substitution integrals, i.e. differential notation. Consider a straight line

$$y = mx + b.$$

Recall that the slope of the line, m, is

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

This relationship can also be written in the form

$$\Delta y = m\Delta x.$$

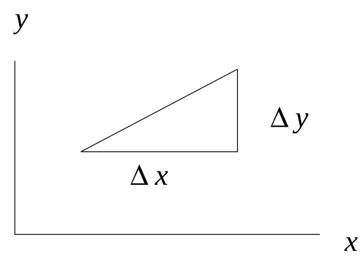

Figure 6.1. *The slope of the line shown here is $m = \Delta y/\Delta x$. This means that the small quantities Δy and Δx are related by $\Delta y = m\Delta x$.*

If we take a very small step along this line in the x direction, call it dx (to remind us of an "infinitesimally small" quantity), then the resulting change in the y direction, (call it dy) is related by

$$dy = m\,dx.$$

Now suppose that we have a curve defined by some arbitrary function, $y = f(x)$ which need not be a straight line.For a given point (x, y) on this curve, a step Δx in the x direction is associated with a step Δy in the y direction. The relationship between the step

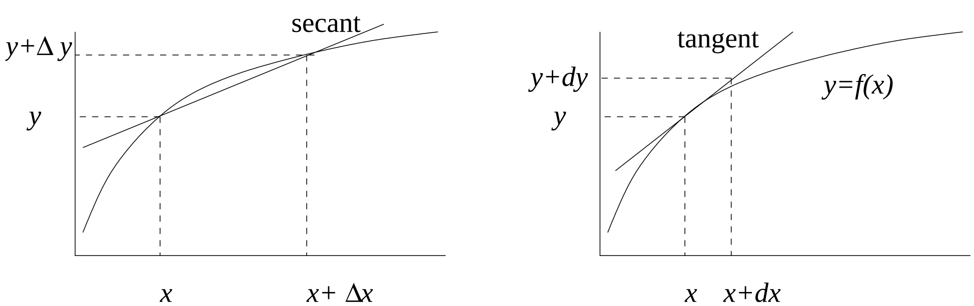

Figure 6.2. *On this figure, the graph of some function is used to illustrate the connection between differentials dy and dx. Note that these are related via the slope of a tangent line, m_t to the curve, in contrast with the relationship of Δy and Δx which stems from the slope of the secant line m_s on the same curve.*

sizes is:

$$\Delta y = m_s \Delta x,$$

where now m_s is the slope of a secant line (shown connecting two points on the curve in Figure 6.2). If the sizes of the steps are small (dx and dy), then this relationship is well approximated by the slope of the tangent line, m_t as shown in Figure 6.2 i.e.

$$dy = m_t\,dx = f'(x)dx.$$

The quantities dx and dy are called **differentials**. In general, they link a small step on the x axis with the resulting small change in height along the tangent line to the curve (shown in Figure 6.2). We might observe that the ratio of the differentials, i.e.

$$\frac{dy}{dx} = f'(x),$$

appears to link our result to the definition of the derivative. We remember, though, that the derivative is actually defined as a limit:

$$f'(x) = \lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x}.$$

When the step size Δx is quite small, it is approximately true that

$$\frac{\Delta y}{\Delta x} \approx \frac{dy}{dx}.$$

This notation will be useful in substitution integrals.

Examples

We give some examples of functions, their derivatives, and the differential notation that goes with them.

1. The function $y = f(x) = x^3$ has derivative $f'(x) = 3x^2$. Thus
$$dy = 3x^2\ dx.$$
2. The function $y = f(x) = \tan(x)$ has derivative $f'(x) = \sec^2(x)$. Therefore
$$dy = \sec^2(x)\ dx.$$
3. The function $y = f(x) = \ln(x)$ has derivative $f'(x) = \frac{1}{x}$ so
$$dy = \frac{1}{x}\ dx.$$

With some practice, we can omit the intermediate step of writing down a derivative and go directly from function to differential notation. Given a function $y = f(x)$ we will often write

$$df(x) = \frac{df}{dx} dx$$

and occasionally, we use just the symbol df to mean the same thing. The following examples illustrate this idea with specific functions.

$$d(\sin(x)) = \cos(x)\ dx, \quad d(x^n) = nx^{n-1}\ dx, \quad d(\arctan(x)) = \frac{1}{1+x^2}\ dx.$$

Moreover, some of the basic rules of differentiation translate directly into rules for handling and manipulating differentials. We give a list of some of these elementary rules below.

Rules for derivatives and differentials

1. $\frac{d}{dx}C = 0, \qquad dC = 0.$
2. $\frac{d}{dx}(u(x) + v(x)) = \frac{du}{dx} + \frac{dv}{dx} \qquad d(u + v) = du + dv.$
3. $\frac{d}{dx}u(x)v(x) = u\frac{dv}{dx} + v\frac{du}{dx} \qquad d(uv) = u\,dv + v\,du.$
4. $\frac{d}{dx}(Cu(x)) = C\frac{du}{dx}, \qquad d(Cu) = C\,du$

6.2 Antidifferentiation and indefinite integrals

In Chapters 2 and 3, we defined the concept of the **definite integral**, which represents a number. It will be useful here to consider the idea of an **indefinite integral**, which is a function, namely an antiderivative.

If two functions, $F(x)$ and $G(x)$, have the same derivative, say $f(x)$, then they differ at most by a constant, that is $F(x) = G(x) + C$, where C is some constant.

Proof

Since $F(x)$ and $G(x)$ have the same derivative, we have

$$\frac{d}{dx}F(x) = \frac{d}{dx}G(x),$$

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

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

This means that the function $F(x) - G(x)$ should be a constant, since its derivative is zero. Thus

$$F(x) - G(x) = C,$$

so

$$F(x) = G(x) + C,$$

as required. $F(x)$ and $G(x)$ are called antiderivatives of $f(x)$, and this confirms, once more, that any two antiderivatives differ at most by a constant.

In another terminology, which means the same thing, we also say that $F(x)$ (or $G(x)$) is the integral of the function $f(x)$, and we refer to $f(x)$ as the *integrand*. We write this as follows:

$$F(x) = \int f(x)\,dx.$$

This notation is sometimes called "an *indefinite integral*" because it does not denote a specific numerical value, nor is an interval specified for the integration range. An indefinite integral is a function with an arbitrary constant. (Contrast this with the definite integral studied in our last chapters: in the case of the definite integral, we specified an interval, and interpreted the result, a number, in terms of areas associated with curves.) We also write

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

if we want to indicate the form of all possible functions that are antiderivatives of $f(x)$. C is referred to as a *constant of integration*.

6.2.1 Integrals of derivatives

Suppose we are given an integral of the form

$$\int \frac{df}{dx} dx,$$

or alternately, the same thing written using differential notation,

$$\int df.$$

How do we handle this? We reason as follows. The df/dx (a quantity that is, itself, a function) is the derivative of the function $f(x)$. That means that $f(x)$ is the antiderivative of df/dx. Then, according to the Fundamental Theorem of Calculus,

$$\int \frac{df}{dx} dx = f(x) + C.$$

We can write this same result using the differential of f, as follows:

$$\int df = f(x) + C.$$

The following examples illustrate the idea with several elementary functions.

Examples

1. $\int d(\cos x) = \cos x + C$.
2. $\int dv = v + C$.
3. $\int d(x^3) = x^3 + C$.

6.3 Simple substitution

In this section, we observe that the forms of some integrals can be simplified by making a judicious substitution, and using our familiarity with derivatives (and the chain rule). The idea rests on the fact that in some cases, we can spot a "helper function"

$$u = f(x),$$

such that the quantity

$$du = f'(x)dx$$

appears in the integrand. In that case, the substitution will lead to eliminating x entirely in favour of the new quantity u, and simplification may occur.

6.3.1 Example: Simple substitution

Suppose we are given the function

$$f(x) = (x+1)^{10}.$$

Then its antiderivative (indefinite integral) is

$$F(x) = \int f(x)\,dx = \int (x+1)^{10}\,dx.$$

We could find an antiderivative by expanding the integrand $(x+1)^{10}$ into a degree 10 polynomial and using methods already known to us; but this would be laborious. Let us observe, however, that if we define

$$u = (x+1),$$

then

$$du = \frac{d(x+1)}{dx}dx = \left(\frac{dx}{dx} + \frac{d(1)}{dx}\right)dx = (1+0)dx = dx.$$

Now replacing $(x+1)$ by u and dx by the equivalent du we get:

$$F(x) = \int u^{10}du.$$

An antiderivative to this can be easily found, namely,

$$F(x) = \frac{u^{11}}{11} = \frac{(x+1)^{11}}{11} + C.$$

In the last step, we converted the result back to the original variable, and included the arbitrary integration constant. A very important point to remember is that we can always check our results by differentiation:

Check

Differentiate $F(x)$ to obtain

$$\frac{dF}{dx} = \frac{1}{11}(11(x+1)^{10}) = (x+1)^{10}.$$

6.3.2 How to handle endpoints

We consider how substitution type integrals can be calculated when we have endpoints, i.e. in evaluating definite integrals. Consider the example:

$$I = \int_1^2 \frac{1}{x+1} dx\,.$$

This integration can be done by making the substitution $u = x+1$ for which $du = dx$. We can handle the endpoints in one of two ways:

Method 1: Change the endpoints

We can change the integral over entirely to a definite integral in the variable u as follows: Since $u = x+1$, the endpoint $x = 1$ corresponds to $u = 2$, and the endpoint $x = 2$ corresponds to $u = 3$, so changing the endpoints to reflect the change of variables leads to

$$I = \int_2^3 \frac{1}{u} du = \ln|u|\Big|_2^3 = \ln 3 - \ln 2 = \ln\frac{3}{2}.$$

In the last steps we have plugged in the new endpoints (appropriate to u).

Method 2: Change back to x before evaluating at endpoints

Alternately, we could rewrite the antiderivative in terms of x.

$$\int \frac{1}{u} du = \ln|u| = \ln|x+1|$$

and then evaluate this function at the original endpoints.

$$\int_1^2 \frac{1}{x+1} dx = \ln|x+1|\Big|_1^2 = \ln\frac{3}{2}$$

Here we plugged in the original endpoints (as appropriate to the variable x).

6.3.3 Examples: Substitution type integrals

Find a simple substitution and determine the antiderivatives (indefinite or definite integrals) of the following functions:

1. $I = \displaystyle\int \frac{2}{x+2}\, dx.$
2. $I = \displaystyle\int_0^1 x^2 e^{x^3}\, dx$
3. $I = \displaystyle\int \frac{1}{(x+1)^2+1}\, dx.$

4. $I = \displaystyle\int (x+3)\sqrt{x^2+6x+10}\,dx.$

5. $I = \displaystyle\int_0^\pi \cos^3(x)\sin(x)\,dx.$

6. $I = \displaystyle\int \frac{1}{ax+b}\,dx$

7. $I = \displaystyle\int \frac{1}{b+ax^2}\,dx.$

Solutions

1. $I = \displaystyle\int \frac{2}{x+2}\,dx$. Let $u = x+2$. Then $du = dx$ and we get
$$I = \int \frac{2}{u}\,du = 2\int \frac{1}{u}\,du = 2\ln|u| = 2\ln|x+2| + C.$$

2. $I = \displaystyle\int_0^1 x^2 e^{x^3}\,dx$. Let $u = x^3$. Then $du = 3x^2\,dx$. Here we use method 2 for handling endpoints.
$$\int e^u \frac{du}{3} = \frac{1}{3}e^u = \frac{1}{3}e^{x^3} + C.$$
Then
$$I = \int_0^1 x^2 e^{x^3}\,dx = \frac{1}{3}e^{x^3}\Big|_0^1 = \frac{1}{3}(e-1).$$
(We converted the antiderivative to the original variable, x, before plugging in the original endpoints.)

3. $I = \displaystyle\int \frac{1}{(x+1)^2+1}\,dx$. Let $u = x+1$, then $du = dx$ so we have
$$I = \int \frac{1}{u^2+1}\,du = \arctan(u) = \arctan(x+1) + C.$$

4. $I = \displaystyle\int (x+3)\sqrt{x^2+6x+10}\,dx$. Let $u = x^2+6x+10$. Then $du = (2x+6)\,dx = 2(x+3)\,dx$. With this substitution we get
$$I = \int \sqrt{u}\,\frac{du}{2} = \frac{1}{2}\int u^{1/2}\,du = \frac{1}{2}\frac{u^{3/2}}{3/2} = \frac{1}{3}u^{3/2} = \frac{1}{3}(x^2+6x+10)^{3/2} + C.$$

5. $I = \displaystyle\int_0^\pi \cos^3(x)\sin(x)\,dx$. Let $u = \cos(x)$. Then $du = -\sin(x)\,dx$. Here we use method 1 for handling endpoints. For $x = 0, u = \cos 0 = 1$ and for $x = \pi, u = \cos\pi = -1$, so changing the integral and endpoints to u leads to
$$I = \int_1^{-1} u^3(-du) = -\frac{u^4}{4}\Big|_1^{-1} = -\frac{1}{4}((-1)^4 - 1^4) = 0.$$

Here we plugged in the new endpoints that are relevant to the variable u.

6. $\displaystyle\int \frac{1}{ax+b}\,dx$. Let $u = ax + b$. Then $du = a\,dx$, so $dx = du/a$. Substitute the above equations into the first equation and simplify to get

$$I = \int \frac{1}{u}\frac{du}{a} = \frac{1}{a}\int \frac{1}{u}\,du = \frac{1}{a}ln|u| + C.$$

Substitute $u = ax + b$ back to arrive at the solution

$$I = \int \frac{1}{ax+b}\,dx = \frac{1}{a}\ln|ax+b| \tag{6.1}$$

7. $\displaystyle I = \int \frac{1}{b+ax^2}\,dx = \frac{1}{b}\int \frac{1}{1+(a/b)x^2}\,dx$. This can be brought to the form of an arctan type integral as follows: Let $u^2 = (a/b)x^2$, so $u = \sqrt{a/b}\,x$ and $du = \sqrt{a/b}\,dx$. Now substituting these, we get

$$I = \frac{1}{b}\int \frac{1}{1+u^2}\,\frac{du}{\sqrt{a/b}} = \sqrt{b/a}\frac{1}{b}\int \frac{1}{1+u^2}\,du$$

$$I = \frac{1}{\sqrt{ba}}\arctan(u)\,du = \frac{1}{\sqrt{ba}}\arctan(\sqrt{a/b}\,x) + C.$$

6.3.4 When simple substitution fails

Not every integral can be handled by simple substitution. Let us see what could go wrong:

Example: Substitution that does not work

Consider the case

$$F(x) = \int \sqrt{1+x^2}\,dx = \int (1+x^2)^{1/2}\,dx.$$

A "reasonable" guess for substitution might be

$$u = (1+x^2).$$

Then

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

and $dx = du/2x$. Attempting to convert the integral to the form containing u would lead to

$$I = \int \sqrt{u}\,\frac{du}{2x}.$$

We have not succeeded in eliminating x entirely, so the expression obtained contains a mixture of two variables. We can proceed no further. This substitution did not simplify the integral and we must try some other technique.

6.3.5 Checking your answer

Finding an antiderivative can be tricky. (To a large extent, methods described in this chapter are a "collection of tricks".) However, it is always possible (and wise) to check for correctness, by differentiating the result. This can help uncover errors.

For example, suppose that (in the previous example) we had incorrectly guessed that the antiderivative of

$$\int (1+x^2)^{1/2}\, dx$$

might be

$$F_{\text{guess}}(x) = \frac{1}{3/2}(1+x^2)^{3/2}.$$

The following check demonstrates the incorrectness of this guess: Differentiate $F_{\text{guess}}(x)$ to obtain

$$F'_{\text{guess}}(x) = \frac{1}{3/2}(3/2)(1+x^2)^{(3/2)-1}\ \cdot 2x = (1+x^2)^{1/2}\cdot 2x$$

The result is clearly not the same as $(1+x^2)^{1/2}$, since an "extra" factor of $2x$ appears from application of the chain rule: this means that the trial function $F_{\text{guess}}(x)$ was not the correct antiderivative. (We can similarly check to confirm correctness of any antiderivative found by following steps of methods here described. This can help to uncover sign errors and other algebraic mistakes.)

6.4 More substitutions

In some cases, rearrangement is needed before the form of an integral becomes apparent. We give some examples in this section. The idea is to reduce each one to the form of an elementary integral, whose antiderivative is known.

Standard integral forms

1. $I = \displaystyle\int \frac{1}{u}\, du = \ln|u| + C.$
2. $I = \displaystyle\int u^n\, du = \frac{u^{n+1}}{n+1}.$
3. $I = \displaystyle\int \frac{1}{1+u^2}\, du = \arctan(u) + C.$

However, finding which of these forms is appropriate in a given case will take some ingenuity and algebra skills. Integration tends to be more of an art than differentiation, and recognition of patterns plays an important role here.

6.4.1 Example: perfect square in denominator

Find the antiderivative for

$$I = \int \frac{1}{x^2-6x+9}\, dx.$$

Solution

We observe that the denominator of the integrand is a perfect square, i.e. that $x^2-6x+9 = (x-3)^2$. Replacing this in the integral, we obtain

$$I = \int \frac{1}{x^2-6x+9}\,dx = \int \frac{1}{(x-3)^2}\,dx.$$

Now making the substitution $u = (x-3)$, and $du = dx$ leads to a power type integral

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

6.4.2 Example: completing the square

A small change in the denominator will change the character of the integral, as shown by this example:

$$I = \int \frac{1}{x^2-6x+10}\,dx.$$

Solution

Here we use "completing the square" to express the denominator in the form $x^2-6x+10 = (x-3)^2+1$. Then the integral takes the form

$$I = \int \frac{1}{1+(x-3)^2}\,dx.$$

Now a substitution $u = (x-3)$ and $du = dx$ will result in

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

Remark: in cases where completing the square gives rise to a constant other than 1 in the denominator, we use the technique illustrated in Example 6.3.3 Eqn. (6.1) to simplify the problem.

6.4.3 Example: factoring the denominator

A change in one sign can also lead to a drastic change in the antiderivative. Consider

$$I = \int \frac{1}{1-x^2}\,dx.$$

In this case, we can factor the denominator to obtain

$$I = \int \frac{1}{(1-x)(1+x)}\,dx.$$

We will show shortly that the integrand can be simplified to the sum of two fractions, i.e. that

$$I = \int \frac{1}{(1-x)(1+x)}\, dx = \int \frac{A}{(1-x)} + \frac{B}{(1+x)}\, dx,$$

where A, B are constants. The algebraic technique for finding these constants, and hence of forming the simpler expressions, called *Partial fractions*, will be discussed in an upcoming section. Once these constants are found, each of the resulting integrals can be handled by substitution.

6.5 Trigonometric substitutions

Trigonometric functions provide a rich set of interconnected functions that show up in many problems. It is useful to remember three very important trigonometric identities that help to simplify many integrals. These are:

Essential trigonometric identities

1. $\sin^2(x) + \cos^2(x) = 1$
2. $\sin(A+B) = \sin(A)\cos(B) + \sin(B)\cos(A)$
3. $\cos(A+B) = \cos(A)\cos(B) - \sin(A)\sin(B)$.

In the special case that $A = B = x$, the last two identities above lead to:

Double angle trigonometric identities

1. $\sin(2x) = 2\sin(x)\cos(x)$.
2. $\cos(2x) = \cos^2(x) - \sin^2(x)$.

From these, we can generate a variety of other identities as special cases. We list the most useful below. The first two are obtained by combining the double-angle formula for cosines with the identity $\sin^2(x) + \cos^2(x) = 1$.

Useful trigonometric identities

1. $\cos^2(x) = \dfrac{1+\cos(2x)}{2}$.
2. $\sin^2(x) = \dfrac{1-\cos(2x)}{2}$.
3. $\tan^2(x) + 1 = \sec^2(x)$.

6.5.1 Example: simple trigonometric substitution

Find the antiderivative of

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

Solution

This integral can be computed by a simple substitution, similar to Example 5 of Section 6.3. We let $u = \cos(x)$ and $du = -\sin(x)dx$ to get the integral into the form

$$I = -\int u^2\, du = \frac{-u^3}{3} = \frac{-\cos^3(x)}{3} + C.$$

We need none of the trigonometric identities in this case. Simple substitution is always the easiest method to use. It should be the first method attempted in each case.

6.5.2 Example: using trigonometric identities (1)

Find the antiderivative of

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

Solution

This is an example in which the "Useful trigonometric identity" 1 leads to a simpler integral. We write

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

Then clearly,

$$I = \frac{1}{2}\left(x + \frac{\sin(2x)}{2}\right) + C.$$

6.5.3 Example: using trigonometric identities (2)

Find the antiderivative of

$$I = \int \sin^3(x)\, dx.$$

Solution

We can rewrite this integral in the form

$$I = \int \sin^2(x)\sin(x)\, dx.$$

Now using the trigonometric identity $\sin^2(x) + \cos^2(x) = 1$, leads to

$$I = \int (1-\cos^2(x))\sin(x)\, dx.$$

This can be split up into

$$I = \int \sin(x)\, dx - \int \sin(x)\cos^2(x)\, dx.$$

The first part is elementary, and the second was shown in a previous example. Therefore we end up with

$$I = -\cos(x) + \frac{\cos^3(x)}{3} + C.$$

Note that it is customary to combine all constants obtained in the calculation into a single constant, C at the end.

Aside from integrals that, themselves, contain trigonometric functions, there are other cases in which use of trigonometric identities, though at first seemingly unrelated, is crucial. Many expressions involving the form $\sqrt{1 \pm x^2}$ or the related form $\sqrt{a \pm bx^2}$ will be simplified eventually by conversion to trigonometric expressions!

6.5.4 Example: converting to trigonometric functions

Find the antiderivative of

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

Solution

The simple substitution $u = 1 - x^2$ will not work, (as shown by a similar example in Section 6.3). However, converting to trigonometric expressions will do the trick. Let

$$x = \sin(u), \quad \text{then} \;\; dx = \cos(u)du.$$

(In Figure 6.3, we show this relationship on a triangle. This diagram is useful in reversing the substitutions after the integration step.) Then $1 - x^2 = 1 - \sin^2(u) = \cos^2(u)$, so the

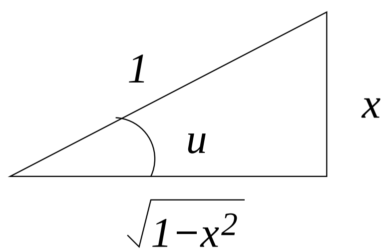

Figure 6.3. *This triangle helps to convert the (trigonometric) functions of u to the original variable x in Example 6.5.4.*

substitutions lead to

$$I = \int \sqrt{\cos^2(u)}\cos(u)\, du = \int \cos^2(u)\, du.$$

From a previous example, we already know how to handle this integral. We find that

$$I = \frac{1}{2}\left(u + \frac{\sin(2u)}{2}\right) = \frac{1}{2}\left(u + \sin(u)\cos(u)\right) + C.$$

(In the last step, we have used the double angle trigonometric identity. We will shortly see why this simplification is relevant.)

We now desire to convert the result back to a function of the original variable, x. We note that $x = \sin(u)$ implies $u = \arcsin(x)$. To convert the term $\cos(u)$ back to an expression depending on x we can use the relationship $1 - \sin^2(u) = \cos^2(u)$, to deduce that

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

It is sometimes helpful to use a Pythagorean triangle, as shown in Figure 6.3, to rewrite the antiderivative in terms of the variable x. The idea is this: We construct the triangle in such a way that its side lengths are related to the "angle" u according to the substitution rule. In this example, $x = \sin(u)$ so the sides labeled x and 1 were chosen so that their ratio ("opposite over hypotenuse" coincides with the sine of the indicated angle, u, thereby satisfying $x = \sin(u)$. We can then determine the length of the third leg of the triangle (using the Pythagorean formula) and thus all other trigonometric functions of u. For example, we note that the ratio of "adjacent over hypotenuse" is $\cos(u) = \sqrt{1 - x^2}/1 = \sqrt{1 - x^2}$. Finally, with these reverse substitutions, we find that,

$$I = \int \sqrt{1 - x^2}\, dx = \frac{1}{2}\left(\arcsin(x) + x\sqrt{1 - x^2}\right) + C.$$

Remark: In computing a definite integral of the same type, we can circumvent the need for the conversion back to an expression involving x by using the appropriate method for handling endpoints. For example, the integral

$$I = \int_0^1 \sqrt{1 - x^2}\, dx$$

can be transformed to

$$I = \int_0^{\pi/2} \sqrt{\cos^2(u)}\cos(u)\, du,$$

by observing that $x = \sin(u)$ implies that $u = 0$ when $x = 0$ and $u = \pi/2$ when $x = 1$. Then this means that the integral can be evaluated directly (without changing back to the variable x) as follows:

$$I = \int_0^{\pi/2} \sqrt{\cos^2(u)}\cos(u)\, du = \frac{1}{2}\left(u + \frac{\sin(2u)}{2}\right)\Bigg|_0^{\pi/2} = \frac{1}{2}\left(\frac{\pi}{2} + \frac{\sin(\pi)}{2}\right) = \frac{\pi}{4}$$

where we have used the fact that $\sin(\pi) = 0$.

Some subtle points about the domains of definition of inverse trigonometric functions will not be discussed here in detail. (See material on these functions in a first term calculus course.) Suffice it to say that some integrals of this type will be undefined if this endpoint conversion cannot be carried out (e.g. if the interval of integration had been $0 \le x \le 2$, we would encounter an impossible relation $2 = \sin(u)$. Since no value of u satisfies this relation, such a definite integral has no meaning, i.e. "does not exist".)

6.5.5 Example: The centroid of a two dimensional shape

We extend the concept of centroid (center of mass) for a region that has uniform density in 2D, but where we consider the distribution of mass along the x (or y) axis. Consider the semicircle shape of uniform thickness, shown in Figure 6.4, and suppose it is balanced along its horizontal edge. Find the x coordinate $\bar{c}$ at which the shape balances.

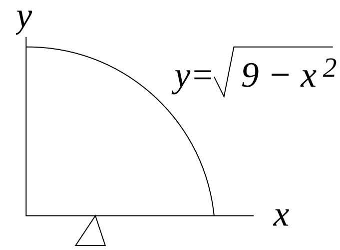

Figure 6.4. *A semicircular shape.*

Solution

The semicircle is one quarter of a circle of radius 3. Its edge is described by the equation

$$y = f(x) = \sqrt{9 - x^2}.$$

We will assume that the density per unit area is uniform. However, the mass per unit length along the x axis is not uniform, due to the shape of the object. We apply the idea of integration: If we cut the shape at increments of Δx along the x axis, we get a collection of pieces whose mass is each proportional to $f(x)\Delta x$. Summing up such contributions and letting the widths $\Delta x \to dx$ get small, we arrive at the integral for mass. The total mass of the shape is thus

$$M = \int_0^3 f(x)\, dx = \int_0^3 \sqrt{9 - x^2}\, dx.$$

Furthermore, if we compute the integral

$$I = \int_0^3 x f(x)\, dx = \int_0^3 x\sqrt{9 - x^2}\, dx,$$

we obtain the x coordinate of the center of mass,

$$\bar{x} = \frac{I}{M}.$$

It is evident that the mass is proportional to the area of one quarter of a circle of radius 3:

$$M = \frac{1}{4}\pi(3)^2 = \frac{9}{4}\pi.$$

(We could also see this by performing a trigonometric substitution integral.) The second integral can be done by simple substitution. Consider

$$I = \int_0^3 x f(x)\, dx = \int_0^3 x\sqrt{9-x^2}\, dx.$$

Let $u = 9 - x^2$. Then $du = -2x\, dx$. The endpoints are converted as follows: $x = 0 \Rightarrow u = 9 - 0^2 = 9$ and $x = 3 \Rightarrow u = 9 - 3^2 = 0$ so that we get the integral

$$I = \int_9^0 \sqrt{u}\, \frac{1}{-2} du.$$

We can reverse the endpoints if we switch the sign, and this leads to

$$I = \frac{1}{2}\int_0^9 u^{1/2}\, du = \left(\frac{1}{2}\right)\left(\frac{u^{3/2}}{3/2}\right)\Big|_0^9.$$

Since $9^{3/2} = (9^{1/2})^3 = 3^3$, we get $I = (3^3)/3 = 3^2 = 9$. Thus the x coordinate of the center of mass is

$$\bar{x} = \frac{I}{M} = \frac{9}{(9/4)\pi} = \frac{4}{\pi}.$$

We can similarly find the y coordinate of the center of mass: To do so, we would express the boundary of the shape in the form $x = f(y)$ and integrate to find

$$\bar{y} = \int_0^3 y f(y)\, dy.$$

For the semicircle, $y^2 + x^2 = 9$, so $x = f(y) = \sqrt{9 - y^2}$. Thus

$$\bar{y} = \int_0^3 y\sqrt{9-y^2}\, dy.$$

This integral looks identical to the one we wrote down for $\bar{x}$. Thus, based on this similarity (or based on the symmetry of the problem) we will find that

$$\bar{y} = \frac{4}{\pi}.$$

6.5.6 Example: $\tan$ and $\sec$ substitution

Find the antiderivative of

$$I = \int \sqrt{1+x^2}\, dx.$$

Solution

We aim for simplification by the identity $1 + \tan^2(u) = \sec^2(u)$, so we set

$$x = \tan(u), \quad dx = \sec^2(u) du.$$

Then the substitution leads to

$$I = \int \sqrt{1+\tan^2(u)}\ \sec^2(u)\ du = \int \sqrt{\sec^2(u)}\ \sec^2(u)\ du = \int \sec^3(u)\ du.$$

This integral will require further work, and will be partly calculated by *Integration by Parts* in Appendix 11.5. In this example, the triangle shown in Figure 6.5 shows the relationship between x and u and will help to convert other trigonometric functions of u to functions of x.

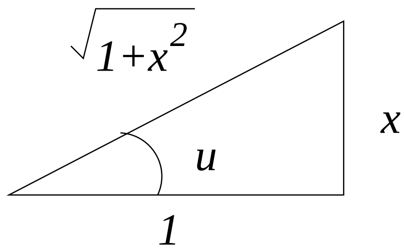

Figure 6.5. *As in Figure 6.3 but for example 6.5.6.*

6.6 Partial fractions

In this section, we show a simple algebraic trick that helps to simplify an integrand when it is in the form of some *rational function* such as

$$f(x) = \frac{1}{(ax+b)(cx+d)}.$$

The idea is to break this up into simpler rational expressions by finding constants A, B such that

$$\frac{1}{(ax+b)(cx+d)} = \frac{A}{(ax+b)} + \frac{B}{(cx+d)}.$$

Each part can then be handled by a simple substitution, as shown in Example 6.3.3, Eqn. (6.1). We give several examples below.

6.6.1 Example: partial fractions (1)

Find the antiderivative of

$$I = \int \frac{1}{x^2-1}.$$

Factoring the denominator, $x^2-1 = (x-1)(x+1)$, suggests breaking up the integrand into the form

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

The two sides are equal provided:

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

This means that

$$1 = A(x-1)+B(x+1)$$

must be true for all x values. We now ask what values of A and B make this equation hold for any x. Choosing two "easy" values, namely $x=1$ and $x=-1$ leads to isolating one or the other unknown constants, A, B, with the results:

$$1 = -2A, \quad 1 = 2B.$$

Thus $B = 1/2, A = -1/2$, so the integral can be written in the simpler form

$$I = \frac{1}{2}\left(\int \frac{-1}{(x+1)}\,dx + \int \frac{1}{(x-1)}\,dx\right).$$

(A common factor of $(1/2)$ has been taken out.) Now a simple substitution will work for each component. (Let $u = x+1$ for the first, and $u = x-1$ for the second integral.) The result is

$$I = \int \frac{1}{x^2-1} = \frac{1}{2}\left(-\ln|x+1| + \ln|x-1|\right) + C.$$

6.6.2 Example: partial fractions (2)

Find the antiderivative of

$$I = \int \frac{1}{x(1-x)}\,dx.$$

This example is similar to the previous one. We set

$$\frac{1}{x(1-x)} = \frac{A}{x} + \frac{B}{(1-x)}.$$

Then

$$1 = A(1-x) + Bx.$$

This must hold for all x values. In particular, convenient values of x for determining the constants are $x = 0,\ 1$. We find that

$$A = 1, \;\; B = 1.$$

Thus

$$I = \int \frac{1}{x(1-x)}\,dx = \int \frac{1}{x}\,dx + \int \frac{1}{1-x}\,dx.$$

Simple substitution now gives

$$I = \ln|x| - \ln|1-x| + C.$$

6.6.3 Example: partial fractions (3)

Find the antiderivative of

$$I = \int \frac{x}{x^2 + x - 2}.$$

The rational expression above factors into $x^2 + x - 2 = (x - 1)(x + 2)$, leading to the expression

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

Consequently, it follows that

$$A(x + 2) + B(x - 1) = x.$$

Substituting the values $x = 1, -2$ into this leads to $A = 1/3$ and $B = 2/3$. The usual procedure then results in

$$I = \int \frac{x}{x^2 + x - 2} = \frac{1}{3} \ln |x - 1| + \frac{2}{3} \ln |x + 2| + C.$$

Another example of the technique of partial fractions is provided in Appendix 11.5.2.

6.7 Integration by parts

The method described in this section is important as an additional tool for integration. It also has independent theoretical stature in many applications in mathematics and physics. The essential idea is that in some cases, we can exchange the task of integrating a function with the job of differentiating it.

The idea rests on the product rule for derivatives. Suppose that $u(x)$ and $v(x)$ are two differentiable functions. Then we know that the derivative of their product is

$$\frac{d(uv)}{dx} = v\frac{du}{dx} + u\frac{dv}{dx},$$

or, in the differential notation:

$$d(uv) = v\ du + u\ dv,$$

Integrating both sides, we obtain

$$\int d(uv) = \int v\ du + \int u\ dv$$

i.e.

$$uv = \int v\ du + \int u\ dv.$$

We write this result in the more suggestive form

$$\int u\ dv = uv - \int v\ du.$$

The idea here is that if we have difficulty evaluating an integral such as $\int u\ dv$, we may be able to "exchange it" for a simpler integral in the form $\int v\ du$. This is best illustrated by the examples below.

Example: Integration by parts (1)

Compute

$$I = \int_1^2 \ln(x)\ dx.$$

Solution

Let $u = \ln(x)$ and $dv = dx$. Then $du = (1/x)\ dx$ and $v = x$.

$$\int \ln(x)\ dx = x\ln(x) - \int x(1/x)\ dx = x\ln(x) - \int\ dx = x\ln(x) - x.$$

We now evaluate this result at the endpoints to obtain

$$I = \int_1^2 \ln(x)\ dx = (x\ln(x) - x)\Big|_1^2 = (2\ln(2) - 2) - (1\ln(1) - 1) = 2\ln(2) - 1.$$

(Where we used the fact that $\ln(1) = 0$.)

Example: Integration by parts (2)

Compute

$$I = \int_0^1 xe^x\ dx.$$

Solution

At first, it may be hard to decide how to assign roles for u and dv. Suppose we try $u = e^x$ and $dv = xdx$. Then $du = e^x\ dx$ and $v = x^2/2$. This means that we would get the integral in the form

$$I = \frac{x^2}{2}e^x - \int \frac{x^2}{2}e^x\ dx.$$

This is certainly *not* a simplification, because the integral we obtain has a higher power of x, and is consequently harder, not easier to integrate. This suggests that our first attempt was not a helpful one. (Note that integration often requires trial and error.)

Let $u = x$ and $dv = e^x\ dx$. This is a wiser choice because when we differentiate u, we reduce the power of x (from 1 to 0), and get a simpler expression. Indeed, $du = dx$, $v = e^x$ so that

$$\int xe^x\ dx = xe^x - \int e^x\ dx = xe^x - e^x + C.$$

To find a definite integral of this kind on some interval (say $0 \le x \le 1$), we compute

$$I = \int_0^1 xe^x\ dx = (xe^x - e^x)\Big|_0^1 = (1e^1 - e^1) - (0e^0 - e^0) = 0 + e^0 = e^0 = 1.$$

Note that all parts of the expression are evaluated at the two endpoints.

Example: Integration by parts (2b)

Compute

$$I_n = \int x^n e^x \, dx.$$

Solution

We can calculate this integral by repeated application of the idea in the previous example. Letting $u = x^n$ and $dv = e^x \, dx$ leads to $du = nx^{n-1}$ and $v = e^x$. Then

$$I_n = x^n e^x - \int nx^{n-1} e^x \, dx = x^n e^x - n \int x^{n-1} e^x \, dx.$$

Each application of integration by parts, reduces the power of the term x^n inside an integral by one. The calculation is repeated until the very last integral has been simplified, with no remaining powers of x. This illustrates that in some problems, integration by parts is needed more than once.

Example: Integration by parts (3)

Compute

$$I = \int \arctan(x) \, dx.$$

Solution

Let $u = \arctan(x)$ and $dv = dx$. Then $du = (1/(1 + x^2)) \, dx$ and $v = x$ so that

$$I = x \arctan(x) - \int \frac{1}{1 + x^2} x \, dx.$$

The last integral can be done with the simple substitution $w = (1 + x^2)$ and $dw = 2x \, dx$, giving

$$I = x \arctan(x) - (1/2) \int (1/w) dw.$$

We obtain, as a result

$$I = x \arctan(x) - \frac{1}{2} \ln(1 + x^2).$$

Example: Integration by parts (3b)

Compute

$$I = \int \tan(x) \, dx.$$

Solution

We might try to fit this into a similar pattern, i.e. let $u = \tan(x)$ and $dv = dx$. Then $du = \sec^2(x)\, dx$ and $v = x$, so we obtain

$$I = x\tan(x) - \int x\sec^2(x)\, dx.$$

This is not really a simplification, and we see that integration by parts will not necessarily work, even on a seemingly related example. However, we might instead try to rewrite the integral in the form

$$I = \int \tan(x)\, dx = \int \frac{\sin(x)}{\cos(x)}\, dx.$$

Now we find that a simple substitution will do the trick, i.e. that $w = \cos(x)$ and $dw = -\sin(x)\, dx$ will convert the integral into the form

$$I = \int \frac{1}{w}\,(-dw) = -\ln|w| = -\ln|\cos(x)|.$$

This example illustrates that we should always try substitution, first, before attempting other methods.

Example: Integration by parts (4)

Compute

$$I_1 = \int e^x \sin(x)\, dx.$$

We refer to this integral as I_1 because a related second integral, that we'll call I_2 will appear in the calculation.

Solution

Let $u = e^x$ and $dv = \sin(x)\, dx$. Then $du = e^x\, dx$ and $v = -\cos(x)\, dx$. Therefore

$$I_1 = -e^x\cos(x) - \int(-\cos(x))e^x\, dx = -e^x\cos(x) + \int \cos(x)e^x\, dx.$$

We now have another integral of a similar form to tackle. This seems hopeless, as we have not simplified the result, but let us not give up! In this case, another application of integration by parts will do the trick. Call I_2 the integral

$$I_2 = \int \cos(x)e^x\, dx,$$

so that

$$I_1 = -e^x\cos(x) + I_2.$$

Repeat the same procedure for the new integral I_2, i.e. Let $u = e^x$ and $dv = \cos(x)\, dx$. Then $du = e^x\, dx$ and $v = \sin(x)\, dx$. Thus

$$I_2 = e^x\sin(x) - \int \sin(x)e^x\, dx = e^x\sin(x) - I_1.$$

This appears to be a circular argument, but in fact, it has a purpose. We have determined that the following relationships are satisfied by the above two integrals:

$$I_1 = -e^x \cos(x) + I_2$$

$$I_2 = e^x \sin(x) - I_1.$$

We can eliminate I_2, obtaining

$$I_1 = -e^x \cos(x) + I_2 = -e^x \cos(x) + e^x \sin(x) - I_1.$$

that is,

$$I_1 = -e^x \cos(x) + e^x \sin(x) - I_1.$$

Rearranging (taking I_1 to the left hand side) leads to

$$2I_1 = -e^x \cos(x) + e^x \sin(x),$$

and thus, the desired integral has been found to be

$$I_1 = \int e^x \sin(x)\, dx = \frac{1}{2}\left(-e^x \cos(x) + e^x \sin(x)\right) = \frac{1}{2} e^x (\sin(x) - \cos(x)) + C.$$

(At this last step, we have included the constant of integration.) Moreover, we have also found that I_2 is related, i.e. using $I_2 = e^x \sin(x) - I_1$ we now know that

$$I_2 = \int \cos(x) e^x\, dx = \frac{1}{2} e^x \left(\sin(x) + \cos(x)\right) + C.$$

6.8 Summary

In this chapter, we explored a number of techniques for computing antiderivatives. We here summarize the most important results:

1. Substitution is the first method to consider. This method works provided the change of variable results in elimination of the original variable and leads to a simpler, more elementary integral.

2. When using substitution on a definite integral, endpoints can be converted to the new variable (Method 1) or the resulting antiderivative can be converted back to its original variable before plugging in the (original) endpoints (Method 2).

3. The integration by parts formula for functions $u(x), v(x)$ is

$$\int u\, dv = uv - \int v\, du.$$

   Integration by parts is useful when u is easy to differentiate (but not easy to integrate). It is also helpful when the integral contains a product of elementary functions such as x^n and a trigonometric or an exponential function. Sometimes more than one application of this method is needed. Other times, this method is combined with substitution or other simplifications.

4. Using integration by parts on a definite integral means that both parts of the formula are to be evaluated at the endpoints.

5. Integrals involving $\sqrt{1 \pm x^2}$ can be simplified by making a trigonometric substitution.

6. Integrals with products or powers of trigonometric functions can sometimes be simplified by application of trigonometric identities or simple substitution.

7. Algebraic tricks, and many associated manipulations are often applied to twist and turn a complicated integral into a set of simpler expressions that can each be handled more easily.

8. Even with all these techniques, the problem of finding an antiderivative can be very complicated. In some cases, we resort to handbooks of integrals, use symbolic manipulation software packages, or, if none of these work, calculate a given definite integral numerically using a spreadsheet.

Table of elementary antiderivatives

1. $\displaystyle\int \frac{1}{u}\, du = \ln|u| + C.$

2. $\displaystyle\int u^n\, du = \frac{u^{n+1}}{n+1} + C$

3. $\displaystyle\int \frac{1}{1+u^2} = \arctan(u) + C$

4. $\displaystyle\int \frac{1}{\sqrt{1-x^2}} = \arcsin(u) + C$

5. $\displaystyle\int \sin(u)\, du = -\cos(u) + C$

6. $\displaystyle\int \cos(u)\, du = \sin(u) + C$

7. $\displaystyle\int \sec^2(u)\, du = \tan(u) + C$

Additional useful antiderivatives

1. $\displaystyle\int \tan(u)\, du = \ln|\sec(u)| + C.$

2. $\displaystyle\int \cot(u)\, du = \ln|\sin(u)| + C$

3. $\displaystyle\int \sec(u) = \ln|\sec(u) + \tan(u)| + C$

Chapter 7

Discrete probability and the laws of chance

7.1 Introduction

In this chapter we lay the groundwork for calculations and rules governing simple discrete probabilities[24]. Such skills are essential in understanding problems related to random processes of all sorts. In biology, there are many examples of such processes, including the inheritance of genes and genetic diseases, the random motion of cells, the fluctuations in the number of RNA molecules in a cell, and a vast array of other phenomena.

To gain experience with probability, it is important to see simple examples. In this chapter, we discuss experiments that can be easily reproduced and tested by the reader.

7.2 Dealing with data

Scientists studying phenomena in the real world, collect data of all kinds, some resulting from experimental measurement or field observations. Data sets can be large and complex. If an experiment is repeated, and comparisons are to be made between multiple data sets, it is unrealistic to compare each and every numerical value. Some shortcuts allow us to summarize trends or descriptions of data sets in simple values such as averages (means), medians, and similar quantities. In doing so we lose the detailed information that the data set contains, in favor of simplicity of one or several "simple" numerical descriptors such as the **mean** and the **median** of a distribution. We have seen related ideas in Chapter 5 in the context of mass distributions. The idea of a center of mass is closely related to that of the mean of a distribution. Here we revisit such ideas in the context of probability. An additional example of real data is described in Appendix 11.6. There, we show how grade distributions on a test can be analyzed by similar methods.

[24]I am grateful to Robert Israel for comments regarding the organization of this chapter

7.3 Simple experiments

7.3.1 Experiment

We will consider "experiments" such as tossing a coin, rolling a die, dealing cards, applying treatment to sick patients, and recording how many are cured. In order for the ideas of probability to apply, we should be able to repeat the experiment as many times as desired under exactly the same conditions. The number of repetitions will often be denoted N.

7.3.2 Outcome

Whenever we perform the experiment, exactly one outcome happens. In this chapter we will deal with discrete probability, where there is a finite list of possible outcomes.

Consider the following experiment: We toss a coin and see how it lands. Here there are only two possible results: "heads" (H) or "tails" (T). A fair coin is one for which these results are equally likely. This means that if we repeat this experiment many many times, we expect that on average, we get H roughly 50% of the time and T roughly 50% of the time. This will lead us to define a probability of 1/2 for each outcome.

Similarly, consider the experiment of rolling a dice: A six-sided die can land on any of its six faces, so that a "single experiment" has six possible outcomes. For a fair die, we anticipate getting each of the results with an equal probability, i.e. if we were to repeat the same experiment many many times, we would expect that, on average, the six possible events would occur with similar frequencies, each 1/6 of the times. We say that the events are random and unbiased for "fair" dice.

We will often be interested in more complex experiments. For example, if we toss a coin five times, an outcome corresponds to a five-letter sequence of "Heads" (H) and "Tails" (T), such as THTHH. We are interested in understanding how to quantify the probability of each such outcome in fair (as well as unfair) coins. If we toss a coin ten times, how probable is it that we get eight out of ten heads? For dice, we could ask how likely are we to roll a 5 *and* a 6 in successive experiments? A 5 *or* a 6? For such experiments we are interested in quantifying how likely it is that a certain event is obtained. Our goal in this chapter is to make more precise our notion of probability, and to examine ways of quantifying and computing probabilities. To motivate this investigation, we first look at results of a real experiment performed in class by students.

7.3.3 Empirical probability

We can arrive at a notion of probability by actually repeating a real experiment N times, and counting how many times each outcome happens. Let us use the notation x_i to refer to the number of times that outcome i was obtained. An example of this sort is illustrated in Section 7.4.1. We define the **empirical probability** p_i of outcome i to be

$$p_i = x_i/N,$$

i.e p_i is the fraction of times that the result i is obtained out of all the experiments. We expect that if we repeated the experiment many more times, this empirical probability would approach, as a limit, the actual probability of the outcome. So if in a coin-tossing experiment, repeated 1000 times, the outcome HHTHH is obtained 25 times, then we would say that the empirical probability p_{HHTHH} is 25/1000.

7.3.4 Theoretical Probability

For theoretical probability, we make some reasonable basic assumptions on which we base a calculation of the probabilities. For example, in the case of a "fair coin", we can argue by symmetry that every sequence of n heads and tails has the same probability as any other. We then use two fundamental rules of probability to calculate the probability as illustrated below.

Rules of probability

1. In discrete probability, $0 \leq p_i \leq 1$ for each outcome i.
2. For discrete probability $\sum_i p_i = 1$, where the sum is over all possible outcomes.

About Rule 1: $p_i = 0$ implies that the given outcome never happens, whereas $p_i = 1$ implies that this outcome is the only possibility (and always happens). Any value inside the range (0,1) means that the outcome occurs some of the time. Rule 2 makes intuitive sense: it means that we have accounted for all possibilities, i.e. the fractions corresponding to all of the outcomes add up to 100% of the results.

In a case where there are M possible outcomes, all with equal probability, it follows that $p_i = 1/M$ for every i.

7.3.5 Random variables and probability distributions

A **random variable** is a numerical quantity X that depends on the outcome of an experiment. For example, suppose we toss a coin n times, and let X be the number of heads that appear. If, say, we toss the coin $n = 4$ times, then the number of heads, X could take on any of the values $\{x_i\} = \{0, 1, 2, 3, 4\}$ (i.e., no heads, one head, . . . four heads). In the case of discrete probability there are a discrete number of possible values for the random variable to take on.

We will be interested in the probability distribution of X. In general if the possible values x_i are listed in increasing order for $i = 0, ..., n$, we would like to characterize their probabilities $p(x_i)$, where $p(x_i)$ =Prob$(X = x_i)$[25].

Even though $p(x_i)$ is a discrete quantity taking on one of a discrete set of values, we should still think of this mathematical object as a function: it associates a number (the probability) p with each allowable value of the random variable x_i for $i = 0, \ldots, n$. In what follows, we will be interested in characterizing such function, termed probability distributions and their properties.

[25]Read: $p(x_i)$ is the probability that the random variable X takes on the value x_i

7.3.6 The cumulative distribution

Given a probability distribution, we can also define a **cumulative function** as follows:

The cumulative function corresponding to the probability distribution $p(x_i)$ is defined as

$$F(x_i) = \mathrm{Prob}(X \le x_i).$$

For a given numerical outcome x_i, the value of $F(x_i)$ is hence

$$F(x_i) = \sum_{j=0}^{i} p(x_j).$$

The function F merely sums up all the probabilities of outcomes up to and including x_i, hence is called "cumulative". This implies that $F(x_n) = 1$ where x_n is the largest value attainable by the random variable. For example, in the rolling of a die, if we list the possible outcomes in ascending order as $\{1, 2, \ldots, 6\}$, then $F(6)$ stands for the probability of rolling a 6 or any lower value, which is clearly equal to 1 for a six-sided die.

7.4 Examples of experimental data

7.4.1 Example1: Tossing a coin

We illustrate ideas with an example of real data obtained by repeating an "experiment" many times. The experiment, actually carried out by each of 121 students in this calculus course, consisted of tossing a coin $n = 10$ times and recording the number, x_i, of "Heads" that came up. Each student recorded one of eleven possible outcomes, $x_i = \{0, 1, 2, \ldots, 10\}$ (i.e. no heads, one, two, etc, up to ten heads out of the ten tosses). By pooling together such data, we implicitly assume that all coins and all tossers are more or less identical and unbiased, so the "experiment" has $N = 121$ replicates (one for each student). Table 7.1 shows the result of this experiment. Here n_i is the number of students who got x_i heads. We refer to this as the **frequency** of the given result. Also, so n_i/N is the fraction of experiments that led to the given result, and we define the empirical probability assigned to x_i as this fraction, that is $p(x_i) = n_i/N$. In column (3) we display the cumulative number of students who got any number up to and including x_i heads, and then in column (5) we compute the cumulative (empirical) probability $F(x_i)$.

In Figure 7.1 we show what this distribution looks like on a bar graph. The horizontal axis is x_i, the number of heads obtained, and the vertical axis is $p(x_i)$. Because in this example, only discrete integer values (0, 1, 2, etc) can be obtained in the experiment, it makes sense to represent the data as discrete points, as shown on the bottom panel in Fig. 7.1. We also show the cumulative function $F(x_i)$, superimposed as an xy-plot on a graph of $p(x_i)$. Observe that F starts with the value 0 and climbs up to value 1, since the probabilities of any of the events (0, 1, 2, etc heads) must add up to 1.

Number of heads x_i	frequency (number of students) n_i	cumulative number $\sum_0^i n_j$	empirical probability $p(x_i) = n_i/N$	cumulative function $F(x_i) = \sum_0^i p(x_j)$
0	0	0	0.00	0.00
1	1	1	0.0083	0.0083
2	2	3	0.0165	0.0248
3	10	13	0.0826	0.1074
4	27	40	0.2231	0.3306
5	26	66	0.2149	0.5455
6	34	100	0.2810	0.8264
7	14	114	0.1157	0.9421
8	7	121	0.0579	1.00
9	0	121	0.00	1.00
10	0	121	0.00	1.00

Table 7.1. *Results of a real coin-tossing experiment carried out by 121 students in this mathematics course. Each student tossed a coin 10 times. We recorded the "frequency", i.e. the number of students n_i who each got $x_i = 0, 1, 2, \ldots, 10$ heads. The fraction of the class that got each outcome, n_i/N, is identified with the (empirical) probability of that outcome, $p(x_i)$. We also compute the cumulative function $F(x_i)$ in the last column. See Figure 7.1 for the same data presented graphically.*

7.4.2 Example 2: grade distributions

Another example of real data is provided in Appendix 11.6. There we discuss distributions of grades on a test. Many of the ideas described here apply in the same way. For space constraints, that example is provided in an Appendix, rather than here.

7.5 Mean and variance of a probability distribution

We next discuss some very important quantities related to the random variable. Such quantities provide numerical descriptions of the average value of the random variable and the fluctuations about that average. We define each of these as follows:

The **mean** (or average or **expected value**), $\bar{x}$ of a probability distribution is

$$\boxed{\bar{x} = \sum_{i=0}^{n} x_i p(x_i) \ .}$$

The expected value is a kind of "average value of x", where values of x are weighted by their frequency of occurrence. This idea is related to the concept of center of mass defined in Section 5.3.1 (x positions weighted by masses associated with those positions).

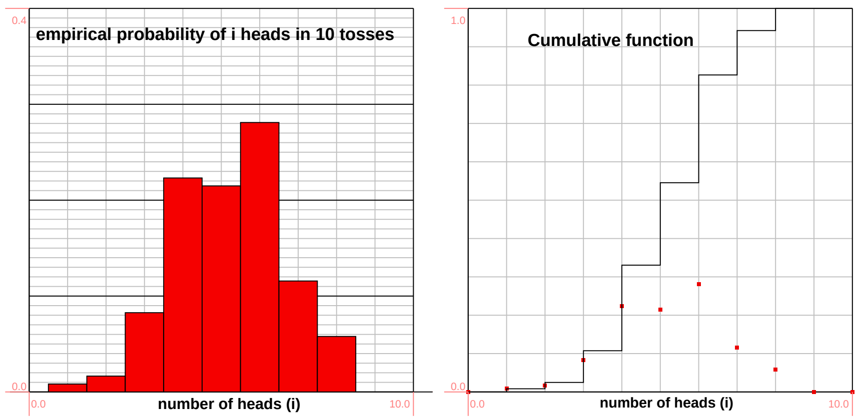

Figure 7.1. *The data from Table 7.1 is shown plotted on this graph. A total of* $N = 121$ *people were asked to toss a coin* $n = 10$ *times. In the bar graph (left), the horizontal axis reflects* i*, the number, of heads (H) that came up during those 10 coin tosses. The vertical axis reflects the fraction* $p(x_i)$ *of the class that achieved that particular number of heads. In the lower graph, the same data is shown by the discrete points. We also show the cumulative function that sums up the values from left to right. Note that the cumulative function is a "step function" .*

The mean is a point on the x axis, representing the "average" outcome of an experiment. (Recall that in the distributions we are describing, the possible outcomes of some observation or measurement process are depicted on the x axis of the graph.) The mean is *not* the same as the average value of a function, discussed in Section 4.6. (In that case, the average is an average y coordinate, or average height of the function.)[26]

We also define quantities that represents the width of the distribution. We define the variance, V and standard deviation, σ as follows:

The **variance**, V, of a distribution is

$$V = \sum_{i=0}^{n} (x_i - \bar{x})^2 p(x_i).$$

where $\bar{x}$ is the mean. The **standard deviation**, σ is

$$\sigma = \sqrt{V}.$$

The variance is related to the square of the quantity represented on the x axis, and since the standard deviation its square root, σ carries the same units as x. For this reason, it is

[26]Note to the instructor: students often mix these two distinct meanings of the word average, and they should be helped to overcome this difficulty with terminology.

common to associate the value of σ, with a typical "width" of the distribution. Having a low value of σ means that most of the experimental results are close to the mean, whereas a large σ signifies that there is a large scatter of experimental values about the mean.

In the problem sets, we show that the variance can also be expressed in the form

$$V = M_2 - \bar{x}^2,$$

where M_2 is the **second moment** of the distribution. Moments of a distribution are defined as the values obtained by summing up products of the probability weighted by powers of x.

The j**'th moment**, M_j of a distribution is

$$M_j = \sum_{i=0}^{n} (x_i)^j p(x_i).$$

Example 7.1 (Rolling a die) Suppose you toss a die, and let the random variable be X be the number obtained on the die, i.e. (1 to 6). If this die is fair, then it is equally likely to get any of the six possible outcomes, so each has probability 1/6. In this case

$$x_i = i, \quad i = 1, 2 \ldots 6 \quad p(x_i) = 1/6.$$

We calculate the various quantities as follows: The mean is

$$\bar{x} = \sum_{i=1}^{6} i \cdot \frac{1}{6} = \frac{1}{6} \cdot \left(\frac{6 \cdot 7}{2} \right) = \frac{7}{2} = 3.5.$$

The second moment, M_2 is

$$M_2 = \sum_{i=1}^{6} i^2 \cdot \frac{1}{6} = \frac{1}{6} \cdot \left(\frac{6 \cdot 7 \cdot 13}{6} \right) = \frac{91}{6}.$$

We can now obtain the variance,

$$V = \frac{91}{6} - \left(\frac{7}{2} \right)^2 = \frac{35}{12},$$

and the standard deviation,

$$\sigma = \sqrt{35/12} \approx 1.7078.$$

Example 7.2 (Expected number of heads (empirical)) For the empirical probability distribution shown in Figure 7.1, the mean (expected value) is calculated from results in Table 7.1 as follows:

$$\bar{x} = \sum_{k=0}^{10} x_i p(x_i) = 0(0)+1(0.0083)+2(0.0165)+\ldots+8(0.0579)+9(0)+10(0) = 5.2149$$

Thus, the mean number of heads in this set of experiments is about 5.2. This is close to what we would expect intuitively in a fair coin, namely that, on average, 5 out of 10 tosses (i.e. 50%) would result in heads. To compute the variance we form the sum

$$V = \sum_{k=0}^{10} (x_k - \bar{x})^2 p(x_k) = \sum_{k=0}^{10} (k - 5.2149)^2 p(k).$$

Here we have used the mean calculated above and the fact that $x_k = k$. We obtain

$$\begin{aligned} V &= (0 - 5.2149)^2(0) + (1 - 5.2149)^2(0.0083) + \ldots + (7 - 5.2149)^2(0.1157) \\ &+ (8 - 5.2149)^2(0.0579) + (9 - 5.2149)^2(0) + (10 - 5.2149)^2(0) = 2.053 \end{aligned}$$

(Because there was no replicate of the experiment that led to 9 or 10 heads out of 10 tosses, these values do not contribute to the calculation.) The standard deviation is then $\sigma = \sqrt{V} = 1.4328$.

7.6 Bernoulli trials

A **Bernoulli trial** is an experiment in which there are two possible outcomes. A typical example, motivated previously, is tossing a coin (the outcome being H or T). Traditionally, we refer to one of the outcomes of a Bernoulli trial as ”success” **S** and the other ”failure”[27], **F**.

Let p be the probability of success and $q = 1 - p$ the probability of failure in a Bernoulli trial. We now consider how to calculate the probability of some number of “successes” in a set of repetitions of a Bernoulli trial. In short, we are interested in the probability of tossing some number of Heads in n coin tosses.

7.6.1 The Binomial distribution

Suppose we repeat a Bernoulli trial n times; we will assume that each trial is identical and independent of the others. This implies that the probability p of success and q of failure is the same in each trial. Let X be the number of successes. Then X is said to have a **Binomial distribution** with parameters n and p.

Let us consider how to calculate the probability distribution of X, i.e. the probability that $X = k$ where k is some number of successes between none ($k = 0$) and all ($k = n$). Recall that the notation for this probability is Prob$(X = k)$ for $k = 0, 1, \ldots, n$. Also note that

$X = k$ means that in the n trials there are k successes and $n - k$ failures. Consider the following example for the case of $n = 3$, where we list all possible outcomes and their probabilities:

In constructing Table 7.2, we use a **multiplication principle** applied to computing the probability of a compound experiment. We state this, together with a useful **addition principle** below.

[27] For example “Heads you win, Tails you lose”.

Result	probability	number of heads
SSS	p^3	$X = 3$
SSF	p^2q	$X = 2$
SFS	p^2q	$X = 2$
SFF	pq^2	$X = 1$
FSS	p^2q	$X = 2$
FSF	pq^2	$X = 1$
FFS	pq^2	$X = 1$
FFF	q^3	$X = 0$

Table 7.2. *A list of all possible results of three repetitions ($n = 3$) of a Bernoulli trial.* **S**=*"success" and* **F**=*"failure. (Substituting H for S, and T for F gives the same results for a coin tossing experiment repeated 3 times).*

Multiplication principle: if $e_1, \dots, e_k$ are independent events, then

$$\text{Prob}(e_1 \text{ and } e_2 \text{ and } \dots e_k) = \text{Prob}(e_1)\text{Prob}(e_2)\dots\text{Prob}(e_k)$$

Addition principle: if $e_1, ..., e_k$ are mutually exclusive events, then

$$\text{Prob}(e_1 \text{ or } e_2 \text{ or } \dots e_k) = \text{Prob}(e_1) + \text{Prob}(e_2) + \dots + \text{Prob}(e_k).$$

Based on the results in Table 7.2 and on the two principles outline above, we can compute the probability of obtaining 0, 1, 2, or 3 successes out of 3 trials. The results are shown in Table 7.3. In constructing Table 7.3, we have considered all the ways of obtaining 0

Probability of X heads
$\text{Prob}(X = 0) = q^3$
$\text{Prob}(X = 1) = 3pq^2$
$\text{Prob}(X = 2) = 3p^2q$
$\text{Prob}(X = 3) = p^3$

Table 7.3. *The probability of obtaining X successes out of 3 Bernoulli trials, based on results in Table 7.2 and the addition principle of probability.*

successes (there is only one such way, namely SSS, and its probability is p^3), all the ways of obtaining only one success (here we must allow for SFF, FSF, FFS, each having the same probability pq^2) etc. Since these results are mutually exclusive (only one such result is possible for any given replicate of the 3-trial experiment), the addition principle is used to compute the probability Prob(SFF or FSF or FFS).

In general, for each replicate of an experiment consisting of n Bernoulli trials, the probability of an outcome that has k successes and $n-k$ failures (in some specific order) is $p^k q^{(n-k)}$. To get the total probability of $X = k$, we need to count how many possible outcomes consist of k successes and $n-k$ failures. As illustrated by the above example, there are, in general, many such ways, since the order in which S and F appear can differ from one outcome to another. In mathematical terminology, there can be many **permutations** (i.e. arrangements of the order) of S and F that have the same number of successes in total. (See Section 11.8 for a review.) In fact, the number of ways that n trials can lead to k successes is $C(n, k)$, the **binomial coefficient**, which is, by definition, the number of ways of choosing k objects out of a collection of n objects. That binomial coefficient is

$$C(n,k) = \text{(n choose k)} = \frac{n!}{(n-k)!k!}.$$

(See Section 11.7 for the definition of factorial notation "!" used here.) We have arrived at the following result for n Bernoulli trials:

> The probability of k successes in n Bernoulli trials is
>
> $$\text{Prob}(X = k) = C(n,k)p^k q^{n-k}.$$

In the above example, with $n = 3$, we find that

$$\text{Prob}(X = 2) = C(3,2)p^2 q = 3p^2 q.$$

7.6.2 The Binomial theorem

The name **binomial coefficient** comes from the **binomial theorem**: which accounts for the expression obtained by expanding a binomial.

$$(a+b)^n = \sum_{k=0}^{n} C(n,k) a^k b^{n-k}.$$

Let us consider a few examples. A familiar example is

$$(a+b)^2 = (a+b)\cdot(a+b) = a^2 + ab + ba + b^2 = a^2 + 2ab + b^2.$$

The coefficients $C(2,2) = 1$, $C(2,1) = 2$, and $C(2,0) = 1$ appear in front of the three terms, representing, respectively, the number of ways of choosing 2 a's, 1 a, and no a's out of the n factors of $(a+b)$. [Respectively, these account for the terms a^2, ab and b^2 in the resulting expansion.] Similarly, the product of three terms is

$$(a+b)^3 = (a+b)\cdot(a+b)\cdot(a+b) = (a+b)^3 = a^3 + 3a^2 b + 3ab^2 + b^3$$

whereby coefficients are of the form $C(3,k)$ for $k = 3, 2, 1, 0$. More generally, an expansion of n terms leads to

$$\begin{aligned}(a+b)^n &= a^n + C(n,1)a^{n-1}b + C(n,2)a^{n-2}b^2 + \ldots + C(n,k)a^k b^{n-k} \\ &\quad + \ldots + C(n,n-2)a^2 b^{n-2} + C(n,n-1)ab^{n-1} + b^n \\ &= \sum_{k=0}^{n} C(n,k) a^k b^{n-k}\end{aligned}$$

$$
\begin{array}{ccccccccccc}
 & & & & & 1 & & & & & \\
 & & & & 1 & & 1 & & & & \\
 & & & 1 & & 2 & & 1 & & & \\
 & & 1 & & 3 & & 3 & & 1 & & \\
 & 1 & & 4 & & 6 & & 4 & & 1 & \\
1 & & 5 & & 10 & & 10 & & 5 & & 1
\end{array}
$$

Table 7.4. *Pascal's triangle contains the binomial coefficients of the $C(n, k)$. Each term in Pascal's triangle is obtained by adding the two diagonally above it. The top of the triangle represents $C(0, 0)$. The next row represents $C(1, 0)$ and $C(1, 1)$. For row number n, terms along the row are the binomial coefficients $C(n, k)$, starting with $k = 0$ at the beginning of the row and and going to $k = n$ at the end of the row.*

The binomial coefficients are symmetric, so that $C(n, k) = C(n, n - k)$. They are entries that occur in **Pascal's triangle**, shown in Table 7.4.

7.6.3 The binomial distribution

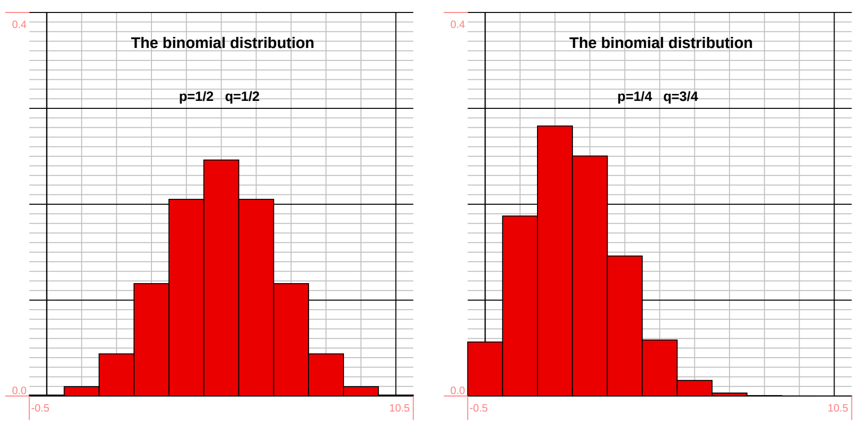

Figure 7.2. *The binomial distribution is shown here for $n = 10$. We have plotted Prob$(X = k)$ versus k for $k = 0, 1, \ldots 10$. This distribution is the same as the probability of getting X heads out of 10 coin tosses for a fair coin. In the first panel, the probability of success and failure are the same, i.e. $p = q = 0.5$. The distribution is then symmetric. In the second panel, the probability of success is $p = 1/4$, so $q = 3/4$ and the resulting distribution is skewed.*

What does the binomial theorem say about the binomial distribution? First, since there are only two possible outcomes in each Bernoulli trial, it follows that

$$p + q = 1, \quad \text{and hence} \quad (p + q)^n = 1.$$

Using the binomial theorem, we can expand the latter to obtain

$$(p+q)^n = \sum_{k=0}^{n} C(n,k)p^k q^{n-k} = \sum_{k=0}^{n} \text{Prob}(X=k) = 1.$$

That is, the sum of these terms represents the sum of probabilities of obtaining $k = 0, 1, \ldots, n$ successes. (And since this accounts for all possibilities, it follows that the sum adds up to 1.)

We can compute the mean and variance of the binomial distribution using the following tricks. We will write out an expansion for a product of the form $(px+q)^n$. Here x will be an abstract quantity introduced for convenience (i.e., for making the trick work):

$$(px+q)^n = \sum_{k=0}^{n} C(n,k)(px)^k q^{n-k} = \sum_{k=0}^{n} C(n,k)p^k q^{n-k} x^k.$$

Taking the derivative of the above with respect to x leads to:

$$n(px+q)^{n-1} \cdot p = \sum_{k=0}^{n} C(n,k)p^k q^{n-k}\, kx^{k-1},$$

which, (plugging in $x = 1$) implies that

$$np = \sum_{k=0}^{n} k \cdot C(n,k)p^k q^{n-k} = \sum_{k=0}^{n} k \cdot \text{Prob}(X=k) = \bar{X}. \tag{7.1}$$

Thus, we have found that

The mean of the binomial distribution is $\bar{X} = np$ where n is the number of trials and p is the probability of success in one trial.

We continue to compute other quantities of interest. Multiply both sides of Eqn. 7.1 by x to obtain

$$nx(px+q)^{n-1}p = \sum_{k=0}^{n} C(n,k)p^k q^{n-k} kx^k.$$

Take the derivative again. The result is

$$n(px+q)^{n-1}p + n(n-1)x(px+q)^{n-2}p^2 = \sum_{k=0}^{n} C(n,k)p^k q^{n-k} k^2 x^{k-1}.$$

Plug in $x = 1$ to get

$$np + n(n-1)p^2 = \sum_{k=0}^{n} k^2 C(n,k)p^k q^{n-k} = M_2.$$

Thereby we have calculated the second moment of the distribution, the variance, and the standard deviation. In summary, we found the following results:

The second moment M_2, the Variance V and the standard deviation σ of a binomial distribution are

$$M_2 = np + n^2p^2 - np^2,$$

$$V = M_2 - \bar{X}^2 = np - np^2 = np(1-p) = npq,$$

$$\sigma = \sqrt{npq}.$$

7.6.4 The normalized binomial distribution

We can "normalize" (i.e. rescale) the binomial random variable so that it has a convenient mean and width. To do so, define the new random variable $\tilde{X}$ to be: $\tilde{X} = X - \bar{X}$. Then $\tilde{X}$ has mean 0 and standard deviation σ. Now define

$$Z = \frac{(X - \bar{X})}{\sigma}$$

Then Z has mean 0 and standard deviation 1. In the limit as $n \to \infty$, we can approximate Z with a continuous distribution, called the standard normal distribution.

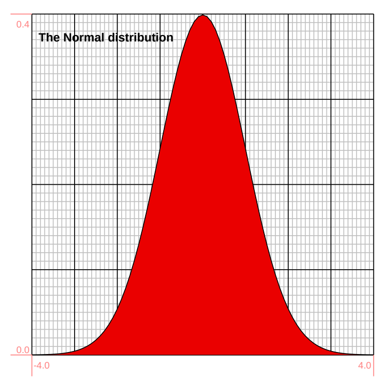

Figure 7.3. *The Normal (or Gaussian) distribution is given by equation (7.2) and has the distribution shown in this figure.*

As the number of Bernoulli trials grows, i.e. as we toss our imaginary coin in longer and longer sets ($n \to \infty$), a remarkable thing happens to the binomial distribution: it becomes smoother and smoother, until it grows to resemble a continuous distribution that looks like a "Bell curve". That curve is known as the **Gaussian** or **Normal distribution**. If we scale this curve vertically and horizontally (stretch vertically and compress horizontally by the factor $\sqrt{N}/2$) and shift its peak to $x = 0$, then we find a distribution that describes the deviation from the expected value of 50% heads. The resulting function is of the form

$$p(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \tag{7.2}$$

We will study properties of this (and other) such continuous distributions in a later section. We show a typical example of the Normal distribution in Figure 7.3. Its cumulative distribution is then shown (without and with the original distribution superimposed) in Figure 7.4.

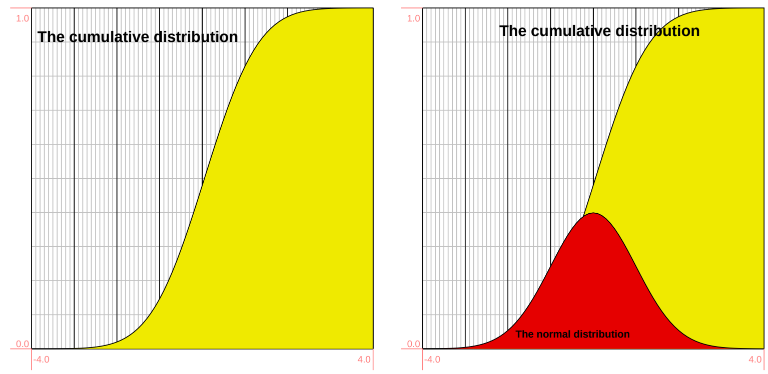

Figure 7.4. *The Normal probability density with its corresponding cumulative function.*

7.7 Hardy-Weinberg genetics

In this section, we investigate how the ideas developed in this chapter apply to genetics. We find that many of the simple concepts presented here will be useful in calculating the probability of inheriting genes from one generation to the next.

Each of us has two entire sets of chromosomes: one set is inherited from our mother, and one set comes from our father. These chromosomes carry genes, the unit of genetic material that “codes” for proteins and ultimately, through complicated biochemistry and molecular biology, determines all of our physical traits.

We will investigate how a single gene (with two “flavors”, called **alleles**) is passed from one generation to the next. We will consider a particularly simple situation, when the single gene determines some physical trait (such as eye color). The trait (say blue or green eyes) will be denoted the **phenotype** and the actual pair of genes (one on each parentally derived chromosome) will be called the **genotype**.

Suppose that the gene for eye color comes in two forms that will be referred to as **A** and **a**. For example, **A** might be an allele for blue eyes, whereas **a** could be an allele for brown eyes. Consider the following “experiment”: select a random individual from the population of interest, and examine the region in one of their chromosomes determining eye colour. Then there are two possible mutually exclusive outcomes, **A** or **a**; according to our previous definition, the experiment just described is a Bernoulli trial.

The actual eye color phenotype will depend on both inherited alleles, and hence, we are interested in a “repeated Bernoulli trial” with $n = 2$. In principle, each chromosome will come with one or the other allele, so each individual would have one of the following pairs of combinations **AA, Aa, aA**, or **aa**. The order Aa or aA is synonymous, so only the

Genotype:	**aA**	**AA**	**aa**	**Aa**
Probability:	pq	p^2	q^2	pq

Genotype:	**aA** or **Aa**	**AA**	**aa**
Probability:	$2pq$	p^2	q^2

Table 7.5. *If the probability of finding allele* **A** *is* p *and the probability of finding allele* **A** *is* q*, then the eye color gene probabilities are as shown in the top table. However, because genotype* **Aa** *is equivalent to genotype* **aA***, we have combined these outcomes in the revised second table.*

number of alleles of type **A** (or equivalently of type **a**) is important.

Suppose we know that the fraction of all genes for eye color of type **A** in the population is p, and the fraction of all genes for eye color of type **a** is q, where $p + q = 1$. (We have used the fact that there are *only* two possibilities for the gene type, of course.) Then we can interpret p and q as probabilities that a gene selected at random from the population will turn out to be type **a** (respectively **A**), i.e., Prob(**A**) = p, Prob(**a**)=q.

Now suppose we draw at random two alleles out of the (large) population. If the population size is N, then, on average we would expect Np^2 individuals of type **AA**, Nq^2 of type **aa** and $2Npq$ individuals of the mixed type. Note that the sum of the probabilities of all the genotypes is

$$p^2 + 2pq + q^2 = (p + q)^2 = 1.$$

(We have seen this before in the discussion of Bernoulli trials, and in the definition of properties of probability.)

7.7.1 Random non-assortative mating

We now examine what happens if mates are chosen randomly and offspring arise from such parents. The father and mother each pass down one or another copy of their alleles to the progeny. We investigate how the proportion of genes of various types is arranged, and whether it changes in the next generation. In Table 7.6, we show the possible genotypes of the mother and father, and calculate the probability that mating of such individuals would occur under the assumption that choice of mate is random - i.e., does not depend at all on "eye color". We assume that the allele donated by the father (carried in his sperm) is independent of the allele found in the mother's egg cell[28]. This means that we can use the multiplicative property of probability to determine the probability of a given combination of parental alleles. (i.e. Prob(x and y)=Prob(x)·Prob(y)).

For example, the probability that a couple chosen at random will consist of a woman of genotype **aA** and a man of genotype **aa** is a product of the fraction of females that are of type **aA** and the fraction of males that are of type **aa**. But that is just $(2pq)(p^2)$, or simply $2p^3q$. Now let us examine the distribution of possible offspring of various parents.

[28]Recall that the sperm and the egg each have one single set of chromosomes, and their union produces the zygote that carries the doubled set of chromosomes.

In Table 7.6, we note, for example, that if the couple are both of type **aA**, each parent can "donate" either **a** or **A** to the progeny, so we expect to see children of types **aa, aA, AA** in the ratio 1:2:1 (regardless of the values of p and q).

We can now group together and summarize all the progeny of a given genotype, with the probabilities that they are produced by one or another such random mating. Using this table, we can then determine the probability of each of the three genotypes in the next generation.

	Mother:	**AA** p^2	**aA** $2pq$	**aa** q^2
Father: **AA** p^2		**AA** p^4	$\frac{1}{2}$**aA** $\frac{1}{2}$**AA** $2pqp^2$	**Aa** p^2q^2
aA $2pq$		$\frac{1}{2}$**aA** $\frac{1}{2}$**AA** $2pqp^2$	$\frac{1}{4}$**aa** $\frac{1}{2}$**aA** $\frac{1}{4}$**AA** $4p^2q^2$	$\frac{1}{2}$**aa** $\frac{1}{2}$**Aa** $2pqq^2$
aa q^2		**Aa** p^2q^2	$\frac{1}{2}$**aA** $\frac{1}{2}$**aa** $2pqq^2$	**aa** q^4

Table 7.6. *The frequency of progeny of various types in Hardy-Weinberg genetics can be calculated as shown in this "mating table". The genotype of the mother is shown across the top and the father's genotype is shown on the left column. The various progeny resulting from mating are shown as entries in bold face. The probabilities of the given progeny are directly under those entries. (We did not simplify the expressions - this is to emphasize that they are products of the original parental probabilities.)*

Example 7.3 (Probability of AA progeny) Find the probability that a random (Hardy Weinberg) mating will give rise to a progeny of type **AA**.

Solution 1

Using Table 7.6, we see that there are only four ways that a child of type **AA** can result from a mating: either both parents are **AA**, or one or the other parent is **Aa**, or both parents are **Aa**. Thus, for children of type **AA** the probability is

$$\text{Prob(child of type } \mathbf{AA}) = p^4 + \frac{1}{2}(2pqp^2) + \frac{1}{2}(2pqp^2) + \frac{1}{4}(4p^2q^2).$$

Simplifying leads to

$$\text{Prob(child of type } \mathbf{AA}) = p^2(p^2 + 2qp + q^2) = p^2(p+q)^2 = p^2.$$

In the problem set, we also find that the probability of a child of type **aA** is $2qp$, the probability of the child being type **aa** is q^2. We thus observe that the frequency of genotypes of the progeny is exactly the same as that of the parents. This type of genetic makeup is termed Hardy-Weinberg genetics.

Alternate solution

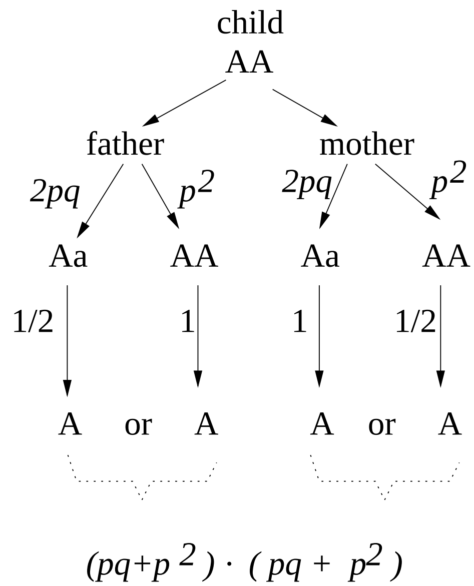

Figure 7.5. *A tree diagram to aid the calculation of the probability that a child with genotype* **AA** *results from random assortative (Hardy Weinberg) mating.*

In Figure 7.5, we show an alternate solution to the same problem using a tree diagram. Reading from the top down, we examine all the possibilities at each branch point. A child **AA** cannot have any parent of genotype aa, so both father and mother's genotype could only have been one of **AA** or **Aa**. Each arrow indicating the given case is accompanied by the probability of that event. (For example, a random individual has probability $2pq$ of having genotype **Aa**, as shown on the arrows from the father and mother to these genotypes.) Continuing down the branches, we ask with what probability the given parent would have contributed an allele of type **A** to the child. For a parent of type **AA**, this is certainly true, so the given branch carries probability 1. For a parent of type **Aa**, the probability that **A** is passed down to the child is only 1/2. The *combined* probability is computed as follows: we determine the probability of getting an **A** from father (of type **AA** OR **Aa**): This is Prob(**A** from father)=$(1/2)2pq + 1 \cdot p^2) = (pq + p^2)$ and multiply it by a similar probability of getting **A** from the mother (of type **AA** OR **Aa**). (We must multiply, since we need **A** from the father AND **A** from the mother for the genotype **AA**.) Thus,

Prob(child of type **AA**) =$(pq + p^2)(pq + p^2) = p^2(q + p)^2 = p^2 \cdot 1 = p^2$.

It is of interest to investigate what happens when one of the assumptions we made is relaxed, for example, when the genotype of the individual has an impact on survival or on the ability to reproduce. While this is beyond our scope here, it forms an important theme in the area of genetics.

7.8 Random walker

In this section we discuss an application of the binomial distribution to the process of a random walk. A shown in Figure 7.6(a), we consider a straight (1 dimensional) path and an erratic walker who takes steps randomly to the left or right. We will assume that the walker never stops. With probability p, she takes a step towards the right, and with probability q she takes a step towards the left. (Since these are the only two choices, it must be true that $p + q = 1$.) In Figure 7.6(b) we show the walker's position, x plotted versus the number of steps (n) she has taken. (We may as well assume that the steps occur at regular intervals of time, so that the horizontal axis of this plot can be thought of as a time axis.)

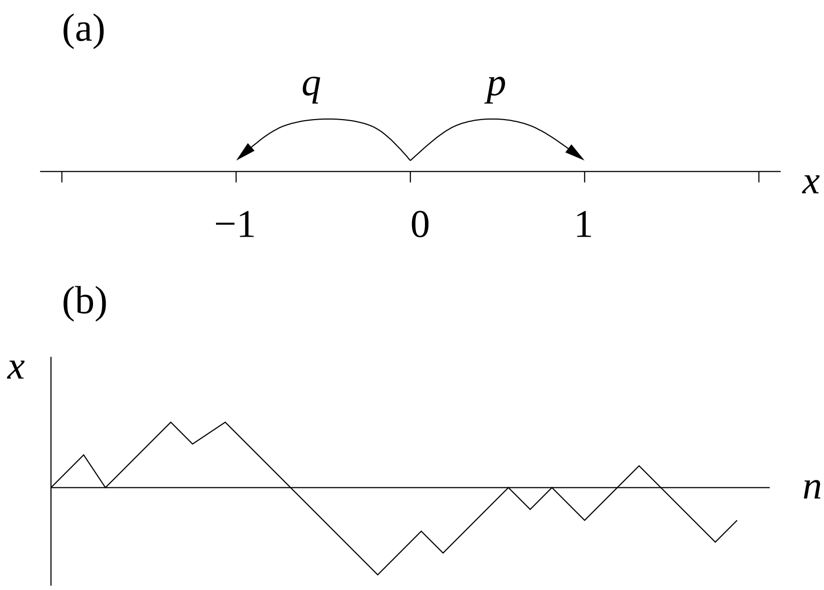

Figure 7.6. *A random walker in 1 dimension takes a step to the right with probability p and a step to the left with probability q.*

The process described here is classic, and often attributed to a drunken wanderer. In our case, we could consider this motion as a 1D simplification of the random tumbles and swims of a bacterium in its turbulent environment. it is usually the case that a goal of this swim is a search for some nutrient source, or possibly avoidance of poor environmental conditions. We shall see that if the probabilities of left and right motion are unequal (i.e. the motion is biased in one direction or another) this swimmer tends to drift along towards a preferred direction.

In this problem, each step has only two outcomes (analogous to a trial in a Bernoulli experiment). We could imagine the walker tossing a coin to determine whether to move right or left. We wish to characterize the probability of the walker being at a certain position at a given time, and to find her expected position after n steps. Our familiarity with Bernoulli trials and the binomial distribution will prove useful in this context.

Example

(a) What is the probability of a run of steps as follows: RLRRRLRLLLL

(b) Find the probability that the walker moves k steps to the right out of a total run of n consecutive steps.

(c) Suppose that $p = q = 1/2$. What is the probability that a walker starting at the origin returns to the origin on her 10'th step?

Solution

(a) The probability of the run RLRRRLRLLL is the product $pqpppqpqqq = p^5q^5$. Note the similarity to the question "What is the probability of tossing HTHHHTHTTT?"

(b) This problem is identical to the problem of k heads in n tosses of a coin. The probability of such an event is given by a term in the binomial distribution:

P(k out of n moves to right)=$C(n,k)p^kq^{n-k}$.

(c) The walker returns to the origin after 10 steps only if she has taken 5 steps to the left (total) and 5 steps to the right (total). The order of the steps does not matter. Thus this problem reduces to the problem (b) with 5 steps out of 10 taken to the right. The probability is thus

P(back at 0 after 10 steps) = P(5 out of 10 steps to right)

=$C(10,5)p^5q^5 = C(10,5)\left(\frac{1}{2}\right)^{10} = \left(\frac{10!}{5!5!}\right)\frac{1}{1024} = 0.24609$

Mean position

We now ask how to determine the expected position of the walker after n steps, i.e. how the mean value of x depends on the number of steps and the probabilities associated with each step. After 1 step, with probability p the position is $x = +1$ and with probability q, the position is $x = -1$. The expected (mean) position after 1 move is thus

$$x_1 = p(+1) + q(-1) = p - q$$

But the process follows a binomial distribution, and thus the mean after n steps is

$$x_n = n(p - q).$$

7.9 Summary

In this chapter, we introduced the notion of discrete probability of elementary events. We learned that a probability is always a number between 0 and 1, and that the sum of (discrete) probabilities of all possible (discrete) outcomes is 1. We then described how to combine probabilities of elementary events to calculate probabilities of compound independent events in a variety of simple experiments. We defined the notion of a Bernoulli trial, such as tossing of a coin, and studied this in detail.

We investigated a number of ways of describing results of experiments, whether in tabular or graphical form, and we used the distribution of results to define simple numerical descriptors. The **mean** is a number that, more or less, describes the location of the "center" of the distribution (analogous to center of mass), defined as follows:

The mean (**expected value**) $\bar{x}$ of a probability distribution is

$$\bar{x} = \sum_{i=0}^{n} x_i p(x_i).$$

The standard deviation is, roughly speaking, the "width" of the distribution.

The **standard deviation**, σ is

$$\sigma = \sqrt{V}$$

where V is the **variance**,

$$V = \sum_{i=0}^{n} (x_i - \bar{x})^2 p(x_i).$$

While the chapter was motivated by results of a real experiment, we then investigated theoretical distributions, including the binomial. We found that the distribution of events in a repetition of a Bernoulli trial (e.g. coin tossed n times) was a binomial distribution, and we computed the mean of that distribution.

Suppose that the probability of one of the events, say event e_1 in a Bernoulli trial is p (and hence the probability of the other event e_2 is $q = 1 - p$), then

$$P(\text{k occurrences of given event out of n trials}) = \frac{n!}{k!(n-k)!} p^k q^{n-k}.$$

This is called the **binomial** distribution. The mean of the binomial distribution, i.e. the mean number of events e_1 in n repeated Bernoulli trials is

$$\bar{x} = np.$$

Chapter 8

Continuous probability distributions

8.1 Introduction

In Chapter 7, we explored the concepts of probability in a discrete setting, where outcomes of an experiment can take on only one of a finite set of values. Here we extend these ideas to continuous probability. In doing so, we will see that quantities such as mean and variance that were previously defined by sums will now become definite integrals. Here again, we will see the concepts of integral calculus in the context of practical examples and applications.

We begin by extending the idea of a discrete random variable to the continuous case. We call x a continuous random variable in $a \leq x \leq b$ if x can take on any value in this interval. An example of a random variable is the height of a person, say an adult male, selected randomly from a population. (This height typically takes on values in the range $0.5 \leq x \leq 3$ meters, say, so $a = 0.5$ and $b = 3$.)

If we select a male subject at random from a large population, and measure his height, we might expect to get a result in the proximity of 1.7-1.8 meters most often - thus, such heights will be associated with a larger value of probability than heights in some other interval of equal length, e.g. heights in the range $2.7 < x < 2.8$ meters, say. Unlike the case of discrete probability, however, the measured height can take on any real number within the interval of interest. This leads us to redefine our idea of a continuous probability, using a continuous function in place of the discrete bar-graph seen in Chapter 7.

8.2 Basic definitions and properties

Here we extend previous definitions from Chapter 7 to the case of continuous probability. One of the most important differences is that we now consider a probability **density**, rather than a value of the probability per se[29]. First and foremost, we observe that now $p(x)$ will no longer be a probability, but rather " a probability *per unit* x". This idea is analo-

[29]This leap from discrete values that are the probability of an outcome (as seen in Chapter 7) to a probability density is challenging for many students. Reinforcing the analogy with discrete masses versus distributed mass density (discussed in Chapter 5) may be helpful.

gous to the connection between the mass of discrete beads and a continuous mass density, encountered previously in Chapter 5.

Definition

A function $p(x)$ is a probability density provided it satisfies the following properties:

1. $p(x) \geq 0$ for all x.
2. $\int_a^b p(x)\,dx = 1$ where the possible range of values of x is $a \leq x \leq b$.

The probability that a random variable x takes on values in the interval $a_1 \leq x \leq a_2$ is defined as

$$\boxed{\int_{a_1}^{a_2} p(x)\,dx.}$$

The transition to probability density means that the quantity $p(x)$ does not carry the same meaning as our previous notation for probability of an outcome x_i, namely $p(x_i)$ in the discrete case. In fact, $p(x)dx$, or its approximation $p(x)\Delta x$ is now associated with the probability of an outcome whose values is “close to x”.

Unlike our previous discrete probability, we will not ask “what is the probability that x takes on some exact value?” Rather, we ask for the probability that x is within some range of values, and this is computed by performing an integral[30].

Having generalized the idea of probability, we will now find that many of the associated concepts have a natural and straight-forward generalization as well. We first define the cumulative function, and then show how the mean, median, and variance of a continuous probability density can be computed. Here we will have the opportunity to practice integration skills, as integrals replace the sums in such calculations.

Definition

For experiments whose outcome takes on values on some interval $a \leq x \leq b$, we define a cumulative function, $F(x)$, as follows:

$$\boxed{F(x) = \int_a^x p(s)\,ds.}$$

Then $F(x)$ represents the probability that the random variable takes on a value in the range (a, x) [31]. The cumulative function is simply the area under the probability density (between the left endpoint of the interval, a, and the point x).

The above definition has several implications:

---

[30]Remark: the probability that x is exactly equal to b is the integral $\displaystyle\int_b^b p(x)\,dx$. But this integral has a value zero, by properties of the definite integral.

[31]By now, the reader should be comfortable with the use of “s” as the “dummy variable” in this formula, where x plays the role of right endpoint of the interval of integration.

Properties of continuous probability

1. Since $p(x) \geq 0$, the cumulative function is an *increasing* function.
2. The connection between the probability density and its cumulative function can be written (using the Fundamental Theorem of Calculus) as
$$p(x) = F'(x).$$
3. $F(a) = 0$. This follows from the fact that
$$F(a) = \int_a^a p(s)\, ds.$$
By a property of the definite integral, this is zero.
4. $F(b) = 1$. This follows from the fact that
$$F(b) = \int_a^b p(s)\, ds = 1$$
by Property 2 of the definition of the probability density, $p(x)$.
5. The probability that x takes on a value in the interval $a_1 \leq x \leq a_2$ is the same as
$$F(a_2) - F(a_1).$$
This follows from the additive property of integrals and the Fundamental Theorem of Calculus:
$$\int_a^{a_2} p(s)\, ds - \int_a^{a_1} p(s)\, ds = \int_{a_1}^{a_2} p(s)\, ds = \int_{a_1}^{a_2} F'(s)\, ds = F(a_2) - F(a_1)$$

Finding the normalization constant

Not every real-valued function can represent a probability density. For one thing, the function must be positive everywhere. Further, the total area under its graph should be 1, by Property 2 of a probability density. Given an arbitrary positive function, $f(x) \geq 0$, on some interval $a \leq x \leq b$ such that
$$\int_a^b f(x)dx = A > 0,$$
we can always define a corresponding probability density, $p(x)$ as
$$p(x) = \frac{1}{A} f(x), \quad a \leq x \leq b.$$
It is easy to check that $p(x) \geq 0$ and that $\int_a^b p(x)dx = 1$. Thus we have converted the original function to a probability density. This process is called **normalization**, and the constant $C = 1/A$ is called the normalization constant[32].

[32] The reader should recognize that we have essentially rescaled the original function by dividing it by the "area" A. This is really what normalization is all about.

8.2.1 Example: probability density and the cumulative function

Consider the function $f(x) = \sin(\pi x/6)$ for $0 \le x \le 6$.

(a) Normalize the function so that it describes a probability density.

(b) Find the cumulative distribution function, $F(x)$.

Solution

The function is positive in the interval $0 \le x \le 6$, so we can define the desired probability density. Let

$$p(x) = C \sin\left(\frac{\pi}{6}x\right).$$

(a) We must find the normalization constant, C, such that Property 2 of continuous probability is satisfied, i.e. such that

$$1 = \int_0^6 p(x)\, dx.$$

Carrying out this computation leads to

$$\int_0^6 C \sin\left(\frac{\pi}{6}x\right)\, dx = C\frac{6}{\pi}\left(-\cos\left(\frac{\pi}{6}x\right)\right)\Big|_0^6 = C\frac{6}{\pi}\left(1 - \cos(\pi)\right) = C\frac{12}{\pi}$$

(We have used the fact that $\cos(0) = 1$ in a step here.) But by Property 2, for $p(x)$ to be a probability density, it must be true that $C(12/\pi) = 1$. Solving for C leads to the desired normalization constant,

$$C = \frac{\pi}{12}.$$

Note that this calculation is identical to finding the area

$$A = \int_0^6 \sin\left(\frac{\pi}{6}x\right)\, dx,$$

and setting the normalization constant to $C = 1/A$.

Once we rescale our function by this constant, we get the probability density,

$$p(x) = \frac{\pi}{12}\sin\left(\frac{\pi}{6}x\right).$$

This density has the property that the total area under its graph over the interval $0 \le x \le 6$ is 1. A graph of this probability density function is shown as the black curve in Figure 8.1.

(b) We now compute the cumulative function,

$$F(x) = \int_0^x p(s)\ ds = \frac{\pi}{12}\int_0^x \sin\left(\frac{\pi}{6}s\right)\ ds$$

Carrying out the calculation[33] leads to

$$F(x) = \frac{\pi}{12}\cdot\frac{6}{\pi}\left(-\cos\left(\frac{\pi}{6}s\right)\right)\Bigg|_0^x = \frac{1}{2}\left(1-\cos\left(\frac{\pi}{6}x\right)\right).$$

This cumulative function is shown as a red curve in Figure 8.1.

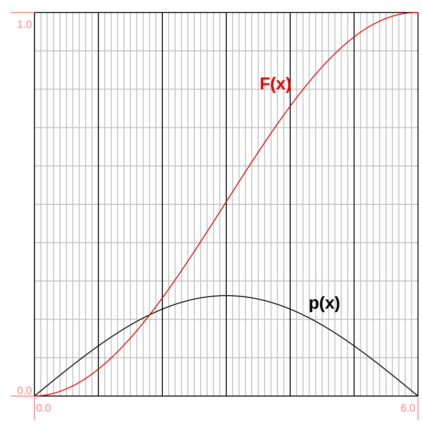

Figure 8.1. *The probability density $p(x)$ (black), and the cumulative function $F(x)$ (red) for Example 8.2.1. Note that the area under the black curve is 1 (by normalization), and thus the value of $F(x)$, which is the cumulative area function is 1 at the right endpoint of the interval.*

8.3 Mean and median

When we are given a distribution, we often want to describe it with simpler numerical values that characterize its "center": the mean and the median both give this type of information. We also want to describe whether the distribution is narrow or fat - i.e. how clustered it is about its "center". The variance and higher moments will provide that type of information.

Recall that in Chapter 5 for mass density $\rho(x)$, we defined a **center of mass**,

$$\bar{x} = \frac{\int_a^b x\rho(x)\ dx}{\int_a^b \rho(x)\ dx}. \tag{8.1}$$

[33]Notice that the integration involved in finding $F(x)$ is the same as the one done to find the normalization constant. The only difference is the ultimate step of evaluating the integral at the variable endpoint x rather than the fixed endpoint $b = 6$.

The mean of a probability density is defined similarly, but the definition simplifies by virtue of the fact that $\int_a^b p(x)\,dx = 1$. Since probability distributions are normalized, the denominator in Eqn. (8.1) is simply 1.Consequently, the **mean** of a probability density is given as follows:

Definition

For a random variable in $a \le x \le b$ and a probability density $p(x)$ defined on this interval, the **mean** or **average** value of x (also called the **expected value**), denoted $\bar{x}$ is given by

$$\boxed{\bar{x} = \int_a^b x p(x)\,dx.}$$

To avoid confusion note the distinction between the mean as an average value of x versus the average value of the function p over the given interval. Reviewing Example 5.3.3 may help to dispel such confusion.

The idea of median encountered previously in grade distributions also has a parallel here. Simply put, the median is the value of x that splits the probability distribution into two portions whose areas are identical.

Definition

The **median** x_{med} of a probability distribution is a value of x in the interval $a \le x_{med} \le b$ such that

$$\boxed{\int_a^{x_{med}} p(x)\,dx = \int_{x_{med}}^b p(x)\,dx = \frac{1}{2}.}$$

It follows from this definition that the median is the value of x for which the cumulative function satisfies

$$\boxed{F(x_{med}) = \frac{1}{2}.}$$

8.3.1 Example: Mean and median

Find the mean and the median of the probability density found in Example 8.2.1.

Solution

To find the **mean** we compute

$$\bar{x} = \frac{\pi}{12}\int_0^6 x \sin\left(\frac{\pi}{6}x\right)\,dx.$$

Integration by parts is required here[34]. Let $u = x,\ dv = \sin\left(\frac{\pi}{6}x\right)\ dx$.

Then $du = dx, v = -\frac{6}{\pi}\cos\left(\frac{\pi}{6}x\right)$. The calculation is then as follows:

$$\bar{x} = \frac{\pi}{12}\left(-x\frac{6}{\pi}\cos\left(\frac{\pi}{6}x\right)\Big|_0^6 + \frac{6}{\pi}\int_0^6 \cos\left(\frac{\pi}{6}x\right)dx\right)$$
$$= \frac{1}{2}\left(-x\cos\left(\frac{\pi}{6}x\right)\Big|_0^6 + \frac{6}{\pi}\sin\left(\frac{\pi}{6}x\right)\Big|_0^6\right)$$
$$= \frac{1}{2}\left(-6\cos(\pi) + \frac{6}{\pi}\sin(\pi) - \frac{6}{\pi}\sin(0)\right) = \frac{6}{2} = 3. \tag{8.2}$$

(We have used $\cos(\pi) = -1, \sin(0) = \sin(\pi) = 0$ in the above.)

To find the **median**, x_{med}, we look for the value of x for which

$$F(x_{med}) = \frac{1}{2}.$$

Using the form of the cumulative function from Example 8.2.1, we find that

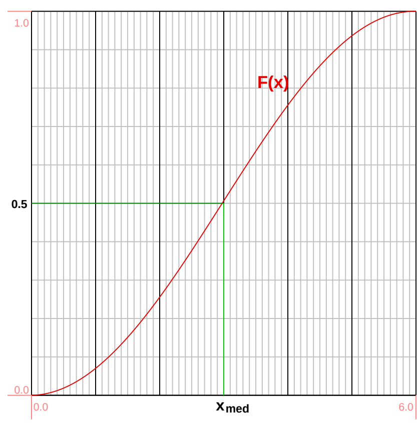

Figure 8.2. *The cumulative function $F(x)$ (red) for Example 8.2.1 in relation to the* **median***, as computed in Example 8.3.1. The median is the value of x at which $F(x) = 0.5$, as shown in green.*

$$\int_0^{x_{med}} \sin\left(\frac{\pi}{6}s\right)\ ds = \frac{1}{2} \quad \Rightarrow \quad \frac{1}{2}\left(1 - \cos\left(\frac{\pi}{6}x_{med}\right)\right) = \frac{1}{2}.$$

[34]Recall from Chapter 6 that $\int u dv = vu - \int v du$. Calculations of the mean in continuous probability often involve Integration by Parts (IBP), since the integrand consists of an expression $xp(x)dx$. The idea of IBP is to reduce the integration to something involving only $p(x)dx$, which is done essentially by "differentiating" the term $u = x$, as we show here.

Here we must solve for the unknown value of x_{med}.

$$1 - \cos\left(\frac{\pi}{6}x_{med}\right) = 1, \quad \Rightarrow \quad \cos\left(\frac{\pi}{6}x_{med}\right) = 0.$$

The angles whose cosine is zero are $\pm\pi/2, \pm 3\pi/2$ etc. We select the angle so that the resulting value of x_{med} will be inside the relevant interval ($0 \le x \le 6$ for this example), i.e. $\pi/2$. This leads to

$$\frac{\pi}{6}x_{med} = \frac{\pi}{2}$$

so the median is

$$x_{med} = 3.$$

In other words, we have found that the point x_{med} subdivides the interval $0 \le x \le 6$ into two subintervals whose probability is the same. The relationship of the median and the cumulative function $F(x)$ is illustrated in Fig 8.2.

Remark

A glance at the original probability distribution should convince us that it is symmetric about the value $x = 3$. Thus we should have anticipated that the mean and median of this distribution would both occur at the same place, i.e. at the midpoint of the interval. This will be true in general for symmetric probability distributions, just as it was for symmetric mass or grade distributions.

8.3.2 How is the mean different from the median?

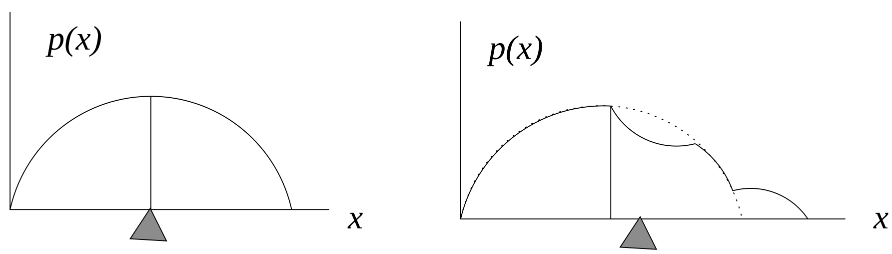

Figure 8.3. *In a symmetric probability distribution (left) the mean and median are the same. If the distribution is changed slightly so that it is no longer symmetric (as shown on the right) then the median may still be the same, which the mean will have shifted to the new "center of mass" of the probability density.*

We have seen in Example 8.3.1 that for symmetric distributions, the mean and the median are the same. Is this always the case? When are the two different, and how can we understand the distinction?

Recall that the *mean* is closely associated with the idea of a center of mass, a concept from physics that describes the location of a pivot point at which the entire "mass" would exactly balance. It is worth remembering that

$$\text{mean of } p(x) = \text{ expected value of } x = \text{ average value of } x.$$

This concept is not to be confused with the average value *of a function*, which is an average value of the y coordinate, i.e., the average height of the function on the given interval.

The *median* simply indicates a place at which the "total mass" is subdivided into two equal portions. (In the case of probability density, each of those portions represents an equal area, $A_1 = A_2 = 1/2$ since the total area under the graph is 1 by definition.)

Figure 8.3 shows how the two concepts of *median* (indicated by vertical line) and *mean* (indicated by triangular "pivot point") differ. At the left, for a symmetric probability density, the mean and the median coincide, just as they did in Example 8.3.1. To the right, a small portion of the distribution was moved off to the far right. This change did not affect the location of the median, since the total areas to the right and to the left of the vertical line are still equal. However, the fact that part of the mass is farther away to the right leads to a shift in the mean of the distribution, to compensate for the change.

Simply put, the mean contains more information about the way that the distribution is arranged spatially. This stems from the fact that the mean of the distribution is a "sum" - i.e. integral - of terms of the form $xp(x)\Delta x$. Thus the location along the x axis, x, not just the "mass", $p(x)\Delta x$, affects the contribution of parts of the distribution to the value of the *mean*.

8.3.3 Example: a nonsymmetric distribution

We slightly modify the function used in Example 8.2.1 to the new expression

$$f(x) = x \sin(\pi x/6) \quad \text{for} \quad 0 \le x \le 6.$$

This results in a nonsymmetric probability density, shown in black in Figure 8.4. Steps in obtaining $p(x)$ would be similar[35], but we have to carry out an integration by parts to find the normalization constant and/or to calculate the cumulative function, $F(x)$. Further, to compute the mean of the distribution we have to integrate by parts twice.

Alternatively, we can carry out all such computations (approximately) using the spreadsheet, as shown in Figure 8.4. We can plot $f(x)$ using sufficiently fine increments Δx along the x axis and compute the approximation for its integral by adding up the quantities $f(x)\Delta x$. The area under the curve A, and hence the normalization constant ($C = 1/A$) will be thereby determined (at the point corresponding to the end of the interval, $x = 6$). It is then an easy matter to replot the revised function $f(x)/A$, which corresponds to the normalized probability density. This is the curve shown in black in Figure 8.4. In the problem sets, we leave as an exercise for the reader how to determine the median and the mean using the same spreadsheet tool for a related (simpler) example.

8.4 Applications of continuous probability

In the next few sections, we explore applications of the ideas developed in this chapter to a variety of problems. We treat the decay of radioactive atoms, consider distribution of

[35] This is good practice, and the reader is encouraged to do this calculation.

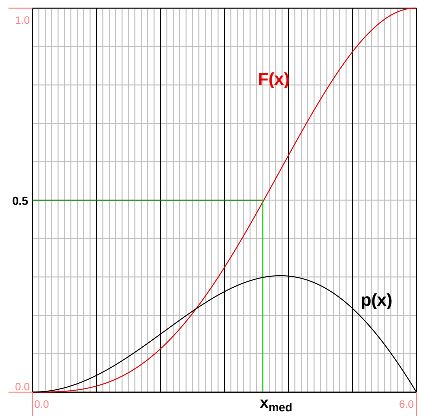

Figure 8.4. *As in Figures 8.1 and 8.2, but for the probability density* $p(x) = (\pi/36)x\sin(\pi x/6)$. *This function is not symmetric, so the mean and median are not the same. From this figure, we see that the median is approximately* $x_{med} = 3.6$. *We do not show the mean (which is close but not identical). We can compute both the mean and the median for this distribution using numerical integration with the spreadsheet. We find that the mean is* $\bar{x} = 3.5679$. *Note that the "most probable value", i.e. the point at which* $p(x)$ *is maximal is at* $x = 3.9$, *which is again different from both the mean and the median.*

heights in a population, and explore how the distribution of radii is related to the distribution of volumes in raindrop drop sizes. The interpretation of the probability density and the cumulative function, as well as the means and medians in these cases will form the main focus of our discussion.

8.4.1 Radioactive decay

Radioactive decay is a probabilistic phenomenon: an atom spontaneously emits a particle and changes into a new form. We cannot predict exactly when a given atom will undergo this event, but we can study a large collection of atoms and draw some interesting conclusions.

We can define a probability density function that represents the probability per unit time that an atom would decay at time t. It turns out that a good candidate for such a function is

$$p(t) = Ce^{-kt},$$

where k is a constant that represents the rate of decay (in units of 1/time) of the specific radioactive material. In principle, this function is defined over the interval $0 \le t \le \infty$; that is, it is possible that we would have to wait a "very long time" to have *all* of the atoms decay. This means that these integrals have to be evaluated "at infinity", leading to an **improper integral**. Using this probability density for atom decay, we can characterize the mean and median decay time for the material.

Normalization

We first find the constant of normalization, i.e. find the constant C such that

$$\int_0^\infty p(t)\,dt = \int_0^\infty Ce^{-kt}\,dt = 1.$$

Recall that an integral of this sort, in which one of the endpoints is at infinity is called an **improper** integral[36]. Some care is needed in understanding how to handle such integrals, and in particular when they "exist" (in the sense of producing a finite value, despite the infinitely long domain of integration). We will delay full discussion to Chapter 10, and state here the definition:

$$I = \int_0^\infty Ce^{-kt}\,dt \equiv \lim_{T\to\infty} I_T \quad \text{where} \quad I_T = \int_0^T Ce^{-kt}\,dt.$$

The idea is to compute an integral over a finite interval $0 \le t \le T$ and then take a limit as the upper endpoint, T goes to infinity ($T \to \infty$). We compute:

$$I_T = C\int_0^T e^{-kt}\,dt = C\left[\frac{e^{-kt}}{-k}\right]\Bigg|_0^T = \frac{1}{k}C(1-e^{-kT}).$$

Now we take the limit:

$$I = \lim_{T\to\infty} I_T = \lim_{T\to\infty}\frac{1}{k}C(1-e^{-kT}) = \frac{1}{k}C(1-\lim_{T\to\infty}e^{-kT}). \tag{8.3}$$

To compute this limit, recall that for $k > 0, T > 0$, the exponential term in Eqn. 8.3 decays to zero as T increases, so that

$$\lim_{T\to\infty} e^{-kT} = 0.$$

Thus, the second term in braces in the integral I in Eqn. 8.3 will vanish as $T \to \infty$ so that the value of the improper integral will be

$$I = \lim_{T\to\infty} I_T = \frac{1}{k}C.$$

To find the constant of normalization C we require that $I = 1$, i.e. $\frac{1}{k}C = 1$, which means that

$$C = k.$$

Thus the (normalized) probability density for the decay is

$$p(t) = ke^{-kt}.$$

This means that the fraction of atoms that decay between time t_1 and t_2 is

$$k\int_{t_1}^{t_2} e^{-kt}\,dt.$$

[36]We have already encountered such integrals in Sections 3.8.5 and 4.5. See also, Chapter 10 for a more detailed discussion of improper integrals.

Cumulative decays

The fraction of the atoms that decay between time 0 and time t (i.e. "any time up to time t" or "by time t - note subtle wording"[37]) is

$$F(t) = \int_0^t p(s)\, ds = k \int_0^t e^{-ks}\, ds.$$

We can simplify this expression by integrating:

$$F(t) = k \left[\frac{e^{-ks}}{-k} \right] \Big|_0^t = -\left[e^{-kt} - e^0\right] = 1 - e^{-kt}.$$

Thus, the probability of the atoms decaying by time t (which means anytime up to time t) is

$$F(t) = 1 - e^{-kt}.$$

We note that $F(0) = 0$ and $F(\infty) = 1$, as expected for the cumulative function.

Median decay time

As before, to determine the median decay time, t_m (the time at which half of the atoms have decayed), we set $F(t_m) = 1/2$. Then

$$\frac{1}{2} = F(t_m) = 1 - e^{-kt_m},$$

so we get

$$e^{-kt_m} = \frac{1}{2}, \quad \Rightarrow \quad e^{kt_m} = 2, \quad \Rightarrow \quad kt_m = \ln 2, \quad \Rightarrow \quad t_m = \frac{\ln 2}{k}.$$

Thus half of the atoms have decayed by this time. (Remark: this is easily recognized as the *half life* of the radioactive process from previous familiarity with exponentially decaying functions.)

Mean decay time

The mean time of decay $\bar{t}$ is given by

$$\bar{t} = \int_0^\infty tp(t)\, dt.$$

We compute this integral again as an improper integral by taking a limit as the top endpoint increases to infinity, i.e. we first find

$$I_T = \int_0^T tp(t)\, dt,$$

[37] Note that the precise English wording is subtle, but very important here. "By time t" means that the event could have happened at any time right up to time t.

and then set

$$\bar{t} = \lim_{T\to\infty} I_T.$$

To compute I_T we use integration by parts:

$$I_T = \int_0^T tke^{-kt}\,dt = k\int_0^T te^{-kt}\,dt.$$

Let $u = t, dv = e^{-kt}\,dt$. Then $du = dt, v = e^{-kt}/(-k)$, so that

$$\begin{aligned} I_T &= k\left[t\frac{e^{-kt}}{(-k)} - \int \frac{e^{-kt}}{(-k)}\,dt\right]\Bigg|_0^T = \left[-te^{-kt} + \int e^{-kt}\,dt\right]\Bigg|_0^T \\ &= \left[-te^{-kt} - \frac{e^{-kt}}{k}\right]\Bigg|_0^T = \left[-Te^{-kT} - \frac{e^{-kT}}{k} + \frac{1}{k}\right] \end{aligned}$$

Now as $T \to \infty$, we have $e^{-kT} \to 0$ so that

$$\bar{t} = \lim_{T\to\infty} I_T = \frac{1}{k}.$$

Thus the mean or expected decay time is

$$\bar{t} = \frac{1}{k}.$$

8.4.2 Discrete versus continuous probability

In Chapter 5.3, we compared the treatment of two types of mass distributions. We first explored a set of discrete masses strung along a "thin wire". Later, we considered a single "bar" with a continuous distribution of density along its length. In the first case, there was an unambiguous meaning to the concept of "mass at a point". In the second case, we could assign a mass to some *section* of the bar between, say $x = a$ and $x = b$. (To do so we had to integrate the mass density on the interval $a \le x \le b$.) In the first case, we talked about the mass of the objects, whereas in the latter case, we were interested in the idea of density (mass per unit distance: Note that the units of mass density are not the same as the units of mass.)

As we have seen so far in this chapter, the same dichotomy exists in the topic of probability. In Chapter 7, we were concerned with the probability of discrete events whose outcome belongs to some finite set of possibilities (e.g. Head or Tail for a coin toss, allele A or a in genetics).

The example below provides some further insight to the connection between continuous and discrete probability. In particular, we will see that one can arrive at the idea of probability density by refining a set of measurements and making the appropriate scaling. We explore this connection in more detail below.

8.4.3 Example: Student heights

Suppose we measure the heights of all UBC students. This would produce about 30,000 data values[38]. We could make a graph and show how these heights are distributed. For example, we could subdivide the student body into those students between 0 and 1.5m, and those between 1.5 and 3 meters. Our bar graph would contain two bars, with the number of students in each height category represented by the heights of the bars, as shown in Figure 8.5(a).

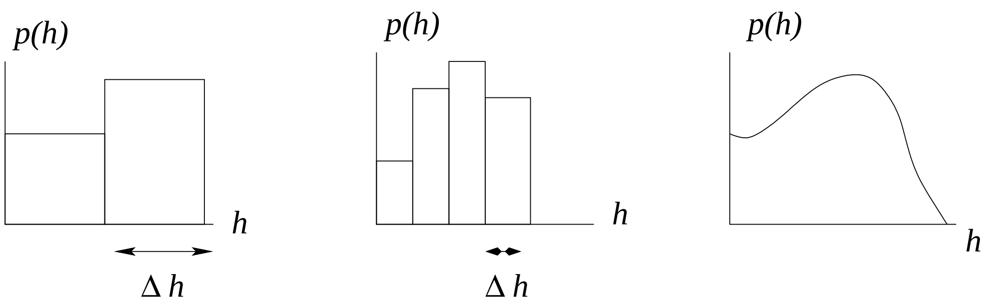

Figure 8.5. *Refining a histogram by increasing the number of bins leads (eventually) to the idea of a continuous probability density.*

Suppose we want to record this distribution in more detail. We could divide the population into smaller groups by shrinking the size of the interval or "bin" into which height is subdivided. (An example is shown in Figure 8.5(b)). Here, by a "bin" we mean a little interval of width Δh where h is height, i.e. a height interval. For example, we could keep track of the heights in increments of 50 cm. If we were to plot the *number* of students in each height category, then as the size of the bins gets smaller, so would the height of the bar: there would be fewer students in each category if we increase the number of categories.

To keep the bar height from shrinking, we might reorganize the data slightly. Instead of plotting the *number* of students in each bin, we might plot

$$\frac{\text{number of students in the bin}}{\Delta h}.$$

If we do this, then both numerator and denominator decrease as the size of the bins is made smaller, so that the shape of the distribution is preserved (i.e. it does not get flatter).

We observe that in this case, the number of students in a given height category is represented by the *area of the bar* corresponding to that category:

$$\text{Area of bin} = \Delta h \left(\frac{\text{number of students in the bin}}{\Delta h}\right) = \text{number of students in the bin.}$$

The important point to consider is that the height of each bar in the plot represents the *number of students per unit height*.

[38] I am grateful to David Austin for developing this example.

This type of plot is precisely what leads us to the idea of a density distribution. As Δh shrinks, we get a continuous graph. If we "normalize", i.e. divide by the total area under the graph, we get a probability density, $p(h)$ for the height of the population. As noted, $p(h)$ represents the fraction of students per unit height[39] whose height is h. It is thus a ***density***, and has the appropriate units. In this case, $p(h)\ \Delta h$ represents the fraction of individuals whose height is in the range $h \leq$ height $\leq h + \Delta h$.

8.4.4 Example: Age dependent mortality

In this example, we consider an age distribution and interpret the meanings of the probability density and of the cumulative function. Understanding the connection between the verbal description and the symbols we use to represent these concepts requires practice and experience. Related problems are presented in the homework.

Let $p(a)$ be a probability density for the probability of mortality of a female Canadian non-smoker at age a, where $0 \leq a \leq 120$. (We have chosen an upper endpoint of age 120 since practically no Canadian female lives past this age at present.) Let $F(a)$ be the cumulative distribution corresponding to this probability density. We would like to answer the following questions:

(a) What is the probability of dying by age a?

(b) What is the probability of surviving to age a?

(c) Suppose that we are told that $F(75) = 0.8$ and that $F(80)$ differs from $F(75)$ by 0.11. Interpret this information in plain English. What is the probability of surviving to age 80? Which is larger, $F(75)$ or $F(80)$?

(d) Use the information in part (c) to estimate the probability of dying between the ages of 75 and 80 years old. Further, estimate $p(80)$ from this information.

Solution

(a) The probability of dying **by** age a is the same as the probability of dying **any time up to** age a. Restated, this is the probability that the age of death is in the interval $0 \leq$ age of death $\leq a$. The appropriate quantity is the cumulative function, for this probability density

$$F(a) = \int_0^a p(x)\ dx.$$

Remark: note that, as customary, x is playing the role of a "dummy variable". We are integrating over all ages between 0 and a, so we do not want to confuse the notation for variable of integration, x and endpoint of the interval a. Hence the symbol x rather than a inside the integral.

[39] Note in particular the units of h^{-1} attached to this probability density, and contrast this with a discrete probability that is a pure number carrying no such units.

(b) The probability of surviving to age a is the same as the probability of **not** dying before age a. By the elementary properties of probability discussed in the previous chapter, this is

$$1 - F(a).$$

(c) $F(75) = 0.8$ means that the probability of dying some time up to age 75 is 0.8. (This also means that the probability of surviving past this age would be 1-0.8=0.2.) From the properties of probability, we know that the cumulative distribution is an *increasing* function, and thus it must be true that $F(80) > F(75)$. Then $F(80) = F(75) + 0.11 = 0.8 + 0.11 = 0.91$. Thus the probability of surviving to age 80 is 1-0.91=0.09. This means that 9% of the population will make it to their 80'th birthday.

(d) The probability of dying between the ages of 75 and 80 years old is exactly

$$\int_{75}^{80} p(x)\, dx.$$

However, we can also state this in terms of the cumulative function, since

$$\int_{75}^{80} p(x)\, dx = \int_{0}^{80} p(x)\, dx - \int_{0}^{75} p(x)\, dx = F(80) - F(75) = 0.11$$

Thus the probability of death between the ages of 75 and 80 is 0.11.

To estimate $p(80)$, we use the connection between the probability density and the cumulative distribution[40]:

$$p(x) = F'(x). \tag{8.4}$$

Then it is approximately true that

$$p(x) \approx \frac{F(x + \Delta x) - F(x)}{\Delta x}. \tag{8.5}$$

(Recall the definition of the derivative, and note that we are approximating the derivative by the slope of a secant line.) Here we have information at ages 75 and 80, so $\Delta x = 80 - 75 = 5$, and the approximation is rather crude, leading to

$$p(80) \approx \frac{F(80) - F(75)}{5} = \frac{0.11}{5} = 0.022 \text{ per year.}$$

Several important points merit attention in the above example. First, information contained in the cumulative function is useful. Differences in values of F between $x = a$ and $x = b$ are, after all, equivalent to an integral of the function $\int_a^b p(x)dx$, and are the probability of a result in the given interval, $a \le x \le b$. Second, $p(x)$ is the derivative of $F(x)$. In the expression (8.5), we approximated that derivative by a small finite difference. Here we see at play many of the themes that have appeared in studying calculus: the connection between derivatives and integrals, the Fundamental Theorem of Calculus, and the relationship between tangent and secant lines.

[40] In Eqn. (8.4) there is no longer confusion between a variable of integration and an endpoint, so we could revert to the notation $p(a) = F'(a)$, helping us to identify the independent variable as age. However, we have avoided doing so simply so that the formula in Eqn. (8.5) would be very recognizable as an approximation for a derivative.

8.4.5 Example: Raindrop size distribution

In this example, we find a rather non-intuitive result, linking the distribution of raindrops of various radii with the distribution of their volumes. This reinforces the caution needed in interpreting and handling probabilities.

During a Vancouver rainstorm, the distribution of raindrop radii is **uniform** for radii $0 \le r \le 4$ (where r is measured in mm) and zero for larger r. By a **uniform distribution** we mean a function that has a constant value in the given interval. Thus, we are saying that the distribution looks like $f(r) = C$ for $0 \le r \le 4$.

(a) Determine what is the probability density for raindrop radii, $p(r)$? Interpret the meaning of that function.

(b) What is the associated cumulative function $F(r)$ for this probability density? Interpret the meaning of that function.

(c) In terms of the volume, what is the cumulative distribution $F(V)$?

(d) In terms of the volume, what is the probability density $p(V)$?

(e) What is the average volume of a raindrop?

Solution

This problem is challenging because one may be tempted to think that the uniform distribution of drop radii should give a uniform distribution of drop volumes. This is not the case, as the following argument shows! The sequence of steps is illustrated in Figure 8.6.

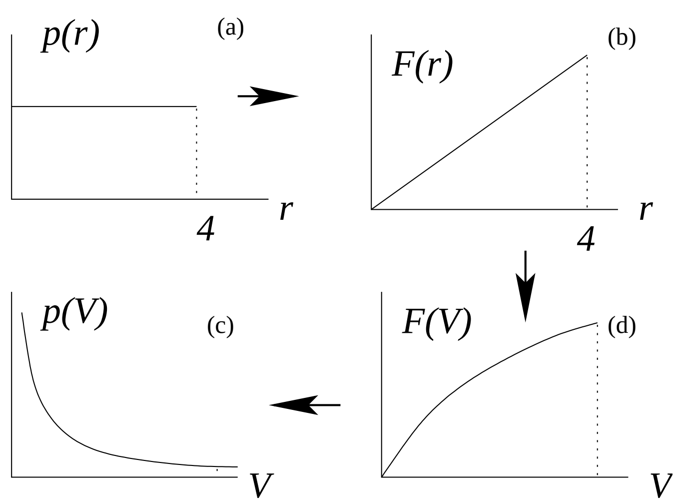

Figure 8.6. *Probability densities for raindrop radius and raindrop volume (left panels) and for the cumulative distributions (right) of each for Example 8.4.5.*

(a) The probability density function is $p(r) = 1/4$ for $0 \le r \le 4$. This means that the probability *per unit radius* of finding a drop of size r is the same for all radii in $0 \le r \le 4$, as shown in Fig. 8.6(a). Some of these drops will correspond to small volumes, and others to very large volumes. We will see that the probability *per unit volume* of finding a drop of given volume will be quite different.

(b) The cumulative function is

$$F(r) = \int_0^r \frac{1}{4} ds = \frac{r}{4}, \quad 0 \le r \le 4. \tag{8.6}$$

A sketch of this function is shown in Fig. 8.6(b).

(c) The cumulative function $F(r)$ is proportional to the radius of the drop. We use the connection between radii and volume of spheres to rewrite that function in terms of the volume of the drop: Since

$$V = \frac{4}{3}\pi r^3 \tag{8.7}$$

we have

$$r = \left(\frac{3}{4\pi}\right)^{1/3} V^{1/3}.$$

Substituting this expression into the formula (8.6), we get

$$F(V) = \frac{1}{4}\left(\frac{3}{4\pi}\right)^{1/3} V^{1/3}.$$

We find the range of values of V by substituting $r = 0,\ 4$ into Eqn. (8.7) to get $V = 0,\ \frac{4}{3}\pi 4^3$. Therefore the interval is $0 \le V \le \frac{4}{3}\pi 4^3$ or $0 \le V \le (256/3)\pi$. The function $F(V)$ is sketched in panel (d) of Fig. 8.6.

(d) We now use the connection between the probability density and the cumulative distribution, namely that p is the derivative of F. Now that the variable has been converted to volume, that derivative is a little more "interesting":

$$p(V) = F'(V)$$

Therefore,

$$p(V) = \frac{1}{4}\left(\frac{3}{4\pi}\right)^{1/3} \frac{1}{3} V^{-2/3}.$$

Thus the probability *per unit volume* of finding a drop of volume V in $0 \le V \le \frac{4}{3}\pi 4^3$ is not at all uniform. This probability density is shown in Fig. 8.6(c) This results from the fact that the differential quantity dr behaves very differently from dV, and reinforces the fact that we are dealing with density, not with a probability per se. We note that this distribution has smaller values at larger values of V.

(e) The range of values of V is

$$0 \le V \le \frac{256\pi}{3},$$

and therefore the mean volume is

$$\bar{V} = \int_0^{256\pi/3} V p(V) dV = \frac{1}{12}\left(\frac{3}{4\pi}\right)^{1/3} \int_0^{256\pi/3} V \cdot V^{-2/3} dV$$

$$= \frac{1}{12}\left(\frac{3}{4\pi}\right)^{1/3} \int_0^{256\pi/3} V^{1/3} dV = \frac{1}{12}\left(\frac{3}{4\pi}\right)^{1/3} \frac{3}{4} V^{4/3}\Big|_0^{256\pi/3}$$

$$= \frac{1}{16}\left(\frac{3}{4\pi}\right)^{1/3}\left(\frac{256\pi}{3}\right)^{4/3} = \frac{64\pi}{3} \approx 67\text{mm}^3.$$

8.5 Moments of a probability density

We are now familiar with some of the properties of probability distributions. On this page we will introduce a set of numbers that describe various properties of such distributions. Some of these have already been encountered in our previous discussion, but now we will see that these fit into a pattern of quantities called **moments** of the distribution.

8.5.1 Definition of moments

Let $f(x)$ be any function which is defined and positive on an interval $[a, b]$. We might refer to the function as a distribution, whether or not we consider it to be a probability density. Then we will define the following **moments** of this function:

$$\text{zero'th moment } M_0 = \int_a^b f(x)\, dx$$

$$\text{first moment } M_1 = \int_a^b x\, f(x)\, dx$$

$$\text{second moment } M_2 = \int_a^b x^2\, f(x)\, dx$$

$$\vdots$$

$$n\text{'th moment } M_n = \int_a^b x^n\, f(x)\, dx.$$

Observe that moments of any order are defined by integrating the distribution $f(x)$ with a suitable power of x over the interval $[a, b]$. However, in practice we will see that usually moments up to the second are usefully employed to describe common attributes of a distribution.

8.5.2 Relationship of moments to mean and variance of a probability density

In the particular case that the distribution is a probability density, $p(x)$, defined on the interval $a \le x \le b$, we have already established the following :

$$M_0 = \int_a^b p(x)\, dx = 1.$$

(This follows from the basic property of a probability density.) Thus *The zero'th moment of any probability density is 1*. Further

$$M_1 = \int_a^b x\, p(x)\, dx = \bar{x} = \mu.$$

That is, *The first moment of a probability density is the same as the mean (i.e. expected value) of that probability density*. So far, we have used the symbol $\bar{x}$ to represent the mean or average value of x but often the symbol μ is also used to denote the mean.

The second moment, of a probability density also has a useful interpretation. From above definitions, the second moment of $p(x)$ over the interval $a \le x \le b$ is

$$M_2 = \int_a^b x^2\, p(x)\, dx.$$

We will shortly see that the second moment helps describe the way that density is distributed about the mean. For this purpose, we must describe the notion of *variance* or *standard deviation*.

Variance and standard deviation

Two children of approximately the same size can balance on a teeter-totter by sitting very close to the point at which the beam pivots. They can also achieve a balance by sitting at the very ends of the beam, equally far away. In both cases, the center of mass of the distribution is at the same place: precisely at the pivot point. However, the mass is distributed very differently in these two cases. In the first case, the mass is clustered close to the center, whereas in the second, it is distributed further away. We may want to be able to describe this distinction, and we could do so by considering higher moments of the mass distribution.

Similarly, if we want to describe how a probability density distribution is distributed about its mean, we consider moments higher than the first. We use the idea of the *variance* to describe whether the distribution is clustered close to its mean, or spread out over a great distance from the mean.

Variance

The *variance* is defined as the average value of the quantity $(distance\ from\ mean)^2$, where the average is taken over the whole distribution. (The reason for the square is that we would not like values to the left and right of the mean to cancel out.) For **discrete probability** with mean, μ we define variance by

$$\boxed{V = \sum (x_i - \mu)^2 p_i.}$$

For a **continuous probability** density, with mean μ, we define the variance by

$$\boxed{V = \int_a^b (x - \mu)^2\, p(x)\, dx.}$$

The standard deviation

The standard deviation is defined as

$$\boxed{\sigma = \sqrt{V}.}$$

Let us see what this implies about the connection between the variance and the moments of the distribution.

Relationship of variance to second moment

From the equation for variance we calculate that

$$V = \int_a^b (x - \mu)^2\, p(x)\, dx \;=\; \int_a^b (x^2 - 2\mu x + \mu^2)\, p(x)\, dx.$$

Expanding the integral leads to:

$$\begin{aligned} V &= \int_a^b x^2\, p(x)dx - \int_a^b 2\mu x\, p(x)\, dx + \int_a^b \mu^2\, p(x)\, dx \\ &= \int_a^b x^2\, p(x)dx - 2\mu \int_a^b x\, p(x)\, dx + \mu^2 \int_a^b p(x)\, dx. \end{aligned}$$

We recognize the integrals in the above expression, since they are simply moments of the probability distribution. Using the definitions, we arrive at

$$V = \; M_2 - 2\mu\, \mu + \mu^2.$$

Thus

$$\boxed{V = M_2 - \mu^2.}$$

Observe that the variance is related to the second moment, M_2 and to the mean, μ of the distribution.

Relationship of variance to second moment

Using the above definitions, the standard deviation, σ can be expressed as

$$\boxed{\sigma \;=\; \sqrt{V} \;=\; \sqrt{M_2 - \mu^2}.}$$

8.5.3 Example: computing moments

Consider a probability density such that $p(x) = C$ is constant for values of x in the interval $[a, b]$ and zero for values outside this interval[41]. The area under the graph of this function for $a \le x \le b$ is $A = C \cdot (b - a) \equiv 1$ (enforced by the usual property of a probability density), so it is easy to see that the value of the constant C should be $C = 1/(b-a)$. Thus

$$p(x) = \frac{1}{b-a}, \quad a \le x \le b.$$

We compute some of the moments of this probability density

$$M_0 = \int_a^b p(x)dx = \frac{1}{b-a}\int_a^b 1\,dx = 1.$$

(This was already known, since we have determined that the zeroth moment of any probability density is 1.) We also find that

$$M_1 = \int_a^b x\,p(x)\,dx = \frac{1}{b-a}\int_a^b x\,dx = \frac{1}{b-a}\left.\frac{x^2}{2}\right|_a^b = \frac{b^2-a^2}{2(b-a)}.$$

This last expression can be simplified by factoring, leading to

$$\mu = M_1 = \frac{(b-a)(b+a)}{2(b-a)} = \frac{b+a}{2}.$$

The value $(b+a)/2$ is a midpoint of the interval $[a, b]$. Thus we have found that the mean μ is in the center of the interval, as expected for a symmetric distribution. The median would be at the same place by a simple symmetry argument: half the area is to the left and half the area is to the right of this point.

To find the variance we calculate the second moment,

$$M_2 = \int_a^b x^2\,p(x)\,dx = \frac{1}{b-a}\int_a^b x^2\,dx = \left(\frac{1}{b-a}\right)\left.\frac{x^3}{3}\right|_a^b = \frac{b^3-a^3}{3(b-a)}.$$

Factoring simplifies this to

$$M_2 = \frac{(b-a)(b^2+ab+a^2)}{3(b-a)} = \frac{b^2+ab+a^2}{3}.$$

[41] As noted before, this is a uniform distribution. It has the shape of a rectangular band of height C and base $(b-a)$.

The variance is then

$$V = M_2 - \mu^2 = \frac{b^2+ab+a^2}{3} - \frac{(b+a)^2}{4} = \frac{b^2-2ab+a^2}{12} = \frac{(b-a)^2}{12}.$$

The standard deviation is

$$\sigma = \frac{(b-a)}{2\sqrt{3}}.$$

8.6 Summary

In this chapter, we extended the discrete probability encountered in Chapter 7 to the case of continuous probability density. We learned that this function is a probability per unit value (of the variable of interest), so that

$$\int_a^b p(x)dx = \text{ probability that x takes a value in the interval } (a,b).$$

We also defined and studied the cumulative function

$$F(x) = \int_a^x p(s)ds = \text{ probability of a value in the interval } (a,x).$$

We noted that by the Fundamental Theorem of Calculus, $F(x)$ is an antiderivative of $p(x)$ (or synonymously, $p'(x) = F(x)$.)

The mean and median are two descriptors for some features of probability densities. For $p(x)$ defined on an interval $a \le x \le b$ and zero outside, the mean, ($\bar{x}$, or sometimes called μ) is

$$\bar{x} = \int_a^b xp(x)dx$$

whereas the median, x_{med} is the value for which

$$F(x_{med}) = \frac{1}{2}.$$

Both mean and median correspond to the "center" of a symmetric distribution. If the distribution is non-symmetric, a long tail in one direction will shift the mean toward that direction more strongly than the median. The variance of a probability density is

$$V = \int_a^b (x-\mu)^2 \, p(x) \, dx,$$

and the standard deviation is

$$\sigma = \sqrt{V}.$$

This quantity describes the "width" of the distribution, i.e. how spread out (large σ) or clumped (small σ) it is.

We defined the n'th moment of a probability density as

$$M_n = \int_a^b x^n p(x)dx,$$

and showed that the first few moments are related to mean and variance of the probability. Most of these concepts are directly linked to the analogous ideas in discrete probability, but in this chapter, we used integration in place of summation, to deal with the continuous, rather than the discrete case.

Chapter 9

Differential Equations

9.1 Introduction

A differential equation is a relationship between some (unknown) function and one of its derivatives. Examples of differential equations were encountered in an earlier calculus course in the context of population growth, temperature of a cooling object, and speed of a moving object subjected to friction. In Section 4.2.4, we reviewed an example of a differential equation for velocity, (4.8), and discussed its solution, but here, we present a more systematic approach to solving such equations using a technique called **separation of variables**. In this chapter, we apply the tools of integration to finding solutions to differential equations. The importance and wide applicability of this topic cannot be overstated.

In this course, since we are concerned only with functions that depend on a single variable, we discuss **ordinary differential equations** (ODE's), whereas later, after a multivariate calculus course where partial derivatives are introduced, a wider class, of **partial differential equations** (PDE's) can be studied. Such equations are encountered in many areas of science, and in any quantitative analysis of systems where rates of change are linked to the state of the system. Most laws of physics are of this form; for example, applying the familiar Newton's law, $F = ma$, links the position of a pendulum's mass to its acceleration (second derivative of position).[42] Many biological processes are also described by differential equations. The rate of growth of a population dN/dt depends on the size of that population at the given time $N(t)$.

Constructing the differential equation that adequately represents a system of interest is an art that takes some thought and experience. In this process, which we call "modeling", many simplifications are made so that the essential properties of a given system are captured, leaving out many complicating details. For example, friction might be neglected in "modeling" a perfect pendulum. The details of age distribution might be neglected in modeling a growing population. Now that we have techniques for integration, we can devise a new approach to computing solutions of differential equations.

Given a differential equation and a starting value, the goal is to make a prediction

[42] Newton's law states that force is proportional to acceleration. For a pendulum, the force is due to gravity, and the acceleration is a second derivative of the x or y coordinate of the bob on the pendulum.

about the future behaviour of the system. This is equivalent to identifying the function that satisfies the given differential equation and initial value(s). We refer to such a function as the **solution** to the **initial value problem** (IVP). In differential calculus, our exploration of differential equations was limited to those whose solution could be guessed, or whose solution was supplied in advance. We also explored some of the fascinating geometric and qualitative properties of such equations and their predictions.

Now that we have techniques of integration, we can find the analytic solution to a variety of simple first-order differential equations (i.e. those involving the first derivative of the unknown function). We will describe the technique of **separation of variables**. This technique works for examples that are simple enough that we can isolate the dependent variable (e.g. y) on one side of the equation, and the independent variable (e.g. time t) on the other side.

9.2 Unlimited population growth

We start with a simple example that was treated thoroughly in the differential calculus semester of this course. We consider a population with per capita birth and mortality rates that are constant, irrespective of age, disease, environmental changes, or other effects. We ask how a population in such ideal circumstances would change over time. We build up a simple model (i.e. a differential equation) to describe this ideal case, and then proceed to find its solution. Solving the differential equation is accomplished by a new technique introduced here, namely separation of variables. This reduces the problem to integration and algebraic manipulation, allowing us to compute the population size at any time t. By going through this process, we essentially convert information about the rate of change and starting level of the population to a detailed prediction of the population at later times.[43]

9.2.1 A simple model for population growth

Let $y(t)$ represent the size of a population at time t. We will assume that at time $t = 0$, the population level is specified, i.e. $y(0) = y_0$ is some given constant. We want to find the population at later times, given information about birth and mortality rates, (both of which are here assumed to be constant over time).

The population changes through births and mortality. Suppose that $b > 0$ is the per capita average birth rate, and $m > 0$ the per capita average mortality rate. The assumption that $b,\ m$ are both constants is a simplification that neglects many biological effects, but will be used for simplicity in this first example.

The statement that the population increases through births and decreases due to mortality, can be restated as

$$\text{rate of change of } \ y = \text{rate of births} - \text{rate of mortality}$$

where the rate of births is given by the product of the per capita average birth rate b and the population size y. Similarly, the rate of mortality is given by my. Translating the rate of

[43]Of course, we must keep in mind that such predictions are based on simplifying assumptions, and are to be taken as an approximation of any real population growth.

change into the corresponding derivative of y leads to

$$\frac{dy}{dt} = by - my = (b - m)y.$$

Let us define the new constant,

$$k = b - m.$$

Then k is the *net per capita growth rate* of the population. We can distinguish two possible cases: $b > m$ means that there are more births then deaths, so we expect the population to grow. $b < m$ means that there are more deaths than births, so that the population will eventually go extinct. There is also a marginal case that $b = m$, for which $k = 0$, where the population does not change at all. To summarize, this simple model of unlimited growth leads to the differential equation and initial condition:

$$\frac{dy}{dt} = ky, \quad y(0) = y_0. \tag{9.1}$$

Recall that a differential equation together with an initial condition is called an initial value problem. To find a solution to such a problem, we look for the function $y(t)$ that describes the population size at any future time t, given its initial size at time $t = 0$.

9.2.2 Separation of variables and integration

We here introduce the technique, **separation of variables**, that will be used in all the examples described in this chapter. Since the differential equation (9.1) is relatively simple, this first example will be relatively straightforward. We would like to determine $y(t)$ given the differential equation

$$\frac{dy}{dt} = ky.$$

Rather than integrating this equation as is[44], we use an alternate approach, considering dt and dy as "differentials" in the sense defined in Section 6.1. We rearrange and rewrite the above equation in the form

$$\frac{1}{y}\,dy = k\,dt, \tag{9.2}$$

This step of putting expressions involving the independent variable t on one side and expressions involving the dependent variable y on the opposite side gives rise to the name "separation of variables".

Now, the LHS of Eqn. (9.2) depends only on the variable y, and the RHS only on t. The constant k will not interfere with any integration step. Moreover, integrating each side of Eqn. (9.2) can be carried out independently.

To determine the appropriate intervals for integration, we observe that when time sweeps over some interval $0 \le t \le T$ (from initial to final time), the value of $y(t)$ will

[44]We may be tempted to integrate both sides of this equation with respect to the independent variable t, e.g. writing $\int \frac{dy}{dt}\,dt = \int ky\,dt + C$, (where C is some constant), but this is not very useful, since the integral on the right hand side (RHS) can only be carried out if we know the function $y = y(t)$, which we are trying to determine.

change over a corresponding interval $y_0 \leq y \leq y(T)$. Here y_0 is the given starting value of y (prescribed by the initial condition in (9.1)). We do not yet know $y(T)$, but our goal is to find that value, i.e to predict the future behaviour of y. Integrating leads to

$$\int_{y_0}^{y(T)} \frac{1}{y}\, dy = \int_0^T k\, dt = k \int_0^T dt,$$

$$\ln|y|\Big|_{y_0}^{y(T)} = kt\Big|_0^T,$$

$$\ln|y(T)| - \ln|y(0)| = k(T-0),$$

$$\ln\left|\frac{y(T)}{y_0}\right| = kT,$$

$$\frac{y(T)}{y_0} = e^{kT},$$

$$y(T) = y_0 e^{kT}.$$

But this result holds for any arbitrary final time, T. In other words, since this is true for any time we chose, we can set $T = t$, arriving at the desired solution

$$y(t) = y_0 e^{kt}. \tag{9.3}$$

The above formula relates the predicted value of y at any time t to its initial value, and to all the parameters of the problem. Observe that plugging in $t = 0$, we get $y(0) = y_0 e^{kt} = y_0 e^0 = y_0$, so that the solution (9.3) satisfies the initial condition. We leave as an exercise for the reader[45] to validate that the function in(9.3) also satisfies the differential equation in (9.1).

By solving the initial value problem (9.1), we have determined that, under ideal conditions, when the net per capita growth rate t is constant, a population will grow exponentially with time. Recall that this validates results that we had encountered in our first calculus course.

9.3 Terminal velocity and steady states

Here we revisit the equation for velocity of a falling object that we first encountered in Section 4.2.4. We wish to derive the appropriate differential equation governing that velocity, and find the solution $v(t)$ as a function of time. We will first reconsider the simplest case of uniformly accelerated motion (i.e. where friction is neglected), as in Section 4.2.3. We then include friction, as in Section 4.2.4 and use the new technique of separation of variables to shortcut the method of solution.

[45]This kind of check is good practice and helps to spot errors. Simply differentiate Eqn. (9.3) and show that the result is the same as k times the original function, as required by the equation (9.1).

9.3.1 Ignoring friction: the uniformly accelerated case

Let $v(t)$ and $a(t)$ be the velocity and the acceleration, respectively of an object falling under the force of gravity at time t. We take the positive direction to be downwards, for convenience. Suppose that at time $t = 0$, the object starts from rest, i.e. the initial velocity of the object is known to be $v(0) = 0$. When friction is neglected, the object will accelerate,

$$a(t) = g,$$

which is equivalent to the statement that the velocity increases at a constant rate,

$$\frac{dv}{dt} = g. \tag{9.4}$$

Because g is constant, *we do not need to use separation of variables*, i.e. we can integrate each side of this equation directly[46]. Writing

$$\int \frac{dv}{dt}\, dt = \int g\, dt + C = g \int\, dt + C,$$

where C is an integration constant, we arrive at

$$v(t) = gt + C. \tag{9.5}$$

Here we have used (on the LHS) that v is the antiderivative of dv/dt. (equivalently, we can simplify the integral $\int \frac{dv}{dt}\, dt = \int\, dv = v$). Plugging in $v(0) = 0$ into Eqn. (9.5) leads to $0 = g \cdot 0 + C = C$, so the constant we need is $C = 0$ and the velocity satisfies

$$v(t) = gt.$$

We have just arrived at a result that parallels Eqn. (4.4) of Section 4.2.3 (in slightly different notation).

9.3.2 Including friction: the case of terminal velocity

When a falling object experiences the force of friction, it cannot accelerate indefinitely. In fact, a frictional force retards the downwards motion. To a good approximation, that force is proportional to the velocity.

A force balance for the falling object leads to

$$ma(t) = mg - \gamma v(t),$$

where γ is the frictional coefficient. For an object of constant mass, we can divide through by m, so

$$a(t) = g - \frac{\gamma}{m} v(t).$$

[46] It is important to note the distinction between this simple example and other cases where separation of variables is required. It would not be *wrong* to use separation of variables to find the solution for Eqn. (9.4), but it would just be “overkill”, since simple integration of the each side of the equation “as is” does the job.

Let $k = \gamma/m$. Then, the velocity at any time satisfies the differential equation and initial condition

$$\frac{dv}{dt} = g - kv, \quad v(0) = 0. \tag{9.6}$$

We can find the solution to this differential equation and predict the velocity at any time t using separation of variables.

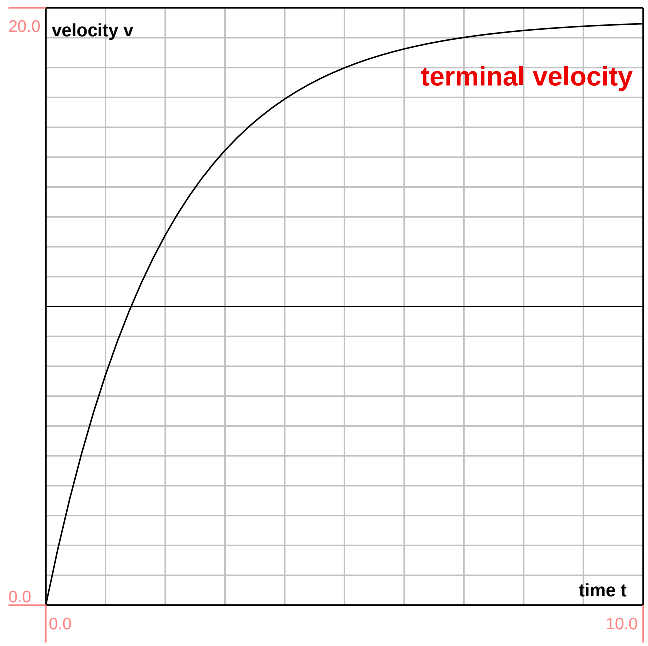

Figure 9.1. *The velocity $v(t)$ as a function of time given by Eqn. (9.7) as found in Section 9.3.2. Note that as time increases, the velocity approaches some constant terminal velocity. The parameters used were $g = 9.8$ m/s^2 and $k = 0.5$.*

Consider a time interval $0 \leq t \leq T$, and suppose that, during this time interval, the velocity changes from an initial value of $v(0) = 0$ to the final value, $v(T)$ at the final time, T. Then using separation of variables and integration, we get

$$\frac{dv}{dt} = g - kv,$$

$$\frac{dv}{g - kv} = dt,$$

$$\int_0^{v(T)} \frac{dv}{g - kv} = \int_0^T dt.$$

Substitute $u = g - kv$ for the integral on the left hand side. Then $du = -kdv$, $dv = (-1/k)du$, so we get an integral of the form

$$-\frac{1}{k}\int \frac{1}{u}\,du = -\frac{1}{k}\ln|u|.$$

After replacing u by $g - kv$, we arrive at

$$-\frac{1}{k}\ln|g - kv|\Big|_0^{v(T)} = t\Big|_0^{T}.$$

We use the fact that $v(0) = 0$ to write this as

$$-\frac{1}{k}\left(\ln|g - kv(T)| - \ln|g|\right) = T,$$

$$-\frac{1}{k}\left(\ln\left|\frac{g - kv(T)}{g}\right|\right) = T,$$

$$\ln\left|\frac{g - kv(T)}{g}\right| = -kT.$$

We are finished with the integration step, but the function we are trying to find, $v(T)$ is still tangled up inside an expression involving the natural logarithm. Extricating it will involve some subtle reasoning about signs because there is an absolute value to contend with. As a first step, we exponentiate both sides to remove the logarithm.

$$\left|\frac{g - kv(T)}{g}\right| = e^{-kT} \quad \Rightarrow \quad |g - kv(T)| = ge^{-kT}.$$

Because the constant g is positive, we could remove absolute values signs from it. To simplify further, we have to consider the sign of the term inside the absolute value in the numerator. In the case we are considering here, $v(0) = 0$. This will mean that the quantity $g - kv(T)$ is always be non-negative (i.e. $g - kv(T) \geq 0$). We will verify this fact shortly. For the moment, supposing this is true, we can write

$$|g - kv(T)| = g - kv(T) = ge^{-kT},$$

and finally solve for $v(T)$ to obtain our final result,

$$v(T) = \frac{g}{k}(1 - e^{-kT}).$$

Here we note that $v(T)$ can never be larger than g/k since the term $(1 - e^{-kT})$ is always ≤ 1. Hence, we were correct in assuming that $g - kv(T) \geq 0$.

As before, the above formula relating velocity to time holds for any choice of the final time T, so we can write, in general,

$$v(t) = \frac{g}{k}(1 - e^{-kt}). \tag{9.7}$$

This is the solution to the initial value problem (9.6). It predicts the velocity of the falling object through time. Note that we have arrived once more at the result obtained in Eqn. (4.11), but using the technique of separation of variables[47].

We graph the expression given in (9.7) in Figure 9.1. Note that as t increases, the term e^{-kt} decreases rapidly, so that the velocity approaches a constant whose value is

$$v(t) \to \frac{g}{k}.$$

We call this the *terminal velocity*[48].

9.3.3 Steady state

We might observe that the terminal velocity can also be found quite simply and directly from the *differential equation itself*: it is the **steady state** of the differential equation, i.e. the value for which no further change takes place. The steady state can be found by setting the derivative in the differential equation, to zero, i.e. by letting

$$\frac{dv}{dt} = 0.$$

When this is done, we arrive at

$$g - kv = 0 \quad \Rightarrow \quad v = \frac{g}{k}.$$

Thus, at steady state, the velocity of the falling object is indeed the same as the terminal velocity that we have just discovered.

9.4 Related problems and examples

The example discussed in Section 9.3.2 belongs to a class of problems that share many common features. Generally, this class is represented by linear differential equations of the form

$$\frac{dy}{dt} = a - by, \tag{9.8}$$

with given initial condition $y(0) = y_0$. Properties of this equation were studied in the context of differential calculus in a previous semester. Now, with the same method as we applied to the problem of terminal velocity, we can integrate this equation by separation of variables, writing

$$\frac{dy}{a - by} = dt$$

and proceeding as in the previous example. We arrive at its solution,

$$y(t) = \frac{a}{b} + \left(y_0 - \frac{a}{b}\right) e^{-bt}. \tag{9.9}$$

[47] It often happens that a differential equation can be solved using several different methods.

[48] A similar plot of the solution of the differential equation (9.6) could be assembled using Euler's method, as studied in differential calculus. That is the numerical method alternative to the analytic technique discussed in this chapter. The student may wish to review results obtained in a previous semester to appreciate the correspondence.

The steps are left as an exercise for the reader.

We observe that the steady state of the above equation is obtained by setting

$$\frac{dy}{dt} = a - by = 0, \quad \text{i.e.} \quad y = \frac{a}{b}.$$

Indeed the solution given in the formula (9.9) has the property that as t increases, the exponential term $e^{-bt} \to 0$ so that the term in large brackets will vanish and $y \to a/b$. This means that from any initial value, y will approach its steady state level.

This equation has a number of important applications that arise in a variety of context. A few of these are mentioned below.

9.4.1 Blood alcohol

Let $y(t)$ be the level of alcohol in the blood of an individual during a party. Suppose that the average rate of drinking is gradual and constant (i.e. small sips are continually taken, so that the rate of input of alcohol is approximately constant). Further, assume that alcohol is detoxified in the liver at a rate proportional to its blood level. Then an equation of the form (9.8) would describe the blood level over the period of drinking. $y(0) = 0$ would signify the absence of alcohol in the body at the beginning of the evening. The constant a would reflect the rate of intake per unit volume of the individual's blood: larger people take longer to "get drunk" for a given amount consumed[49]. The constant b represents the rate of decay of alcohol per unit time due to degradation by the liver, assumed constant[50]; young healthy drinkers have a higher value of b than those who can no longer metabolize alcohol as efficiently.

The solution (9.9) has several features of note: it illustrates the fact that alcohol would increase from the initial level, but only up to a maximum of a/b, where the intake and degradation balance. Indeed, the level $y = a/b$ represents a steady state level (as long as drinking continues). Of course, this level could be toxic to the drinker, and the assumptions of the model may break down in that region! In the phase of "recovery", after drinking stops, the above differential equation no longer describes the level of blood alcohol. Instead, the process of recovery is represented by

$$\frac{dy}{dt} = -by, \quad y(0) = y_0. \tag{9.10}$$

The level of blood alcohol then decays exponentially with rate b from its level at the moment that drinking ends. We show this typical pattern in Figure 9.2.

9.4.2 Chemical kinetics

The same ideas apply to any chemical substance that is formed at a constant rate (or supplied at a constant rate) a, and then breaks down with rate proportional to its concentration. We then call the constant b the "decay rate constant".

[49]Of course, we are here assuming a constant intake rate, as though the alcohol is being continually sipped all evening at a uniform rate. Most people do not drink this way, instead quaffing a few large drinks over some hour(s). It is possible to describe this, but we will not do so in this chapter.

[50]This is also a simplifying assumption, as the rate of metabolism can depend on other factors, such as food intake.

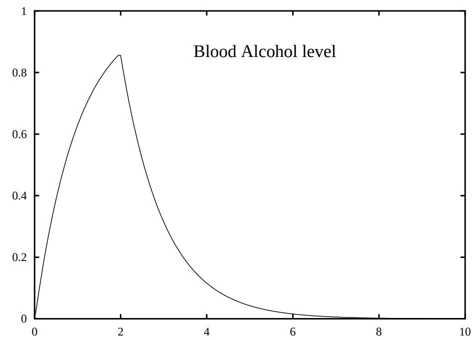

Figure 9.2. *The level of alcohol in the blood is described by Eqn. (9.8) for the first two hours of drinking. At $t = 2h$, the drinking stopped (so $a = 0$ from then on). The level of alcohol in the blood then decays back to zero, following Eqn. (9.10).*

The variable $y(t)$ represents the concentration of chemical at time t, and the same differential equation describes this chemical process. As above, given any initial level of the substance, $y(t) = y_0$, the level of y will eventually approach the steady state, $y = a/b$.

9.5 Emptying a container

In this section we investigate a new problem in which the differential equation that describes a process will be derived from basic physical principles[51]. We will look at the flow of fluid leaking out of a container, and use mass balance to derive a differential equation model. When this is done, we will also use separation of variables to predict how long it takes for the container to be emptied.

We will assume that the container has a small hole at its base. The rate of emptying of the container will depend on the height of fluid in the container above the hole[52]. We can derive a simple differential equation that describes the rate that the height of the fluid changes using the following physical argument.

9.5.1 Conservation of mass

Suppose that the container is a cylinder, with a constant cross sectional area $A > 0$, as shown in Fig. 9.3. Suppose that the area of the hole is a. The rate that fluid leaves through the hole must balance with the rate that fluid decreases in the container. This principle is called **mass balance**. We will here assume that the density of water is constant, so that we can talk about the net changes in volume (rather than mass).

[51] This example is particularly instructive. First, it shows precisely how physical laws can be combined to formulate a model, then it shows how the problem can be recast as a single ODE in one dependent variable. Finally, it illustrates a slightly different integral.

[52] As we have assumed that the hole is at $h = 0$, we henceforth consider the height of the fluid surface, $h(t)$ to be the same as ”the height of fluid above the hole”.

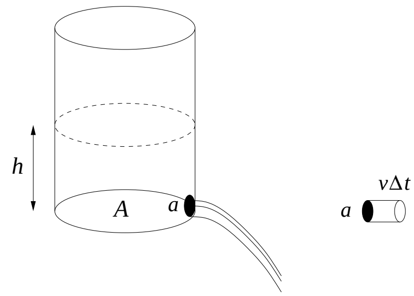

Figure 9.3. *We investigate the time it takes to empty a container full of fluid by deriving a differential equation model and solving it using the methods developed in this chapter. A is the cross-sectional area of the cylindrical tank, a is the cross-sectional area of the hole through which fluid drains, $v(t)$ is the velocity of the fluid, and $h(t)$ is the time dependent height of fluid remaining in the tank (indicated by the dashed line). The volume of fluid leaking out in a time span Δt is $av\Delta t$ - see small cylindrical volume indicated on the right.*

We refer to $V(t)$ as the volume of fluid in the container at time t. Note that for the cylindrical container, $V(t) = Ah(t)$ where A is the cross-sectional area and $h(t)$ is the height of the fluid at time t. The rate of change of V is

$$\frac{dV}{dt} = -(\text{rate volume lost as fluid flows out}).$$

(The minus sign indicates that the volume is decreasing).

At every second, some amount of fluid leaves through the hole. Suppose we are told that the velocity of the water molecules leaving the hole is precisely $v(t)$ in units of cm/sec. (We will find out how to determine this velocity shortly.) Then in one second, those particles have moved a distance v cm/sec $\cdot$ 1 sec $= v$ cm. In fact, all the particles in a little cylinder of length v behind these molecules have also left the hole. Indeed, if we know the area of the hole, we can determine precisely what volume of water exits through the hole each second, namely

$$\text{rate volume lost as fluid flows out} = va.$$

(The small inset in Fig. 9.3 shows a little "cylindrical unit" of fluid that flows out of the hole per second. The area is a and the length of that little volume is v. Thus the volume leaving per second is va.)

So far we have a relationship between the volume of fluid in the tank and the velocity of the water exiting the hole:

$$\frac{dV}{dt} = -av.$$

Now we need to determine the velocity v of the flow to complete the formulation of the problem.

9.5.2 Conservation of energy

The fluid "picks up speed" because it has "dropped" by a height h from the top of the fluid surface to the hole. In doing so, a small mass of water has simply exchanged some potential energy (due to its relative height above the hole) for kinetic energy (expressed by how fast it is moving). Potential energy of a small mass of water (m) at height h will be mgh, whereas when the water flows out of the hole, its kinetic energy is given by $(1/2)mv^2$ where v is velocity. Thus, for these to balance (so that total energy is conserved) we have

$$\frac{1}{2}mv^2 = mgh.$$

(Here $v = v(t)$ is the instantaneous velocity of the fluid leaving the hole and $h = h(t)$ is the height of the water column.) This allows us to relate the velocity of the fluid leaving the hole to the height of the water in the tank, i.e.

$$v^2 = 2gh \quad \Rightarrow \quad v = \sqrt{2gh}. \tag{9.11}$$

In fact, both the height of fluid and its exit velocity are constantly changing as the fluid drains, so we might write $[v(t)]^2 = 2gh(t)$ or $v(t) = \sqrt{2gh(t)}$. We have arrived at this result using an **energy balance** argument.

9.5.3 Putting it together

We now combine the various pieces of information to arrive at the model, a differential equation for a single (unknown) function of time. There are three time-dependent variables that were discussed above, the volume $V(t)$, the height $h(t)$, of the velocity $v(t)$. It proves convenient to express everything in terms of the height of water in the tank, $h(t)$, though this choice is to some extent arbitrary. Keeping units in an equation consistent is essential. Checking for unit consistency can help to uncover errors in equations, including differential equations.

Recall that the volume of the water in the tank, $V(t)$ is related to the height of fluid $h(t)$ by

$$V(t) = Ah(t),$$

where $A > 0$ is a constant, the cross-sectional area of the tank. We can simplify as follows:

$$\frac{dV}{dt} = \frac{d(Ah(t))}{dt} = A\frac{d(h(t))}{dt}.$$

But by previous steps and Eqn. (9.11)

$$\frac{dV}{dt} = -av = -a\sqrt{2gh}.$$

Thus

$$A\frac{d(h(t))}{dt} = -a\sqrt{2gh},$$

or simply put,

$$\frac{dh}{dt} = -\frac{a}{A}\sqrt{2gh} = -k\sqrt{h}. \tag{9.12}$$

where k is a constant that depends on the size and shape of the cylinder and its hole:

$$k = \frac{a}{A}\sqrt{2g}.$$

If the area of the hole is very small relative to the cross-sectional area of the tank, then k will be very small, so that the tank will drain very slowly (i.e. the rate of change in h per unit time will not be large). On a planet with a very high gravitational force, the same tank will drain more quickly. A taller column of water drains faster. Once its height has been reduced, its rate of draining also slows down. We comment that Equation (9.12) has a minus sign, signifying that the height of the fluid decreases.

Using simple principles such as conservation of mass and conservation of energy, we have shown that the height $h(t)$ of water in the tank at time t satisfies the differential equation (9.12). Putting this together with the initial condition (height of fluid h_0 at time $t = 0$), we arrive at initial value problem to solve:

$$\frac{dh}{dt} = -k\sqrt{h}, \quad h(0) = h_0. \tag{9.13}$$

Clearly, this equation is valid only for h non-negative. We also remark that Eqn. (9.13) is **nonlinear**[53] as it involves the variable h in a nonlinear term, $\sqrt{h}$. Next, we use separation of variables to find the height as a function of time.

9.5.4 Solution by separation of variables

The equation (9.13) shows how height of fluid is related to its rate of change, but we are interested in an explicit formula for fluid height h versus time t. To obtain that relationship, we must determine the solution to this differential equation. We do this using separation of variables. (We will also use the initial condition $h(0) = h_0$ that accompanies Eqn. (9.13).) As usual, rewrite the equation in the separated form,

$$\frac{dh}{\sqrt{h}} = -kdt,$$

We integrate from $t = 0$ to $t = T$, during which the height of fluid that started as h_0 becomes some new height $h(T)$ to be determined.

$$\int_{h_0}^{h(T)} \frac{1}{\sqrt{h}}\, dh = -k \int_0^T \, dt.$$

Now integrate both sides and simplify:

$$\left.\frac{h^{1/2}}{(1/2)}\right|_{h_0}^{h(T)} = -kT$$

$$2\left(\sqrt{h(T)} - \sqrt{h_0}\right) = -kT$$

[53] In many cases, nonlinear differential equations are more challenging than linear ones. However, examples chosen in this chapter are simple enough that we will not experience the true challenges of such nonlinearities.

$$\sqrt{h(T)} = -k\frac{T}{2} + \sqrt{h_0}$$

$$h(T) = \left(\sqrt{h_0} - k\frac{T}{2}\right)^2 .$$

Since this is true for any time t, we can also write the form of the solution as

$$h(t) = \left(\sqrt{h_0} - k\frac{t}{2}\right)^2 . \tag{9.14}$$

Eqn. (9.14) predicts fluid height remaining in the tank versus time t. In Fig. 9.4 we show some of the "solution curves"[54], i.e. functions of the form Eqn. (9.14) for a variety of initial fluid height values h_0. We can also use our results to predict the emptying time, as shown in the next section.

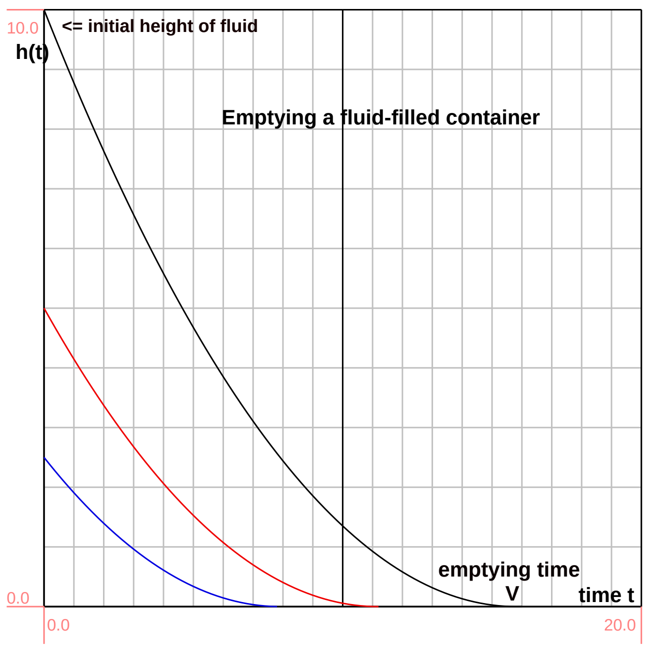

Figure 9.4. *Solution curves obtained by plotting Eqn. (9.14) for three different initial heights of fluid in the container,* $h_0 = 2.5, 5, 10$. *The parameter* $k = 0.4$ *in each case. The "V" points to the time it takes the tank to empty starting from a height of* $h(t) = 10$.

[54] As before, this figure was produced by plotting the analytic solution (9.14). A numerical method alternative would use Euler's Method and the spreadsheet to obtain the (approximate) solution directly from the initial value problem (9.13).

9.5.5 How long will it take the tank to empty?

The tank will be empty when the height of fluid is zero. Setting $h(t) = 0$ in Eqn. 9.14

$$\left(\sqrt{h_0} - k\frac{t}{2}\right)^2 = 0.$$

Solving this equation for the emptying time t_e, we get

$$k\frac{t_e}{2} = \sqrt{h_0} \quad \Rightarrow \quad t_e = \frac{2\sqrt{h_0}}{k}.$$

The time it takes to empty the tank depends on the initial height of water in the tank. Three examples are shown in Figure 9.4 for initial heights of $h_0 = 2.5,\ 5,\ 10$. The emptying time depends on the square-root of the initial height. This means, for instance, that doubling the height of fluid initially in the tank only increases the time it takes by a factor of $\sqrt{2} \approx 1.41$. Making the hole smaller has a more direct "proportional" effect, since we have found that $k = (a/A)\sqrt{2g}$.

9.6 Density dependent growth

The simple model discussed in Section 9.2 for population growth has an unrealistic feature of unlimited explosive exponential growth. To correct for this unrealistic feature, a common assumption is that the rate of growth is "density dependent". In this section, we consider a revised differential equation that describes such growth, and use the new tools to analyze its predictions. In place of our previous notation we will now use N to represent the size of a population.

9.6.1 The logistic equation

The logistic equation is the simplest density dependent growth equation, and we study its behaviour below.

Let $N(t)$ be the size of a population at time t. Clearly, we expect $N(t) \geq 0$ for all time t, since a population cannot be negative. We will assume that the initial population is known, $N(0) = N_0$. The logistic differential equation states that the rate of change of the population is given by

$$\frac{dN}{dt} = rN\left(\frac{K-N}{K}\right). \tag{9.15}$$

Here $r > 0$ is called the **intrinsic growth rate** and $K > 0$ is called the **carrying capacity**. K reflects that size of the population that can be sustained by the given environment. We can understand this equation as a modified growth law in which the "density dependent" term, $r(K - N)/K$, replaces the previous constant net growth rate k.

9.6.2 Scaling the equation

The form of the equation can be simplified if we measure the population in units of the carrying capacity, instead of "numbers of individuals". i.e. if we define a new quantity

$$y(t) = \frac{N(t)}{K}.$$

This procedure is called **scaling**. To see this, consider dividing each side of the logistic equation (9.15) by the constant K. Then

$$\frac{1}{K}\frac{dN}{dt} = \frac{r}{K}N\left(\frac{K-N}{K}\right).$$

We now group terms conveniently, forming

$$\frac{d(\frac{N}{K})}{dt} = r\left(\frac{N}{K}\right)\left(1-\left(\frac{N}{K}\right)\right).$$

Replacing (N/K) by y in each case, we obtain the scaled equation and initial condition given by

$$\frac{dy}{dt} = ry(1-y), \quad y(0) = y_0. \tag{9.16}$$

Now the variable $y(t)$ measures population size in "units" of the carrying capacity, and $y_0 = N_0/K$ is the scaled initial population level. Here again is an initial value problem, like Eqn. (9.13), but unlike Eqn. (9.1), the logistic differential equation is nonlinear. That is, the variable y appears in a nonlinear expression (in fact a quadratic) in the equation.

9.6.3 Separation of variables

Here we will solve Eqn. (9.16) by separation of variables. The idea is essentially the same as our previous examples, but is somewhat more involved. To show an alternative method of handling the integration, we will treat both sides as indefinite integrals. Separating the variables leads to

$$\frac{1}{y(1-y)}\,dy = r\,dt$$

$$\int \frac{1}{y(1-y)}\,dy = \int r\,dt + K.$$

The integral on the right will lead to $rt + K$ where K is some constant of integration that we need to incorporate since we do not have endpoints on our integrals. But we must work harder to evaluate the integral on the left. We can do so by partial fractions, the technique described in Section 6.6. Details are given in Section 9.6.4.

9.6.4 Application of partial fractions

Let

$$I = \int \frac{1}{y(1-y)}\,dy.$$

Then for some constants A, B we can write

$$I = \int \frac{A}{y} + \frac{B}{1-y}\, dy = A \ln|y| - B \ln|1-y|.$$

(The minus sign in front of B stems from the fact that letting $u = 1-y$ would lead to $du = -dy$.) We can find A, B from the fact that

$$\frac{A}{y} + \frac{B}{1-y} = \frac{1}{y(1-y)},$$

so that

$$A(1-y) + By = 1.$$

This must be true for all y, and in particular, substituting in $y = 0$ and $y = 1$ leads to $A = 1, B = 1$ so that

$$I = \ln|y| - \ln|1-y| = \ln\left|\frac{y}{1-y}\right|.$$

9.6.5 The solution of the logistic equation

We now have to extract the quantity y from the equation

$$\left(\ln\left|\frac{y}{1-y}\right|\right) = rt + K.$$

That is, we want y as a function of t. After exponentiating both sides we need to remove the absolute value. We will now assume that y is initially smaller than 1, and show that it remains so. In that case, everything inside the absolute value is positive, and we can write

$$\frac{y(t)}{(1-y(t))} = e^{rt+K} = e^K e^{rt} = Ce^{rt}.$$

In the above step, we have simply renamed the constant, e^K by the new name C for simplicity. $C > 0$ is now also an arbitrary constant whose value will be determined from the initial conditions. Indeed, if we substitute $t = 0$ into the most recent equation, we find that

$$\frac{y(0)}{(1-y(0))} = Ce^0 = C,$$

so that

$$C = \frac{y_0}{(1-y_0)}.$$

We will use this fact shortly. What remains now is some algebra to isolate the desired function $y(t)$

$$y(t) = (1-y(t))Ce^{rt}.$$

$$y(t)\left(1 + Ce^{rt}\right) = Ce^{rt}.$$

$$y(t) = \frac{Ce^{rt}}{(1 + Ce^{rt})} = \frac{1}{(1/C)e^{-rt} + 1}.$$

The desired function is now expressed in terms of the time t, and the constants r, C. We can also express it in terms of the initial value of y, i.e. y_0, by using what we know to be true about the constant C, i.e. $C = y_0/(1 - y_0)$. When we do so, we arrive at

$$y(t) = \frac{1}{\frac{1+y_0}{y_0}e^{-rt} + 1} = \frac{y_0}{(y_0 + (1 - y_0)e^{-rt})}. \tag{9.17}$$

Some typical solution curves of the logistic equation are shown in Fig. 9.5.

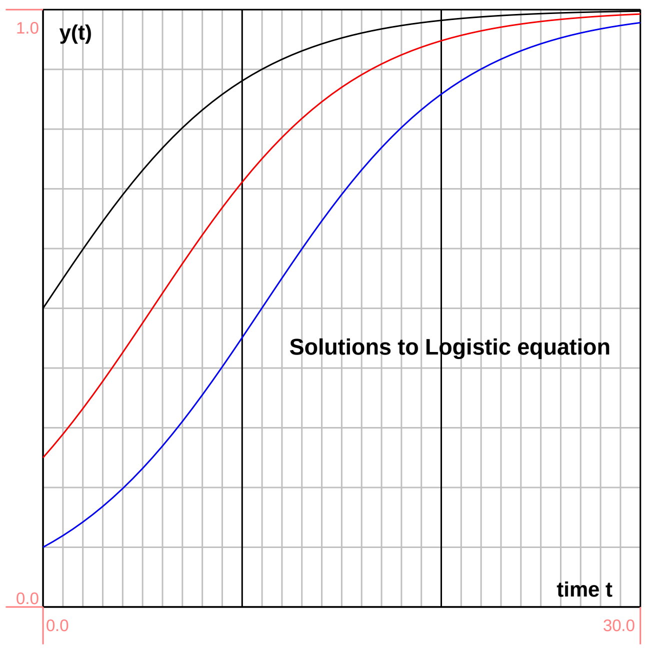

Figure 9.5. *Solution curves for $y(t)$ in the scaled form of the logistic equation based on (9.18). We show the predicted behaviour of $y(t)$ as given by Eqn. (9.17) for three different initial conditions, $y_0 = 0.1, 0.25, 0.5$. Note that all solutions approach the value $y = 1$.*

9.6.6 What this solution tells us

We have arrived at the function that describes the scaled population as a function of time as predicted by the scaled logistic equation, (9.16). The level of population (in units of the carrying capacity K) follows the time-dependent function

$$y(t) = \frac{y_0}{(y_0 + (1 - y_0)e^{-rt})}. \tag{9.18}$$

We can convert this result to an equivalent expression for the unscaled total population $N(t)$ by recalling that $y(t) = N(t)/K$. Substituting this for $y(t)$, and noting that $y_0 = N_0/K$ leads to

$$N(t) = \frac{N_0}{(N_0 + (K - N_0)e^{-rt})}. \tag{9.19}$$

It is left as an exercise for the reader to check this claim.

Now recall that $r > 0$. This means that e^{-rt} is a decreasing function of time. Therefore, (9.18) implies that, after a long time, the term e^{-rt} in the denominator will be negligibly small, and so

$$y(t) \to \frac{y_0}{y_0} = 1,$$

so that y will approach the value 1. This means that

$$(N/K) \to 1 \quad \text{or simply} \quad N(t) \to K.$$

The population will thus settle into a constant level, i.e., a **steady state**, at which no further change will occur.

As an aside, we observe that this too, could have been predicted directly from the differential equation. By setting $dy/dt = 0$, we find that

$$0 = ry(1 - y),$$

which suggests that $y = 1$ is a steady state. (This is also true for the less interesting case of no population, i.e. $y = 0$ is also a steady state.) Similarly, this could have been found by setting the derivative to zero in Eqn. (9.15), the original, unscaled logistic differential equation. Doing so leads to

$$\frac{dN}{dt} = 0 \Rightarrow rN\left(\frac{K - N}{K}\right) = 0.$$

If $r > 0$, the only values of N satisfying this steady state equation are $N = 0$ or $N = K$. This implies that either $N = 0$ or $N = K$ are steady states. The former is not too interesting. It states the obvious fact that if there is no population, then there can be no population growth. The latter reflects that $N = K$, the carrying capacity, is the population size that will be sustained by the environment.

In summary, we have shown that the behaviour of the logistic equation for population growth is more realistic than the simpler exponential growth we studied earlier. We saw in Figure 9.5, that a small population will grow, but only up to some constant level (the carrying capacity). Integration, and in particular the use of partial fractions allowed us to make a full prediction of the behaviour of the population level as a function of time, given by Eqn. (9.19).

9.7 Extensions and other population models: the "Law of Mortality"

There are many variants of the logistic model that are used to investigate the growth or mortality of a population. Here we extend tools to another example, the gradual decline of a group of individuals born at the same time. Such a group is called a “cohort”.[55]. In 1825, Gompertz suggested that the rate of mortality, m would depend on the age of the individuals. Because we consider a group of people who were born at the same time, we can trade ”age” for ”time”. Essentially, Gompertz assumed that mortality is not constant: it is low at first, and increase as individuals age. Gompertz argued that mortality increases exponentially. This turns out to be equivalent to the assumption that the logarithm of mortality increases linearly with time.[56] It is easy to see that these two statements are equivalent: Suppose we assume that for some constants $A > 0, \mu > 0$,

$$\ln(m(t)) = A + \mu t. \tag{9.20}$$

Then Eqn. (9.20) means that

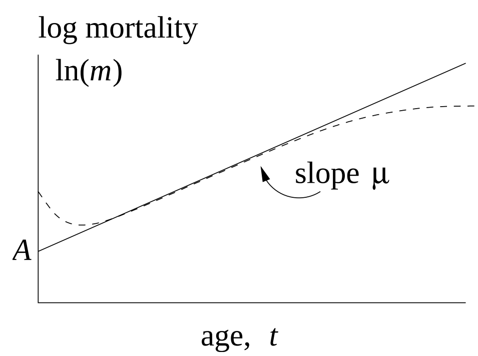

Figure 9.6. *In the Gompertz Law of Mortality, it is assumed that the log of mortality increases linearly with time, as depicted by Eqn. 9.20 and by the solid curve in this diagram. Here the slope of* $\ln(m)$ *versus time (or age) is* μ*. For real populations, the mortality looks more like the dashed curve.*

$$m(t) = e^{A+\mu t} = e^A e^{\mu t}$$

Since A is constant, so is e^A. For simplicity we define Let us define $m_0 = e^A$. ($m_0 = m(0)$ is the so-called “birth mortality” i.e. value of m at age 0.) Thus, the time-dependent mortality is

$$m = m(t) = m_0 e^{\mu t}. \tag{9.21}$$

9.7.1 Aging and Survival curves for a cohort:

We now study a population model having Gompertz mortality, together with the following additional assumptions.

[55]This section was formulated with help from Lu Fan

[56]In actual fact, this is likely true for some range of ages. Infant mortality is generally higher than mortality for young children, whereas mortality levels off or even decreases slightly for those oldest old who have survived past the average lifespan.

1. All individuals are assumed to be identical.

2. There is "natural" mortality, but no other type of removal. This means we ignore the mortality caused by epidemics, by violence and by wars.

3. We consider a single cohort, and assume that no new individuals are introduced (e.g. by immigration)[57].

We will now study the size of a "cohort", i.e. a group of people who were born in the same year. We will denote by $N(t)$ the number of people in this group who are alive at time t, where t is time since birth, i.e. age. Let $N(0) = N_0$ be the initial number of individuals in the cohort.

9.7.2 Gompertz Model

All the people in the cohort were born at time (age) $t = 0$, and there were N_0 of them at that time. That number changes with time due to mortality. Indeed,

$$\begin{aligned}\text{The rate of change of cohort size} &= -[\text{number of deaths per unit time}]\\ &= -[\text{mortality rate}]\cdot[\text{cohort size}]\end{aligned}$$

Translating to mathematical notation, we arrive at the differential equation

$$\frac{dN(t)}{dt} = -m(t)N(t),$$

and using information about the size of the cohort at birth leads to the initial condition, $N(0) = N_0$. Together, this leads to the initial value problem

$$\frac{dN(t)}{dt} = -m(t)N(t), \quad N(0) = N_0.$$

Note similarity to Eqn. (9.1), but now mortality is time-dependent.

In the Problem set, we apply separation of variables and integrate over the time interval $[0, T]$: to show that the remaining population at age t is

$$N(t) = N_0 e^{-\frac{m_0}{\mu}(e^{\mu t}-1)}.$$

9.8 Summary

In this chapter, we used integration methods to find the analytical solutions to a variety of differential equations where initial values were prescribed.

We investigated a number of **population growth** models:

1. Exponential growth, given by $\frac{dy}{dt} = ky$, with initial population level $y(0) = y_0$ was investigated (Eqn. (9.1)). This model had an unrealistic feature that growth is unlimited.

[57]Note that new births would contribute to other cohorts.

2. The Logistic equation $\frac{dN}{dt} = rN\left(\frac{K-N}{K}\right)$ was analyzed (Eqn. (9.15)), showing that density-dependent growth can correct for the above unrealistic feature.

3. The Gompertz equation, $\frac{dN(t)}{dt} = -m(t)N(t)$, was solved to understand how age-dependent mortality affects a cohort of individuals.

In each of these cases, we used separation of variables to "integrate" the differential equation, and predict the population as a function of time.

We also investigated several other **physical models** in this chapter, including the velocity of a falling object subject to drag force. This led us to study a differential equation of the form $\frac{dy}{dt} = a - by$. By slight reinterpretation of terms in this equation, we can use results to understand chemical kinetics and blood alcohol levels, as well as a host of other scientific applications.

Section 9.5, the "centerpiece" of this chapter, illustrated the detailed steps that go into the formulation of a differential equation model for flow of liquid out of a container. Here we saw how conservation statements and simplifying assumptions are interpreted together, to arrive at a differential equation model. Such ideas occur in many scientific problems, in chemistry, physics, and biology.

Chapter 10

Infinite series, improper integrals, and Taylor series

10.1 Introduction

This chapter has several important and challenging goals. The first of these is to understand how concepts that were discussed for finite series and integrals can be meaningfully extended to infinite series and improper integrals - i.e. integrals of functions over an infinite domain. In this part of the discussion, we will find that the notion of **convergence** and **divergence** will be important.

A second theme will be that of approximation of functions in terms of power series, also called **Taylor series**. Such series can be described informally as infinite polynomials (i.e. polynomials containing infinitely many terms). Understanding when these objects are meaningful is also related to the issue of convergence, so we use the background assembled in the first part of the chapter to address such concepts arising in the second part.

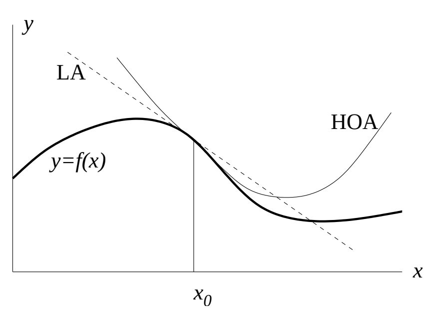

Figure 10.1. *The function $y = f(x)$ (solid heavy curve) is shown together with its linear approximation (LA, dashed line) at the point x_0, and a better "higher order" approximation (HOA, thin solid curve). Notice that this better approximation stays closer to the graph of the function near x_0. In this chapter, we discuss how such better approximations can be obtained.*

The theme of approximation has appeared often in our calculus course. In a previous semester, we discussed a **linear approximation** to a function. The idea was to approximate the value of the function close to a point on its graph using a straight line (the tangent line). We noted in doing so that the approximation was good only close to the point of tangency. Further away, the graph of the functions curves away from that straight line. This leads naturally to the question: can we do better in making this approximation if we include other terms to describe this "curving away"? Here we extend such linear approximation methods. Our goal is to increase the accuracy of the linear approximation by including higher order terms (quadratic, cubic, etc), i.e. to find a polynomial that approximates the given function. This idea forms an important goal in this chapter.

We first review the idea of series introduced in Chapter 1.

10.2 Convergence and divergence of series

Recall the geometric series discussed in Section 1.6.

The sum of a **finite geometric series**,

$$S_n = 1 + r + r^2 + \ldots + r^n = \sum_{k=0}^{n} r^k, \quad \text{is} \quad S_n = \frac{1 - r^{n+1}}{1 - r}. \tag{10.1}$$

We also review definitions discussed in Section 1.7

Definition: Convergence of infinite series

An infinite series that has a finite sum is said to be **convergent**. Otherwise it is **divergent**.

Definition: Partial sums and convergence

Suppose that S is an (infinite) series whose terms are a_k. Then the **partial sums**, S_n, of this series are

$$S_n = \sum_{k=0}^{n} a_k.$$

We say that the sum of the infinite series is S, and write

$$S = \sum_{k=0}^{\infty} a_k, \quad \text{provided that} \quad S = \lim_{n\to\infty} S_n = \lim_{n\to\infty} \sum_{k=0}^{n} a_k.$$

That is, we consider the infinite series as the limit of partial sums S_n as the number of terms n is increased. If this limit exists, we say that the infinite series **converges**[58] to S. This leads to the following conclusion:

[58] If the limit does not exist, we say that the series diverges.

The sum of an **infinite geometric series**,

$$S = 1 + r + r^2 + \ldots + r^k + \ldots = \sum_{k=0}^{\infty} r^k = \frac{1}{1-r}, \quad \text{provided } |r| < 1. \tag{10.2}$$

If this inequality is not satisfied, then we say that this sum does not exist (meaning that it is not finite).

It is important to remember that an infinite series, i.e. a sum with infinitely many terms added up, can exhibit either one of these two very different behaviours. It may converge in some cases, as the first example shows, or **diverge** (fail to converge) in other cases. We will see examples of each of these trends again. It is essential to be able to distinguish the two. Divergent series (or series that diverge under certain conditions) must be handled with particular care, for otherwise, we may find contradictions or “seemingly reasonable” calculations that have meaningless results.

We can think of convergence or divergence of series using a geometric analogy. If we start on the number line at the origin and take a sequence of steps $\{a_0, a_1, a_2, \ldots, a_k, \ldots\}$, we can think of $S = \sum_{k=0}^{\infty} a_k$ as the total distance we have travelled. S converges if that distance remains finite and if we approach some fixed number.

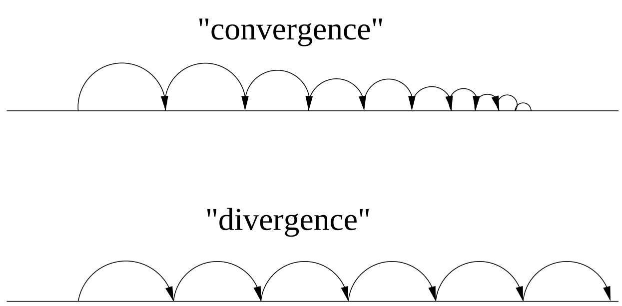

Figure 10.2. *An informal schematic illustrating the concept of convergence and divergence of infinite series. Here we show only a few terms of the infinite series: from left to right, each step is a term in the series. In the top example, the sum of the steps gets closer and closer to some (finite) value. In the bottom example, the steps lead to an ever increasing total sum.*

In order for the sum of ‘infinitely many things’ to add up to a finite number, the terms have to get smaller. But just getting smaller is not, in itself, enough to guarantee convergence. (We will show this later on by considering the harmonic series.) There are rigorous mathematical tests which help determine whether a series converges or not. We discuss some of these tests in Appendix 11.9.

10.3 Improper integrals

We will see that there is a close connection between certain infinite series and improper integrals, i.e. integrals over an infinite domain. We have already encountered an example of an improper integral in Section 3.8.5 and in the context of radioactive decay in Section 8.4. Recall the following definition:

Definition

An **improper** integral is an integral performed over an infinite domain, e.g.

$$\int_a^\infty f(x)\, dx.$$

The value of such an integral is understood to be a limit, as given in the following definition:

$$\boxed{\int_a^\infty f(x)\, dx = \lim_{b\to\infty} \int_a^b f(x)\, dx.}$$

i.e. we evaluate an improper integral by first computing a definite integral over a finite domain $a \le x \le b$, and then taking a limit as the endpoint b moves off to larger and larger values. The definite integral can be interpreted as an area under the graph of the function. The essential question being addressed here is whether that area remains bounded when we include the “infinite tail” of the function (i.e. as the endpoint b moves to larger values.) For some functions (whose values get small enough fast enough) the answer is “yes”.

Definition

When the limit shown above exists, we say that the improper integral **converges**. Other wise we say that the improper integral **diverges**.

With this in mind, we compute a number of classic integrals:

10.3.1 Example: A decaying exponential: convergent improper integral

Here we recall that the improper integral of a decaying exponential converges. (We have seen this earlier, in Section 3.8.5, and again in applications in Sections 4.5 and 8.4.1. Here we recap this important result in the context of our discussion of improper integrals.) Suppose that $r > 0$ and let

$$I = \int_0^\infty e^{-rt}\, dt \equiv \lim_{b\to\infty} \int_0^b e^{-rt}\, dt.$$

Then note that $b > 0$ so that

$$I = \lim_{b\to\infty} -\frac{1}{r} e^{-rt}\Big|_0^b = -\frac{1}{r} \lim_{b\to\infty} (e^{-rb} - e^0) = -\frac{1}{r}(\underbrace{\lim_{b\to\infty} e^{-rb}}_{0} - 1) = \frac{1}{r}.$$

We have used the fact that $\lim_{b\to\infty} e^{-rb} = 0$ since (for $r, b > 0$) the exponential function is decreasing with increasing b. Thus *the limit exists (is finite)* and the integral **converges**. In fact it converges to the value $I = 1/r$.

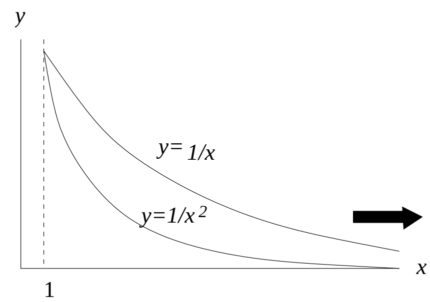

Figure 10.3. *In Sections 10.3.2 and 10.3.3, we consider two functions whose values decrease along the x axis, $f(x) = 1/x$ and $f(x) = 1/x^2$. We show that one, but not the other encloses a finite (bounded) area over the interval $(1, \infty)$. To do so, we compute an improper integral for each one. The heavy arrow is meant to remind us that we are considering areas over an unbounded domain.*

10.3.2 Example: The improper integral of $1/x$ diverges

We now consider a classic and counter-intuitive result, and *one of the most important results in this chapter*. Consider the function

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

Examining the graph of this function for positive x, e.g. for the interval $(1, \infty)$, we know that values decrease to zero as x increases[59]. The function is not only itself bounded, but also falls to arbitrarily small values as x increases. Nevertheless, this is *not enough* to guarantee that the enclosed area remains finite! We show this in the following calculation.

$$I = \int_1^{\infty} \frac{1}{x}\, dx = \lim_{b\to\infty} \int_a^b \frac{1}{x}\, dx = \lim_{b\to\infty} \ln(x)\Big|_1^b = \lim_{b\to\infty} (\ln(b) - \ln(1))$$

$$I = \lim_{b\to\infty} \ln(b) = \infty$$

The fact that we get an infinite value for this integral follows from the observation that $\ln(b)$ increases without bound as b increases, that is *the limit does not exist (is not finite)*. Thus the area under the curve $f(x) = 1/x$ over the interval $1 \le x \le \infty$ is infinite. We say that the improper integral of $1/x$ **diverges** (or does not converge). We will use this result again in Section 10.4.1.

[59]We do not chose the interval $(0, \infty)$ because this function is undefined at $x = 0$. We want here to emphasize the behaviour at infinity, not the blow up that occurs close to $x = 0$.

10.3.3 Example: The improper integral of $1/x^2$ converges

Now consider the related function

$$y = f(x) = \frac{1}{x^2}, \quad \text{and the corresponding integral} \quad I = \int_1^\infty \frac{1}{x^2}\,dx$$

Then

$$I = \lim_{b\to\infty} \int_1^b x^{-2}\,dx. = \lim_{b\to\infty} (-x^{-1})\Big|_1^b = -\lim_{b\to\infty}\left(\frac{1}{b} - 1\right) = 1.$$

Thus, *the limit exists*, and, in fact, $I = 1$, so, in contrast to the Example 10.3.2, this integral **converges**.

We observe that the behaviours of the improper integrals of the functions $1/x$ and $1/x^2$ are very different. The former diverges, while the latter converges. The only difference between these functions is the power of x. As shown in Figure 10.3, that power affects how rapidly the graph "falls off" to zero as x increases. The function $1/x^2$ decreases much faster than $1/x$. (Consequently $1/x^2$ has a sufficiently "slim" infinite "tail", that the area under its graph does not become infinite - not an easy concept to digest!) This observations leads us to wonder what power p is needed to make the improper integral of a function $1/x^p$ converge. We answer this question below.

10.3.4 When does the integral of $1/x^p$ converge?

Here we consider an arbitrary power, p, that can be any real number. We ask when the corresponding improper integral converges or diverges. Let

$$I = \int_1^\infty \frac{1}{x^p}\,dx.$$

For $p = 1$ we have already established that this integral diverges (Section 10.3.2), and for $p = 2$ we have seen that it is convergent (Section 10.3.3). By a similar calculation, we find that

$$I = \lim_{b\to\infty} \frac{x^{1-p}}{(1-p)}\Big|_1^b = \lim_{b\to\infty}\left(\frac{1}{1-p}\right)\left(b^{1-p} - 1\right).$$

Thus this integral converges provided that the term b^{1-p} does not "blow up" as b increases. For this to be true, we require that the exponent $(1-p)$ should be negative, i.e. $1 - p < 0$ or $p > 1$. In this case, we have

$$I = \frac{1}{p-1}.$$

To summarize our result,

$$\boxed{\int_1^\infty \frac{1}{x^p}\,dx \qquad \text{converges if } p > 1, \quad \text{diverges if } p \le 1.}$$

Examples: $\int 1/x^p$ that do or do not converge

1. The integral

$$\int_1^\infty \frac{1}{\sqrt{x}}\,dx, \quad \text{diverges.}$$

   We see this from the following argument: $\sqrt{x} = x^{\frac{1}{2}}$, so $p = \frac{1}{2} < 1$. Thus, by the general result, this integral diverges.

2. The integral

$$\int_1^\infty x^{-1.01}\,dx, \quad \text{converges.}$$

   Here $p = 1.01 > 1$, so the result implies convergence of the integral.

10.3.5 Integral comparison tests

The integrals discussed above can be used to make comparisons that help us to identify when other improper integrals converge or diverge[60]. The following important result establishes how these comparisons work:

Suppose we are given two functions, $f(x)$ and $g(x)$, both continuous on some infinite interval $[a, \infty)$. Suppose, moreover, that, at all points on this interval, the first function is smaller than the second, i.e.

$$0 \le f(x) \le g(x).$$

Then the following conclusions[a] can be made:

1. $\displaystyle\int_a^\infty f(x)\,dx \le \int_a^\infty g(x)\,dx$. (This means that the area under $f(x)$ is smaller than the area under $g(x)$.)
2. If $\displaystyle\int_a^\infty g(x)\,dx$ converges, then $\displaystyle\int_a^\infty f(x)\,dx$ converges. (If the larger area is finite, so is the smaller one)
3. If $\displaystyle\int_a^\infty f(x)\,dx$ diverges, then $\displaystyle\int_a^\infty g(x)\,dx$ diverges. (If the smaller area is infinite, so is the larger one.)

[a] These statements have to be carefully noted. What is assumed and what is concluded works “one way”. That is the order “if..then” is important. Reversing that order leads to a common error.

[60] The reader should notice the similarity of these ideas to the comparisons made for infinite series in the Appendix 11.9.2. This similarity stems from the fact that there is a close connection between series and integrals, a recurring theme in this course.

Example: comparison of improper integrals

We can determine that the integral

$$\int_1^\infty \frac{x}{1+x^3}\,dx \quad \text{converges}$$

by noting that for all $x > 0$

$$0 \le \frac{x}{1+x^3} \le \frac{x}{x^3} = \frac{1}{x^2}.$$

Thus

$$\int_1^\infty \frac{x}{1+x^3}\,dx \le \int_1^\infty \frac{1}{x^2}\,dx.$$

Since the larger integral on the right is known to converge, so does the smaller integral on the left.

10.4 Comparing integrals and series

The convergence of infinite series was discussed earlier, in Section 1.7 and here again in Section 10.2. Many tests for convergence are provided in the Appendix 11.9, and will not be discussed in detail due to time and space constraints. However, an interesting connection exists between convergence of series and integrals. This is the topic we examine here.

Convergence of series and convergence of integrals is related. We can use the convergence/divergence of an integral/series to determine the behaviour of the other. Here we give an example of this type by establishing a connection between the harmonic series and a divergent improper integral.

10.4.1 The harmonic series

The harmonic series is a sum of terms of the form $1/k$ where $k = 1, 2, \ldots$. At first appearance, this series might seem to have the desired qualities of a convergent series, simply because the successive terms being added are getting smaller and smaller, but this appearance is deceptive and actually wrong[61].

The harmonic series

$$\sum_{k=1}^{\infty} \frac{1}{k} = 1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \ldots + \frac{1}{k} + \ldots \qquad \text{diverges}$$

We establish that the harmonic series diverges by comparing it to the improper integral of the related function[62].

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

[61] We have already noticed a similar surprise in connection with the improper integral of $1/x$. These two "surprises" are closely related, as we show here using a comparison of the series and the integral.

[62] This function is "related" since for integer values of x, the function takes on values that are the same as successive terms in the series, i.e. if $x = k$ is an integer, then $f(x) = f(k) = 1/k$

Figure 10.4. *The harmonic series is a sum that corresponds to the area under the staircase shown above. Note that we have purposely shown the stairs arranged so that they are higher than the function. This is essential in drawing the conclusion that the sum of the series is infinite: It is* larger *than an area under* $1/x$ *that we already know to be infinite, by Section 10.3.2.*

In Figure 10.4 we show on one graph a comparison of the area under this curve, and a staircase area representing the first few terms in the harmonic series. For the area of the staircase, we note that the width of each step is 1, and the heights form the sequence

$$\{1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \ldots\}$$

Thus the area of (infinitely many) of these steps can be expressed as the (infinite) harmonic series,

$$A = 1 \cdot 1 + 1 \cdot \frac{1}{2} + 1 \cdot \frac{1}{3} + 1 \cdot \frac{1}{4} + \ldots = 1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \ldots = \sum_{k=1}^{\infty} \frac{1}{k}.$$

On the other hand, the area under the graph of the function $y = f(x) = 1/x$ for $0 \le x \le \infty$ is given by the improper integral

$$\int_1^{\infty} \frac{1}{x}\, dx.$$

We have seen previously in Section 10.3.2 that *this integral diverges*!

From Figure 10.4 we see that the areas under the function, A_f and under the staircase, A_s, satisfy

$$0 < A_f < A_s.$$

Thus, since the smaller of the two (the improper integral) is infinite, so is the larger (the sum of the harmonic series). We have established, using this comparison, that the the sum of the harmonic series cannot be finite, so that this series **diverges**.

Other comparisons: The "p" series

More generally, we can compare series of the form

$$\sum_{k=1}^{\infty} \frac{1}{k^p} \quad \text{to the integral} \quad \int_1^{\infty} \frac{1}{x^p}\,dx$$

in precisely the same way. This leads to the conclusion that

The "p" series,

$$\sum_{k=1}^{\infty} \frac{1}{k^p} \qquad \text{converges if } p > 1, \text{ diverges if } p \leq 1.$$

For example, the series

$$\sum_{k=1}^{\infty} \frac{1}{k^2} = 1 + \frac{1}{4} + \frac{1}{9} + \frac{1}{16} + \dots$$

converges, since $p = 2 > 1$. Notice, however, that the comparison does not give us a value to which the sum converges. It merely indicates that the series does converge.

10.5 From geometric series to Taylor polynomials

In studying calculus, we explored a variety of functions. Among the most basic are polynomials, i.e. functions such as

$$p(x) = x^5 + 2x^2 + 3x + 2.$$

Our introduction to differential calculus started with such functions for a reason: these functions are convenient and simple to handle. We found long ago that it is easy to compute derivatives of polynominals. The same can be said for integrals. One of our first examples, in Section 3.6.1 was the integral of a polynomial. We needed only use a power rule to integrate each term. An additional convenience of polynomials is that "evaluating" the function, (i.e. plugging in an x value and determining the corresponding y value) can be done by simple multiplications and additions, i.e. by basic operations easily handled by an ordinary calculator. This is not the case for, say, trigonometric functions, exponential functions, or for that matter, most other functions we considered[63]. For this reason, being able to *approximate* a function by a polynomial is an attractive proposition. This forms our main concern in the sections that follow.

We can arrive at connections between several functions and their polynomial approximations by exploiting our familiarity with the **geometric series**. We use both the results for convergence of the geometric series (from Section 10.2) and the formula for the sum of that series to derive a number of interesting, (somewhat haphazard) results[64].

Recall from Sections 1.7.1 and 10.2 that *the sum of an infinite geometric series is*

$$S = 1 + r + r^2 + \ldots + r^k + \ldots = \sum_{k=0}^{\infty} r^k = \frac{1}{1-r}, \quad \text{provided} \quad |r| < 1. \tag{10.3}$$

To connect this result to a statement about a function, we need a "variable". Let us consider the behaviour of this series when we vary the quantity r. To emphasize that now r is our variable, it will be helpful to change notation by substituting $r = x$ into the above equation, while remembering that the formula in Eqn (10.3) hold only provided $|r| = |x| < 1$.

10.5.1 Example 1: A simple expansion

Substitute the variable $x = r$ into the series (10.3). Then formally, rewriting the above with this substitution, leads to the conclusion that

$$\frac{1}{1-x} = 1 + x + x^2 + \ldots \tag{10.4}$$

We can think of this result as follows: Let

$$f(x) = \frac{1}{1-x} \tag{10.5}$$

Then for every x in $-1 < x < 1$, it is true that $f(x)$ can be approximated by terms in the polynomial

$$p(x) = 1 + x + x^2 + \ldots \tag{10.6}$$

In other words, by (10.3), for $|x| < 1$ the two expressions "are the same", in the sense that the polynomial converges to the value of the function. We refer to this $p(x)$ as an (infinite) Taylor polynomial[65] or simply a **Taylor series** for the function $f(x)$. The usefulness of this kind of result can be illustrated by a simple example.

Example 10.1 (Using the Taylor Series (10.6) to approximate the function (10.5)) Compute the value of the function $f(x)$ given by Eqn. (10.5) for $x = 0.1$ without using a calculator.

[63]For example, to find the decimal value of $\sin(2.5)$ we would need a scientific calculator. These days the distinction is blurred, since powerful hand-held calculators are ubiquitous. Before such devices were available, the ease of evaluating polynomials made them even more important.

[64]We say "haphazard" here because we are not yet at the point of a systematic procedure for computing a Taylor Series. That will be done in Section 10.6. Here we "take what we can get" using simple manipulations of a geometric series.

[65]A Taylor polynomial contains finitely many terms, n, whereas a Taylor series has $n \to \infty$.

Solution: Plugging in the value $x = 0.1$ into the function directly leads to $1/(1-0.1) = 1/0.9$, whose evaluation with no calculator requires long division[66]. Using the polynomial representation, we have a simpler method:

$$p(0.1) = 1 + 0.1 + 0.1^2 + \ldots = 1 + 0.1 + 0.01 + \ldots = 1.11\ldots$$

We provide a few other examples based on substitutions of various sorts using the geometric series as a starting point.

10.5.2 Example 2: Another substitution

We make the substitution $r = -t$, then $|r| < 1$ means that $|-t| = |t| < 1$, so that the formula (10.3) for the sum of a geometric series implies that:

$$\frac{1}{1-(-t)} = 1 + (-t) + (-t)^2 + (-t)^3 + \ldots$$
$$\frac{1}{1+t} = 1 - t + t^2 - t^3 + t^4 + \ldots \quad \text{provided } |t| < 1$$

This means we have produced a series expansion for the function $1/(1+t)$. We can go farther with this example by a new manipulation, whereby we integrate both sides to arrive at a new function and its expansion, shown next.

10.5.3 Example 3: An expansion for the logarithm

We will use the results of Example 10.5.2, but we follow our substitution by integration. On the left, we integrate the function $f(t) = 1/(1+t)$ (to arrive at a logarithm type integral) and on the right we integrate the power terms of the expansion. We are permitted to integrate the power series term by term *provided that the series converges*. This is an important restriction that we emphasize: *Manipulation of infinite series by integration, differentiation, addition, multiplication, or any other "term by term" computation makes sense only so long as the original series converges.*

Provided $|t| < 1$, we have that

$$\int_0^x \frac{1}{1+t}\, dt = \int_0^x (1 - t + t^2 - t^3 + t^4 - \ldots)\, dt$$
$$\ln(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \ldots$$

This procedure has allowed us to find a series representation for a new function, $\ln(1+x)$.

$$\ln(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \ldots = \sum_{k=1}^{\infty} (-1)^{k+1} \frac{x^k}{k}. \tag{10.7}$$

[66] This example is slightly "trivial", in the sense that evaluating the function itself is not very difficult. However, in other cases, we will find that the polynomial expansion is the *only* way to find the desired value.

The formula appended on the right is just a compact notation that represents the pattern of the terms. Recall that in Chapter 1, we have gotten thoroughly familiar with such summation notation[67].

Example 10.2 (Evaluating the logarithm for $x = 0.25$) An expansion for the logarithm is definitely useful, in the sense that (without a scientific calculator or log tables) it is not possible to easily calculate the value of this function at a given point. For example, for $x = 0.25$, we cannot find $\ln(1 + 0.25) = \ln(1.25)$ using simple operations, whereas the value of the first few terms of the series are computable by simple multiplication, division, and addition ($0.25 - \frac{0.25^2}{2} + \frac{0.25^3}{3} \approx 0.2239$). (A scientific calculator gives $\ln(1.25) \approx 0.2231$, so the approximation produced by the series is relatively good.)

When is the series for $\ln(1 + x)$ in (10.7) expected to converge? Retracing our steps from the beginning of Example 10.5.2 we note that the value of t is not permitted to leave the interval $|t| < 1$ so we need also $|x| < 1$ in the integration step[68]. We certainly cannot expect the series for $\ln(1 + x)$ to converge when $|x| > 1$. Indeed, for $x = -1$, we have $\ln(1 + x) = \ln(0)$ which is undefined. Also note that for $x = -1$ the right hand side of (10.7) becomes

$$-\left(1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \dots\right).$$

This is the recognizable harmonic series (multiplied by -1). But we already know from Section 10.4.1 that the harmonic series diverges. Thus, we must avoid $x = -1$, since the expansion will not converge there, and neither is the function defined. *This example illustrates that outside the interval of convergence, the series and the function become "meaningless"*.

Example 10.3 (An expansion for $\ln(2)$) Strictly speaking, our analysis does not predict what happens if we substitute $x = 1$ into the expansion of the function found in Section 10.5.3, because this value of x is outside of the permitted range $-1 < x < 1$ in which the Taylor series can be guaranteed to converge. It takes some deeper mathematics (Abel's theorem) to prove that the result of this substitution actually makes sense, and converges, i.e. that

$$\ln(2) = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \dots$$

We state without proof here that the *alternating harmonic series converges to* $\ln(2)$.

10.5.4 Example 4: An expansion for arctan

Suppose we make the substitution $r = -t^2$ into the geometric series formula, and recall that we need $|r| < 1$ for convergence. Then

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

[67] The summation notation is not crucial and should certainly not be memorized. We are usually interested only in the first few terms of such a series in any approximation of practical value.

[68] Strictly speaking, we should have ensured that we are inside this interval of convergence before we computed the last example.

$$\frac{1}{1+t^2} = 1 - t^2 + t^4 - t^6 + t^8 + \ldots = \sum_{k=0}^{\infty} (-1)^n t^{2n}$$

This series will converge provided $|t| < 1$. Now integrate both sides, and recall that the antiderivative of the function $1/(1+t^2)$ is $\arctan(t)$. Then

$$\int_0^x \frac{1}{1+t^2}\, dt = \int_0^x (1 - t^2 + t^4 - t^6 + t^8 + \ldots)\, dt$$

$$\arctan(x) = x - \frac{x^3}{3} + \frac{x^5}{5} - \frac{x^7}{7} + \ldots = \sum_{k=1}^{\infty} (-1)^{k+1} \frac{x^{(2k-1)}}{(2k-1)}. \qquad (10.8)$$

Example 10.4 (An expansion for π) For a particular application of this expansion, consider plugging in $x = 1$ into Equation (10.8). Then

$$\arctan(1) = 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \ldots$$

But $\arctan(1) = \pi/4$. Thus we have found a way of computing the irrational number π, namely

$$\pi = 4\left(1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \ldots\right) = 4\left(\sum_{k=1}^{\infty} (-1)^{k+1} \frac{1}{(2k-1)}\right).$$

While it is true that this series converges, the convergence is slow. (This can be seen by adding up the first 100 or 1000 terms of this series with a spreadsheet.) This means that it is not practical to use such a series as an approximation for π. (There are other series that converge to π very rapidly that are used in any practical application.)

10.6 Taylor Series: a systematic approach

In Section 10.5, we found a variety of Taylor series expansions directly from the formula for a geometric series. Here we ask how such Taylor series can be constructed more systematically, if we are given a function that we want to approximate [69].

Suppose we have a function which we want to represent by a power series,

$$f(x) = a_0 + a_1 x + a_2 x^2 + a_3 x^3 + \ldots = \sum_{k=0}^{\infty} a_k x^k.$$

Here we will use the function to directly determine the coefficients a_k. To determine a_0, let $x = 0$ and note that

$$f(0) = a_0 + a_1 0 + a_2 0^2 + a_3 0^3 + \ldots = a_0.$$

We conclude that

$$a_0 = f(0).$$

[69] The development of this section was motivated by online notes by David Austin.

By differentiating both sides we find the following:

$$\begin{aligned}
f'(x) &= a_1 + 2a_2x + 3a_3x^2 + \ldots + ka_kx^{k-1} + \ldots \\
f''(x) &= 2a_2 + 2\cdot 3a_3x + \ldots + (k-1)ka_kx^{k-2} + \ldots \\
f'''(x) &= 2\cdot 3a_3 + \ldots + (k-2)(k-1)ka_kx^{k-3} + \ldots \\
f^{(k)}(x) &= 1\cdot 2\cdot 3\cdot 4\ldots ka_k + \ldots
\end{aligned}$$

Here we have used the notation $f^{(k)}(x)$ to denote the k'th derivative of the function. Now evaluate each of the above derivatives at $x = 0$. Then

$$\begin{aligned}
f'(0) &= a_1, & \Rightarrow a_1 &= f'(0) \\
f''(0) &= 2a_2, & \Rightarrow a_2 &= \tfrac{f''(0)}{2} \\
f'''(0) &= 2\cdot 3a_3, & \Rightarrow a_3 &= \tfrac{f'''(0)}{2\cdot 3} \\
f^{(k)}(0) &= k!a_k, & \Rightarrow a_k &= \tfrac{f^{(k)}(0)}{k!}
\end{aligned}$$

This gives us a recipe for calculating all coefficients a_k. This means that if we can compute all the derivatives of the function $f(x)$, then we know the coefficients of the Taylor series as well. Because we have evaluated all the coefficients by the substitution $x = 0$, we say that the resulting power series is the *Taylor series of the function about* $x = 0$.

10.6.1 Taylor series for the exponential function, e^x

Consider the function $f(x) = e^x$. All the derivatives of this function are equal to e^x. In particular,

$$f^{(k)}(x) = e^x \quad \Rightarrow \quad f^{(k)}(0) = 1.$$

So that the coefficients of the Taylor series are

$$a_k = \frac{f^{(k)}(0)}{k!} = \frac{1}{k!}.$$

Therefore the Taylor series for e^x about $x = 0$ is

$$a_0+a_1x+a_2x^2+a_3x^3+\ldots+a_kx^k+\ldots = 1+x+\frac{x^2}{2}+\frac{x^3}{6}+\frac{x^4}{24}+\ldots+\frac{x^k}{k!}+\ldots = \sum_{k=0}^{\infty}\frac{x^k}{k!}$$

This is a very interesting series. We state here without proof that this series converges for all values of x. Further, the function defined by the series is in fact equal to e^x that is,

$$\boxed{e^x = 1 + x + \frac{x^2}{2} + \frac{x^3}{6} + \ldots = \sum_{k=0}^{\infty}\frac{x^k}{k!}}$$

The implication is that the function e^x is completely determined (for all x values) by its behaviour (i.e. derivatives of all orders) at $x = 0$. In other words, the value of the function at $x = 1,000,000$ is determined by the behaviour of the function around $x = 0$. This means that e^x is a very special function with superior "predictable features". If a function $f(x)$ agrees with its Taylor polynomial on a region $(-a, a)$, as was the case here, we say that f is **analytic** on this region. It is known that e^x is analytic for all x.

We can use the results of this example to establish the fact that the exponential function grows "faster" than any power function x^n. That is the same as saying that the ratio of e^x to x^n (for any power n) increases with x. We leave this as an exercise for the reader.

We can also easily obtain a Taylor series expansion for functions related to e^x, without assembling the derivatives. We start with the result that

$$e^u = 1 + u + \frac{u^2}{2} + \frac{u^3}{6} + \ldots = \sum_{k=0}^{\infty} \frac{u^k}{k!}$$

Then, for example, the substitution $u = x^2$ leads to

$$e^{x^2} = 1 + x^2 + \frac{(x^2)^2}{2} + \frac{(x^2)^3}{6} + \ldots = \sum_{k=0}^{\infty} \frac{(x^2)^k}{k!}$$

10.6.2 Taylor series of trigonometric functions

In this example we determine the Taylor series for the sine function. The function and its derivatives are

$$f(x) = \sin x, f'(x) = \cos x, f''(x) = -\sin x, f'''(x) = -\cos x, f^{(4)}(x) = \sin x, \ldots$$

After this, the cycle repeats. This means that

$$f(0) = 0, f'(0) = 1, f''(0) = 0, f'''(0) = -1, \ldots$$

and so on in a cyclic fashion. In other words,

$$a_0 = 0, a_1 = 1, a_2 = 0, a_3 = -\frac{1}{3!}, a_4 = 0, a_5 = \frac{1}{5!}, \ldots$$

Thus,

$$\boxed{\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \ldots = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{(2n+1)!}.}$$

We state here without proof that the function $\sin(x)$ is **analytic**, so that the expansion converges to the function for all x.

It is instructive to demonstrate how successive terms in a Taylor series expansion lead to approximations that improve. *Doing this kind of thing will be the subject of the last computer laboratory exercise in this course.*

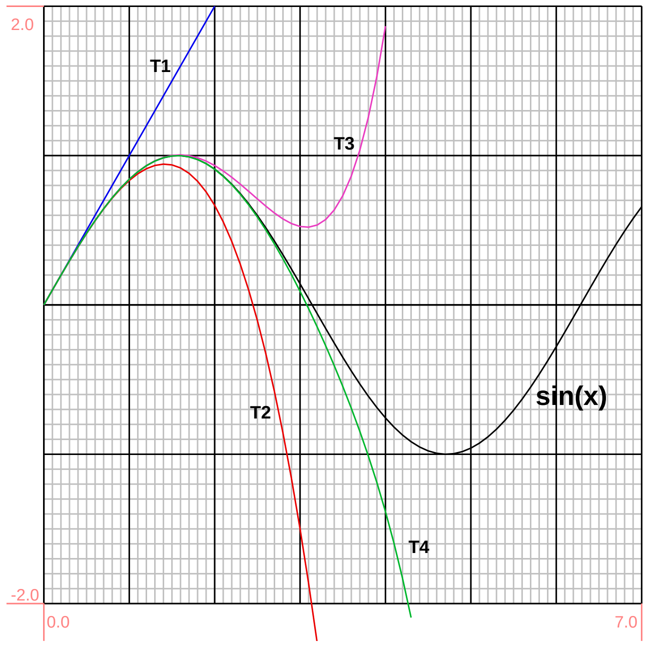

Figure 10.5. *An approximation of the function $y = \sin(x)$ by successive Taylor polynomials, T_1, T_2, T_3, T_4. The higher Taylor polynomials do a better job of approximating the function on a larger interval about $x = 0$.*

Here we demonstrate this idea with the expansion for the function $\sin(x)$ that we just obtained. To see this, consider the sequence of polynomials

$$\begin{aligned} T_1(x) &= x, \\ T_2(x) &= x - \frac{x^3}{3!}, \\ T_3(x) &= x - \frac{x^3}{3!} + \frac{x^5}{5!}, \\ T_4(x) &= x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!}. \end{aligned}$$

Then these polynomials provide a better and better approximation to the function $\sin(x)$ close to $x = 0$. The first of these is just a linear (or tangent line) approximation that we had studied long ago. The second improves this with a quadratic approximation, etc. Figure 10.5 illustrates how the first few Taylor polynomials approximate the function $\sin(x)$ near $x = 0$. Observe that as we keep more terms, n, in the polynomial $T_n(x)$, the approximating curve "hugs" the graph of $\sin(x)$ over a longer and longer range. The student will be asked to use the spreadsheet, together with some calculations as done in this section, to produce a composite graph similar to Fig. 10.5 for some other function.

Example 10.5 (The error in successive approximations) For a given value of x close to the base point (at $x = 0$), the **error** in the approximation between the polynomials and the function is the vertical distance between the graphs of the polynomial and the function $\sin(x)$ (shown in black). For example, at $x = 2$ radians $\sin(2) = 0.9093$ (as found on a scientific calculator). The approximations are: $T_1(2) = 2$, which is very inaccurate, $T_2(2) = 2 - 2^3/3! \approx 0.667$ which is too small, $T_3(2) \approx 0.9333$ that is much closer and $T_4(2) \approx .9079$ that is closer still. In general, we can approximate the size of the error using the next term that would occur in the polynomial if we kept a higher order expansion. The details of estimating such errors is omitted from our discussion.

We also note that all polynomials that approximate $\sin(x)$ contain only odd powers of x. This stems from the fact that $\sin(x)$ is an odd function, i.e. its graph is symmetric to rotation about the origin, a concept we discussed in an earlier term.

The Taylor series for $\cos(x)$ could be found by a similar sequence of steps. But in this case, this is unnecessary: We already know the expansion for $\sin(x)$, so we can find the Taylor series for $\cos(x)$ by simple differentiation term by term. (Since $\sin(x)$ is analytic, this is permitted for all x.) We leave as an exercise for the reader to show that

$$\boxed{\cos(x) = 1 - \frac{x^2}{2} + \frac{x^4}{4!} - \frac{x^6}{6!} + \ldots = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n}}{(2n)!}.}$$

Since $\cos(x)$ has symmetry properties of an even function, we find that its Taylor series is composed of even powers of x only.

10.7 Application of Taylor series

In this section we illustrate some of the applications of Taylor series to problems that may be difficult to solve using other conventional methods. Some functions do not have an antiderivative that can be expressed in terms of other simple functions. Integrating these functions can be a problem, as we cannot use the Fundamental Theorem of Calculus specifies. In some cases, we can approximate the value of the definite integral using a Taylor series. We show this in Section 10.7.1.

Another application of Taylor series is to compute an approximate solution to a differential equation. We provide one example of that sort in Section 10.7.2 and another in Appendix 11.11.

10.7.1 Example 1: using a Taylor series to evaluate an integral

Evaluate the definite integral

$$\int_0^1 \sin(x^2)\, dx.$$

A simple substitution (e.g. $u = x^2$) will not work here, and we cannot find an antiderivative. Here is how we might approach the problem using Taylor series: We know that the series expansion for $\sin(t)$ is

$$\sin t = t - \frac{t^3}{3!} + \frac{t^5}{5!} - \frac{t^7}{7!} + \dots$$

Substituting $t = x^2$, we have

$$\sin(x^2) = x^2 - \frac{x^6}{3!} + \frac{x^{10}}{5!} - \frac{x^{14}}{7!} + \dots$$

In spite of the fact that we cannot antidifferentiate the function, we can antidifferentiate the Taylor series, just as we would a polynomial:

$$\begin{aligned}\int_0^1 \sin(x^2)\,dx &= \int_0^1 (x^2 - \frac{x^6}{3!} + \frac{x^{10}}{5!} - \frac{x^{14}}{7!} + \dots)\,dx \\ &= \left(\frac{x^3}{3} - \frac{x^7}{7\cdot 3!} + \frac{x^{11}}{11\cdot 5!} - \frac{x^{15}}{15\cdot 7!} + \dots\right)\Big|_0^1 \\ &= \frac{1}{3} - \frac{1}{7\cdot 3!} + \frac{1}{11\cdot 5!} - \frac{1}{15\cdot 7!} + \dots\end{aligned}$$

This is an alternating series so we know that it converges. If we add up the first four terms, the pattern becomes clear: the series converges to 0.31026.

10.7.2 Example 2: Series solution of a differential equation

We are already familiar with the differential equation and initial condition that describes unlimited exponential growth.

$$\frac{dy}{dx} = y, \qquad y(0) = 1.$$

Indeed, from previous work, we know that the solution of this differential equation and initial condition is $y(x) = e^x$, but we will pretend that we do not know this fact in illustrating the usefulness of Taylor series. In some cases, where separation of variables does not work, this option would have great practical value.

Let us express the "unknown" solution to the differential equation as

$$y = a_0 + a_1 x + a_2 x^2 + a_3 x^3 + a_4 x^4 + \dots$$

Our task is to determine values for the coefficients a_i

Since this function satisfies the condition $y(0) = 1$, we must have $y(0) = a_0 = 1$. Differentiating this power series leads to

$$\frac{dy}{dx} = a_1 + 2a_2 x + 3a_3 x^2 + 4a_4 x^3 + \dots$$

But according to the differential equation, $\frac{dy}{dx} = y$. Thus, it must be true that the two Taylor series match, i.e.

$$a_0 + a_1x + a_2x^2 + a_3x^3 + a_4x^4 + \ldots = a_1 + 2a_2x + 3a_3x^2 + 4a_4x^3 + \ldots$$

This equality hold for all values of x. This can only happen if the coefficients of like terms are the same, i.e. if the constant terms on either side of the equation are equal, if the terms of the form Cx^2 on either side are equal, and so on for all powers of x. Equating coefficients, we obtain:

$$\begin{aligned}
&a_0 = a_1 = 1, && \Rightarrow a_1 = 1, \\
&a_1 = 2a_2, && \Rightarrow a_2 = \tfrac{a_1}{2} = \tfrac{1}{2}, \\
&a_2 = 3a_3, && \Rightarrow a_3 = \tfrac{a_2}{3} = \tfrac{1}{2\cdot 3}, \\
&a_3 = 4a_4, && \Rightarrow a_4 = \tfrac{a_3}{4} = \tfrac{1}{2\cdot 3\cdot 4}, \\
&a_{n-1} = na_n, && \Rightarrow a_n = \tfrac{a_{n-1}}{n} = \tfrac{1}{1\cdot 2\cdot 3\ldots n} = \tfrac{1}{n!}.
\end{aligned}$$

This means that

$$y = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \ldots + \frac{x^n}{n!} + \ldots = e^x,$$

which, as we have seen, is the expansion for the exponential function. This agrees with the solution we have been expecting. In the example here shown, we would hardly need to use series to arrive at the right conclusion, but in the next example, we would not find it as easy to discover the solution by other techniques discussed previously.

We provide an example of a more complicated differential equation and its series solution in Appendix 11.11.

10.8 Summary

The main points of this chapter can be summarized as follows:

1. We reviewed the definition of an improper integral

$$\int_a^\infty f(x)\, dx = \lim_{b\to\infty} \int_a^b f(x)\, dx.$$

2. We computed some examples of improper integrals and discussed their convergence or divergence. We recalled (from earlier chapters) that

$$I = \int_0^\infty e^{-rt}\, dt \quad \text{converges,}$$

whereas

$$I = \int_1^\infty \frac{1}{x}\, dx \quad \text{diverges.}$$

3. More generally, we showed that

$$\int_1^{\infty} \frac{1}{x^p}\, dx \qquad \text{converges if } p > 1, \quad \text{diverges if } p \le 1.$$

4. Using a comparison between integrals and series we showed that the harmonic series,

$$\sum_{k=1}^{\infty} \frac{1}{k} = 1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \ldots + \frac{1}{k} + \ldots \qquad \text{diverges.}$$

5. More generally, our results led to the conclusion that the "p" series,

$$\sum_{k=1}^{\infty} \frac{1}{k^p} \qquad \text{converges if } p > 1, \text{ diverges if } p \le 1.$$

6. We studied Taylor series and showed that some can be found using the formula for convergent geometric series. Two examples of Taylor series that were obtained in this way are

$$\ln(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \ldots \quad \text{for } |x| < 1$$

and

$$\arctan(x) = x - \frac{x^3}{3} + \frac{x^5}{5} - \frac{x^7}{7} + \ldots \quad \text{for } |x| < 1$$

7. In discussing Taylor series, we considered some of the following questions: (a) For what range of values of x can we expect the series to converges? (b) Suppose we approximate the function on the right by a finite number of terms on the left. How good is that approximation? Another interesting question is: (c) If we include more and more such terms, does that approximation get better and better? (i.e., does the series converge to the function?) (d) Is the convergence rate rapid? Some of these questions occupy the attention of career mathematicians, and are beyond the scope of this introductory calculus course.

8. More generally, we showed that the Taylor series for a function about $x = 0$,

$$f(x) = a_0 + a_1 x + a_2 x^2 + a_3 x^3 + \ldots = \sum_{k=0}^{\infty} a_k x^k.$$

can be found by computing the coefficients

$$a_k = \frac{f^{(k)}(0)}{k!}$$

9. We discussed some of the applications of Taylor series. We used Taylor series to approximate a function, to find an approximation for a definite integral of a function, and to solve a differential equation.

Chapter 11

Appendix

11.1 How to prove the formulae for sums of squares and cubes

Mathematicians are concerned with rigorously establishing formulae such as sums of squared (or cubed) integers. While it is not hard to see that these formulae "work" for a few cases, determining that they work in general requires more work. Here we provide a taste of how such careful arguments works. We give two examples. The first, based on *mathematical induction* provides a general method that could be used in many similar kinds of proofs. The second argument, also for purposes of illustration uses a "trick". Devising such tricks is not as straightforward, and depends to some degree on serendipity or experience with numbers.

Proof by induction (optional)

Here, we prove the formula for the sum of square integers,

$$\sum_{k=1}^{N} k^2 = \frac{N(N+1)(2N+1)}{6},$$

using a technique called *induction*. The idea of the method is to check that the formula works for one or two simple cases (e.g. the "sum" of just one or just two terms of the series), and then show that whenever it works for one case (summing up to N), it has to also work for the next case (summing up to $N+1$).

First, we verify that this formula works for a few test cases:

$N = 1$**:** If there is only one term, then clearly, by inspection,

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

The formula indicates that we should get

$$S = \frac{1(1+1)(2\cdot 1+1)}{6} = \frac{1(2)(3)}{6} = 1,$$

so this case agrees with the prediction.

$N = 2$**:**

$$\sum_{k=1}^{2} k^2 = 1^2 + 2^2 = 1 + 4 = 5.$$

The formula would then predict that

$$S = \frac{2(2+1)(2\cdot 2+1)}{6} = \frac{2(3)(5)}{6} = 5.$$

So far, elementary computation matches with the result predicted by the formula. Now we show that if this formula holds for any one case, e.g. for the sum of the first N squares, then it is also true for the next case, i.e. for the sum of $N+1$ squares. So we will *assume* that someone has checked that for some particular value of N it is true that

$$S_N = \sum_{k=1}^{N} k^2 = \frac{N(N+1)(2N+1)}{6}.$$

Now the sum of the first $N+1$ squares will be just a bit bigger: it will have one more term added to it:

$$S_{N+1} = \sum_{k=1}^{N+1} k^2 = \sum_{k=1}^{N} k^2 + (N+1)^2.$$

Thus

$$S_{N+1} = \frac{N(N+1)(2N+1)}{6} + (N+1)^2.$$

Combining terms, we get

$$S_{N+1} = (N+1)\left[\frac{N(2N+1)}{6} + (N+1)\right],$$

$$S_{N+1} = (N+1)\frac{2N^2+N+6N+6}{6} = (N+1)\frac{2N^2+7N+6}{6}.$$

Simplifying and factoring the last term leads to

$$S_{N+1} = (N+1)\frac{(2N+3)(N+2)}{6}.$$

We want to check that this still agrees with what the formula predicts. To make the notation simpler, we will let M stand for $N+1$. Then, expressing the result in terms of the quantity $M = N+1$ we get

$$S_M = \sum_{k=1}^{M} k^2 = (N+1)\frac{[2(N+1)+1][(N+1)+1]}{6} = M\frac{[2M+1][M+1]}{6}.$$

This is the same formula as we started with, only written in terms of M instead of N. Thus we have verified that the formula works. By *Mathematical Induction* we find that the result has been proved.

Another method using a trick[70]

There is another method for determining the sums $\sum_{k=1}^{n} k^2$ or $\sum_{k=1}^{n} k^3$. Write

$$(k+1)^3 - (k-1)^3 = 6k^2 + 2,$$

so

$$\sum_{k=1}^{n} \left((k+1)^3 - (k-1)^3\right) = \sum_{k=0}^{n} (6k^2 + 2).$$

But looking more carefully at the left hand side (LHS), we see that

$$\sum_{k=1}^{n} ((k+1)^3 - (k-1)^3) = 2^3 - 0^3 + 3^3 - 1^3 + 4^3 - 2^3 + 5^3 - 3^3 ... + (n+1)^3 - (n-1)^3.$$

most of the terms cancel, leaving only $-1 + n^3 + (n+1)^3$, so this means that

$$-1 + n^3 + (n+1)^3 = 6\sum_{k=1}^{n} k^2 + \sum_{k=1}^{n} 2,$$

so

$$\sum_{k=1}^{n} k^2 = (-1 + n^3 + (n+1)^3 - 2n)/6 = (2n^3 + 3n^2 + n)/6.$$

Similarly, the formula for $\sum_{k=1}^{n} k^3$, can be obtained by starting with

$$(k+1)^4 - (k-1)^4 = 4k^3 + 4k.$$

11.2 Riemann Sums: Extensions and other examples

We take up some issues here that were not yet considered in the context of our examples of Riemann sums in Chapter 2 . First, we consider an arbitrary interval $a \le x \le b$. Then we comment on other ways of constructing the rectangular strip approximation (that eventually lead to the same limit when the true area is computed.)

11.2.1 A general interval: $a \le x \le b$

Example 2: (Lu Fan)

Find the area under the graph of the function

$$y = f(x) = x^2 + 2x + 1 \qquad a \le x \le b.$$

[70] I want to thank Robert Israel for contributing this material

Here the interval is $a \le x \le b$. Let us leave the values of a, b general for a moment, and consider how the calculation is set up in this case. Then we have

$$\begin{aligned}
&\text{length of interval} = b - a\\
&\text{number of segments} = N\\
&\text{width of rectangular strips} = \Delta x = \frac{b-a}{N}\\
&\text{the } k\text{'th } x \text{ value} = x_k = a + k\frac{(b-a)}{N}\\
&\text{height of } k\text{'th rectangular strip} = f(x_k) = x_k^2 + 2x_k + 1
\end{aligned}$$

Combining the last two steps, the height of rectangle k is:

$$f(x_k) = \left(a + \frac{k(b-a)}{N}\right)^2 + 2\left(a + \frac{k(b-a)}{N}\right) + 1$$

and its area is

$$a_k = f(x_k) \times \Delta x = f(x_k) \times \left(\frac{b-a}{N}\right).$$

We use the last two equations to express a_k in terms of k (and the quantities a, b, N), then sum over k as before ($A = \Sigma A_k$). Some algebra is needed to simplify the sums so that summation formulae can be applied. The details are left as an exercise for the reader (see homework problems). Evaluating the limit $N \to \infty$, we finally obtain

$$A = \lim_{N\to\infty} \sum_{k=1}^{N} a_k = (a+1)^2(b-a) + (a+1)(b-a)^2 + \frac{(b-a)^3}{3}.$$

as the area under the function $f(x) = x^2 + 2x + 1$, over the interval $a \le x \le b$. Observe that the solution depends on a, and b. (The endpoints of the interval influence the total area under the curve.) For example, if the given interval happens to be $2 \le x \le 4$. then, substituting $a = 2, b = 4$ into the above result for A, leads to

$$A = (2+1)^2(4-2) + (2+1)(4-2)^2 + \frac{4-2}{3} = 18 + 12 + \frac{2}{3} = \frac{32}{3}$$

In the next chapter, we will show that the tools of integration will lead to the same conclusion.

11.2.2 Using left (rather than right) endpoints

So far, we used the right endpoint of each rectangular strip to assign its height using the given function (see Figs. 2.2, 2.3, 2.4). Restated, we "glued" the top right corner of the rectangle to the graph of the function. This is the so called **right endpoint approximation**. We can just as well use the left corners of the rectangles to assign their heights (**left endpoint approximation**). A comparison of these for the function $y = f(x) = x^2$ is shown in Figs. 11.1 and 11.2. In the case of the left endpoint approximation, we evaluate

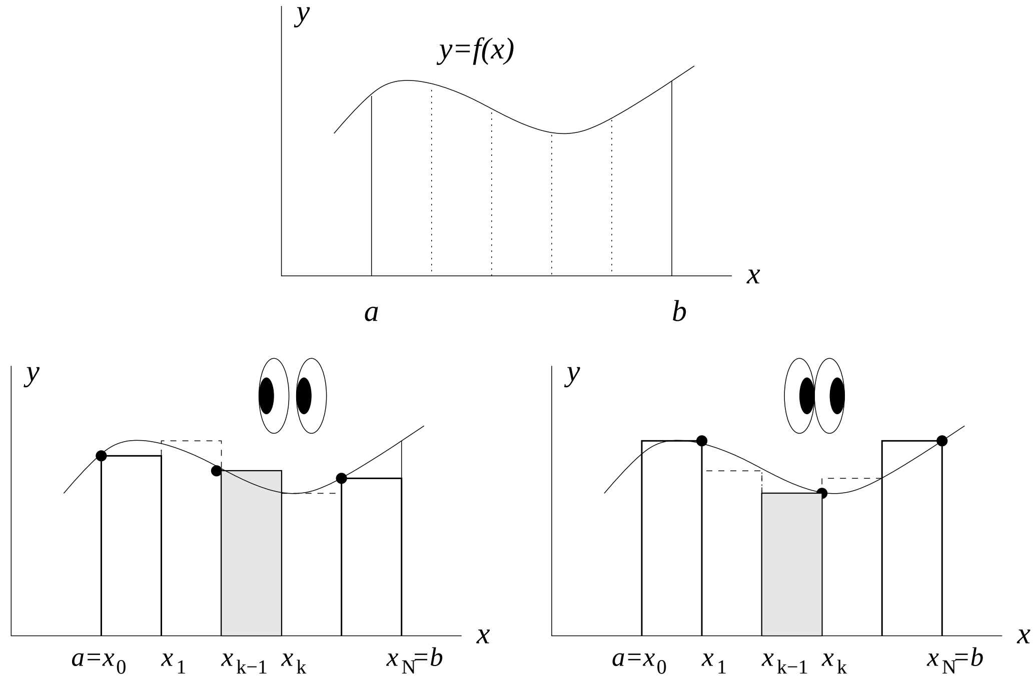

Figure 11.1. *The area under the curve $y = f(x)$ over an interval $a \leq x \leq b$ could be computed by using either a left or right endpoint approximation. That is, the heights of the rectangles are adjusted to match the function of interest either on the right or on their left corner. Here we compare the two approaches. Usually both lead to the same result once a limit is computer to arrive at the "true " area.*

the heights of the rectangles starting at x_0 (instead of x_1, and ending at x_{N-1} (instead of x_N). There are still N rectangles. To compare, sum of areas of the rectangles in the left versus the right endpoint approximation is

$$\text{Right endpoints:} \quad A_{N \text{ strips}} = \sum_{k=1}^{N} f(x_k)\Delta x.$$

$$\text{Left endpoints:} \quad A_{N \text{ strips}} = \sum_{k=0}^{N-1} f(x_k)\Delta x.$$

Details of one such computation is given in the box.

Example of left endpoint calculation

We here look again at a simple example, using the quadratic function,

$$f(x) = x^2, \quad 0 \le x \le 1,$$

We now compare the right and left endpoint approximation. These are shown in panels of Figure 11.2. Note that

$$\Delta x = \frac{1}{N}, \quad x_k = \frac{k}{N},$$

The area of the k'th rectangle is

$$a_k = f(x_k) \times \Delta x = (k/N)^2\,(1/N)\,,$$

but now the sum starts at $k = 0$ so

$$A_{N \text{ strips}} = \sum_{k=0}^{N-1} f(x_k)\Delta x = \sum_{k=0}^{N-1} \left(\frac{k}{N}\right)^2 \left(\frac{1}{N}\right) = \left(\frac{1}{N^3}\right) \sum_{k=0}^{N-1} k^2.$$

The first rectangle corresponds to $k = 0$ in the left endpoint approximation (rather than $k = 1$ in the right endpoint approximation). But the $k = 0$ rectangle makes no contribution (as its area is zero in this example) and we have one less rectangle at the right endpoint of the interval, since the N'th rectangle is $k = N - 1$. Then the sum is

$$A_{N \text{ strips}} = \left(\frac{1}{N^3}\right) \frac{(2(N-1)+1)(N-1)(N)}{6} = \frac{(2N-1)(N-1)}{6N^2}.$$

The area, obtained by taking a limit for $N \to \infty$ is

$$A = \lim_{N\to\infty} A_{N \text{ strips}} = \lim_{N\to\infty} \frac{(2N-1)(N-1)}{6N^2} = \frac{2}{6} = \frac{1}{3}.$$

We see that, after computing the limit, the result for the "true area" under the curve is exactly the same as we found earlier in this chapter using the right endpoint approximation.

11.3 Physical interpretation of the center of mass

We defined the idea of a center of mass in Chapter 5. The center of mass has a physical interpretation for a real mass distribution. Loosely speaking, it is the position at which the mass "balances" without rotating to the left or right. In physics, we say that there is no net torque. The analogy with children sitting on a teeter-totter is relevant: many children may sit along the length of the frame of a teeter totter, but if they distribute themselves in a way that the center of mass is at the fulcrum of the teeter totter, they will remain precariously balanced (until one of them fidgets or gets off!). Notice that both the mass and the position of each child is important - a light child sitting on the very edge of the teeter totter can balance a heavier child sitting closer to the fulcrum (middle). The center of mass need not be the same as the median position. As we have see, the median is a position that

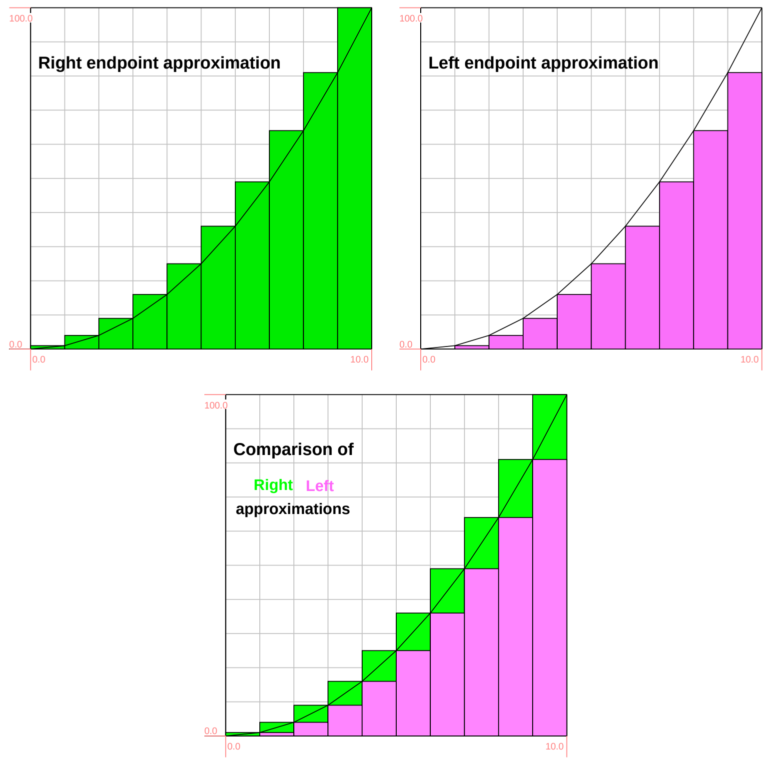

Figure 11.2. *Rectangles with left or right corners on the graph of $y = x^2$ are compared in this picture. The approximation shown in pink is "missing" the largest rectangle shown in green. However, in the limit as the number of rectangles, $N \to \infty$, the true area obtained is the same.*

subdivides the distribution into two equal masses (or, more generally, produces equal sized areas under the graph of the density function.) The center of mass assigns a greater weight to parts of the distribution that are "far away" in the same sense. (However, for symmetric distributions, the median and the mean are the same.)

In physics, we speak of the "moment of mass" of a distribution about a point. This quantity is related to the tendency of the mass to contribute a torque, i.e. to make the object rotate. Suppose we are interested in a particular point of reference x. In a discrete mass distribution, for example, the moment of mass of each of the beads relative to point x is given by the product of the mass and its distance away from the point - as with the teeter totter, beads farther away will contribute more torque than beads closer to point x, and heavier beads (i.e. greater mass) will contribute more torque than lighter beads. For example, mass 1 contributes an amount $m_1(x - x_1)$ to the total moment of mass of the distribution about the point x. Altogether the moment of mass of the distribution about the point x is defined as

$$M_1(x) = \sum_{i=1}^{n} m_i \ (x - x_i).$$

The center of mass is a special point $\bar{x}$ such that the moment of mass about that point is zero. (Loosely speaking the tendency to rotate to the left or the right are the same: thus the distribution would be balanced if it "rested on that point".)

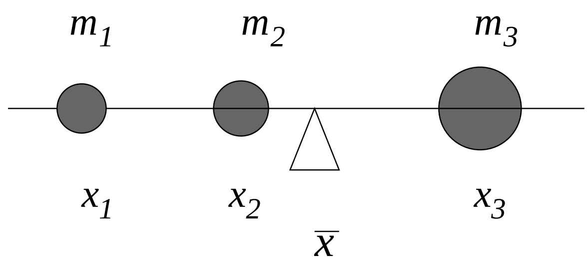

Figure 11.3. *A discrete set of masses m_1, m_2, m_3 is distributed at positions x_1, x_2, x_3. The center of mass of the distribution is the position at which the given mass distribution would balance, here represented by the white triangle.*

Thus, we identify the center of mass as the point at which

$$M_1(\bar{x}) = 0,$$

or

$$\sum_{i=1}^{n} m_i(\bar{x} - x_i) = 0.$$

Now expanding the sum, we rewrite the above as

$$\left(\sum_{i=1}^{n} m_i \bar{x}\right) - \left(\sum_{i=1}^{n} m_i x_i\right) = 0,$$

$$\bar{x} \sum_{i=1}^{n} m_i - \left(\sum_{i=1}^{n} m_i x_i\right) = 0.$$

But we already know that the first summation above is just the total mass, so that

$$\bar{x} M - \left(\sum_{i=1}^{n} m_i x_i\right) = 0,$$

so, taking the second term to the other side and dividing by M leads to

$$\bar{x} = \frac{1}{M} \sum_{i=1}^{n} m_i x_i.$$

We have recovered precisely the definition of the center of mass or "average x coordinate".

11.4 The shell method for computing volumes

In Chapter 5, we used dissection into small disks to compute the volume of solids of revolution. Here we show use an alternative dissection into shells.

11.4.1 Example: Volume of a cone using the shell method

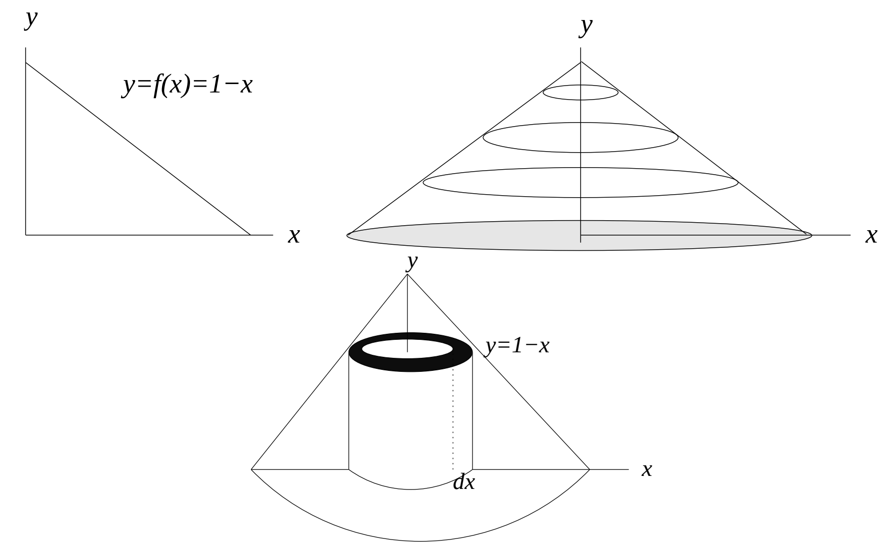

Figure 11.4. *Top: The curve that generates the cone (left) and the shape of the cone (right). Bottom: the cone showing one of the series of shells that are used in this example to calculate its volume.*

We use the shell method[71] to find the volume of the cone formed by rotating the curve

$$y = 1 - x$$

about the y axis.

Solution

We show the cone and its generating curve in Figure 11.4, together with a representative shell used in the calculation of total volume. The volume of a cylindrical shell of radius r, height h and thickness τ is

$$V_{\text{shell}} = 2\pi r h \tau.$$

We will place these shells one inside the other so that their radii are parallel to the x axis (so $r = x$). The heights of the shells are determined by their y value (i.e. $h = y = 1 - x =$

[71] Note to the instructor: This material may be skipped in the interest of time. It presents an alternative to the disk method, but there may not be enough time to cover this in detail.

$1-r$). For the tallest shell $r=0$, and for the flattest shell $r=1$. The thickness of the shell is Δr. Therefore, the volume of one shell is

$$V_{\text{shell}} = 2\pi r(1-r)\,\Delta r.$$

The volume of the object is obtained by summing up these shell volumes. In the limit, as $\Delta r \to dr$ gets infinitesimally small, we recognize this as a process of integration. We integrate over $0 \le r \le 1$, to obtain:

$$V = 2\pi \int_0^1 r(1-r)\,dr = 2\pi \int_0^1 (r-r^2)\,dr.$$

We find that

$$V = 2\pi \left(\frac{r^2}{2} - \frac{r^3}{3} \right)\Bigg|_0^1 = 2\pi \left(\frac{1}{2} - \frac{1}{3} \right) = \frac{\pi}{3}.$$

11.5 More techniques of integration

11.5.1 Secants and other “hard integrals”

In a previous section, we encountered the integral

$$I = \int \sec^3(x)\,dx.$$

This integral can be simplified to some extent by integration by parts as follows: Let $u = \sec(x)$, $dv = \sec^2(x)\,dx$. Then $du = \sec(x)\tan(x)dx$ while $v = \int \sec^2(x)\,dx = \tan(x)$. The integral can be transformed to

$$I = \sec(x)\tan(x) - \int \sec(x)\tan^2(x)\,dx.$$

The latter can be rewritten:

$$I_1 = \int \sec(x)\tan^2(x)\,dx = \int \sec(x)(\sec^2(x)-1).$$

where we have use a trigonometric identity for $\tan^2(x)$. Then

$$I = \sec(x)\tan(x) - \int \sec^3(x)\,dx + \int \sec(x)\,dx = \sec(x)\tan(x) - I + \int \sec(x)\,dx$$

so (taking both I’s to the left hand side, and dividing by 2)

$$I = \frac{1}{2}\left(\sec(x)\tan(x) + \int \sec(x)\,dx \right).$$

We are now in need of an antiderivative for $\sec(x)$. No “obvious substitution” or further integration by parts helps here, but it can be checked by differentiation that

$$\int \sec(x)\,dx = \ln|\sec(x)+\tan(x)| + C$$

Then the final result is

$$I = \frac{1}{2}\left(\sec(x)\tan(x) + \ln|\sec(x)+\tan(x)|\right) + C$$

11.5.2 A special case of integration by partial fractions

Evaluate this integral[72]:

$$\int_1^2 \frac{7x+4}{6x^2+7x+2}\,dx$$

This integral involves a rational function (that is, a ratio of two polynomials). The denominator is a degree 2 polynomial function that has two roots and that can be factored easily; the numerator is a degree 1 polynomial function. In this case, we can use the following strategy. First, factor the denominator:

$$6x^2+7x+2=(2x+1)(3x+2)$$

Assign A and B in the following way:

$$\frac{A}{2x+1}+\frac{B}{3x+2}=\frac{7x+4}{(2x+1)(3x+2)}=\frac{7x+4}{6x^2+7x+2}$$

(Remember, this is how we define A and B.)

Next, find the common denominator and rewrite it as a single fraction in terms of A and B.

$$\frac{A}{2x+1}+\frac{B}{3x+2}=\frac{3Ax+2A+2Bx+B}{(2x+1)(3x+2)}$$

Group like terms in the numerator, and note that this has to match the original fraction, so:

$$\frac{3Ax+2A+2Bx+B}{(2x+1)(3x+2)}=\frac{(3A+2B)x+(2A+B)}{(2x+1)(3x+2)}=\frac{7x+4}{(2x+1)(3x+2)}$$

The above equation should hold true for all x values; therefore:

$$3A+2B=7, \quad 2A+B=4$$

Solving the system of equations leads to $A=1, B=2$. Using this result, we rewrite the original expression in the form:

$$\frac{7x+4}{6x^2+7x+2}=\frac{7x+4}{(2x+1)(3x+2)}=\frac{A}{2x+1}+\frac{B}{3x+2}=\frac{1}{2x+1}+\frac{2}{3x+2}$$

Now we are ready to rewrite the integral:

$$I=\int_1^2 \frac{7x+4}{6x^2+7x+2}\,dx=\int_1^2\left(\frac{1}{2x+1}+\frac{2}{3x+2}\right)\,dx$$

Simplify:

$$I=\int_1^2 \frac{1}{2x+1}\,dx+2\int_1^2\frac{1}{3x+2}\,dx$$

Now the integral becomes a simple natural log integral that follows the pattern of Eqn. 6.1. Simplify:

$$I=\frac{1}{2}\ln|2x+1|\Big|_1^2+\frac{2}{3}\ln|3x+2|\Big|_1^2.$$

[72]This section was contributed by Lu Fan

Simplify further:

$$I = \frac{1}{2}(\ln 5 - \ln 3) + \frac{2}{3}(\ln 8 - \ln 5) = -\frac{1}{6}\ln 5 - \frac{1}{2}\ln 3 + \frac{2}{3}\ln 8.$$

This method can be used to solve *any* integral that contain a fraction with a degree 1 polynomial in the numerator and a degree 2 polynomial (that has two roots) in the denominator.

11.6 Analysis of data: a student grade distribution

We study the distribution of student grades on a test written by 76 students and graded out of a maximum of 50 points.

11.6.1 Defining an average grade

Let N be the size of the class,and y_k the grade of student k. Here k is the number of the student from 1 to N, and y_k takes any value between 0 and 50 points). Then the average grade $\bar{Y}$ is computed by adding up the scores of all students and dividing by the number of students as follows:

$$\bar{Y} = \frac{1}{N}\sum_{k=1}^{N} y_k.$$

For example, for a class of 76 students, we would have the sum

$$\bar{Y} = \frac{1}{76}\sum_{k=1}^{76} y_k.$$

11.6.2 Fraction of students that scored a given grade

Suppose that the number of students who got the grade x_i is p_i. If the class consists of a total of N students, then it follows that

$$N = \sum_{i=1}^{10} p_i.$$

This is just saying that the sum of the number of students in every one of the categories has to add up to the total class size. The fraction of the class that scored grade x_i is

$$\frac{p_i}{N}.$$

(Dividing by N has normalized the distribution. The value p_i/N is the empirical probability of getting grade x_i.) The mean or average grade is:

$$\bar{X} = \frac{1}{N}\sum_{i=0}^{50} x_i\, p_i.$$

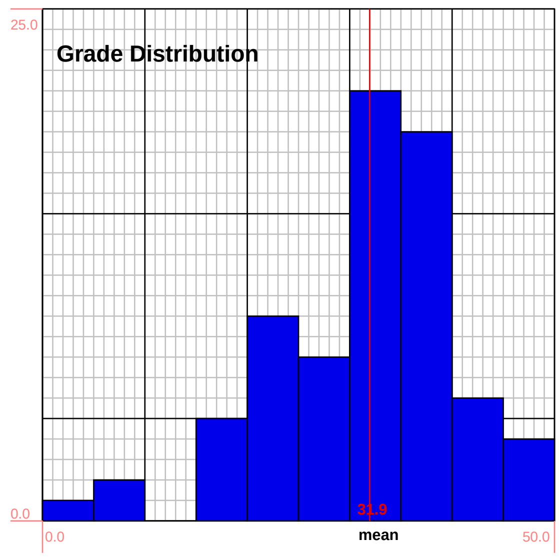

Figure 11.5. *Distributions of grades on a test with 50 point maximum. There were a total of 76 students writing the test. The mean grade 31.9 is shown.*

11.6.3 Frequency distribution

It is difficult to visualize all the data if we list all the grades obtained. We "lump together" scores into various categories (or "bins") and create a distribution. For example, test scores might be divided into ranges of bins in increments of 5 points: (1-5, 6-10, 11-15, etc). We could represent grades in each bin by some value up to a specified level of accuracy. For example, grades in the the range 16-20 can be described by the score18 up to an accuracy of ± 2. This is what we have done in Table 11.1.

We will now reinterpret our notation somewhat. We will refer to $\tilde{x}_i$ as the score and p_i the number of students whose test score fell within the range represented by $\tilde{x}_i \pm$accuracy. (The notation $\tilde{x}_i$ is meant to remind us that we are approximating the grade value.) For example, consider 10 "bins" or grade categories. In that case, the index i takes on values $i = 1, 2, \ldots 10$. The, e.g., $\tilde{x}_4$ represents all grades in the fourth "bin", i.e. grades between 16-20. A plot of p_i against $\tilde{x}_i$ is called a *frequency distribution*. The bar graph shown in Figure 11.5 represents this distribution. Table 11.1 shows the data that produced that bar graph.

11.6.4 Average/mean of the distribution

The frequency distribution can also be used to compute an average value: each (approximate) grade value $\tilde{x}_i$ is achieved by p_i students, which is a fraction (p_i/N) of the whole class. When we form the multiple $(p_i/N)\tilde{x}_i$, we assign a "weight" to each of the cate-

i	grade $\tilde{x}_i$	number p_i	$\sum p_i$	$\sum \tilde{x}_i p_i$	$(1/N)\sum \tilde{x}_i p_i$
0	0	0	0	0.0	0.0
1	3±2	1	1	3.0	0.0395
2	8±2	2	3	19	0.25
3	13±2	0	3	19	0.25
4	18±2	5	8	109	1.4342
5	23±2	10	18	339	4.4605
6	28±2	8	26	563	7.4079
7	33±2	21	47	1256	16.5263
8	38±2	19	66	1978	26.0263
9	43±2	6	72	2236	29.4211
10	48±2	4	76	2428	31.9474

Table 11.1. *Distribution of grades (out of* 50*) for a class of* 76 *students. The mean grade for this class is* 31.9474*.*

gories according to the proportion of the class that was in that category. (The terminology *weighted average* is sometimes used.)

We define the *mean* or *average* grade in the distribution by

$$\bar{x} = \sum_{i=1}^{M} \tilde{x}_i \frac{p_i}{N} \ . \tag{11.1}$$

Where M is the number of bins. An equivalent way of expressing the mean (average) is:

$$\bar{x} = \frac{1}{N}\sum_{i=1}^{M} \tilde{x}_i p_i = \frac{\sum_{i=1}^{M} \tilde{x}_i p_i}{\sum_{i=1}^{M} p_i}. \tag{11.2}$$

The sum in the denominator of this last fraction is simply the total class size.

In Table 11.1, we show steps in the calculation of the mean grade for the class. This calculation is easily handled on the same spreadsheet that recorded the frequency of grades and that was used to plot the bar graph of that distribution. Equations 11.1 and 11.2 are saying the same thing. We will see the second of these again in the context of a more general probability distribution in Chapter 8.

11.6.5 Cumulative function

We can calculate a "running total" as shown on Figure 11.6, where we plot for each grade category, the total number of students whose grade was in the given range.

We define the *cumulative function*, F_i to be:

$$\boxed{F_i = \sum_{k=1}^{i} p_k \ .}$$

Then F_i is the number of students whose grade x_k was between x_1 and x_i ($x_1 \leq x_k \leq x_i$). Of course, when we add up all the way to the last category, we arrive at the total number of students in the class (assuming each student wrote the test and received a grade). Thus

$$F_m = \sum_{k=1}^{M} p_k = N,$$

Where as before, M stands for the number of "bins" used to represent the grade distribution. (Note that each student has been counted in one of the categories corresponding to the grade he or she achieved.) Another way of saying the same thing is that

$$\sum_{k=1}^{m} \frac{p_k}{N} = 1.$$

In Figure 11.6 we show the cumulative function, i.e. we plot $\tilde{x}_i$ vs F_i. Note that this graph is a **step function**. That is,the function takes on a set of discrete values with jumps at every 5th integer[73].

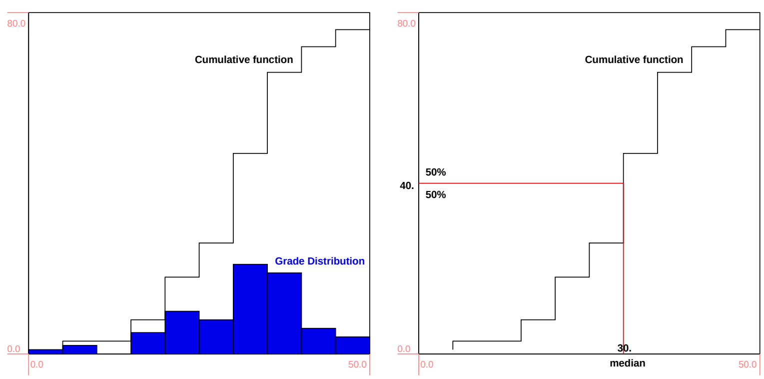

Figure 11.6. *Top: The same grade distribution as in Figure 11.5, but showing the cumulative function. The grid has been removed for easier visualization of that step function. Bottom:The cumulative function is used to determine an approximate median grade.*

[73]Note: ideally, this graph should be discontinuous, with horizontal segments only. The vertical"jumps" cannot correspond to values of a function. However the spreadsheet tool used to plot this function does not currently allow this graphing option.

11.6.6 The median

We can use the cumulative function and its features to come up with new ways of summarizing the distribution or comparing the performance of two sections. Suppose we subdivide a given class into exactly two equal groups based on performance on the test. Then there would be some grade that was achieved or surpassed by the top half of the class only; the rest of the students (i.e. the other half of the class) got scores below that level. We call that grade the *median* of the distribution.

To find the median grade using a cumulative function, we must ask what grade level corresponds to a cumulative 1/2 of the class, i.e. to $N/2$ students. To determine that level, we draw a horizontal line corresponding to $N/2$. As shown in Figure 11.6, because the function f is discontinuous, we only have an approximate median of 30. We observe that the median is not in general equal to the mean computed earlier.

11.7 Factorial notation

Let n be an integer, $n \geq 0$. Then $n!$, called "n factorial", is defined as the following product of integers:

$$\boxed{n! = n(n-1)(n-2)\dots(2)(1)}$$

Example

$$1! = 1$$

$$2! = 2 \cdot 1 = 2$$

$$3! = 3 \cdot 2 \cdot 1 = 6$$

$$4! = 4 \cdot 3 \cdot 2 \cdot 1 = 24$$

$$5! = 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1 = 120$$

We also define

$$\boxed{0! = 1}$$

11.8 Appendix: Permutations and combinations

11.8.1 Permutations

A *permutation* is a way of arranging objects, where the order of appearance of the objects is important.

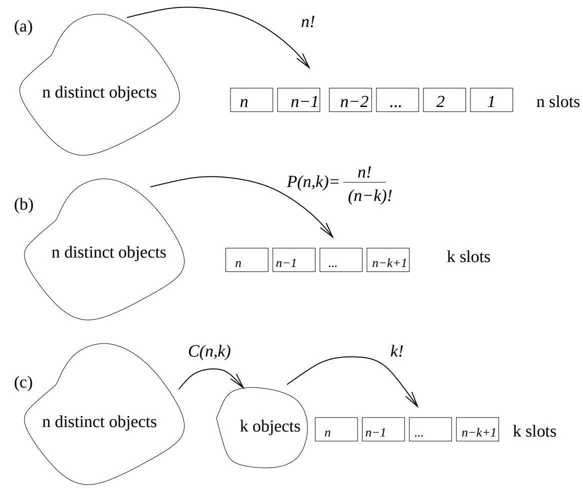

Figure 11.7. *This diagram illustrates the meanings of permutations and combinations. (a) The number of permutations (ways of arranging) n objects into n slots. There are n choices for the first slot, and for each of these, there are $n-1$ choices for the second slot, etc. In total there are $n!$ ways of arranging these objects. (Note that the order of the objects is here important.) (b) The number of permutations of n objects into k slots, $P(n,k)$, is the product $n \cdot (n-1) \cdot (n-2) \dots (n-k+1)$ which can also be written as a ratio of factorials. (c) The number of combinations of n objects in groups of k is called $C(n,k)$ (shown as the first arrow in part c). Here order is not important. The step shown in (b) is equivalent o the two steps shown in (c). This means that there is a relationship between $P(n,k)$ and $C(n,k)$, namely, $P(n,k) = k!C(n,k)$.*

11.9 Appendix: Tests for convergence of series

In order for the sum of ‘infinitely many things’ to add up to a finite number, the terms have to get smaller. But just getting smaller is not, in itself, enough to guarantee convergence. (We will show this later on by considering the harmonic series.)

There are rigorous mathematical tests which help determine whether a series converges or not. We discuss some of these tests here[74].

[74] Recall that $\Rightarrow$ means “implies that”. This is a one-way implication: $A \Rightarrow B$ says that “A implies B” and cannot be used to conclude that B implies A. $\Leftrightarrow$ means that each statement implies the other, a two-way implication. Just as it is important to “obey traffic signs” and avoid “driving the wrong way” on a one-way street, it is also important to be careful about use of these mathematical statements.

11.9.1 The ratio test:

If $\sum_{k=0}^{\infty} a_k$ is a series with $a_n > 0$ and $\lim_{k\to\infty} \frac{a_{k+1}}{a_k} = L$, then

(a) $L < 1 \Rightarrow$ the series converges,

(a) $L > 1 \Rightarrow$ the series diverges,

(a) $L = 1 \Rightarrow$ the test is inconclusive.

Example 1: Reciprocal factorial series

Recall that if $k > 0$ is an integer then the notation $k!$ (read "k factorial") means

$$k! = k \cdot (k-1) \cdot (k-2) \ldots 3 \cdot 2 \cdot 1.$$

Consider the series

$$S = \sum_{k=1}^{\infty} \frac{1}{k!} = 1 + \frac{1}{2 \cdot 1} + \frac{1}{3 \cdot 2 \cdot 1} + \ldots + \frac{1}{k(k-1)\ldots 1},$$

then

$$a_{k+1} = \frac{1}{(k+1)!}, \quad a_k = \frac{1}{k!},$$

$$\frac{a_{k+1}}{a_k} = \lim_{k\to\infty} \frac{\frac{1}{(k+1)!}}{\frac{1}{k!}} = \lim_{k\to\infty} \frac{k!}{(k+1)!} = \lim_{k\to\infty} \frac{1}{k+1} = 0.$$

Thus $L = 0, L < 1$ so this series converges by the ratio test. Later, we will see a second method (comparison) to arrive at the same conclusion.

Example 2: Harmonic series

Does the following converge?

$$S = \sum_{k=1}^{\infty} \frac{1}{k} = 1 + \frac{1}{2} + \frac{1}{3} + \ldots + \frac{1}{k} + \ldots,$$

This series is the *Harmonic Series*. To apply the ratio test, we note that

$$a_{k+1} = \frac{1}{k+1}, \quad a_k = \frac{1}{k},$$

$$L = \lim_{k\to\infty} \frac{a_{k+1}}{a_k} = \lim_{k\to\infty} \frac{\frac{1}{k+1}}{\frac{1}{k}} = \lim_{k\to\infty} \frac{k}{k+1} = 1.$$

Since $L = 1$, in this case, the test is inconclusive. In fact, we show in Section 10.4 that *the harmonic series diverges*.

Example 3: Geometric series

Apply the ratio test to determine the condition for convergence of the geometric series,

$$S = \sum_{k=0}^{\infty} r^k.$$

Here

$$a_{k+1} = r^{k+1}, a_k = r^k, \frac{a_{k+1}}{a_k} = r,$$

$$L = \lim_{k\to\infty} \frac{a_{k+1}}{a_k} = r.$$

So, by the ratio test, if $L = r < 1$ then the geometric series converges (confirming a fact we have already established).

11.9.2 Series comparison tests

We can sometimes use the convergence (or divergence) of a known series to conclude whether a second series converges (or diverges).

> Suppose we have two series,
>
> $$S_a = \sum_{k=0}^{\infty} a_k \quad \text{and} \quad S_b = \sum_{k=0}^{\infty} b_k,$$
>
> such that terms of one series are always smaller than terms of the other, i.e. satisfy
>
> $$0 < a_k < b_k \quad \text{for all } k = 0, 1, \ldots .$$
>
> Then
>
> $$\sum b_k \text{ converges} \Rightarrow \sum a_k \text{ converges},$$
>
> $$\sum a_k \text{ diverges} \Rightarrow \sum b_k \text{ diverges}.$$

The idea behind the first of these statements is that the "smaller" series $\sum a_k$ is "squeezed in" between 0 (the lower bound) and the sum of the larger series (which we know must exist, since $\sum b_k$ converges.) This means that the smaller series cannot become unbounded. For the second statement, we have that the smaller of the two series is known to diverge, forcing the larger also to be unbounded. One must carefully observe that "$\Rightarrow$" applies only in one direction. (For example, if the smaller series converges, we cannot conclude anything about the larger series.)

Example: Comparison with geometric series

Does the series below converge or diverge?

$$S = \sum_{k=0}^{\infty} \frac{1}{2^k + 1}.$$

Solution: We compare terms in this series to a terms in a geometric series with $r = \frac{1}{2}$. i.e. consider

$$a_k = \frac{1}{2^k + 1}, \quad b_k = \frac{1}{2^k}.$$

Then clearly

$$0 < a_k < b_k \quad \text{for every } k$$

(since the denominator in a_k is larger). But we know that $\sum \frac{1}{2^k}$ converges. Therefore, so does $\sum \frac{1}{2^k+1}$.

11.9.3 Alternating series

An alternating series is a series in which the signs of successive terms alternate. An example of this type is the series

$$1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \ldots = \sum (-1)^{n+1} \frac{1}{n}$$

We will show that this series converges (essentially because terms nearly cancel out), and in fact, we show in Section 10.5.3 that it converges to the number $\ln(2) \approx 0.693$. More generally, we have the following result.

If S is an alternating series,

$$S = \sum_{k=1}^{\infty} (-1)^k a_k = a_1 - a_2 + a_3 - a_4 + \ldots$$

with $a_k > 0$ and such that (1) $|a_1| \geq |a_2| \geq |a_3| \geq \ldots$ etc. and (2) $\lim_{k\to\infty} a_k = 0$, then the series converges. (This was established by Leibniz in 1705.)

11.10 Adding and multiplying series

We first comment that arithmetic operations on infinite series only make sense if the series are convergent. In this discussion, we will deal only with series of the convergent type. When this is true, then (and only then) is it true that we can exchange the order of operations as discussed below.

If $\sum a_k$ and $\sum b_k$ both converge and $\sum a_k = S \quad \sum b_k = T$, then

(a) $\sum (a_k + b_k)$ converges and $\sum (a_k + b_k) = \sum a_k + \sum b_k = S + T$.

(b) $\sum c a_k = c \sum a_k = cS$, where c is any constant.

(c) The product $(\sum a_k) \cdot (\sum b_k) = \sum_{k=0}^{\infty} \sum_{i=0}^{k} a_i b_{k-i} = S \cdot T$.

Example:

$$\sum\left(\frac{1}{2}\right)^k \cdot \sum\left(\frac{1}{3}\right)^j = \left(1+\frac{1}{2}+\frac{1}{4}+\dots\right)\left(1+\frac{1}{3}+\frac{1}{9}+\dots\right).$$

Both series converge, so we can write

$$\sum_{j=0}^{\infty}\sum_{k=0}^{\infty}\left(\frac{1}{2}\right)^k\left(\frac{1}{3}\right)^j = \frac{1}{1-\frac{1}{2}}\cdot\frac{1}{1-\frac{1}{3}} = 2\cdot\frac{3}{2} = 3.$$

11.11 Using series to solve a differential equation

Airy's equation arises in the study of optics, and (with initial conditions) is as follows:

$$y'' = xy, \quad y(0) = 1, \quad y'(0) = 0.$$

As before, we will write the solution as a series:

$$y = a_0 + a_1x + a_2x^2 + a_3x^3 + a_4x^4 + a_5x^5 + \dots$$

Using the information from the initial conditions, we get $y(0) = a_0 = 1$ and $y'(0) = a_1 = 0$. Now we can write down the derivatives:

$$\begin{aligned} y' &= a_1 + 2a_2x + 3a_3x^2 + 4a_4x^3 + 5a_5x^4 + \dots \\ y'' &= 2a_2 + 2\cdot 3x + 3\cdot 4x^2 + 4\cdot 5x^3 + \dots \end{aligned}$$

The equation then gives

$$\begin{aligned} y'' &= xy \\ 2a_2 + 2\cdot 3a_3x + 3\cdot 4a_4x^2 + 4\cdot 5a_5x^3 + \dots &= x(a_0 + a_1x + a_2x^2 + a_3x^3 + \dots) \\ 2a_2 + 2\cdot 3a_3x + 3\cdot 4a_4x^2 + 4\cdot 5a_5x^3 + \dots &= a_0x + a_1x^2 + a_2x^3 + a_3x^4 + \dots \end{aligned}$$

Again, we can equate the coefficients of x, and use $a_0 = 1$ and $a_1 = 0$, to obtain

$$\begin{array}{ll} 2a_2 = 0 & \Rightarrow a_2 = 0, \\ 2\cdot 3a_3 = a_0 & \Rightarrow a_3 = \frac{1}{2\cdot 3}, \\ 3\cdot 4a_4 = a_1 & \Rightarrow a_4 = 0, \\ 4\cdot 5a_5 = a_2 & \Rightarrow a_5 = 0, \\ 5\cdot 6a_6 = a_3 & \Rightarrow a_6 = \frac{1}{2\cdot 3\cdot 5\cdot 6}. \end{array}$$

This gives us the first few terms of the solution:

$$y = 1 + \frac{x^3}{2\cdot 3} + \frac{x^6}{2\cdot 3\cdot 5\cdot 6} + \dots$$

If we continue in this way, we can write down many terms of the series.

Index