MATH1131 Mathematics 1A

and

MATH1141 Higher Mathematics 1A

CALCULUS NOTES

CRICOS Provider No: 00098G

Preface

Please read carefully.

These Notes form the basis for the calculus strand of MATH1131 and MATH1141. However, not all of the material in these Notes is included in the MATH1131 or MATH1141 calculus syllabuses. A detailed syllabus will be uploaded to Moodle.

In using these Notes, you should remember the following points:

1. Most courses at university present new material at a faster pace than you will have been accustomed to in high school, so it is essential that you start working right from the beginning of the session and continue to work steadily throughout the session. Make every effort to keep up with the lectures and to do problems relevant to the current lectures.

2. These Notes are **not** intended to be a substitute for attending lectures or tutorials. The lectures will expand on the material in the notes and help you to understand it.

3. These Notes may seem to contain a lot of material but not all of this material is equally important. One aim of the lectures will be to give you a clearer idea of the relative importance of the topics covered in the Notes.

4. Use the tutorials for the purpose for which they are intended, that is, to ask questions about both the theory and the problems being covered in the current lectures.

5. Some of the material in these Notes is more difficult than the rest. This extra material is marked with the symbol [**H**]. Material marked with an [**X**] is intended for students in MATH1141.

6. Some of the problems are marked [**V**]. These have a video solution available from Moodle.

7. It is **essential** for you to do **problems** which are given at the end of each chapter. If you find that you do not have time to attempt all of the problems, you should at least attempt a representative selection of them. You will find advice about this on Moodle. You should also work through the Online Tutorals that you will find on Moodle.

8. You will be expected to use the computer algebra package Maple in tests and understand Maple syntax and output for the end of term examination.

Note.

This version of the Calculus Notes has been prepared by Robert Taggart and Peter Brown. They build on notes first developed by Tony Dooley and subsequently edited by several members of the School of Mathematics and Statistics. The main editors include Mike Banner, Ian Doust and V. Jeyakumar.

Contents

CALCULUS SYLLABUS

The calculus course for both MATH1131 and MATH1141 is based on these MATH1131/MATH1141 Calculus Notes that are included in the Course Pack. A detailed syllabus and lecture schedule will be posted on Moodle.

The computer package Maple will be used in the calculus course. An introduction to Maple is included in the booklet titled *First Year Maple Notes*.

CALCULUS PROBLEM SETS

The Calculus problems are located at the end of each chapter of the Calculus Notes booklet. They are also available from the course module on the UNSW Moodle server. Some of the problems are very easy, some are less easy but still routine and some are quite hard. To help you decide which problems to try first, each problem is marked with an [**R**], an [**H**] or an [**X**]. The problems marked [**R**] form a basic set of problems which you should try first. Problems marked [**H**] are harder and can be left until you have done the problems marked [**R**]. Problems marked [**V**] have a video solution available on Moodle.

You *do* need to make an attempt at the [**H**] problems because problems of this type will occur on tests and in the exam. If you have difficulty with the [**H**] problems, ask for help in your tutorial. The problems marked [**X**] are intended for students in MATH1141 – they relate to topics which are only covered in MATH1141. Extra problem sheets for MATH1141 may be issued in lectures.

Remember that working through a wide range of problems is the key to success in mathematics. However, *solving problems* and *writing mathematics clearly* are two separate skills that need to be developed through practice. We recommend that you keep a workbook to practice *writing* solutions to mathematical problems.

Chapter 1

Sets, inequalities and functions

In the days of the Roman empire, the word 'calculus' denoted a pebble that was used for counting and gambling. As time progressed, the word 'calculare' came to mean 'to compute.' In the second half of the seventeenth century, two mathematicians, the Englishman Isaac Newton and the German Gottfried Leibniz, independently invented methods for

- calculating gradients of tangents to curves,
- calculating instantaneous acceleration and velocity,
- calculating the area of regions with a curved boundary,
- calculating the volume of solids with curved boundary,
- calculating the length of a curve,
- calculating the work done by a force, and
- calculating the centre of mass of a general solid.

These methods were developed by combining algebra, geometry and trigonometry with the *limiting process* and became known as *the calculus*.

Calculus (as it is known today) has many applications to engineering, physics, chemistry, biology, geology, surveying, sociology, economics and statistics. It includes two major branches: the *differential calculus* (introduced in Chapter 4) and the *integral calculus* (introduced in Chapter 8). These two branches are related by *the fundamental theorem of calculus* (see Theorem 8.5.1). The underlying tool used to develop these branches is the concept of the *limit* (introduced in Chapter 2), which is applied to *functions* (introduced in Chapter 1).

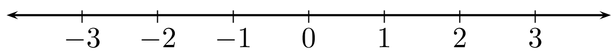
The main goal of this chapter is to introduce functions. Although functions will be familiar to students from high school, important points that deserve students' attention are emphasised here. Since the functions we study in this course take real numbers as inputs and give real numbers of outputs, we begin the chapter with a review of different sets that consist of real numbers. We also devote a little time to revising inequalities and absolute values, in preparation for our discussion of limits in Chapter 2.

1.1 Sets of numbers

(Ref: SH10 §1.2)

A *set* is a collection of distinct objects. The objects in a set are called the *elements* or *members* of the set. Some commonly used sets of numbers are listed below.

- The set $\mathbb{N}$ of *natural numbers* is given by
$$\mathbb{N} = \{0, 1, 2, 3, 4, \ldots\}.$$
- The set $\mathbb{Z}$ of *integers* is given by
$$\mathbb{Z} = \{\ldots, -3, -2, -1, 0, 1, 2, 3, \ldots\}.$$
- The set $\mathbb{Q}$ of *rational numbers* is the collection of all numbers of the form $\frac{p}{q}$, where p and q are integers and $q \neq 0$.
- The set $\mathbb{R}$ of *real numbers* may be represented as the collection of points lying on the number line.



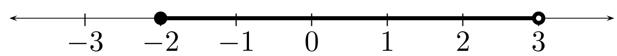
If A is a set of numbers and the number x is a member of the set A, then we write
$$x \in A.$$
If x is not a member of A then we write
$$x \notin A.$$
The next example illustrates this notation.

Example 1.1.1. The following statements are true:
$$2 \in \mathbb{N}, \qquad -2 \notin \mathbb{N}, \qquad \frac{3}{4} \notin \mathbb{Z}, \qquad \frac{3}{4} \in \mathbb{Q}, \qquad \sqrt{2} \in \mathbb{R}, \qquad \sqrt{2} \notin \mathbb{Q}.$$

Many other sets of numbers can be described using $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$ and inequalities.

Example 1.1.2. (a) The set
$$\{n \in \mathbb{Z} : n > 0\}$$
is the set of all numbers n belonging to the integers such that n is greater than 0. In other words, it is the set of *positive integers*. (Note that the colon ':' above is read as 'such that'.)

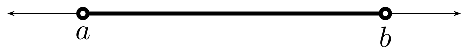
(b) The set
$$\{x \in \mathbb{R} : -2 \leq x < 3\}$$
is represented by the following interval on the real line.



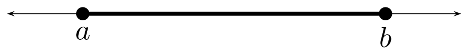
(c) The set
$$\{x \in \mathbb{R} : x \notin \mathbb{Q}\}$$
is the set of *irrational numbers*.

Sets that are represented by intervals of the real line occur so frequently that we introduce a special notation for them. Suppose that a and b are real numbers and that $a < b$. Then the intervals (a, b), $[a, b]$, $(a, b]$ and $[a, b)$ are given by

- $(a, b) = \{x \in \mathbb{R} : a < x < b\}$

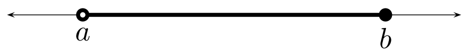


- $[a, b] = \{x \in \mathbb{R} : a \le x \le b\}$



- $[a, b) = \{x \in \mathbb{R} : a \le x < b\}$

- $(a, b] = \{x \in \mathbb{R} : a < x \le b\}$.

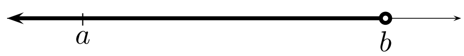


In each case, the numbers a and b are called the *endpoints* of the interval. An interval $[a, b]$ that includes its endpoints is called a *closed interval*, while an interval (a, b) that excludes its endpoints is called an *open interval*. The intervals $[a, b)$ and $(a, b]$ are neither open nor closed.

Interval notation can be extended to describe rays of the real line by using the symbol ∞ for infinity. Thus we have

- $[a, \infty) = \{x \in \mathbb{R} : a \le x\}$

- $(-\infty, b) = \{x \in \mathbb{R} : x < b\}$



- $(-\infty, \infty) = \mathbb{R}$.

We end by introducing an important set relation.

> **Definition 1.1.3.** Suppose that A and B are two sets. We say that A is a *subset* of B if $x \in A$ implies that $x \in B$. If A is a subset of B then we also say that B *contains* the set A.

The above definition is illustrated by the following examples:

- $\{0,1,2,3\}$ is a subset of $\{0,1,2,3,4\}$,
- $\mathbb{N}$ is a subset of $\mathbb{Z}$,
- $[-1,2]$ is a subset of $(-3,\infty)$,
- $\mathbb{R}$ is a subset of $\mathbb{R}$, and
- $[1,4)$ is not a subset of $(1,4]$.

1.2 Solving inequalities

(Ref: SH10 §1.3)

When solving an inequality, one must remember that if the inequality is multiplied by a negative number, the sign of the inequality reverses.

Two types of inequalities deserve special attention: quadratic inequalities and inequalities with an unknown quantity in the denominator.

Example 1.2.1. Solve the inequality

$$x^2 > 6 - x.$$

Solution. Rearranging the inequality gives

$$x^2 + x - 6 > 0.$$

The values of x for which $x^2 + x - 6$ is positive are the same as those for which the graph of

$$y = x^2 + x - 6$$

lies above the x-axis.

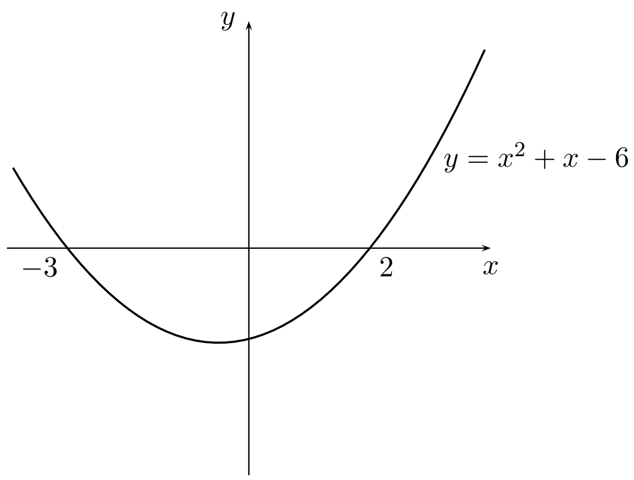
To graph

$$y = x^2 + x - 6,$$

we factorise the right-hand side to obtain

$$y = (x-2)(x+3).$$

Hence $y = 0$ when $x = 2$ and when $x = -3$. We can now easily sketch its graph.



Since the graph lies *above* the x-axis when $x < -3$ and when $x > 2$, the solution to the inequality is

$$x < -3 \qquad \text{or} \qquad x > 2.$$

$\square$

Example 1.2.2. Solve the inequality

$$\frac{1}{3} < \frac{1}{x-1}.$$

Solution. We begin by noting that $x \neq 1$.

To solve the inequality, we cannot simply multiply the inequality through by $x - 1$, since we do not know whether $x - 1$ is positive or negative. (Taking the reciprocal of both sides is similarly problematic.) Instead, we multiply through by the *positive* number $(x - 1)^2$ to obtain

$$\frac{1}{3}(x-1)^2 < x - 1.$$

Multiplication by 3 and rearrangement gives

$$(x-1)^2 - 3(x-1) < 0$$

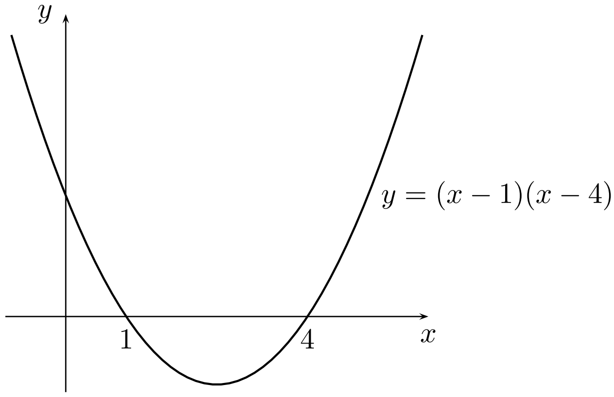
and by factorisation we obtain

$$(x-1)[(x-1)-3] = (x-1)(x-4) < 0.$$

So proceding in a manner similar to the previous example, we seek to identify the values of x for which the graph of

$$y = (x-1)(x-4)$$

lies *below* the x-axis.



Hence the solution to the inequality is given by

$$1 < x < 4.$$

$\square$

The next example is slightly more difficult, but uses techniques illustrated in the previous two examples.

Example 1.2.3. Solve the inequality

$$x + 2 \geq \frac{4}{x - 1}$$

Solution. First observe that $x \neq 1$.

Multiplying the inequality through by $(x-1)^2$ gives

$$(x-1)^2(x+2) \geq 4(x-1).$$

Hence

$$(x-1)^2(x+2) - 4(x-1) \geq 0.$$

Now

$$\begin{aligned}(x-1)^2(x+2) - 4(x-1) &= (x-1)[(x-1)(x+2) - 4] \\ &= (x-1)[x^2 + x - 6] \\ &= (x-1)(x-2)(x+3).\end{aligned}$$

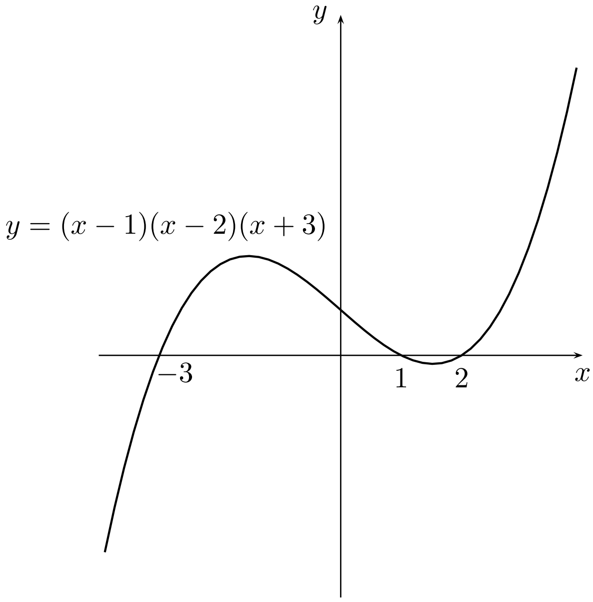
So the original inequality is equivalent to

$$(x-1)(x-2)(x+3) \geq 0, \qquad x \neq 1.$$

We now sketch the graph of

$$y = (x-1)(x-2)(x+3)$$

and see for which values of x the graph lies on or above the x-axis.



Hence the solution to the inequality is

$$-3 \leq x < 1 \qquad \text{or} \qquad x \geq 2.$$

□

Remark 1.2.4. A variety of methods (some good and others bad) exist for solving inequalities where an unknown lies in the denominator. It is strongly recommended that students use the method shown in this section, or that shown by their first year calculus lecturer.

1.3 Absolute values

(Ref: SH10 §§1.2, 1.3)

The magnitude, or *absolute value*, of a real number x is denoted by $|x|$ and defined by

$$|x| = \begin{cases} x & \text{if } x \geq 0 \\ -x & \text{if } x < 0. \end{cases} \tag{1.1}$$

Absolute values of real numbers interact in the following way.

Proposition 1.3.1. *Suppose that x and y are real numbers. Then*

(i) $|-x| = |x|$,

(ii) $|xy| = |x||y|$,

(iii) $\left|\dfrac{x}{y}\right| = \dfrac{|x|}{|y|}$, *provided that* $y \neq 0$, *and*

(iv) $|x+y| \leq |x| + |y|$.

Property (iv) is known as *the triangle inequality.*

Three other useful facts about absolute values are noted below.

(F1) For every real number x,

$$|x| = \sqrt{x^2} \qquad \text{and} \qquad |x|^2 = x^2. \tag{1.2}$$

(F2) If a and x are real numbers then $|x-a|$ is equal to the distance between x and a on the real number line.

(F3) For any positive real number a,

$$|y| < a \qquad \text{if and only if} \qquad -a < y < a$$

and

$$|y| > a \qquad \text{if and only if} \qquad y < -a \quad \text{or} \quad y > a.$$

These facts give different methods for solving inequalities involving absolute values.

Example 1.3.2. Solve the following inequalities:

(a) $|x-3| \leq 5$

(b) $|2x+5| > 8$

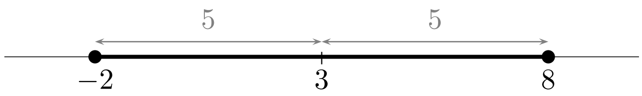
(c) $\left|\dfrac{x-3}{x-1}\right| < 1.$

Solution. (a) *An algebraic solution.* Using (F3) we have

$$\begin{gathered} |x-3| \leq 5 \\ -5 \leq x-3 \leq 5 \\ -2 \leq x \leq 8. \end{gathered}$$

(a) *A geometric solution.* By (F2) the inequality $|x-3| \leq 5$ is equivalent to saying that 'the distance between x and 3 is less than or equal to 5.'

The points x on the number line that have this property are shown below.



Hence $-2 \le x \le 8$.

(b) *An algebraic solution.* Using (F3) the inequality is equivalent to

$$2x + 5 < -8 \qquad \text{or} \qquad 2x + 5 > 8.$$

Hence

$$2x < -13 \qquad \text{or} \qquad 2x > 3,$$

which gives the solution

$$x < -\frac{13}{2} \qquad \text{or} \qquad x > \frac{3}{2}.$$

(c) *An algebraic solution.* By Prop.1.3.1 (iii), we can write, for $x \neq 1$,

$$\left|\frac{x-3}{x-1}\right| = \frac{|x-3|}{|x-1|}$$

and since the denominator is positive, the inequality can be written as

$$|x - 3| < |x - 1|.$$

Squaring both (positive) sides),and using (F1), we have

$$(x - 3)^2 < (x - 1)^2.$$

Expanding and solving, we have $x > 2$.

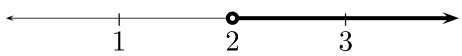
(c) *A geometric solution.* As above, the inequality is equivalent to

$$|x - 3| < |x - 1|, \qquad x \neq 1.$$

By (F2) the inequality is interpreted geometrically as

'the distance from x to 3 is less than the distance from x to 1.'

In other words, x is closer to 3 than to 1. Points that have this property are shown on the numberline below.



Hence $x > 2$. $\square$

1.4 Functions

(Ref: SH10 §1.5)

A function $f : A \to B$ is a rule which assigns to every element x belonging to a set A exactly one element $f(x)$ belonging to a set B. The set A is called the *domain* of the function f and the set B is called the *codomain* of f. In this course, A and B are always sets of real numbers.

Example 1.4.1. A function $f : [0, \infty) \to \mathbb{R}$ is given by

$$f(x) = \sqrt{x}$$

for all x in $[0, \infty)$. Informally, this means that f takes a number x from $[0, \infty)$ as input and gives the number $\sqrt{x}$ as output. More formally, we say that f *maps* an element x of $[0, \infty)$ to $\sqrt{x}$. We offer a few comments on terminology employed to describe f.

- The expression $f : [0, \infty) \to \mathbb{R}$ says that $[0, \infty)$ is the domain of f and that $\mathbb{R}$ is the codomain of f. We write
$$\mathrm{Dom}(f) = [0, \infty) \qquad \text{and} \qquad \mathrm{Codom}(f) = \mathbb{R}.$$

- The domain of f is the set of all inputs for the function.

- The codomain of f is a set that contains all the output values of the function. Since the output values are all real numbers, we say that f is a *real-valued function.* While all the outputs must lie in $\mathrm{Codom}(f)$, not every number in $\mathrm{Codom}(f)$ need be an output value.

- The expression $f(x)$ (read as 'f of x') is the *value of* f at the point x. That is, $f(x)$ is the unique number in $\mathbb{R}$ that corresponds to the input x. To emphasise the point: $f(x)$ is a real number, *not* a function.

- The statement
$$\text{`}f(x) = \sqrt{x} \text{ for all } x \text{ in } [0, \infty)\text{'}$$
can be abbreviated as
$$f(x) = \sqrt{x} \qquad \forall x \in [0, \infty).$$
The symbol $\forall$ means 'for all.'

In high school, functions are often described by specifying a rule without specifying a domain. For example, the following sentence was taken from an HSC study guide:

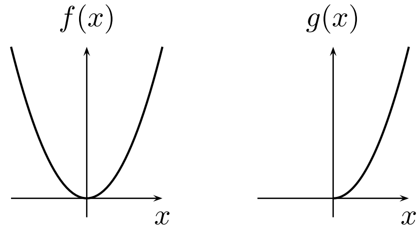
'Consider the function $f(x) = x^2$.'

There are two problems with this sentence. First, neither $f(x)$ nor $f(x) = x^2$ is a function. As mentioned above, $f(x)$ is a real number, and $f(x) = x^2$ is an equation. (This may seem picky, but being able to distinguish between a function f and its value $f(x)$ at the point x is crucial to understanding key concepts in both algebra and calculus.) Second, the quoted sentence does not specify the domain of the function. The next example gives two different functions that obey the same squaring rule.

Example 1.4.2. Consider the functions f and g, defined by

$$\begin{aligned} &f : \mathbb{R} \to \mathbb{R} && & &g : [0, \infty) \to \mathbb{R} \\ &f(x) = x^2 \qquad \forall x \in \mathrm{Dom}(f) && \text{and} & &g(x) = x^2 \qquad \forall x \in \mathrm{Dom}(g). \end{aligned}$$

Even though they are both squaring functions, f and g are *not* the same function. For example, f can square negative numbers while g does not. On the other hand, g is *invertible* (this important and desirable property is discussed in Chapter 6) while f is not. Their graphs, shown below, are also noticeably different.



If, for whatever reason, the domain of a function is not specified, but the function rule is, then the default domain, known as the *maximal* or *natural domain*, is the largest possible domain for which the rule makes sense. We give two examples.

- If
$$f(x) = \sqrt{x-1}$$
then the rule makes sense only if $x - 1 \geq 0$, that is, if $x \geq 1$. Therefore the maximal domain of the function f is $[1, \infty)$.

- If
$$f(x) = \frac{2}{(x-2)(x-8)}$$
then the rule only makes sense if $x \neq 2$ or $x \neq 8$. So the maximal domain is the set
$$\{x \in \mathbb{R} : x \neq 2, x \neq 8\}.$$

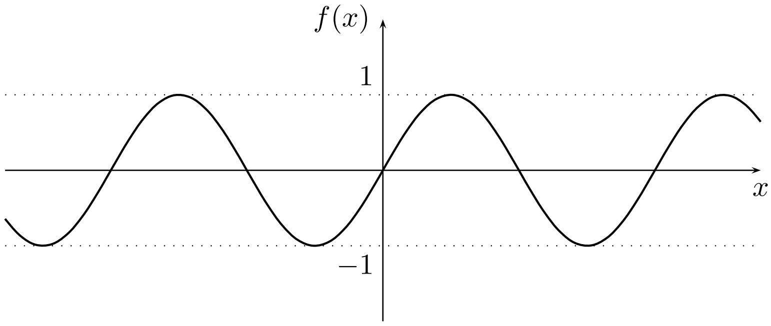
Another set associated to a function is its *range*.

Definition 1.4.3. Suppose that $f : A \to B$ is a function. The *range* of f, denoted by $\mathrm{Range}(f)$, is defined by
$$\mathrm{Range}(f) = \{f(x) \in B : x \in A\}.$$

In other words, the range of f is the set of all output values. Equivalently, if the rule of a function is described by the equation $y = f(x)$, then the range of f is the set of all corresponding y-values. The range is always a subset of the codomain, but the range need not equal the codomain. The difference between range and codomain is illustrated by the following examples.

Example 1.4.4. Suppose that $f : \mathbb{R} \to \mathbb{R}$ is given by the rule
$$f(x) = \sin x \qquad \forall x \in \mathbb{R}.$$

Then $\mathrm{Dom}(f) = \mathbb{R}$ and $\mathrm{Codom}(f) = \mathbb{R}$ (this is obtained from the expression $f : \mathbb{R} \to \mathbb{R}$). To determine the range, consider the graph of f is sketched below.

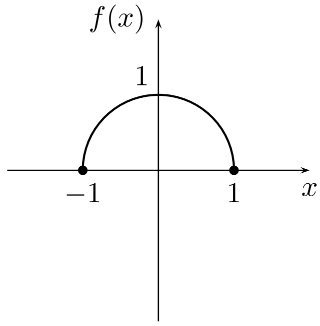


The set of all output values (equivalently, the set of all corresponding y-values) is $[-1, 1]$. Hence $\text{Range}(f) = [-1, 1]$.

Example 1.4.5. Suppose that $f : [-1, 1] \to \mathbb{R}$ is given by the rule

$$f(x) = \sqrt{1 - x^2} \qquad \forall x \in [-1, 1].$$

Then $\text{Dom}(f) = [-1, 1]$ and $\text{Codom}(f) = \mathbb{R}$ (again, we are simply reading this from $f : [-1, 1] \to \mathbb{R}$). To determine the range, we observe that the graph of f is a semicircle.



The set of all outputs (or y-values) is $[0, 1]$. Hence $\text{Range}(f) = [0, 1]$.

If f and g are two functions with the same domain, then one can combine f and g to form new functions.

Definition 1.4.6. Suppose that $f : A \to B$ and $g : A \to B$ are real-valued functions. Then the functions $f + g$, $f - g$, $f.g$ and f/g are defined by the rules

$$\begin{aligned} (f+g)(x) &= f(x) + g(x) && \forall x \in A \\ (f-g)(x) &= f(x) - g(x) && \forall x \in A \\ (f.g)(x) &= f(x)g(x) && \forall x \in A \\ (f/g)(x) &= \frac{f(x)}{g(x)} && \forall x \in A, \text{ provided that } g(x) \neq 0. \end{aligned}$$

Another way of constructing new functions is given below.

Definition 1.4.7. Suppose that $f : C \to D$ and $g : A \to B$ are functions such that Range(g) is a subset of Dom(f). Then the *composition* $f \circ g : A \to D$ is defined by the rule

$$(f \circ g)(x) = f(g(x)) \qquad \forall x \in A.$$

Example 1.4.8. Suppose that $f : \mathbb{R} \to \mathbb{R}$ and $g : \mathbb{R} \to \mathbb{R}$ are given by the rules

$$f(x) = x^2 + 1 \quad \forall x \in \mathbb{R} \qquad \text{and} \qquad g(x) = x - 2 \quad \forall x \in \mathbb{R}.$$

Find $(f + g)(5)$ and $(f \circ g)(x)$. Is 2 in the domain of the function f/g? Give reasons.

Solution. First,

$$\begin{aligned}(f + g)(5) &= f(5) + g(5)\\ &= (5^2 + 1) + (5 - 2)\\ &= 29.\end{aligned}$$

Second, if x in $\mathbb{R}$ then

$$\begin{aligned}(f \circ g)(x) &= f(g(x))\\ &= (x - 2)^2 + 1\\ &= x^2 - 4x + 5.\end{aligned}$$

Finally, $g(2) = 0$ so 2 is not in the domain of f/g. □

The final example illustrates the importance of bearing in mind domain and range when composing two functions.

Example 1.4.9. Suppose that $f : [0, \infty) \to \mathbb{R}$ and $g : \mathbb{R} \to \mathbb{R}$ are given by

$$f(x) = \sqrt{x} \quad \forall x \in [0, \infty) \qquad \text{and} \qquad g(x) = \sin x - 2 \quad \forall x \in \mathbb{R}.$$

Find, if they exist, $(f \circ g)(x)$ and $(g \circ f)(x)$.

Solution. Consider first $f \circ g$. For this function to exist, we require that Range(g) is a subset of Dom(f). Now

$$-1 \le \sin x \le 1 \qquad \forall x \in \mathbb{R}$$

so

$$-3 \le \sin x - 2 \le -1 \qquad \forall x \in \mathbb{R};$$

that is, Range(g) $= [-3, -1]$. Since Dom(f) $= [0, \infty)$, Range(g) is not a subset of Dom(f) and hence $f \circ g$ does not exist. (The point here is that, if

$$(f \circ g)(x) = f(g(x)) = \sqrt{\sin x - 2},$$

then the expression under the square root sign is always negative.)

Now consider $g \circ f$. Since f is a real-valued function and Dom(g) $= \mathbb{R}$, it is clear that Range(f) is a subset of Dom(g). (The fact that Range(f) $= [0, \infty)$ is not needed because Dom(g) is so large.) Hence

$$(g \circ f)(x) = g(f(x)) = \sin(\sqrt{x}) - 2 \qquad \forall x \in \text{Dom}(f) = [0, \infty).$$

□

1.5 Polynomials and rational functions

(Ref: SH10 §1.6)

In the next three sections we discuss some important classes of functions. Since these were introduced in high school, we give a brief survey only.

Definition 1.5.1. A function $f : \mathbb{R} \to \mathbb{R}$ is called a *polynomial* if

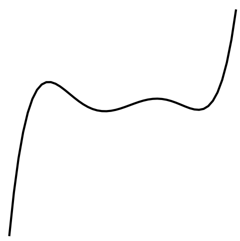
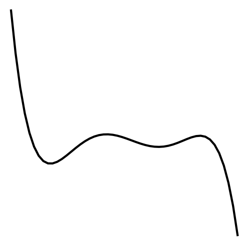
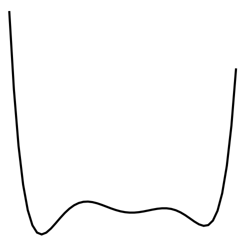
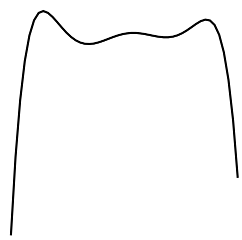
$$f(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_2 x^2 + a_1 x + a_0 \qquad \forall x \in \mathbb{R},$$

where n is a natural number, the constants $a_0, a_1, \ldots, a_n$ are real numbers and $a_n \neq 0$. We call n the *degree of* f, and the constants $a_0, a_1, \ldots, a_n$ are called the *coefficients* of f. The constant a_n is called the *leading coefficient* of f.

Polynomials of small degree have specials names and properties.

- If $n = 0$ then $f(x) = a_0$. Hence f is a constant function and the graph of f is a horizontal straight line.
- If $n = 1$ then $f(x) = a_1 x + a_0$. The graph of f is a straight line with gradient a_1 and y-axis intercept a_0.
- If $n = 2$ then f is called a *quadratic function* and its graph is a parabola.
- If n is 3, 4 or 5 then f is called a *cubic*, *quartic* or *quintic*, polynomial respectively.

The general shape of the graph of a polynomial function of degree greater than 1 is indicated by the table below.

	Positive leading coefficient	Negative leading coefficient
Odd degree		
Even degree		

The number of turning points in the graph is always strictly less than the degree of the polynomial. Exactly how many turning points the graph has depends on the coefficients of the polynomial, and can be determined using calculus (see, for example, Chapters 4 and 5).

By dividing two polynomials, one obtains a new type of function.

Definition 1.5.2. Suppose that p and q are polynomials. A function f is called a *rational function* if

$$\mathrm{Dom}(f) = \{x \in \mathbb{R} : q(x) \neq 0\}$$

and

$$f(x) = \frac{p(x)}{q(x)}, \qquad \forall x \in \mathrm{Dom}(f).$$

Hence the expressions

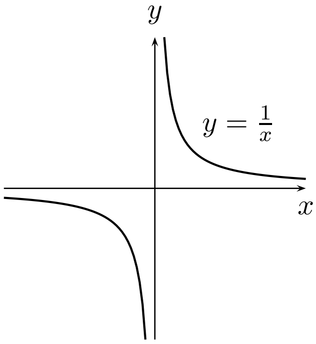
$$f(x) = \frac{x}{x^2+2} \qquad \text{and} \qquad f(x) = \frac{x^3 - 4x^2 + 1}{x-2}$$

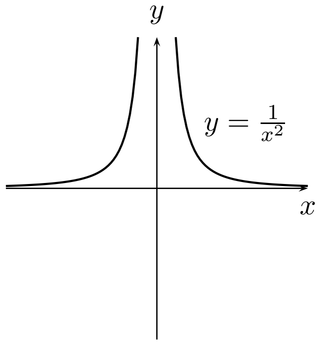
both give rise to rational functions, as does

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

(to see why, rewrite right-hand side with a common denominator of $x^2 + 3$).

The graphs of two simple rational functions are shown below.





The graphs of more sophisticated rational functions are studied in Chapter 7.

1.6 The trigonometric functions

(Ref: SH10 §1.6)

The trigonometric functions will already be familiar to you. This section gives a brief summary of their definitions and some of their properties.

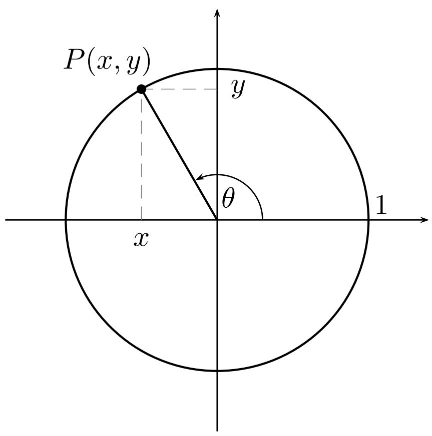
Remark 1.6.1. In this course, angles are always measured in *radians* rather than degrees. (Radian measure is a more natural way of measuring angles; moreover, important formulae used in calculus, such as

$$\frac{d}{d\theta}(\sin\theta) = \cos\theta,$$

only make sense geometrically when θ is interpreted as an angle measured in radians rather than degrees.) We remind readers that

$$2\pi \text{ radians } = 360 \text{ degrees.}$$

Consider an angle θ and the corresponding point $P(x, y)$ that lies on the unit circle centred at the origin, as shown below.



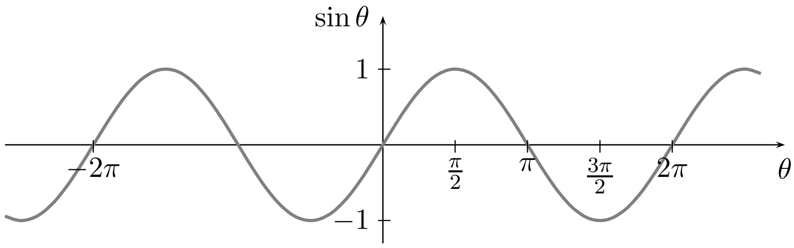
The functions $\cos : \mathbb{R} \to \mathbb{R}$ and $\sin : \mathbb{R} \to \mathbb{R}$ are defined by

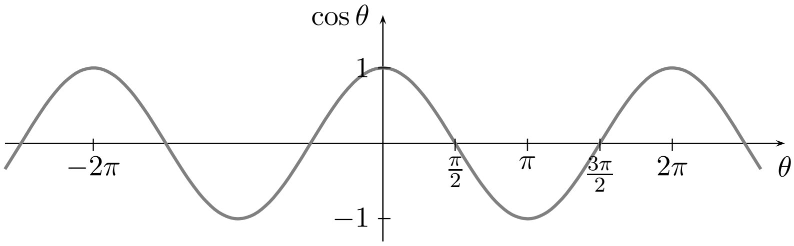
$$\cos\theta = x \qquad \text{and} \qquad \sin\theta = y.$$

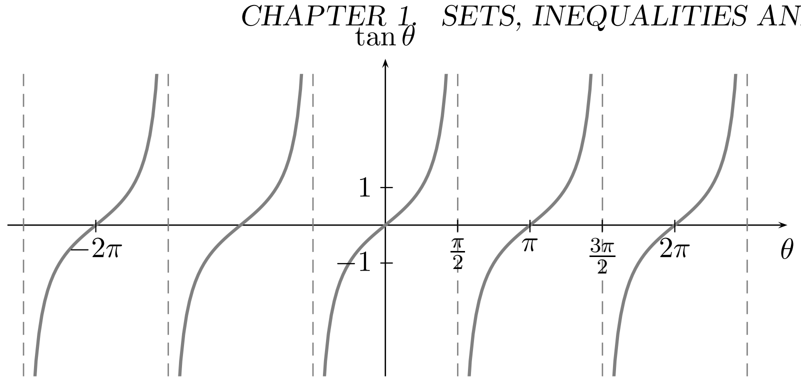
Other trig functions are defined in terms of the sine and cosine functions:

$$\begin{aligned}
\tan\theta &= \frac{\sin\theta}{\cos\theta}, && \text{provided that } \cos\theta \neq 0\\
\sec\theta &= \frac{1}{\cos\theta}, && \text{provided that } \cos\theta \neq 0\\
\operatorname{cosec}\theta &= \frac{1}{\sin\theta}, && \text{provided that } \sin\theta \neq 0\\
\cot\theta &= \frac{\cos\theta}{\sin\theta}, && \text{provided that } \sin\theta \neq 0.
\end{aligned}$$

The graphs of sin, cos and tan are shown below.







The sine and cosine functions are 2π-periodic, while the tangent function is π-periodic, which means that

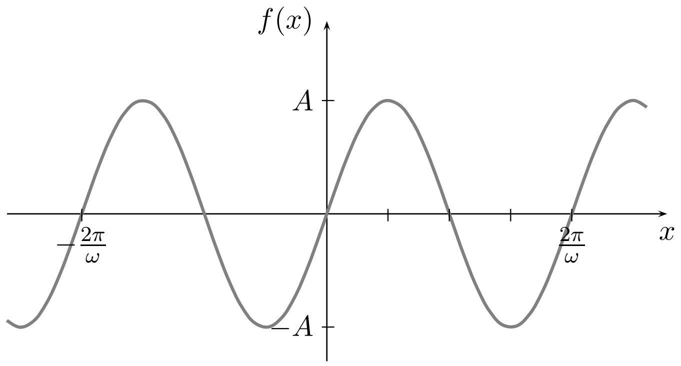
$$\sin(\theta + 2\pi) = \sin\theta, \quad \cos(\theta + 2\pi) = \cos\theta \quad \text{and} \quad \tan(\theta + \pi) = \tan\theta$$

whenever θ is in the domain of the respective function. Sine and cosine both have an amplitude of 1.

It is not hard to obtain functions whose graphs have the same general shape as the sine (or cosine) curve with a period other than 2π and amplitude other than 1. For example, if A and ω are positive real numbers then the function $f : \mathbb{R} \to \mathbb{R}$, given by

$$f(x) = A\sin(\omega x) \qquad \forall x \in \mathbb{R},$$

has an amplitude of A and a period of $\frac{2\pi}{\omega}$. Its graph is shown below.



The six trigonometric functions are related to one another by various identities and formulae:

- complementary identities

$$\sin\left(\frac{\pi}{2} - x\right) = \cos x$$
$$\cos\left(\frac{\pi}{2} - x\right) = \sin x$$

- Pythagorean identities

$$\cos^2 x + \sin^2 x = 1$$
$$1 + \tan^2 x = \sec^2 x$$
$$\cot^2 x + 1 = \operatorname{cosec}^2 x$$

- the sum and difference formulae

$$\sin(x \pm y) = \sin x \cos y \pm \cos x \sin y$$
$$\cos(x \pm y) = \cos x \cos y \mp \sin x \sin y$$
$$\tan(x \pm y) = \frac{\tan x \pm \tan y}{1 \mp \tan x \tan y}$$

- double-angle formulae

$$\sin(2x) = 2 \sin x \cos x$$
$$\cos(2x) = \cos^2 x - \sin^2 x$$
$$\tan(2x) = \frac{2 \tan x}{1 - \tan^2 x}.$$

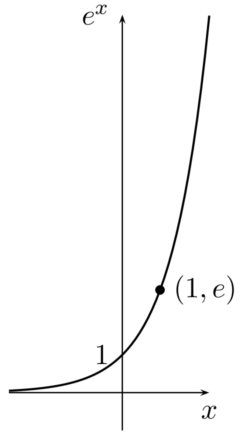
The above formulae, especially the double-angle formulae for $\sin(2x)$ and $\cos(2x)$, should be learnt carefully.

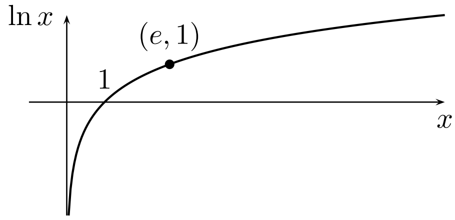
The inverse trigonometric functions will be discussed in Chapter 6.

1.7 The elementary functions

(Ref: SH10 §1.6)

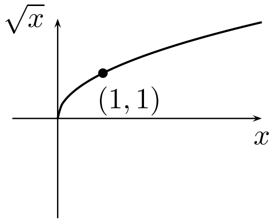
Other important functions that have not yet been mentioned include the exponential and logarithm functions. Their graphs are shown below.

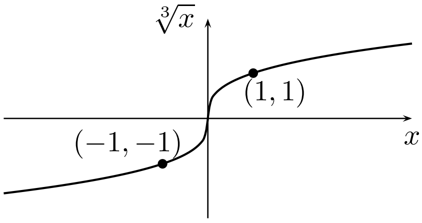




These functions will be studied in greater depth in Chapter 9.

Functions arising from the study of roots of polynomials are also important to mathematics, especially those functions f of the form $f(x) = \sqrt[n]{x}$, where n is a positive integer. It is important that students are familiar with the graphs of the square root and cube root functions, shown below.





The *elementary functions* are all those functions that can be constructed by combining a finite number of polynomials, exponentials, logarithms, roots and trigonometric functions (including the inverse trigonometric functions) via function composition, addition, subtraction, multiplication and division. Hence the following expressions give rise to elementary functions:

$$\begin{aligned} f(x) &= e^{\sin x} + x^2 \\ g(x) &= \frac{\ln x - \tan x}{\sqrt{x}} \\ h(x) &= \sqrt[3]{x^4 - 2x^2 + 5}. \end{aligned}$$

It also follows that every rational function is an elementary function.

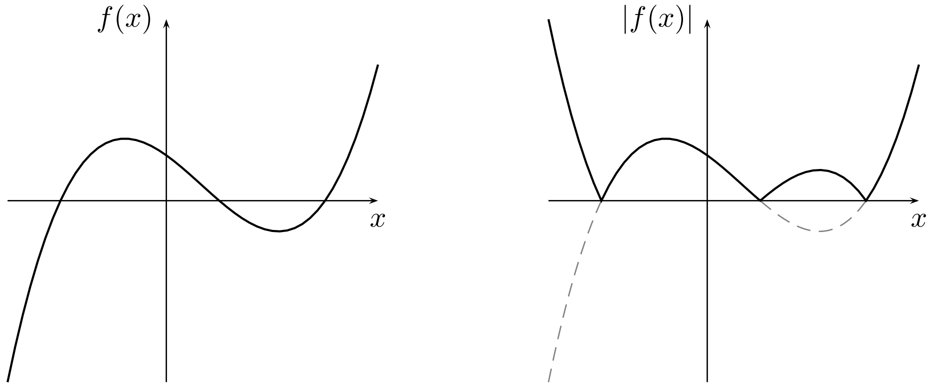
The absolute values function, given by

$$\begin{aligned} &\mathrm{abs} : \mathbb{R} \to \mathbb{R} \\ &\mathrm{abs}(x) = |x| \qquad \forall x \in \mathbb{R}, \end{aligned}$$

is an elementary function by equation (1.2). It follows that if f is an elementary function then $\mathrm{abs} \circ f$ is also an elementary function. It is standard to denote function $\mathrm{abs} \circ f$ by $|f|$. Thus

$$|f|(x) = (\mathrm{abs} \circ f)(x) = |f(x)| \qquad \forall x \in \mathrm{Dom}(f).$$

The graph of $|f|$ is obtained by reflecting in the x-axis any part of the graph of f that lies below the x-axis. This is illustrated below.



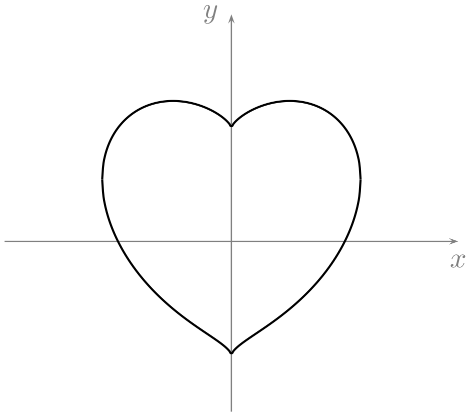
Most, if not all functions studied in high school were elementary functions. In this course we shall meet some useful functions that are not elementary functions. (These are defined by integrals; see Section 8.12.)

1.8 Implicitly defined functions

Many curves on the plane can be described as all those points (x, y) on the plane that satisfy some equation involving x and y. For example, consider the equation

$$(x^2 + y^2 - 1)^3 - x^2y^3 = 0. \tag{1.3}$$

The set of points (x, y) satisfying this equation are shown on the graph below.



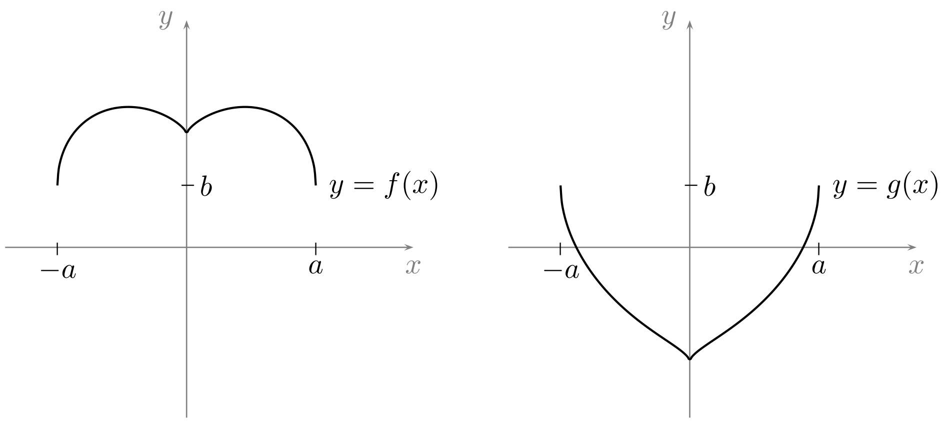
Note that the curve cannot be described by one function since some x-values have two corresponding y-values. However, the curve can be described by combining the graphs of two functions $f : [-a, a] \to \mathbb{R}$ and $g : (-a, a) \to \mathbb{R}$, which are defined by the rules

$$\begin{cases} y = f(x) \\ (x^2 + y^2 - 1)^3 - x^2y^3 = 0 \\ y \geq b \end{cases}$$

and

$$\begin{cases} y = g(x) \\ (x^2 + y^2 - 1)^3 - x^2y^3 = 0 \\ y < b \end{cases}$$

(where $a \approx 1.139$ and $b \approx 0.545$), as illustrated below.



We say that f and g are *implicitly* defined by equation (1.3).

Example 1.8.1. The set of points (x, y) that satisfy the equation

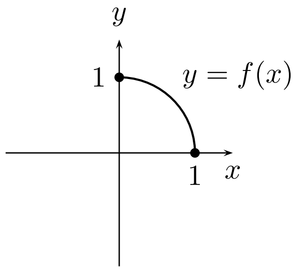
$$x^2 + y^2 = 1$$

describes a circle of radius 1 centred at the origin. The above equation could be used to implicitly define many different functions. We give two examples.

- The function $f : [0,1] \to \mathbb{R}$ is defined by the rule

$$\begin{cases} y = f(x) \\ x^2 + y^2 = 1 \\ y \geq 0. \end{cases}$$

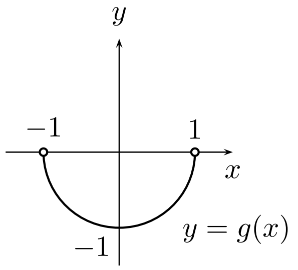
Its graph is shown below.



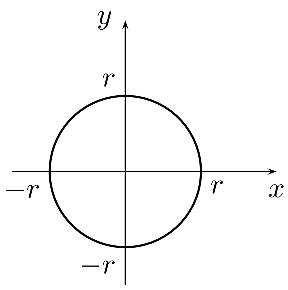
- The function $g : (-1,1) \to \mathbb{R}$ is defined by the rule

$$\begin{cases} y = f(x) \\ x^2 + y^2 = 1 \\ y \leq 0. \end{cases}$$

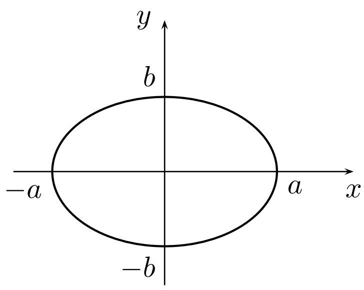
Its graph is shown below.



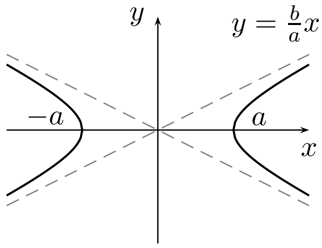
Many heavenly bodies (such as planets and comets) trace out paths that are roughly elliptic, parabolic or hyperbolic. (The fact that planets trace out elliptic paths around the sun was discovered by Johannes Kepler at the beginning of the sixteenth century, and was later used by Newton, in conjunction with calculus, to establish his universal law of gravitation.) Circles, ellipses, parabolas and hyperbolas are known as *conic sections* because they can be obtained by intersecting a cone with a plane. Equations that describe conic sections, accompanied by corresponding diagrams, are shown below.



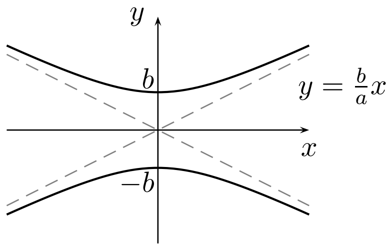
Circle: $x^2 + y^2 = r^2$



Ellipse: $\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} = 1$



Hyperbola: $\dfrac{x^2}{a^2} - \dfrac{y^2}{b^2} = 1$



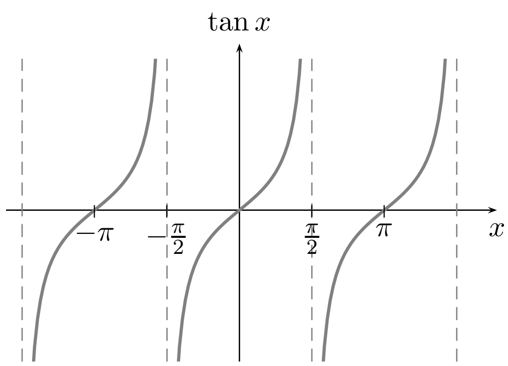
Hyperbola: $\dfrac{y^2}{b^2} - \dfrac{x^2}{a^2} = 1$

Students should be able to recognise such curves.

1.9 Continuous functions

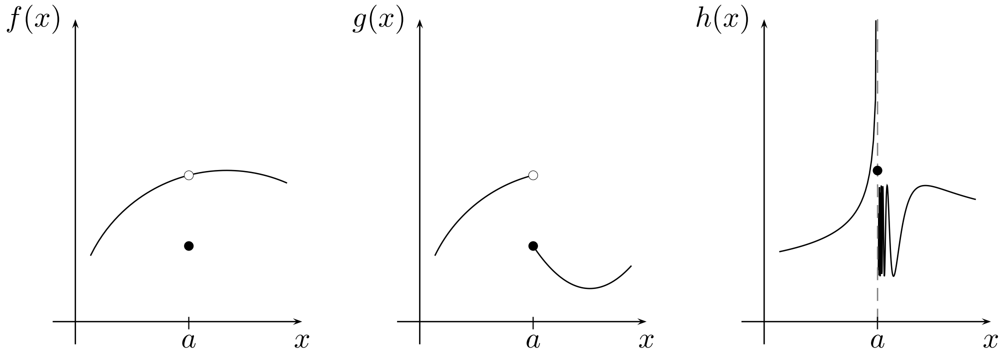
The sine, cosine, exponential, logarithmic and polynomial functions are examples of *continuous functions*. Intuitively, a function f is continuous on its domain if small changes in the input variable x produce small changes in $f(x)$ (assuming, of course, that x always remains within $\text{Dom}(f)$).

The intuitive notion that a function is continuous if 'its graph can be drawn without lifting the pencil off the page' works well for some functions, but not for others. For example, the tangent function is a continuous function, even though its graph (part of which is shown below) cannot be drawn without lifting the pencil off the page.



The break in the graph when, say, $x = -\pi/2$ or $x = \pi/2$, is a consequence of a break in the domain of the function rather than a discontinuity of the function. Similarly, the trigonometric functions sec, cosec and cot are continuous, as are the rational functions. The notion of continuity shall be discussed in more precise terms in Chapters 2 and 3.

Intuitively, a function f has a *discontinuity* at a point a if $a \in \text{Dom}(f)$ and there is a break in the graph of f at a. It is important to realise that there are different types of discontinuities. We compare three examples.



The functions f, g and h each have a discontinuity at a.

- For f, the discontinuity can be easily removed by redefining the value of $f(a)$ to 'plug up the hole' at a. This is an example of a *removable discontinuity*.

- For g, the discontinuity is a result of a finite jump in the graph of g at the point a. The discontinuity for g, which is an example of a *jump discontinuity*, is worse than that for f but not nearly as bad as that for h.

- For h, the discontinuity is a result of rapid oscillation on the right-hand side of a and an infinite jump on the left-hand side of a. This is an example of an *essential discontinuity*.

While students are not expected to learn or use the terms removable, jump and essential discontinuity, it is expected that they will appreciate that some discontinuities are worse than others.

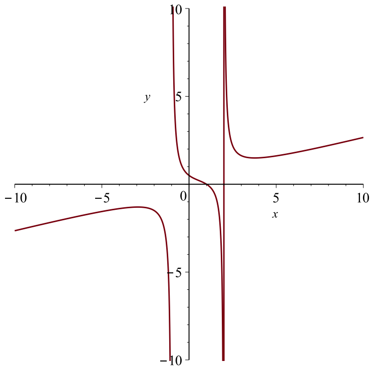
1.10 Maple notes

Many chapters in these notes conclude with a summary of relevant Maple commands. Other sources of information are the First Year Computing Notes and the online Maple `help` command. While Maple can relieve you of the tedium of calculations, it is not a substitute for understanding the mathematics behind the calculations, which is what you need in order to interpret Maple output intelligently.

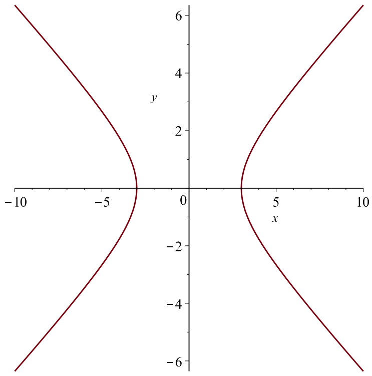
The following Maple command is relevant to the material of this chapter:

`plot(f(x),x=a..b);` draws a (two-dimensional) plot of the graph of $y = f(x)$ for $a \le x \le b$. To find out about the many options available, use `?plot`. Implicitly defined functions can be plotted using the `smartplot` command. For example,

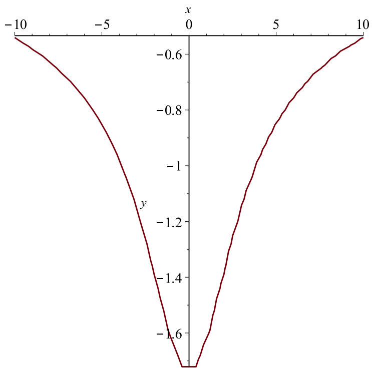
```
> # draw a rational function

> f:=x^5-2*x^4+3*x^3-6*x^2-4*x+8:

> g:=4*x^4-8*x^3-12*x^2+16*x+16:

> plot(f/g, x=-10..10,y=-10..10,numpoints=1000,discont=true);
```



```
> # draw a hyperbola, via implicit definition
> smartplot(x^2/9-y^2/4=1);
```



```
> # use the Maple plot command implicitplot --- need more plotting tools
> with(plots):
> implicitplot(y^5+x^2*y^3+16,x=-10..10,y=-2..0);
```



Problems for Chapter 1

Questions marked with [**R**] are routine, with [**H**] are harder and with [**X**] are for MATH1141 only. Questions marked with [**V**] have a video solution available from Moodle. You should make sure that you can do the easier questions before you tackle the more difficult questions.

Problems 1.1 : Sets of numbers

1. [**R**] Express the following sets in words. Graph the sets on the number line (if possible).

   a) $\{x \in \mathbb{Z} : -\pi < x < \pi\}$
   b) $\{x \in \mathbb{R} : x^2 - x - 1 < 0\}$
   c) $\{x \in \mathbb{Q} : x^2 = 2\}$

2. [**R**] Graph on the number line the following sets.

   a) $[3, \infty), \quad (-\infty, 3), \quad (-\infty, \infty), \quad (-3, 3]$
   b) $\{x : |x - 2| < 5\}$
   c) $\{x : x^2 + 4x - 5 > 0\}$

3. [**R**] Sketch the set of points (x, y) which satisfy the following relations.

   a) $0 \le y \le 2x$ and $0 \le x \le 2$ b) $y/2 \le x \le 2$ and $0 \le y \le 4$

Problems 1.2 : Solving inequalities and 1.3 : Absolute values

4. [**R**] Solve the following inequalities.

   a) $x(x-1) > 0$ b) $(x-1)(x-2) < 0$ c) $\dfrac{1}{x} > -\dfrac{1}{2}$
   d) $\dfrac{1}{1-x} > \dfrac{1}{2}$ e) $x \ge \dfrac{6}{x-1}$

5. [**R**] Solve the following inequalities.

   a) $|x+1| < 3$ b) $|x+2| > 3$ c) $|3x+2| < 1$ d) $\left|\dfrac{x-1}{x+1}\right| < 1$

6. [**R**] [**V**]

   a) By expanding $(x-y)^2$, prove that $x^2 + y^2 \ge 2xy$ for all real numbers x and y.
   b) Deduce that $\dfrac{a+b}{2} \ge \sqrt{ab}$ for all non-negative real numbers a and b. When does equality hold?
   c) Use the result above to find the minimum value of $y = x^2 + \dfrac{1}{x^2}$.

7. [**R**] True or false:

   a) If $a > b$ then $\frac{1}{a} < \frac{1}{b}$.
   b) If $a < b$ then $a^2 < b^2$.
   c) If $0 < a < b$ then $a^2 < b^2$.
   d) If $a^2 + b^2 = 0$ then $a = b = 0$.
   e) If $-1 < a < b$ then $a^2 < b^2$.

8. [**H**] Prove that $(x+y)^2 \geq 4xy$ and hence deduce that $\frac{1}{x^2} + \frac{1}{y^2} \geq \frac{4}{x^2+y^2}$.

9. [**H**] [**V**]

   a) Prove that $f(x) = 1 + x + x^2$ is positive for all real numbers x.
   b) By considering cases (or otherwise) prove that $1 + x + x^2 + x^3 + x^4$ is always positive.
   c) Generalise the above results.

Problems 1.4 : Functions

10. [**R**] Determine the (maximal) domain and corresponding range for each function f described below.

   a) $f(x) = \sqrt{5 - x^2}$
   b) $f(x) = \sqrt{x^2 - 5}$
   c) $f(x) = (x-8)^{-1/3}$
   d) $f(x) = \sqrt{x-1}$
   e) $f(x) = \dfrac{1}{\sqrt{x-1}}$
   f) $f(x) = \sqrt{\sin x}$
   g) $f(x) = \sqrt{1 - 2\sin x}$
   h) $f(x) = 1 + \tan^2 x$
   i) $f(x) = \begin{cases} \cos x & \text{if } x < 0 \\ \sqrt{1-x} & \text{if } 0 \leq x \leq 1 \\ |x| & \text{if } x > 1 \end{cases}$

11. [**R**] If $f(x) = x + 5$ and $g(x) = x^2 - 3$ then find

   a) $(g \circ f)(0)$ b) $(g \circ f)(x)$ c) $(f \circ g)(2)$ d) $(f \circ g)(x)$.

12. [**R**] If $f(x) = x - 1$ and $g(x) = \dfrac{1}{\sqrt{x-1}}$, then give the explicit forms of

   a) $(f+g)(x)$ b) $(fg)(x)$ c) $\left(\dfrac{f}{g}\right)(x)$ d) $(f \circ g)(x)$.

Problems 1.5 : Polynomials and rational functions, 1.6 : The trigonometric functions and 1.7 : The elementary functions

13. [**R**] Draw neat sketches (preferably without using calculus) of the graphs given by the following equations.

a) $y = x^2 - 5x + 6$ b) $y = 2x^3 - 16$ c) $y = \dfrac{4}{x-3}$

d) $y = 2e^{x-1}$ e) $y = \dfrac{1}{x^2+4}$ f) $y = 3\sin 2x$

g) $y = \sqrt{x} - 1$

14. [**R**]

  a) Sketch the graph of $y = \sqrt{x+1}$ and use your graph to sketch (on the same diagram) $y = \dfrac{1}{\sqrt{x+1}}$.
  b) Repeat for $y = x^2 - 4x + 3$.

15. [**R**] Sketch the graph of $y = x^2 - 7x - 8$ and hence sketch the graph of $y = |x^2 - 7x - 8|$.

16. [**R**] What range of values will $x^2 + 4$ take if $-2 \le x \le 3$?

17. [**R**] Use a graphical approach to solve $|2x - 5| = x + 2$.

18. [**R**]

  a) Show that if p and q are polynomials then $p \circ q$ is again a polynomial.
  b) [**H**] Is the same true for rational functions?

Problems 1.8 : Implicitly defined functions

19. [**R**] Sketch the graphs given by the following equations.

a) $\dfrac{x^2}{9} + \dfrac{y^2}{4} = 1$ b) $\dfrac{x^2}{9} - \dfrac{y^2}{4} = 1$

c) $4x^2 + 9y^2 = 36$ d) $\dfrac{y^2}{9} - \dfrac{x^2}{4} = 1$

Chapter 2

Limits

In many situations, it is desirable to know what the long term state of a system will be. For example, suppose that an initially unpolluted lake containing 10^9 litres of water has a river flowing through it at a rate of 10^6 litres per day. A factory is built next to the lake and discharges 10^4 litres of pollution per day. By making some simple assumptions, it can be shown that the amount $P(t)$ of pollution in the lake after t days of the factory's operation is given by

$$P(t) = \frac{10^9}{101}\bigl(1 - e^{-101t/10^5}\bigr). \tag{2.1}$$

Environmental authorities want to know what the long term level of the pollution will be. To see whether the pollution in the lake eventually stabilises, and what the level of pollution will be, one studies the behaviour of $P(t)$ as $t \to \infty$.

Understanding the limiting behaviour of functions at infinity is important for practical and mathematical reasons. That is what we study in the next four sections.

2.1 Limits of functions at infinity

(Ref: SH10 §4.7)

In this section, we examine some techniques for calculating limits of the form $\lim\limits_{x\to\infty} f(x)$.

2.1.1 Basic rules for limits

There are a few basic things to bear in mind when working with limits.

- If $f(x)$ gets arbitrarily close to some real number L as x tends to infinity, then we write

  $$\lim_{x\to\infty} f(x) = L.$$

  We can also write $f(x) \to L$ as $x \to \infty$.

- If $f(x)$ gets arbitrarily large (that is, approaches ∞) as x tends to ∞ then we write

  $$f(x) \to \infty \text{ as } x \to \infty.$$

  We do *not* write $\lim\limits_{x\to\infty} f(x) = \infty$ since ∞ is not a real number. (The lim notation is only used when the limit exists and is a real number.)

- If $f(x) \to \infty$ as $x \to \infty$ then
$$\lim_{x\to\infty} \frac{1}{f(x)} = 0.$$
This is intuitively obvious and gives limits such as
$$\lim_{x\to\infty} \frac{1}{x} = 0, \qquad \lim_{x\to\infty} \frac{5}{4x^6} = 0, \qquad \lim_{x\to\infty} e^{-x} = 0$$
(for the last example, recall that $e^{-x} = \frac{1}{e^x}$).

The following proposition shows that the arithmetic of limits behaves nicely in many situations.

Proposition 2.1.1. *Suppose that* $\lim_{x\to\infty} f(x)$ *and* $\lim_{x\to\infty} g(x)$ *exist and are finite real numbers. Then*

(i) $\lim_{x\to\infty} [f(x) + g(x)] = \lim_{x\to\infty} f(x) + \lim_{x\to\infty} g(x)$

(ii) $\lim_{x\to\infty} [f(x) - g(x)] = \lim_{x\to\infty} f(x) - \lim_{x\to\infty} g(x)$

(iii) $\lim_{x\to\infty} [f(x)g(x)] = [\lim_{x\to\infty} f(x)][\lim_{x\to\infty} g(x)]$

(iv) $\lim_{x\to\infty} \dfrac{f(x)}{g(x)} = \dfrac{\lim_{x\to\infty} f(x)}{\lim_{x\to\infty} g(x)}$, *provided that* $\lim_{x\to\infty} g(x) \neq 0$.

Part of this proposition will be proved in Section 2.4. For now, we demonstrate how the proposition is used to calculate limits.

Example 2.1.2. Calculate
$$\lim_{x\to\infty} \frac{3 - 1/x^2}{2 + e^{-x}}$$
if the limit exists. Justify each step of your working with reference to Proposition 2.1.1.

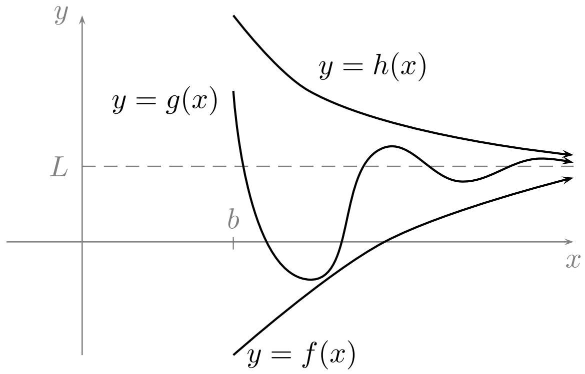
Solution.
$$\begin{aligned}
\lim_{x\to\infty} \frac{3 - 1/x^2}{2 + e^{-x}} &= \frac{\lim_{x\to\infty} (3 - 1/x^2)}{\lim_{x\to\infty} (2 + e^{-x})} && \text{(rule (iv))} \\
&= \frac{\lim_{x\to\infty} 3 - \lim_{x\to\infty} 1/x^2}{\lim_{x\to\infty} 2 + \lim_{x\to\infty} e^{-x}} && \text{(rules (i) and (ii))} \\
&= \frac{3 - 0}{2 + 0} \\
&= \frac{3}{2}.
\end{aligned}$$

□

In practice, you may use the rules of Proposition 2.1.1 without referring to them, and leave out some of the steps shown above.

2.1.2 The pinching theorem

The idea behind the following theorem is intuitively obvious. Suppose, as illustrated in the diagram below, that the graph of a function h always lies above the graph of a function f. Suppose also that f and h have the same limit as $x \to \infty$. If the graph of a function g always lies between the graphs of h and f (so that g is 'pinched' between f and h) then g has the same limit also.



This idea is stated precisely below.

Theorem 2.1.3 (The pinching theorem)**.** *Suppose that f, g and h are all defined on the interval (b, ∞), where $b \in \mathbb{R}$. If*

$$f(x) \le g(x) \le h(x) \qquad \forall x \in (b, \infty)$$

and

$$\lim_{x\to\infty} f(x) = \lim_{x\to\infty} h(x) = L$$

then

$$\lim_{x\to\infty} g(x) = L.$$

One can modify the theorem to cover the case when $x \to -\infty$.

Example 2.1.4. Use the pinching theorem to find $\displaystyle\lim_{x\to\infty} \frac{\sin x}{x}$.

Solution. We begin with the basic inequality

$$-1 \le \sin x \le 1,$$

which is valid for every real number x. If we restrict x to the interval $(0, \infty)$ then we have

$$\frac{-1}{x} \le \frac{\sin x}{x} \le \frac{1}{x}.$$

Now

$$\lim_{x\to\infty} -\frac{1}{x} = \lim_{x\to\infty} \frac{1}{x} = 0,$$

and so

$$\lim_{x\to\infty} \frac{\sin x}{x} = 0$$

by the pinching theorem. $\square$

2.1.3 Limits of the form $f(x)/g(x)$

Suppose that we want to calculate a limit of the form

$$\lim_{x\to\infty} \frac{f(x)}{g(x)}$$

where both $f(x)$ and $g(x)$ tend to infinity as $x \to \infty$. In this situation, we cannot apply the rules of Proposition 2.1.1 directly because the numerator and denominator don't have finite limits. Instead, we find an equivalent form of $\frac{f(x)}{g(x)}$ for which Proposition 2.1.1 can be applied. The key idea is to divide both f and g by the *fastest growing term* appearing in the denominator g.

Example 2.1.5. Evaluate $\displaystyle\lim_{x\to\infty} \frac{4x^2-5}{2x^2+3x}$, if it exists.

Solution. There are two terms appearing in the denominator: $2x^2$ and $3x$. As $x \to \infty$, the term which grows fastest is the one involving x^2. So we divide the numerator and denominator by x^2 to evaluate the limit:

$$\begin{aligned} \frac{4x^2-5}{2x^2+3x} &= \frac{4-5/x^2}{2+3/x} \\ &\to \frac{4-0}{2+0} \end{aligned}$$

as $x \to \infty$. Therefore

$$\lim_{x\to\infty} \frac{4x^2-5}{2x^2+3x} = 2.$$

□

In general, we divide the numerator and denominator by the highest power of x in the *denominator.*

Example 2.1.6. Find $\displaystyle\lim_{x\to\infty} \frac{5x^3+6x^2-4\sin x}{\cos 3x + 5x - 2x^3}$.

Example 2.1.7. Find $\displaystyle\lim_{x\to\infty} \frac{x^2+3x}{\sqrt{2x^4+3}-4x}$.

2.1.4 Limits of the form $\sqrt{f(x)} - \sqrt{g(x)}$

Consider the limit of $\sqrt{x+5} - \sqrt{x+2}$ as $x \to \infty$. Since both $\sqrt{x+5}$ and $\sqrt{x+2}$ tend to infinity (rather than a finite number) as $x \to \infty$, one cannot apply Proposition 2.1.1. As the next example shows, a simple algebraic trick is used to overcome this difficulty.

Example 2.1.8. Evaluate $\displaystyle\lim_{x\to\infty} \left(\sqrt{x+5} - \sqrt{x+2}\right)$, if it exists.

Solution. The algebraic trick is to multiply both the numerator and denominator by $\sqrt{x+5} + \sqrt{x+2}$ and expand the numerator as a difference of squares:

$$\begin{aligned} \sqrt{x+5} - \sqrt{x+2} &= \frac{(\sqrt{x+5} - \sqrt{x+2})(\sqrt{x+5} + \sqrt{x+2})}{\sqrt{x+5} + \sqrt{x+2}} \\ &= \frac{(x+5)-(x+2)}{\sqrt{x+5} + \sqrt{x+2}} \\ &= \frac{3}{\sqrt{x+5} + \sqrt{x+2}} \\ &\to 0 \end{aligned}$$

as $x \to \infty$. $\square$

One should not be fooled into thinking, on the basis of the above example, that all limits of this type are zero.

Example 2.1.9. Show that $\lim_{x\to\infty}\left(\sqrt{x^2+x}-x\right) = \dfrac{1}{2}$.

2.1.5 Indeterminate forms

We have already mentioned that the rules of Proposition 2.1.1 only apply when the limits involved are finite. In Section 2.1.3, we studied limits of the type

$$\lim_{x\to\infty} \frac{f(x)}{g(x)},$$

when $f(x) \to \infty$ and $g(x) \to \infty$ as $x \to \infty$. We say that such a limit is of the form $\frac{\infty}{\infty}$. We cannot say in advance whether or not a limit of the form $\frac{\infty}{\infty}$ exists, and if it does exist, what its value is.

For example, while the following limits have the form $\frac{\infty}{\infty}$, each displays very different limiting behaviour as $x \to \infty$:

- $\frac{x^2}{x} \to \infty$ as $x \to \infty$
- $\frac{x}{x^2} \to 0$ as $x \to \infty$
- $\frac{2x^2}{x^2} \to 2$ as $x \to \infty$
- $\frac{x^2}{2x^2} \to \frac{1}{2}$ as $x \to \infty$.

Because we cannot determine in advance what kind of limiting behaviour something of the form $\frac{\infty}{\infty}$ has, we say that $\frac{\infty}{\infty}$ is an *indeterminate form*.

Other types of indeterminate forms (for example, limits of the form $\frac{0}{0}$ and $\infty-\infty$) arise naturally in many applications. Techniques for evaluating such limits will be developed at various stages of the course.

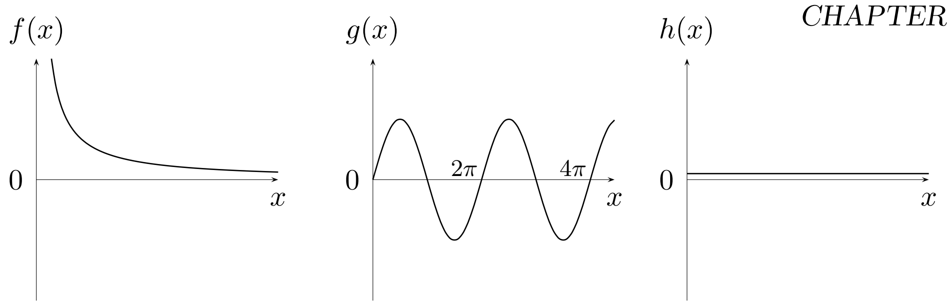
2.2 The definition of $\lim_{x\to\infty} f(x)$

Up to this point, we have treated limits in a rather intuitive way. It is now time to establish a rigorous basis for our treatment of limits. That is, we seek a definition for $\lim_{x\to\infty} f(x) = L$ that is mathemetically precise and agrees with our intuitive notion of limits. How do we do this? A good strategy is to think deeply about a few simple examples before considering the general case.

Consider the real-valued functions f, g and h defined by

$$f(x) = \frac{1}{x}, \qquad g(x) = \sin x \qquad \text{and} \qquad h(x) = 0.05$$

whenever $x > 0$. Their graphs are sketched below.



Intuitively, we see that $f(x)$ tends to 0 as x tends to infinity, while neither g nor h have this property. What is it about f that distinguishes its limiting behaviour from that of g or h?

The above graph shows that the distance between $f(x)$ and its limit 0 is small whenever x is large. However, to get to the core difference between f and the other two functions, we must be more precise.

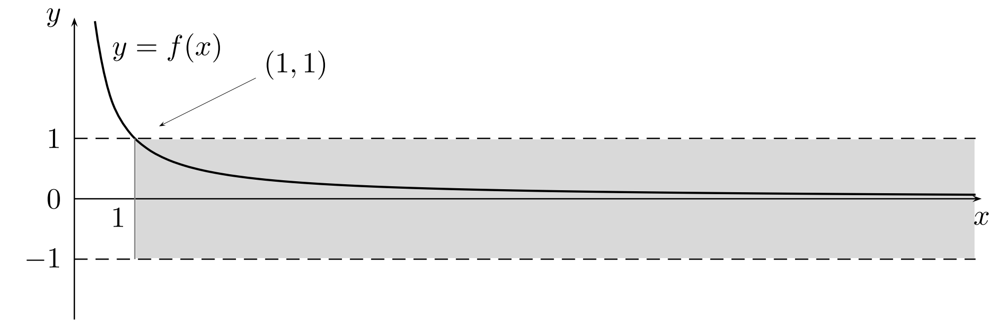
- The distance between $h(x)$ and 0 is also small, but its limit is not 0. The reason why $f(x)$ tends to 0 as $x \to \infty$ is that the distance between $f(x)$ and 0 can be made *as small as we like* whenever x is large enough.

- The distance between $g(x)$ and 0 can also be made as small as we like, by taking x sufficiently close to an integer multiple of π. However the limit of $g(x)$ is not 0. The difference between f and g is that the distance between $f(x)$ and 0 *remains* small for *all* sufficiently large values of x. The same is not true for g.

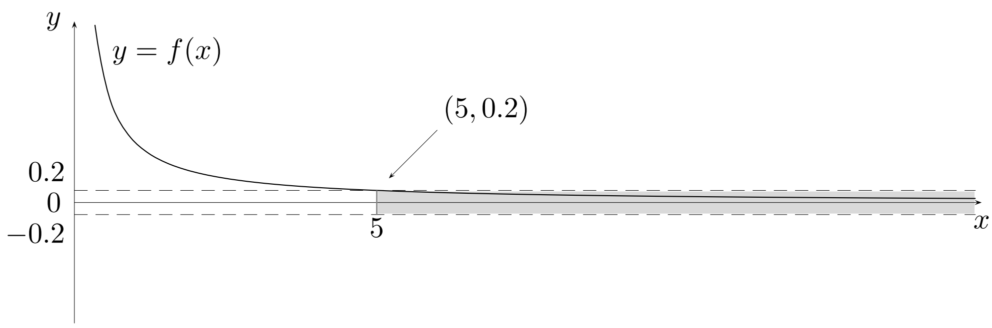
In summary, $f(x)$ tends to 0 as $x \to \infty$ because the distance between $f(x)$ and 0 can be made *as small as we like* for *all* values of x that are sufficiently large.

The next step in crafting a good definition is to express phrases like 'can be made as small as we like' and 'sufficiently large' in concrete mathematical language. The following list is one attempt to do this. (As a reminder, $f(x) = 1/x$.)

- The distance between $f(x)$ and 0 is less than 1 whenever $x > 1$.



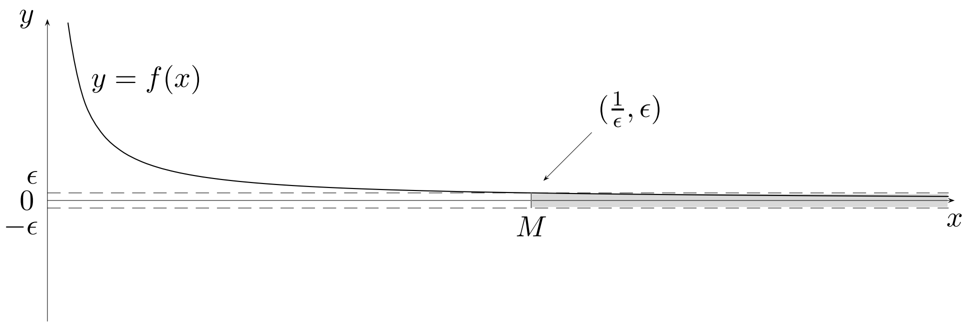
- The distance between $f(x)$ and 0 is less than 0.2 whenever $x > 5$.



- The distance between $f(x)$ and 0 is less than 0.1 whenever $x > 10$.
- The distance between $f(x)$ and 0 is less than 0.01 whenever $x > 100$.
- The distance between $f(x)$ and 0 is less than 0.0001 whenever $x > 10000$.
- *etc.*

Of course, to express the fact that 'the distance between $f(x)$ and 0 can be made as small as we like' in this way would require an infinite list. Instead, we say that, for *every* small positive number ϵ, there is real number M such that

- the distance between $f(x)$ and 0 is less than ϵ whenever $x > M$.



By looking at the graph we can see that $M = \frac{1}{\epsilon}$. (The symbol ϵ is the Greek letter 'epsilon' and is traditionally used by mathematicians in this context.)

So far we have agreed that $f(x) \to 0$ as $x \to \infty$ precisely because

> for every positive real number ϵ, there is a real number M such that the distance between $f(x)$ and 0 is less than ϵ whenever $x > M$.

This statement forms a basis for defining what is meant by $\lim_{x\to\infty} f(x) = 0$. A general definition for what is meant by $\lim_{x\to\infty} f(x) = L$ is now obtained by

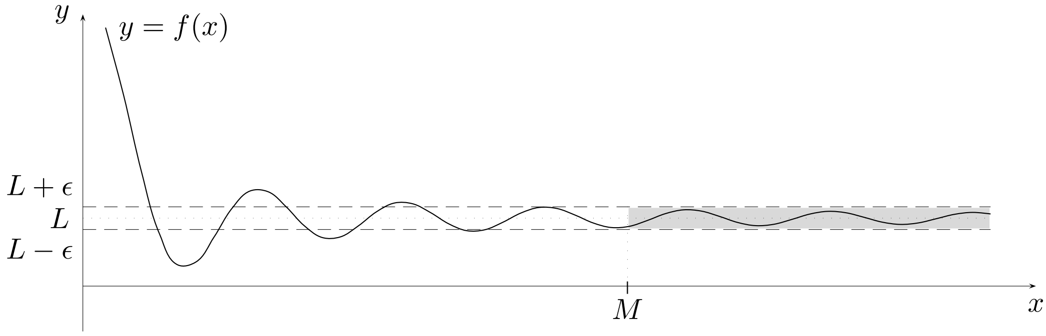
- replacing the limit 0 with the limit L, and then
- replacing 'the distance between $f(x)$ and L' with $|f(x) - L|$ (see Section 1.3).

We are now ready to give a general definition for the limit of a function at infinity.

Definition 2.2.1. Suppose that L is a real number and f is a real-valued function defined on some interval (b, ∞). We say that $\lim_{x\to\infty} f(x) = L$ if

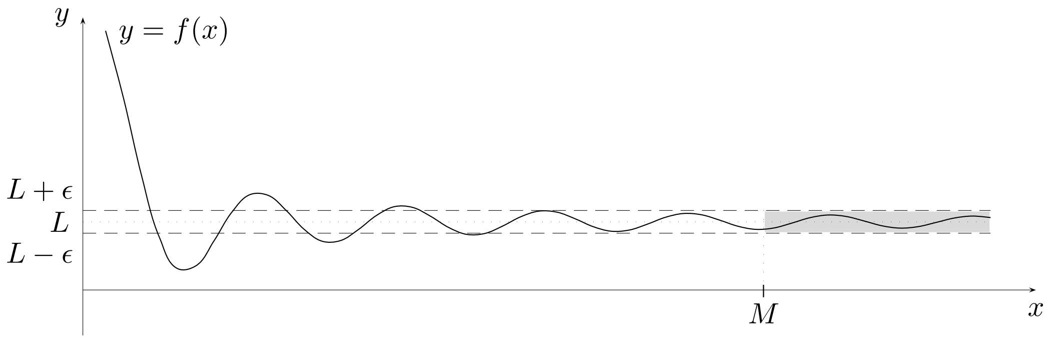
> for every positive real number ϵ, there is a real number M such that if $x > M$ then $|f(x) - L| < \epsilon$.

Remark 2.2.2. Suppose that $\lim_{x\to\infty} f(x) = L$. Definition 2.2.1 can be interpreted geometrically in the following way. For every small number ϵ, you can find a marker M on the x-axis such that the distance between $f(x)$ and the limit L is less than ϵ whenever x lies to the right of M.



Remark 2.2.3. In Definition 2.2.1, the number M depends on ϵ. In general, the smaller the value of ϵ, the larger the value of M. We note, however, that for any particular value of ϵ the choice of M is not unique. For example, the M chosen for the above diagram is different from the M chosen for the diagram below, but both choices guarantee that

$$|f(x) - L| < \epsilon \qquad \text{whenever} \qquad x > M.$$



Remark 2.2.4. Students who find Definition 2.2.1 difficult to understand can take comfort in the fact that a rigorous formulation of limits evaded mathematicians for over 2000 years. The problems and paradoxes involving limiting processes tabled around 450 BC by the ancient Greek philosopher Zeno of Elea were only satisfactorily resolved in the nineteenth century, when a rigorous definition of the limit was given by Bolzano (in 1817) and Weierstrass (in the 1850s).

2.3 Proving that $\lim_{x\to\infty} f(x) = L$ using the limit definition

In the previous section we gave a definition for $\lim_{x\to\infty} f(x) = L$. In this section this definition is used to prove limits that were previously taken for granted.

Example 2.3.1. Prove from Definition 2.2.1 that $\lim_{x\to\infty} \dfrac{4}{x^2} = 0$.

Proof. Suppose that $f(x) = 4/x^2$, $L = 0$ and $\epsilon > 0$. We begin by calculating the distance between the function and its limit:

$$\begin{aligned} |f(x) - L| &= \left| \frac{4}{x^2} - 0 \right| \\ &= \frac{4}{x^2}. \end{aligned} \tag{2.2}$$

We need to find a condition on x such that $|f(x) - L| < \epsilon$. By (2.2), we require that

$$\frac{4}{x^2} < \epsilon.$$

By rearrangement, we see that this is satisfied whenever

$$x > \frac{2}{\sqrt{\epsilon}}.$$

So far we have shown that $|f(x) - L| < \epsilon$ whenever $x > \frac{2}{\sqrt{\epsilon}}$. Hence if

$$M = \frac{2}{\sqrt{\epsilon}}. \tag{2.3}$$

then $|f(x) - L| < \epsilon$ whenever $x > M$. This completes the proof. □

The above proof can be streamlined by omitting some explanatory commentary. In the next example, explanatory comments (which are not part of the proof) are inserted inside square brackets.

Example 2.3.2. Prove that

$$\lim_{x\to\infty} \frac{5x}{x+3} = 5$$

by using Definition 2.2.1.

Proof. Suppose that $\epsilon > 0$. For convenience, suppose also that $x > 0$. [This is allowed since we are considering the behaviour of $f(x)$ for large positive values of x.] Now

$$\begin{aligned}
|f(x) - L| &= \left|\frac{5x}{x+3} - 5\right| \\
&= \left|\frac{5x - 5(x+3)}{x+3}\right| \\
&= \left|\frac{-15}{x+3}\right| \\
&= \frac{15}{x+3} && \text{[since } x > 0\text{]} \\
&< \frac{15}{x} && \text{[to make algebra simpler later on].}
\end{aligned}$$

In summary,

$$|f(x) - L| < \frac{15}{x}.$$

[This inequality gives an *upper bound* for the distance between $f(x)$ and L.] Hence $|f(x) - L| < \epsilon$ whenever

$$\frac{15}{x} < \epsilon.$$

This condition is equivalent to

$$x > \frac{15}{\epsilon}.$$

Hence if $M = \frac{15}{\epsilon}$ then

$$|f(x) - L| < \epsilon \qquad \text{whenever} \qquad x > M,$$

as required. $\square$

Before attempting one more proof, we reflect on the general strategy employed. Given ϵ, we need to find a number M such that

$$|f(x) - L| < \epsilon \qquad \text{whenever} \qquad x > M.$$

The number M can be found by following the procedure below.

1. Find a good upper bound for $|f(x) - L|$.
2. Find a simple condition on x such that this upper bound is less than ϵ.
3. Use this condition to state an appropriate value for M (in terms of ϵ).

Example 2.3.3. Prove that

$$\lim_{x\to\infty} \frac{\sin x}{9x^2 + e^{3x}} = 0$$

by using Definition 2.2.1.

Proof. Suppose that $\epsilon > 0$. Now

$$\begin{aligned} |f(x) - L| &= \left| \frac{\sin x}{9x^2 + e^{3x}} - 0 \right| \\ &= \frac{|\sin x|}{|9x^2 + e^{3x}|} \\ &\le \frac{1}{9x^2 + e^{3x}} && \text{[using the fact that } -1 \le \sin x \le 1] \\ &< \frac{1}{9x^2} && \text{[decreasing the denominator increases the RHS].} \end{aligned}$$

In summary,

$$|f(x) - L| < \frac{1}{9x^2}.$$

[This is our upper bound for $|f(x) - L|$.] Now

$$\frac{1}{9x^2} < \epsilon$$

whenever

$$x > \frac{1}{3\sqrt{\epsilon}}.$$

Set

$$M = \frac{1}{3\sqrt{\epsilon}}.$$

Then $|f(x) - L| < \epsilon$ whenever $x > M$, as required. $\square$

Remark 2.3.4. As pointed out in Remark 2.2.3, the value of M is not unique. For instance, the upper bound for $|f(x) - L|$ in our solution to Example 2.3.3 was $1/(9x^2)$ and the corresponding value for M was $1/(3\sqrt{\epsilon})$. If instead we use the upper bound $1/e^{3x}$ then the corresponding value for M is $\frac{1}{3} \ln \frac{1}{\epsilon}$.

2.4 Proofs of basic limit results (MATH1141 only)

By using Definition 2.2.1, one can now prove the limit rules stated in Section 2.1. The selection of proofs given in this section is designed to give MATH1141 students a taste of how such results are proved.

Proof of Proposition 2.1.1 (i). Suppose that

$$\lim_{x\to\infty} f(x) = L_1 \qquad \text{and} \qquad \lim_{x\to\infty} g(x) = L_2.$$

We need to show, using Definition 2.2.1, that

$$\lim_{x\to\infty} \big(f(x) + g(x)\big) = L_1 + L_2.$$

We begin by fixing a positive real number ϵ.

Since $\lim\limits_{x\to\infty} f(x) = L$, there is a positive M_1 such that

$$|f(x) - L_1| < \frac{\epsilon}{2} \qquad \text{whenever} \qquad x > M_1.$$

Similarly, there is a positive M_2 such that

$$|g(x) - L_2| < \frac{\epsilon}{2} \qquad \text{whenever} \qquad x > M_2.$$

Hence, by the triangle inequality,

$$\begin{aligned}\big|(f(x) + g(x)) - (L_1 + L_2)\big| &\le |f(x) - L_1| + |g(x) - L_2| \\ &< \frac{\epsilon}{2} + \frac{\epsilon}{2} \\ &= \epsilon\end{aligned}$$

whenever both $x > M_1$ and $x > M_2$. Now set $M = \max\{M_1, M_2\}$ so that

$$\big|(f(x) + g(x)) - (L_1 + L_2)\big| < \epsilon \qquad \text{whenever} \qquad x > M.$$

This proves that $\lim\limits_{x\to\infty} \big(f(x) + g(x)\big) = L_1 + L_2$. □

The other rules in Proposition 2.1.1 can be proved similarly, except that sometimes the desired inequalities are harder to obtain. We move on to prove the pinching theorem.

Proof of Theorem 2.1.3. Suppose that

$$f(x) \le g(x) \le h(x) \qquad \forall x \in (b, \infty)$$

and that

$$\lim_{x\to\infty} f(x) = \lim_{x\to\infty} h(x) = L.$$

Fix a positive real number ϵ.

Since $\lim\limits_{x\to\infty} f(x) = \lim\limits_{x\to\infty} h(x) = L$, there are real numbers M_1 and M_2, both greater than b, such that

$$|f(x) - L| < \epsilon \qquad \text{whenever} \qquad x > M_1$$

and

$$|h(x) - L| < \epsilon \qquad \text{whenever} \qquad x > M_2.$$

In other words,

$$L - \epsilon < f(x) < L + \epsilon \qquad \text{whenever} \qquad x > M_1$$

and

$$L - \epsilon < h(x) < L + \epsilon \qquad \text{whenever} \qquad x > M_2.$$

Hence

$$L - \epsilon < f(x) \le g(x) \le h(x) < L + \epsilon$$

whenever $x > \max\{M_1, M_2\}$. This shows that

$$|g(x) - L| < \epsilon \qquad \text{whenever} \qquad x > \max\{M_1, M_2\},$$

completing the proof. ☐

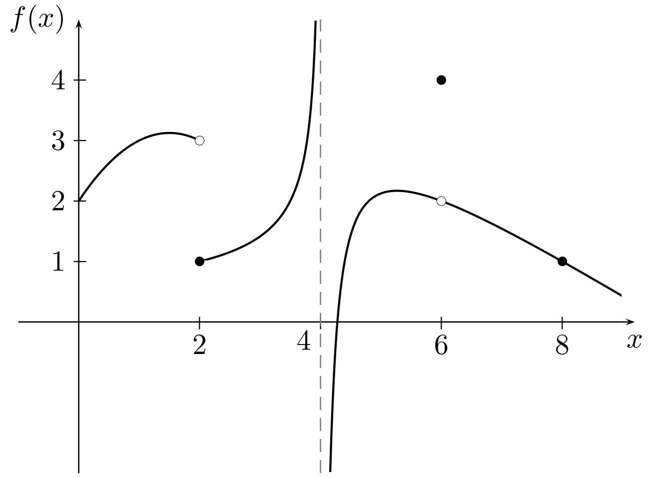
2.5 Limits of functions at a point

(Ref: SH10 §§2.1–2.3, 2.5)

So far we have only discussed limits of the form $\lim_{x\to\infty} f(x)$. In this section we study limits of the form $\lim_{x\to a} f(x)$, where a is a finite real number.

2.5.1 Left-hand, right-hand and two-sided limits

Consider the function f whose graph is shown below.



With reference to this graph, we will discuss the behaviour of $f(x)$ when x is near the points 2, 4, 6 and 8.

- As x approaches 2 from the left-hand side, $f(x)$ approaches 3. In this situation, we write

$$\lim_{x\to 2^-} f(x) = 3$$

and say that 'the limit of $f(x)$ as x approaches 2 from the left is 3.' (Note that $f(2) = 1 \neq 3$. One should not confuse a limit with a function value.) As x approaches 8 from the left-hand side, $f(x)$ approaches 1 and hence we write

$$\lim_{x \to 8^-} f(x) = 1.$$

Both of these limits are called *left-hand limits.*

- As x approaches 2 from the right-hand side, $f(x)$ approaches 1. In this situation, we write

$$\lim_{x \to 2^+} f(x) = 1$$

and say that 'the limit of $f(x)$ as x approaches 2 from the right is 1.' As x approaches 6 from the right-hand side, $f(x)$ approaches 2 and hence we write

$$\lim_{x \to 6^+} f(x) = 2.$$

Both of these limits are called *right-hand limits.*

- As x approaches 4 from the left-hand side, $f(x)$ is positive and grows large without bound. In this case, $\lim_{x \to 4^-} f(x)$ does not exist because $f(x)$ does not approach a real number. However, we can write

$$f(x) \to \infty \qquad \text{as} \qquad x \to 4^-.$$

- As x approaches 4 from the right-hand side, $f(x)$ is negative and grows large without bound. In this case, $\lim_{x \to 4^+} f(x)$ does not exist but we can we write

$$f(x) \to -\infty \qquad \text{as} \qquad x \to 4^+.$$

Once an understanding of left- and right-hand limits is grasped, we can talk about *two-sided limits.*

Definition 2.5.1. If the left-hand limit $\lim_{x \to a^-} f(x)$ and the right-hand limit $\lim_{x \to a^+} f(x)$ both exist and equal the same real number L, then we say that *the limit of $f(x)$ as $x \to a$ exists and is equal to L*, and we write

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

If any one of these conditions fails then we say that $\lim_{x \to a} f(x)$ does *not* exist.

Recall the graph of the function f introduced at the beginning of this section.

- The two-sided limit $\lim_{x \to 2} f(x)$ does not exist because

$$\lim_{x \to 2^-} f(x) \neq \lim_{x \to 2^+} f(x).$$

- The two-sided limit $\lim_{x \to 4} f(x)$ does not exist because $\lim_{x \to 4^-} f(x)$ does not exist.

- Since the left-hand and right-hand limits both exist at 6 and

$$\lim_{x\to 6^-} f(x) = \lim_{x\to 6^+} f(x) = 2,$$

the two-sided limit $\lim_{x\to 6} f(x)$ exists and

$$\lim_{x\to 6} f(x) = 2.$$

Note that $f(6) = 4$ and so

$$\lim_{x\to 6} f(x) \neq f(6).$$

That is, the two-sided limit of a function at a point need not equal the value of the function at that point.

- Since the left-hand and right-hand limits at 8 both exist and equal 1, the two-sided limit $\lim_{x\to 8} f(x)$ exists and

$$\lim_{x\to 8} f(x) = 1.$$

In this case, the two-sided limit of f at 8 equals the value of the function at 8. That is,

$$\lim_{x\to 8} f(x) = f(8).$$

The last two points emphasise an important feature of limits and functions. If $\lim_{x\to a} f(x) = L$ then $f(x)$ approaches L as x tends to a. The function f may not be defined at a, or if it is defined at a, we can have the situation where $f(a) = L$ or where $f(a) \neq L$. In other words, the value of $f(a)$ does not determine the value of $\lim_{x\to a} f(x)$. As far as the limit is concerned, all that matters is how f behaves when it is *very close* to the point a.

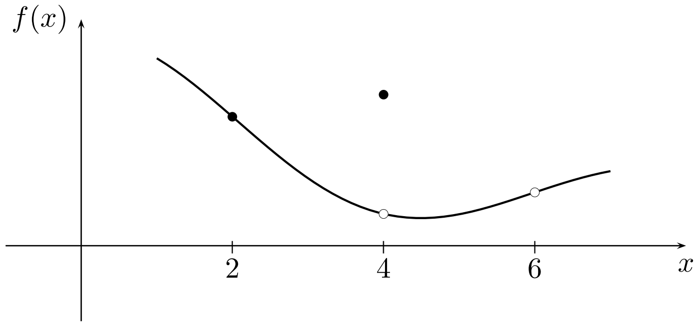
Remark 2.5.2. In this section we have not given a proper definition for limits of the form $\lim_{x\to a^-} f(x)$ or $\lim_{x\to a^+} f(x)$. At this stage it is enough to know that such a rigorous definition exists and is similar to the definition for limits of the form $\lim_{x\to\infty} f(x)$ (see Definition 2.2.1). This definition can be used to prove the basic limit results presented in the remainder of this chapter.

2.5.2 Limits and continuous functions

Using limits we can now give a definition for continuity.

Definition 2.5.3. Suppose that f is defined on some open interval containing the point a. If $\lim_{x\to a} f(x) = f(a)$, then we say that f is *continuous* at a; otherwise, we say that f is *discontinuous* at a.

To illustrate, consider the function f whose graph is sketch below.



The function f is continuous at 2, discontinuous at 4 and undefined at 6. (We do not say that f is discontinuous at 6 since $6 \notin \mathrm{Dom}(f)$.)

We now extend our formal notion of continuity at a point to continuity on the real line.

Definition 2.5.4. If $f : \mathbb{R} \to \mathbb{R}$ is continuous at every point a in $\mathbb{R}$ then we say that f is *continuous everywhere.*

If a function $f : \mathbb{R} \to \mathbb{R}$ is continuous everywhere then it has the property that its graph can be drawn on the Cartesian plane as an unbroken curve. This coincides with intuitive notions about continuity gained in high school.

As mentioned in Section 1.9, polynomials, rational functions, the trigonometric functions, exponentials and logarithms are continuous at every point in the respective domains. Moreover, if f is continuous at a point a then $|f|$ is also continuous at a (see Section 1.7 for the definition of $|f|$). The proof of these facts requires a rigorous definition of the limit at a point, and, as already mentioned, such a definition is not presented in this course. However, we will discuss aspects of the proof in Chapter 3.

A consequence of the above definition is that both one-sided and two-sided limits of f at a point a are easy to evaluate provided that (a) $a \in \mathrm{Dom}(f)$ and (b) we know that f is a continuous at a. Hence, in view the facts mentioned above,

- $\lim_{x \to 2} \sin(x) = \sin(2)$
- $\lim_{x \to 3^+} (x^2 - x + 1) = \lim_{x \to 3} (x^2 - x + 1) = 3^2 - 3 + 1 = 7$
- $\lim_{x \to 4} |x - 10| = |4 - 10| = 6.$

Remark 2.5.5. Continuity is a deep property for a function to have. Contrary to the impression that many students form from their study of functions at school, most functions are *not* continuous. Three interesting functions with discontinuities are mentioned below.

- The function $f : \mathbb{R} \to \mathbb{R}$, defined by

  $$f(x) = \begin{cases} 1 & \text{if } x \in \mathbb{Q} \\ 0 & \text{if } x \notin \mathbb{Q}, \end{cases}$$

  is discontinuous at every point in its domain.

- The function $f : \mathbb{R} \to \mathbb{R}$, defined by

$$f(x) = \begin{cases} x & \text{if } x \in \mathbb{Q} \\ -x & \text{if } x \notin \mathbb{Q}, \end{cases}$$

  is continuous at 0 but discontinuous everywhere else.

- It is not too difficult to construct a function $f : \mathbb{R} \to \mathbb{R}$ which is continuous at every irrational point but discontinuous at every rational point.

Hence, one should not assume that the graph of a general function $f : \mathbb{R} \to \mathbb{R}$ is continuous at even a majority of points in its domain. Continuous functions are quite special and deserve much attention. We will study them in greater depth in Chapter 3.

2.5.3 Rules for limits at a point

The following proposition, which is analogous to Proposition 2.1.1, shows that limits at a point have nice arithmetic properties.

Proposition 2.5.6. *Suppose that $a \in \mathbb{R}$ and that $\lim_{x\to a} f(x)$ and $\lim_{x\to a} g(x)$ exist and are finite real numbers. Then*

(i) $\lim_{x\to a}(f(x) + g(x)) = \lim_{x\to a} f(x) + \lim_{x\to a} g(x)$

(ii) $\lim_{x\to a}(f(x) - g(x)) = \lim_{x\to a} f(x) - \lim_{x\to a} g(x)$

(iii) $\lim_{x\to a} f(x)g(x) = \lim_{x\to a} f(x) \lim_{x\to a} g(x)$

(iv) $\displaystyle\lim_{x\to a} \frac{f(x)}{g(x)} = \frac{\lim_{x\to a} f(x)}{\lim_{x\to a} g(x)}$, *provided that* $\lim_{x\to a} g(x) \neq 0$.

A similar set of rules hold for left- and right-hand limits.

Proposition 2.5.6 says that limits interact nicely with function addition, subtraction, multiplication and division. The next proposition shows that limits interact nicely with function composition.

Proposition 2.5.7. *If $\lim_{x\to a} f(x) = L$ and g is continuous at L then*

$$\lim_{x\to a} g(f(x)) = g\Big(\lim_{x\to a} f(x)\Big).$$

If the functions f and g are continuous everywhere, then Proposition 2.5.7 implies that

$$\lim_{x\to a} g(f(x)) = g\Big(\lim_{x\to a} f(x)\Big)$$

for any point a in $\mathbb{R}$.

Example 2.5.8. Evaluate the limit

$$\lim_{x\to\sqrt{\pi/3}} \cos\left(x^2 + \tfrac{\pi}{6}\right),$$

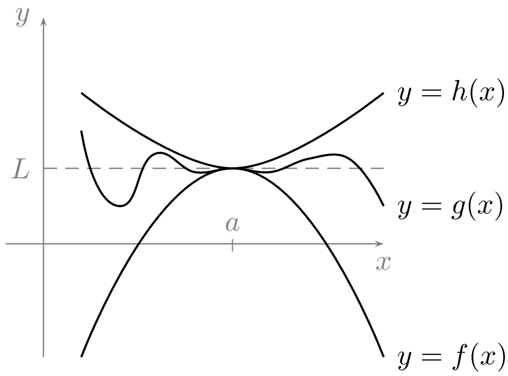
justifying each step of your calculation.

Solution. Since the cosine function and polynomials are continuous everywhere,

$$\begin{aligned}
\lim_{x\to\sqrt{\pi/3}} \cos\left(x^2+\frac{\pi}{6}\right) &= \cos\left(\lim_{x\to\sqrt{\pi/3}} x^2+\frac{\pi}{6}\right) && \text{(by Proposition 2.5.7)}\\
&= \cos\left(\frac{\pi}{3}+\frac{\pi}{6}\right) && \text{(since polynomials are continuous)}\\
&= \cos\left(\frac{\pi}{2}\right)\\
&= 0.
\end{aligned}$$

□

A version of the pinching theorem for two-sided limits is given below. Similar versions exist for left- and right-hand limits.



Theorem 2.5.9 (The pinching theorem)**.** *Let I be an open interval containing the point a. Suppose that f, g and h are all defined on I except possibly at the point a. If*

$$f(x) \le g(x) \le h(x) \qquad \forall x \in I$$

and

$$\lim_{x\to a} f(x) = \lim_{x\to a} h(x) = L$$

then

$$\lim_{x\to a} g(x) = L.$$

Remark 2.5.10. The pinching theorem can be used to prove the well-known limit

$$\lim_{\theta\to 0} \frac{\sin\theta}{\theta} = 1.$$

See the problem set of Chapter 2 for details.

We close this chapter with four examples involving limits of the form $\lim_{x\to a} f(x)$, where f is not necessarily continuous at the point a. Determining whether or not the limit exists, and if so what its value is, requires special analysis of the function f.

Example 2.5.11. Determine whether or not the limit

$$\lim_{x \to 0} x \sin(1/x)$$

exists. If it exists, find its value.

Solution. We begin with the inequality

$$-1 \leq \sin\theta \leq 1$$

which is valid for all θ in $\mathbb{R}$. In particular, if $\theta = 1/x$ then

$$-1 \leq \sin(1/x) \leq 1 \qquad \text{whenever } x \neq 0.$$

Case 1: Suppose that $x > 0$. Then we have

$$-x \leq x \sin(1/x) \leq x.$$

Since $\lim_{x \to 0^+} -x = \lim_{x \to 0^+} x = 0$, it follows from the pinching theorem that $\lim_{x \to 0^+} x \sin(1/x) = 0$.

Case 2: Suppose that $x < 0$. Then

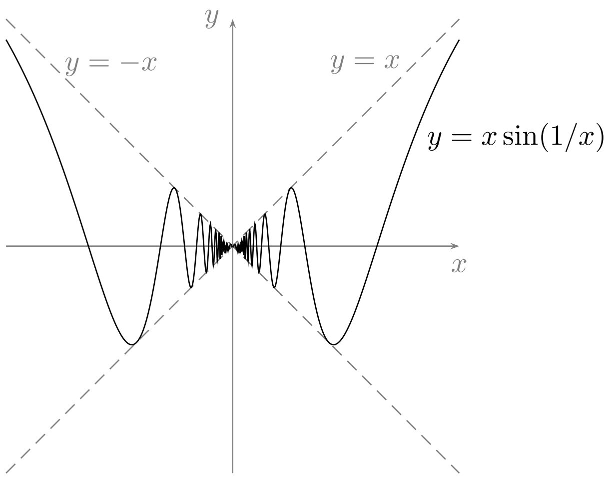
$$-x \geq x \sin(1/x) \geq x.$$

Since $\lim_{x \to 0^-} -x = \lim_{x \to 0^-} x = 0$, it follows from the pinching theorem that $\lim_{x \to 0^-} x \sin(1/x) = 0$.

Conclusion: Since the left- and right-hand limits exist and are both equal to 0, the two-sided limit exists and

$$\lim_{x \to 0} x \sin(1/x) = 0.$$

This limit is illustrated graphically below.



□

Example 2.5.12. Suppose that $f : \mathbb{R} \to \mathbb{R}$ is defined by the formula

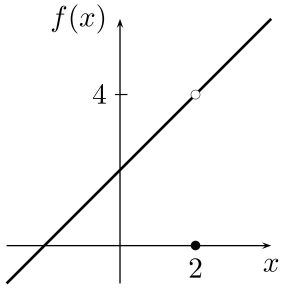
$$f(x) = \begin{cases} \dfrac{x^2 - 4}{x - 2} & \text{if } x \neq 2 \\ 0 & \text{if } x = 2. \end{cases}$$

Discuss the limiting behaviour of $f(x)$ as x approaches 2.

Solution. Since we are discussing limiting behaviour, the value of f at 2 is irrelevant. All that matters is how $f(x)$ behaves *near* 2. So suppose that $x \neq 2$. Then

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

The graph of f is now easy to sketch.



Since

$$\lim_{x\to 2^-} f(x) = \lim_{x\to 2^+} f(x) = 4,$$

the two-sided limit exists and we have $\lim_{x\to 2} f(x) = 4$. (Note that f is not continuous at 2 since $\lim_{x\to 2} f(x) \neq f(2)$.) □

Example 2.5.13. Suppose that $f : \mathbb{R} \to \mathbb{R}$ is defined by the formula

$$f(x) = \begin{cases} \dfrac{|x^2-9|}{x-3} & \text{if } x \neq 3 \\ 2 & \text{if } x = 3. \end{cases}$$

Discuss the limiting behaviour of $f(x)$ as x approaches 3.

Example 2.5.14. Suppose that

$$f(x) = \sin\left(\frac{\pi}{x}\right), \qquad x \neq 0.$$

Discuss the limiting behaviour of $f(x)$ as x approaches 0.

Solution. Since f is an odd function, we initially restrict our attention to positive values of x. Now $\sin\theta = 0$ whenever $\theta \in \{\pi, 2\pi, 3\pi, \ldots\}$. If $\theta = \pi/x$ then we have $\sin(\pi/x) = 0$ whenever

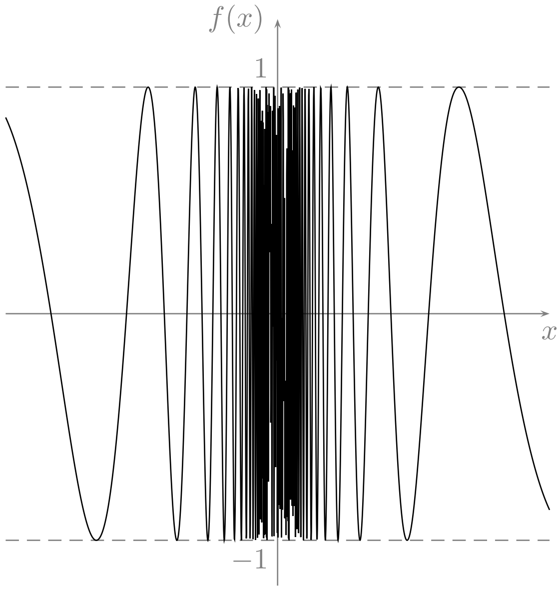
$$x = \frac{1}{1}, \frac{1}{2}, \frac{1}{3}, \ldots.$$

Similarly, $\sin\theta = 1$ whenever $\theta \in \{\pi/2, 5\pi/2, 9\pi/2, 13\pi/2, \ldots\}$ so $\sin(\pi/x) = 1$ whenever

$$x = \frac{2}{1}, \frac{2}{5}, \frac{2}{9}, \frac{2}{13}, \ldots.$$

Consequently, no matter how close x is to 0 on the right, we can *always* find two closer points x_1 and x_2 such that $f(x_1) = 0$ and $f(x_2) = 1$. We conclude that f has no right-hand limit at 0. Since f is odd, it follows that f has no left-hand limit at 0. Hence $\lim_{x\to 0^+} f(x)$ does not exist.

The graph of f near the origin is shown below.



The $f(x)$ oscillates wildly between 1 and -1 as x tends to 0, and $\lim_{x\to 0^+} f(x)$ does not exist. □

2.6 Maple notes

The following Maple command is relevant to the material of this chapter:
`limit(f(x),x=a);` finds the (two-sided) limit of $f(x)$ as $x \to a$. One-sided limits can be specified using `limit(f(x),x=a,right);` or `limit(f(x),x=a,left);`. For example,

```
>  limit(sin(x)/x, x=0);
```

$$1$$

```
>  limit(sin(x)/x, x=infinity);
```

$$0$$

```
>  # the next limit needs methods from chapter 5 for hand calculations
>  limit((1-cos(x))/x^2, x=0, left);
```

$$1/2$$

```
>  # Maple produces results that are not in strict agreement with our definitions
as x goes to plus or minus infinity
>  limit(x^2, x=-infinity);
```

$$\infty$$

```
>  limit(sin(x), x=-infinity);
```

$$-1 \ldots 1$$

```
>  limit(x^2*sin(x), x=-infinity);
```

$$\textit{undefined}$$

Problems for Chapter 2

Problems 2.1 : Limits of functions at infinity

1. [**R**] [**V**] Evaluate the following limits if they exist.

   a) $\displaystyle\lim_{x\to\infty} \frac{x^2-1}{x^2+1}$ b) $\displaystyle\lim_{x\to\infty} \frac{2x^2+x-1}{x^2+4x-3}$

   c) $\displaystyle\lim_{x\to\infty} \frac{2x^2+5x-1}{x^3+x}$ d) $\displaystyle\lim_{x\to\infty} \frac{x^5+5x+1}{x^4+3}$

   e) $\displaystyle\lim_{x\to\infty} \frac{5x^2-3x+\cos 7x}{4+\sin 2x+x^2}$ f) $\displaystyle\lim_{x\to\infty} \sin x$

2. [**R**] Use the pinching theorem to find the following limits.

   a) $\displaystyle\lim_{x\to\infty} \frac{\sin x}{x}$ b) $\displaystyle\lim_{x\to\infty} \frac{\cos x}{x^2}$

3. [**R**] [**V**]

   a) Prove that $\displaystyle\lim_{x\to\infty} (\sqrt{x+1}-\sqrt{x}) = 0.$

   b) Show that $\displaystyle\lim_{x\to\infty} (\sqrt{x^2+x}-x) = \frac{1}{2}.$

Problems 2.2 : The definition of $\lim_{x\to\infty} f(x)$ and 2.3 : Proving that $\lim_{x\to\infty} f(x) = L$ using the limit definition

4. [**R**]

   a) Write down the formal definition for the statement
   $$\lim_{x\to\infty} f(x) = L.$$

   b) Evaluate $\displaystyle\lim_{x\to\infty} \frac{1}{2x^2}.$

   c) Verify from the formal definition that your answer in (b) is correct.

5. [**R**]

   a) Evaluate $\displaystyle\lim_{x\to\infty} \frac{x^2+1}{x^2}.$

   b) Find a real number M such that the distance between $\dfrac{x^2+1}{x^2}$ and its limit is less than 0.01 whenever $x > M$.

   c) Suppose that $\epsilon > 0$. Find a real number M (expressed in terms of ϵ) such that the distance between $\dfrac{x^2+1}{x^2}$ and its limit is less than ϵ whenever $x > M$.

6. **[R]** **[V]** For each of the following, find the limit of $f(x)$ as x tends to infinity and prove from the definition that your answer is correct.

a) $f(x) = \dfrac{4x}{x+7}$ b) $f(x) = \dfrac{x-3}{x^2+3}$ c) $f(x) = e^{-2x}$

d) $f(x) = \dfrac{\sin x}{x}$ e) $f(x) = \dfrac{\sin 3x}{x^2+4}$

7. **[X]**

a) With ϵ in $(0,1)$, Sarah solves the inequality $|f(x)-4| < \epsilon$ and finds that the required x values satisfy
$$x \in \left(\frac{1}{\epsilon}, \infty\right).$$
Does $\lim_{x\to\infty} f(x)$ exist? Give reasons for your answer.

b) With ϵ in $(0,1)$, Lyndal solves the inequality $|g(x)-5| < \epsilon$ and finds that the inequality holds for all x satisfying
$$x \in \left(\frac{1}{\epsilon}, \infty\right).$$
Does $\lim_{x\to\infty} g(x)$ exist? Give reasons for your answer.

8. **[R]** **[V]** A parcel is dropped from an aeroplane. A simple model, taking into account gravity and air resistance, suggests that the parcel's velocity $v(t)$ (in metres per second) is given by $v(t) = 50(1 - e^{-t/5})$, where t is the number of seconds since leaving the plane.

a) Calculate the terminal velocity of the parcel (that is, find $\lim_{t\to\infty} v(t)$).

b) The parcel never attains its terminal velocity. How long does it take to come within 1 metre per second of its terminal velocity?

Problems 2.4 : Proofs of basic limit results (MATH1141 only)

9. **[X]** For each question below, give reasons for your answer. [In some cases a single example will be sufficient while in other cases a general proof will be required. As a reminder, if $f(x) \to \infty$ as $x \to \infty$ then $\lim_{x\to\infty} f(x)$ does not exist.]

a) If $\lim_{x\to\infty} f(x)$ and $\lim_{x\to\infty} g(x)$ do not exist, can $\lim_{x\to\infty} [f(x)+g(x)]$ or $\lim_{x\to\infty} f(x)g(x)$ exist?

b) If $\lim_{x\to\infty} f(x)$ exists and $\lim_{x\to\infty} [f(x)+g(x)]$ exists, must $\lim_{x\to\infty} g(x)$ exist?

c) If $\lim_{x\to\infty} f(x)$ exists and $\lim_{x\to\infty} g(x)$ does not exist, can $\lim_{x\to\infty} [f(x)+g(x)]$ exist?

d) If $\lim_{x\to\infty} f(x)$ exists and $\lim_{x\to\infty} f(x)g(x)$ exists, does it follow that $\lim_{x\to\infty} g(x)$ exists?

Problems 2.5 : Limits of functions at a point

10. **[R]** Evaluate the following limits.

a) $\lim_{x\to 3} 2x+4$ b) $\lim_{x\to 2} \dfrac{x^2-4}{x-2}$ c) $\lim_{x\to 1} \dfrac{x^3-1}{x-1}$ d) $\lim_{x\to 3} \dfrac{\frac{1}{x}-\frac{1}{3}}{x-3}$

11. [**R**] [**V**]

    a) Find the left-hand limit $\lim\limits_{x \to 2^-} \dfrac{|x-2|}{x-2}$.

    b) Find the right-hand limit $\lim\limits_{x \to 2^+} \dfrac{|x-2|}{x-2}$.

    c) Does $\lim\limits_{x \to 2} \dfrac{|x-2|}{x-2}$ exist?

12. [**R**] By finding the left- and right-hand limits first, decide whether or not each of the following limits exist and if so find their values.

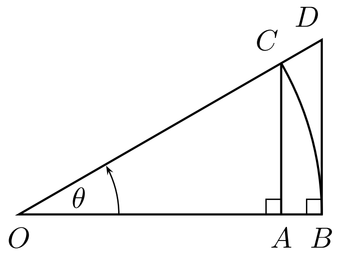
    a) $\lim\limits_{x \to 0} \dfrac{x}{|x|}$ b) $\lim\limits_{x \to 2} \dfrac{|x^2-4|}{x-2}$ c) $\lim\limits_{x \to 4} \dfrac{x-4}{|x-4|}$ d) $\lim\limits_{x \to 0} \dfrac{4}{x}$

13. [**R**] [**V**]

    a) Use the pinching theorem to find $\lim\limits_{x \to 0} x \sin \frac{1}{x}$.

    b) Repeat for $\lim\limits_{x \to 0} x^2 \sin \dfrac{1}{2x}$.

14. [**R**] [**V**] Suppose that θ is a (positive) angle measured in radians and consider the diagram below.



The curve segment CB is the arc of a circle of radius 1 centre O.

    a) Write down, in terms of θ, the length of arc CB and the lengths of the line segments CA and DB.

    b) By considering areas, deduce that $\sin\theta\cos\theta \le \theta \le \tan\theta$ whenever $0 < \theta < \frac{\pi}{2}$.

    c) Use the pinching theorem to show that $\lim\limits_{\theta \to 0^+} \dfrac{\theta}{\sin\theta} = 1$.

    d) Deduce that $\lim\limits_{\theta \to 0} \dfrac{\sin\theta}{\theta} = 1$.

15. [**H**] Discuss the limiting behaviour of $\cos\frac{1}{x}$ as $x \to 0$.

Chapter 3

Properties of continuous functions

In the world of functions, there is a small civilised country inhabited by the race of continuous functions. Citizens of this country, such as the polynomials and exponentials, are studied (and perhaps even loved) by students in high schools throughout the human world. However, outside this small civilised country is a vast untamed universe of all sorts of discontinuous functions. From this perspective, the continuous functions are a very rare breed.

However, continuous functions are far from endangered. They are among the most useful functions for modelling real-life phenomena, particularly when the quantity that is being modelled occurs in nature and changes smoothly with time. For example,

- the distance from the earth to the sun at time t,
- the height of Mount Everest at time t,
- the volume of water in a dam at time t,
- the mass of an earthworm at time t, and
- the temperature at the back of a lecture theatre at time t

can all be modelled by continuous functions of the variable t.

In addition to their usefulness for modelling naturally occurring quantities, continuous functions have important mathematical characteristics. The combination of the usefulness of continuous functions and their rarity among the world of functions means that continuity is a deep property. In this chapter we begin to explore the ramifications of continuity.

3.1 Combining continuous functions

(Ref: SH10 §2.4)

If two continuous functions are combined via function addition, subtraction, multiplication, division or composition, when is the resulting function also continuous? The two next propositions answer this question.

Proposition 3.1.1. *Suppose that the functions f and g are continuous at a point a. Then $f+g$, $f-g$ and fg are continuous at a. If $g(a) \neq 0$ then f/g is also continuous at a.*

Proof. Suppose that f and g are continuous at a. Then, by the definition of continuity at a point (see Definition 2.5.3), $\lim_{x\to a} f(x) = f(a)$ and $\lim_{x\to a} g(x) = g(a)$. Therefore

$$\begin{aligned}
\lim_{x\to a}(f+g)(x) &= \lim_{x\to a}(f(x)+g(x)) && \text{(by the definition of } f+g\text{)}\\
&= \lim_{x\to a} f(x) + \lim_{x\to a} g(x) && \text{(by Proposition 2.5.6)}\\
&= f(a)+g(a) && \text{(since } f \text{ and } g \text{ are continuous)}\\
&= (f+g)(a) && \text{(by the definition of } f+g\text{).}
\end{aligned}$$

Hence $f+g$ is continuous at a.

The proofs that the functions $f-g$, fg and f/g are continuous at a are similar. □

Proposition 3.1.2. *Suppose that f is continuous at a and that g is continuous at $f(a)$. Then $g \circ f$ is continuous at a.*

Proof. Combine Definition 2.5.3 with Proposition 2.5.7. □

In Chapters 1 and 2 it was stated (without proof) that the polynomials, rational functions and trigonometric functions are continuous at every point in their respective domains. The next example shows that the continuity of these functions can be deduced from the continuity of

- the constant functions
- the sine function, and
- the function $f : \mathbb{R} \to \mathbb{R}$ given by $f(x) = x$.

Example 3.1.3. Suppose that a is a point in $\mathbb{R}$. By assuming that the constant functions, the sine function and the function $f : \mathbb{R} \to \mathbb{R}$, given by $f(x) = x$, are continuous everywhere, show that

(a) every polynomial is continuous at a,

(b) every rational function r is continuous at a (provided that $a \in \mathrm{Dom}(r)$), and

(c) the cosine and tangent functions are continuous at a (provided that $a \in \mathrm{Dom}(\tan)$).

Solution. (a) Every polynomial can be constructed from constant functions and the function f by a finite number of function multiplications and additions. (For example,

$$x^3 - 4x^2 + 5 = (x \times x \times x) + ((-4) \times x \times x) + 5$$

for all x in $\mathbb{R}$.) Since the constant functions and f are continuous at a, it follows from Proposition 3.1.1 that every polynomial is continuous at a.

(b) Every rational function r can be written in the form p/q, where p and q are polynomials. By the result of part (a), p and q are continuous a. It follows from Proposition 3.1.1 that p/q is also continuous at a, provided that $q(a) \neq 0$.

(c) Recall that

$$\cos x = \sin(\pi/2 - x) \qquad \forall x \in \mathbb{R}.$$

That is, $\cos x = h(g(x))$ where $h(x) = \sin x$ and $g(x) = \pi/2 - x$. Since the polynomial g is continuous at a and the sine function h is continuous everywhere, the cosine function must also be continuous at a by Proposition 3.1.2.

Finally, since

$$\tan x = \frac{\sin x}{\cos x},$$

the tangent function is also continuous at a (provided that $a \in \mathrm{Dom}(\tan)$) by Proposition 3.1.1. $\square$

The next example justifies an assertion made in Section 2.5.2. We continue working under the assumption that the constant functions and the function $f : \mathbb{R} \to \mathbb{R}$, given by $f(x) = x$, are continuous everywhere.

Example 3.1.4. Suppose that $g : \mathbb{R} \to \mathbb{R}$ is continuous at the point a and let $|g|$ denote the function given by

$$|g|(x) = |g(x)| \qquad \forall x \in \mathbb{R}.$$

Show that $|g|$ is continuous at the point a.

Remark 3.1.5. Throughout this section we have taken for granted that the constant functions, the sine function and the function $f : \mathbb{R} \to \mathbb{R}$, given by $f(x) = x$, are continuous everywhere. To prove this, one needs a rigorous definition of the limit of a function at a point. Such a definition is similar in flavour to the definition for the limit of a function at infinity, but will not be presented in these notes.

Example 3.1.6. Suppose that a and b are real numbers and consider the function $f : \mathbb{R} \to \mathbb{R}$ given by

$$f(x) = \begin{cases} e^{ax+b} & \text{if } x \geq 0 \\ \cos x & \text{if } x < 0. \end{cases}$$

For what values of a and b will the function f be continuous at 0?

Solution. We require that

$$\lim_{x\to 0^-} f(x) = f(0) = \lim_{x\to 0^+} f(x).$$

Now

$$\lim_{x\to 0^-} f(x) = \lim_{x\to 0^-} \cos x = 1, \qquad f(0) = e^b \qquad \text{and} \qquad \lim_{x\to 0^+} f(x) = \lim_{x\to 0^+} e^{ax+b} = e^b.$$

So we require that $e^b = 1$, which implies that $b = 0$. There is no restriction on a. $\square$

3.2 Continuity on intervals

(Ref: SH10 §2.4)

In section 2.5.2, a definition for continuity at a point was given. We now extend this definition to define what is meant by continuity on an interval.

Definition 3.2.1. Suppose that f is a real-valued function defined on an open interval (a, b). We say that f is a *continuous on* (a, b) if f is continuous at every point in the interval (a, b).

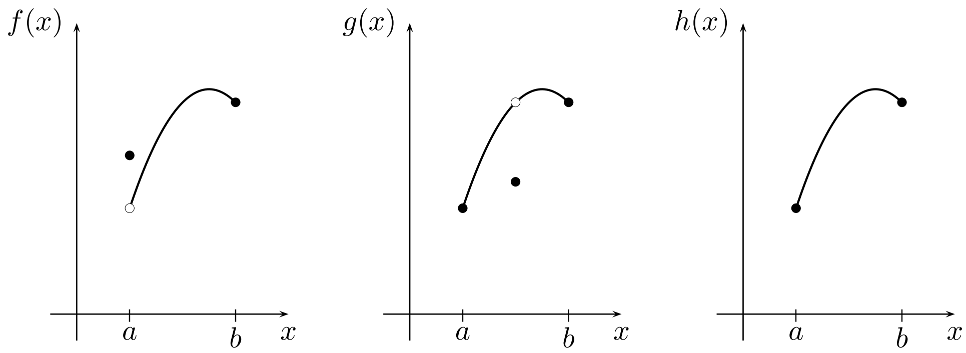
Definition 3.2.2. Suppose that f is a real-valued function defined on a closed interval $[a, b]$. We say that

(a) f is continuous at the endpoint a if $\lim_{x \to a^+} f(x) = f(a)$,

(b) f is continuous at the endpoint b if $\lim_{x \to b^-} f(x) = f(b)$, and

(c) f is continuous on the closed interval $[a, b]$ if f is continuous on the open interval (a, b) and at each of the endpoints a and b.

To illustrate these definitions, consider the functions f, g and h, whose graphs are shown below.



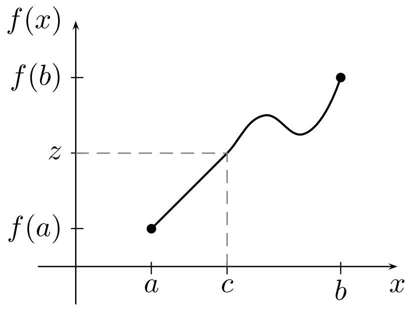
All three functions are defined on the interval $[a, b]$. We see that

- f is continuous on the open interval (a, b) and at the endpoint b;
- g is continuous at the endpoints a and b but not continuous on the open interval (a, b); and
- h is continuous on the closed interval $[a, b]$ (and, by implication, on the open interval (a, b) and at both endpoints a and b).

3.3 The intermediate value theorem

(Ref: SH10 §§2.6, B.1)

The idea of this section is very simple. Suppose that a function f is continuous on some closed interval $[a, b]$ and that z is any real number that lies between $f(a)$ and $f(b)$, as illustrated below.

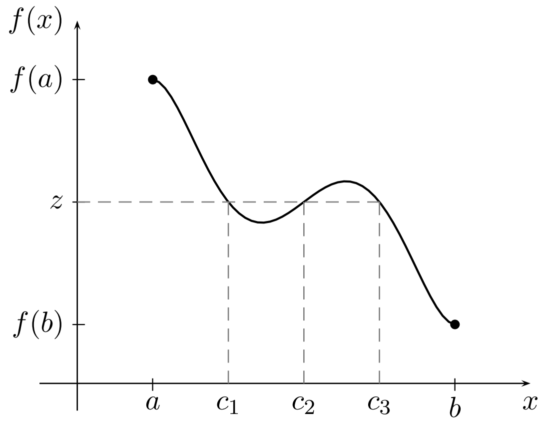


It is clear from the graph that one can find a real number c in $[a, b]$ such that $f(c) = z$. The *intermediate value theorem*, which is stated below, says that this can always be done for any function that is continuous on $[a, b]$.

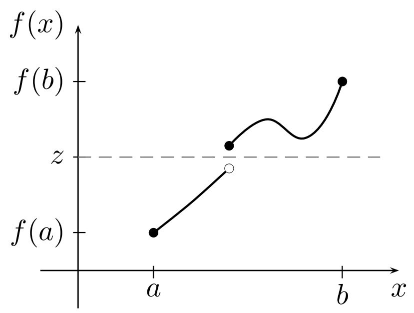
Theorem 3.3.1 (The intermediate value theorem). *Suppose that f is continuous on the closed interval $[a, b]$. If z lies between $f(a)$ and $f(b)$ then there is at least one real number c in $[a, b]$ such that $f(c) = z$.*

The intermediate value theorem is proven using *the least upper bound property* of the real numbers. (It is this property that distinguishes the real numbers from the rational numbers.) We shall not examine the proof here.

We note three important points about the theorem. First, there may be more than one number c in $[a, b]$ that satisfies the conclusion of the intermediate value theorem. The diagram below illustrates this possibility.



Second, the condition in Theorem 3.3.1 that f is continuous cannot be relaxed. For example, consider the function $f : [a, b] \to \mathbb{R}$ whose graph is shown below.



Because of the discontinuity, z is not in the range of f, even though z lies between $f(a)$ and $f(b)$. That is, there is no real number c in $[a, b]$ such that $f(c) = z$.

Third, the intermediate value theorem is a theorem about continuous functions defined on the *real numbers*. There is no analogous theorem for continuous functions defined on the rational numbers. For example, consider the functions $f : \mathbb{R} \to \mathbb{R}$ and $g : \mathbb{Q} \to \mathbb{Q}$, given by

$$f(x) = x^2 - 2 \qquad \text{and} \qquad g(x) = x^2 - 2.$$

Both functions are continuous on their domains. Also,

$$f(0) = g(0) = -2 \qquad \text{and} \qquad f(2) = g(2) = 2.$$

However, while there is a number c in $[0, 2]$ such that $f(c) = 0$, there is no number c in $[0, 2]$ such that $g(c) = 0$ (since $\sqrt{2} \notin \mathrm{Dom}(g)$).

One of the chief applications of the intermediate value theorem is determining whether an equation has a solution. Moreover, if a solution exists, then the theorem helps us determine an approximate location of this solution. This is particularly helpful when the equation cannot be solved by simple algebra.

Example 3.3.2. Show that the equation

$$e^{2x} = \sin x + 4 \tag{3.1}$$

has at least one positive solution.

Solution. Consider the function f given by

$$f(x) = e^{2x} - \sin x - 4.$$

Now

$$f(0) = e^0 - \sin 0 - 4 = -3$$

while

$$f(\pi/2) = e^{\pi} + 1 - 4 \approx 20.14.$$

Since f is continuous on the closed interval $[0, \pi/2]$ and 0 lies between $f(0)$ and $f(\pi/2)$, the intermediate value theorem implies that there is a real number c in $[0, \pi/2]$ such that $f(c) = 0$. That is,

$$e^{2c} = \sin c + 4$$

for some c in $[0, \pi/2]$. Moreover, one can check that $c \neq 0$. Hence the equation has at least one positive solution c. $\square$

Note that in equation (3.1), one cannot isolate x to find an explicit solution. However, now that it is known a solution exists in the interval $[0, \pi/2]$, one could use Newton's method to find a good approximation to this solution.

3.4 The maximum-minimum theorem

(Ref: SH10 §§2.6, B.2)

In many situations, it is desirable to know the maximum (or minimum) value of a quantity that varies over an interval of time. Examples include the maximum temperature in a 24 hour period or the minimum stock price over a seven day period. Depending on the quantity and the time interval, a maximum or minimum value may not be attained. The *maximum-minimum theorem* gives conditions under which a maximum and minimum is attained. To clarify discussion, we begin with a definition.

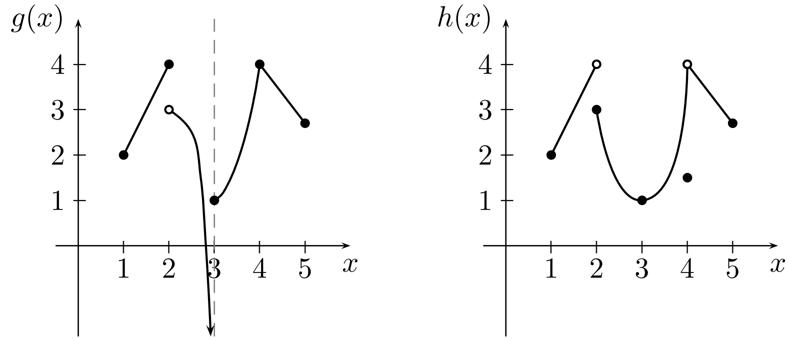
Definition 3.4.1. Suppose that f is defined on a closed interval $[a, b]$.

(a) We say that a point c in $[a, b]$ is an *absolute minimum point* for f on $[a, b]$ if $f(c) \le f(x)$ for all x in $[a, b]$. The corresponding value $f(c)$ is called the *absolute minimum value* of f on $[a, b]$. If f has an absolute minimum point on $[a, b]$ then we say that f *attains its minimum on* $[a, b]$.

(b) We say that a point d in $[a, b]$ is an *absolute maximum point* for f on $[a, b]$ if $f(x) \le f(d)$ for all x in $[a, b]$. The corresponding value $f(d)$ is called the *absolute maximum value* of f on $[a, b]$. If f has an absolute maximum point on $[a, b]$ then we say that f *attains its maximum on* $[a, b]$.

An absolute maximum point and an absolute minimum point are sometimes referred to as a *global maximum point* and a *global minimum point.*

Example 3.4.2. Consider the functions g and h, whose graphs are illustrated below.



The absolute minimim and maximum points of g and h on $[1, 5]$ are recorded in the following table.

	g	h
Absolute minimum points	none	3
Absolute minimum value	n.a.	1
Absolute maximum points	2, 4	none
Absolute maximum value	4	n.a.

The above example shows that a function $f : [a, b] \to \mathbb{R}$ need not have an absolute maximum point (or an absolute minimum point) on a closed interval $[a, b]$. However, if f is continuous on $[a, b]$ then such points always exist.

Theorem 3.4.3 (The maximum-minimum theorem)**.** *If f is continuous on a closed interval $[a, b]$ then f attains its minimum and maximum on $[a, b]$. That is, there exist points c and d in $[a, b]$ such that*

$$f(c) \le f(x) \le f(d)$$

for all x in $[a, b]$.

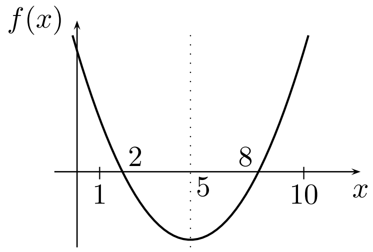
As with the intermediate value theorem, the least upper bound property of the real numbers is used to prove the maximum-minimum theorem. We omit the proof.

Example 3.4.4. Suppose that the function $f : \mathbb{R} \to \mathbb{R}$ is given by the rule

$$f(x) = x^2 - 10x + 16.$$

Find the absolute maximum and absolute minimum values of f on the interval $[1, 10]$.

Solution. Since $f(x) = (x - 2)(x - 8)$, the function f has zeros at 2 and 8.



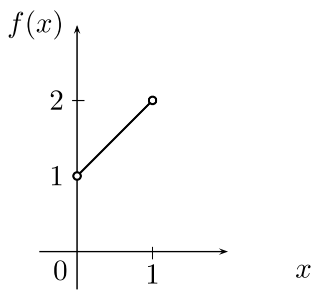
By symmetry, we know that the axis of the parabola is the line $x = 5$. Hence 5 is an absolute minimum point for f on $[1, 10]$. Since $f(5) = -9$, the absolute minimum value of f on $[1, 10]$ is -9.

It is clear from the sketch of f that at least one of the endpoints of the interval $[1, 10]$ must be an absolute maximum point for f on $[1, 10]$. Now $f(1) = 7$ while $f(10) = 16$. So 10 is an absolute maximum point for f on $[1, 10]$ and the corresponding maximum value of f is 16. $\square$

Remark 3.4.5. While continuity on a closed interval guarantees the existence of at least one absolute maximum point and absolute minimum point, locating such points may not be easy. A systematic approach to finding maximum and minimum points will be developed in Chapters 4 and 5.

Remark 3.4.6. All the hypotheses stated in the maximum-minimum theorem must be satisfied for the conclusion of the theorem to be valid. If the interval is closed but the function is not continuous, then an absolute maximum (or minimum) point may not exist (see Example 3.4.2). The next example shows that if the function is continuous on the interval but the interval is not closed then an absolute maximum (or minimum) point may not exist. It is expected that students will know the precise hypotheses of the maximum-minimum theorem (and other important theorems, such as the intermediate value theorem) and know how to apply the theorem correctly.

Example 3.4.7. Consider the function $f : \mathbb{R} \to \mathbb{R}$ given by $f(x) = x + 1$. Since f is a polynomial, it is continuous on the open interval $(0, 1)$. However, f does not have an absolute maximum point on $(0, 1)$. To see why, consider the graph of f on $(0, 1)$.



As $x \to 1^-$, $f(x) \to 2$ but there is no real number d in $(0, 1)$ such that $f(d) = 2$. By a similar observation, f has no absolute minimum point on $(0, 1)$.

The final result of this chapter is a corollary of maximum-minimum theorem. We begin with a definition.

Definition 3.4.8. A function f is said to be *bounded* on an interval I if there is some positive number M such that $|f(x)| \leq M$ for all x in I.

In other words, f is bounded if the y-values of its graph lie between $-M$ and M for some positive number M. The *sine* and *cosine* functions are obvious examples of functions that are bounded on the whole of the real line.

Corollary 3.4.9. *If $f : [a, b] \to \mathbb{R}$ is continuous on $[a, b]$ then f is bounded on $[a, b]$.*

Proof. Suppose that f is continuous on $[a, b]$. Then the function $|f|$ is also continuous on $[a, b]$ (see Example 3.1.4). By the maximum-minimum theorem, $|f|$ attains its maximum on $[a, b]$. If M denotes the absolute maximum value of $|f|$ on $[a, b]$ then

$$|f(x)| \leq M \qquad \forall x \in [a, b].$$

Hence f is bounded. $\square$

Problems for Chapter 3

Problems 3.1 : Combining continuous functions and 3.2 : Continuity on intervals

1. [**R**] Suppose that $f : \mathbb{R} \to \mathbb{R}$ is defined by $f(x) = |x|$.

   a) Show that f is continuous at 0.

   b) Is f continuous everywhere? Give brief reasons for your answer.

2. [**R**] Determine at which points each function $f : \mathbb{R} \to \mathbb{R}$ is continuous. Give reasons.

   a) $f(x) = \begin{cases} e^{2x} & \text{if } x < 0 \\ \cos x & \text{if } x \geq 0 \end{cases}$

   b) $f(x) = \begin{cases} e^{-2x} & \text{if } x < 0 \\ \sin x + 1 & \text{if } 0 \leq x \leq \pi/2 \\ 2x - \pi & \text{if } x > \pi/2 \end{cases}$

3. [**R**] Suppose that

$$f(x) = \begin{cases} \frac{x^2-16}{x-4} & \text{if } x \neq 4 \\ k & \text{if } x = 4, \end{cases}$$

   where k is a real number. For which values of k (if any) will f be continuous everywhere?

4. [**H**] Use the pinching theorem for limits to show that if f, g and h are three functions defined on an open interval I, such that

   - $f(x) \leq g(x) \leq h(x)$ for all $x \in I$,
   - $f(a) = g(a) = h(a)$ for some $a \in I$, and
   - f and h are continuous at a,

   then g is also continuous at a.

Problems 3.3 : The intermediate value theorem

5. [**R**] Show that the function f, given by $f(x) = x^3 - 5x + 3$, has a zero in each of the intervals $[-3, -2]$, $[0, 1]$ and $[1, 2]$.

6. [**R**] [**V**] Use the intermediate value theorem to show that the equation $e^x = 2\cos x$ has at least one positive real solution.

7. [**H**] Suppose that f is continuous on $[0, 1]$ and that $\text{Range}(f)$ is a subset of $[0, 1]$. By using $g(x) = f(x) - x$, prove that there is a real number c in $[0, 1]$ such that $f(c) = c$.

8. [**X**] Suppose that f is a continuous function such that $f(0) = 1$ and $\lim_{x\to\infty} f(x) = -1$. Show that f has a zero somewhere in $(0, \infty)$.

Problems 3.4 : The maximum-minimum theorem

9. [**R**] In each case, determine whether or not f attains a maximum on the given interval. Give reasons for your answer.

a) $f(x) = x^2 - 4$ on $[-3, 5]$ b) $f(x) = \left|\sin(e^x) + \dfrac{\ln x}{x^2 - 1}\right|$ on $[2, 4]$

c) $f(x) = x^2 - 4$ on $(-3, 5)$ d) $f(x) = -(x^2 - 4)$ on $(-3, 5)$

10. [**H**] [**V**] Suppose that f is a continuous function on $\mathbb{R}$ and that $\lim_{x\to\infty} f(x) = \lim_{x\to-\infty} f(x) = 0$.

a) Give an example of such a function which has both a maximum value and a minimum value.

b) Give an example of such a function which has a minimum value but no maximum value.

c) [**X**] Show that if there is a real number ξ such that $f(\xi) > 0$ then f attains a maximum value on $\mathbb{R}$. [Note that the maximum-minimum theorem only applies to *finite* closed intervals $[a, b]$.]

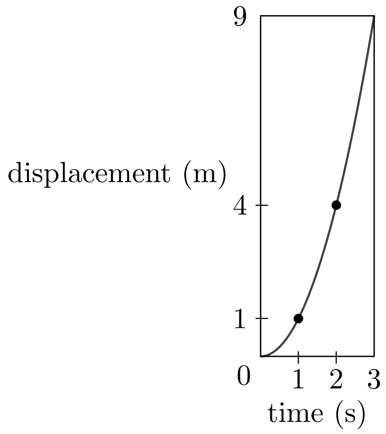
Chapter 4

Differentiable functions

A cyclist accelerates from rest in such a way that her displacement s (in metres) from her starting position after t seconds is given by

$$s = t^2 \qquad \forall t \in [0, 3].$$

The corresponding displacement–time graph for the first three seconds is shown below.



What is the cyclist's instantaneous velocity when $t = 1$?

At present, we do not have a formula for calculating instantaneous velocity. However, we can calculate the average velocity $\overline{v}$ of the cyclist over any time interval by using the formula

$$\overline{v} = \frac{\Delta s}{\Delta t},$$

where Δs denotes the change in displacement corresponding to a change Δt in time. For example, the average velocity over the time interval $[1, 1.5]$ is 2.5 metres per second, as calculated below:

$$\overline{v}_{[1,1.5]} = \frac{\Delta s}{\Delta t} = \frac{(1.5)^2 - (1)^2}{1.5 - 1} = 2.5. \qquad (4.1)$$

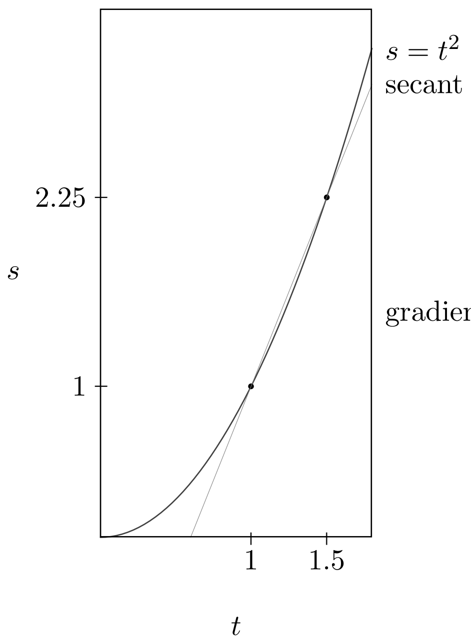
Average velocities over small time intervals are recorded in the following table.

Time interval	Δt	Average velocity
$[1, 1.5]$	0.5	2.5
$[1, 1.4]$	0.4	2.4
$[1, 1.3]$	0.3	2.3
$[1, 1.2]$	0.2	2.2
$[1, 1.1]$	0.1	2.1
$[1, 1.01]$	0.01	2.01
$[1, 1.001]$	0.001	2.001

The table suggests that the average velocity approaches 2 meters per second as $\Delta t \to 0$. Hence it appears that instantaneous velocity of the cyclist when $t = 1$ is 2 meters per second.

One of the goals of this chapter is put these kinds of ideas on a *rigorous* footing in a *general* context (where they can be applied to understanding rates of change of any quantity, not just displacement). Hence we reframe these ideas in terms of functions and analytic geometry. More precisely, we shall now show that the problem of determining instantaneous velocity is equivalent to the problem of finding the gradient of a tangent line to the graph of a function.

To begin, the problem of calculating average velocity is equivalent to the problem of calculating the gradient of a secant to the graph of a function. (A secant is a straight line that intersects a curve at least twice.) For example, compare the calculation of average velocity in (4.1) with the calculation of the gradient of the secant shown in the following diagram.



gradient of secant $= \frac{\Delta s}{\Delta t} = \frac{(1.5)^2-(1)^2}{1.5-1} = 2.5$

In the same way, the above table of average velocities can be reinterpreted geometrically as a table of gradients corresponding to secants of the curve $s = t^2$.

Intersection points	Δt	Gradient of secant
$t = 1, t = 1.5$	0.5	2.5
$t = 1, t = 1.4$	0.4	2.4
$t = 1, t = 1.3$	0.3	2.3
$t = 1, t = 1.2$	0.2	2.2
$t = 1, t = 1.1$	0.1	2.1

The secants used in the table are illustrated over the page in Figure 4.1. The figure suggests that as $\Delta t \to 0$, the gradient of the corresponding secant approaches the gradient of the tangent line to the curve. Thus the problem of determining instantaneous velocity is equivalent to the problem of determining the gradient of a tangent to the graph of a function. It is to this general problem that we now turn.

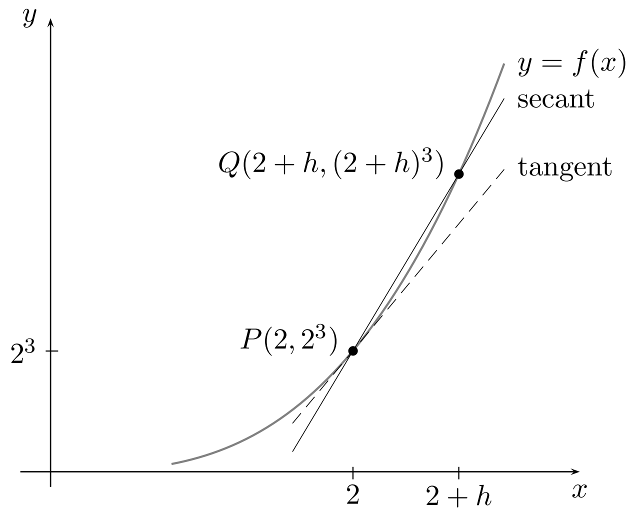
4.1 Gradients of tangents and derivatives

(Ref: SH10 §3.1)

Suppose that a function f has a 'smooth' graph. In this section, we give a general approach to calculating the gradient of tangents to f. This leads the concept of a gradient function, also known as a *derivative* (since it is derived from f), whose values give the gradient of the tangent to f at any point. We begin with an example.

Example 4.1.1. Suppose that $f : \mathbb{R} \to \mathbb{R}$ is given by $f(x) = x^3$. Find the gradient of the tangent to the graph of f when $x = 2$.

Solution. The tangent to the graph of f passes through the point $P(2, 2^3)$. We will approximate the gradient of the tangent with the gradient of a secant passing through the points $P(2, 2^3)$ and $Q(2 + h, (2 + h)^3)$, where h is a very small real number. (See the illustration below for the case when $h > 0$.)



Using the 'rise over run' formula, the gradient of the secant is

$$\frac{(2 + h)^3 - 2^3}{h}.$$

Now, as h approaches 0, the point Q moves along the curve towards P, and hence the gradient of

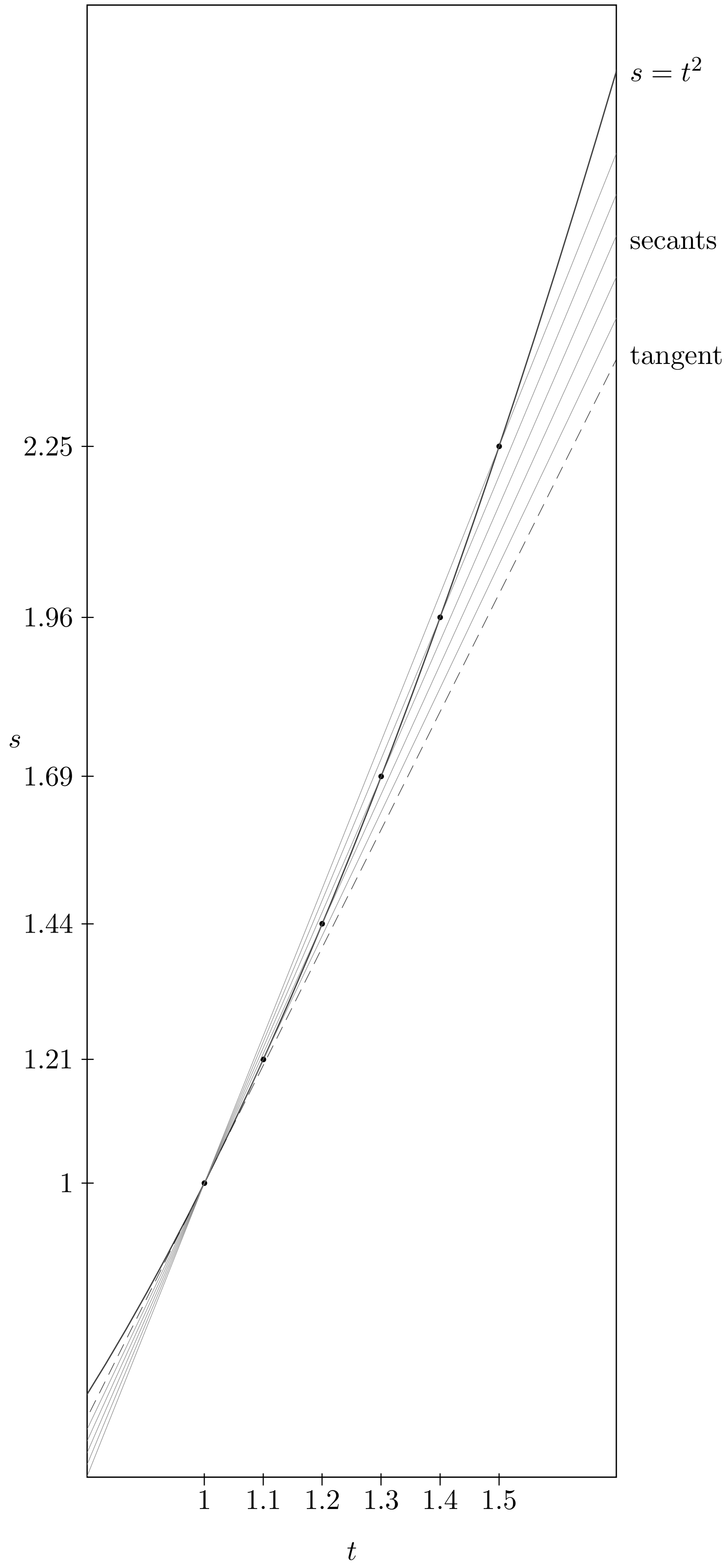


Figure 4.1: Using secants to determine the gradient of a tangent.

the secant approaches the gradient of the tangent. In other words,

$$\begin{aligned}
\text{(gradient of tangent)} &= \lim_{h\to 0} \frac{(2+h)^3 - 2^3}{h} \\
&= \lim_{h\to 0} \frac{8 + 12h + 6h^2 + h^3 - 8}{h} \\
&= \lim_{h\to 0} \frac{12h + 6h^2 + h^3}{h} \\
&= \lim_{h\to 0} 12 + 6h + h^2 \\
&= 12.
\end{aligned}$$

Thus the gradient of the tangent to f at the point 2 is 12. This gives a quantitative description of the rate of change of f at the point 2. ☐

There is nothing special about the point 2 in the previous example; we could calculate the gradient of the tangent to f at any point x. The calculation would go something like this:

$$\begin{aligned}
\text{(gradient of tangent at } x) &= \lim_{h\to 0} \frac{(x+h)^3 - x^3}{h} \\
&= \lim_{h\to 0} \frac{x^3 + 3x^2h + 3xh^2 + h^3 - x^3}{h} \\
&= \lim_{h\to 0} \frac{3x^2h + 3xh^2}{h} \\
&= \lim_{h\to 0} 3x^2 + 3xh \\
&= 3x^2.
\end{aligned}$$

So the gradient of the tangent to f when $x = 5$ is 75 (since $3 \times 5^2 = 75$), while the gradient of the tangent when $x = -4$ is 48. The gradient function of f, denoted by f' and given by

$$f'(x) = 3x^2 \qquad \forall x \in \mathbb{R},$$

is called the *derivative* of f.

We can repeat the same procedure for functions other than $f(x) = x^3$. The gradient function (or derivative) may be calculated using the limit given in the next definition.

Definition 4.1.2. Suppose that f is defined on some open interval containing the point x. We say that f is *differentiable* at x if

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

exists. If the limit exists, we denote it by $f'(x)$. We call $f'(x)$ the derivative of f at x.

Other notation for $f'(x)$ includes $\frac{df}{dx}(x)$ and $\frac{d}{dx}f(x)$. If $y = f(x)$ then the derivative is often denoted by y' or $\frac{dy}{dx}$. The ratio

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

is called the *difference quotient* for f at the point x.

Remark 4.1.3. The notation $\frac{dy}{dx}$ originates from the gradient formula $\frac{\Delta y}{\Delta x}$, where Δy represents a change in y corresponding to a small change Δx in x. (The symbol Δ represents the Greek letter 'delta', which is equivalent to our capital 'D'.) Some calculus texts prefer the use of delta notation and write the difference quotient for a function f as

$$\frac{f(x+\Delta x)-f(x)}{\Delta x}.$$

Example 4.1.4. Suppose that $f:[0,\infty)\to\mathbb{R}$ is given by $f(x)=\sqrt{x}$. Find the equation of the tangent to f when $x=16$.

Solution. First we shall calculate the derivative of f (if it exists) at 16. Note that, in the working below, the limit as h approaches 0 is difficult to evaluate immediately, so we use a standard trick (see, for example, Subsection 2.1.4):

$$\begin{aligned}\frac{f(16+h)-f(16)}{h} &= \frac{\sqrt{16+h}-\sqrt{16}}{h}\\ &= \frac{\sqrt{16+h}-\sqrt{16}}{h}\times\frac{\sqrt{16+h}+\sqrt{16}}{\sqrt{16+h}+\sqrt{16}}\\ &= \frac{16+h-16}{h(\sqrt{16+h}+\sqrt{16})}\\ &= \frac{h}{h(\sqrt{16+h}+\sqrt{16})}\\ &= \frac{1}{\sqrt{16+h}+\sqrt{16}}\\ &\to \frac{1}{8}\end{aligned}$$

as $h\to 0$. So the gradient of the tangent is $1/8$. Using the point-gradient formula, the equation of the tangent is

$$\begin{aligned}y-y_1 &= m(x-x_1)\\ y-4 &= \frac{1}{8}(x-16)\\ 8y-x-16 &= 0.\end{aligned}$$

$\square$

Example 4.1.5. Suppose that $f:\mathbb{R}\to\mathbb{R}$ is given by $f(x)=|x|$. Determine whether or not f is differentiable at the point 0.

Solution. We calculate the difference quotient of f at 0:

$$\begin{aligned}\frac{f(0+h)-f(0)}{h} &= \frac{|h|}{h}\\ &= \begin{cases}\frac{h}{h} & \text{if } h>0\\ \frac{-h}{h} & \text{if } h<0.\end{cases}\\ &= \begin{cases}1 & \text{if } h>0\\ -1 & \text{if } h<0.\end{cases}\end{aligned}$$

Because of the 'split formula,' we must consider the left- and right-hand limits as h approaches 0 separately. Now

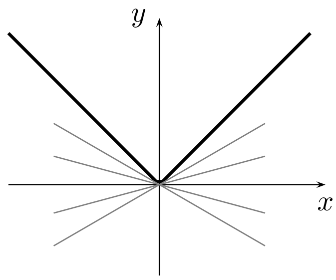
$$\lim_{h\to 0^-} \frac{f(0+h)-f(0)}{h} = -1$$

while

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

Since the left- and right-hand limits are different, the two-sided limit doesn't exist and hence f is *not* differentiable at 0.

Note that this conclusion makes sense geometrically, since the graph of f has a 'vertex' at 0 and there is no unique tangent that touches the graph at this point.



□

Example 4.1.6. Suppose that $f : \mathbb{R} \to \mathbb{R}$ is defined by

$$f(x) = \begin{cases} x\sin\frac{1}{x} & \text{if } x \neq 0 \\ 0 & \text{if } x = 0. \end{cases}$$

Determine whether or not f is differentiable at 0.

4.2 Rules for differentiation

(Ref: SH10 §§3.2, 3.5, 3.6)

So far we have calculated derivatives working directly from the difference quotient. In this section we see that there are fast ways of calculating derivatives without referring to the difference quotient and limits at all. This fast method is part of what is known as 'the calculus.' We begin with the derivatives of some standard functions, and then show how derivatives of other functions are obtained by applying some simple rules.

The following table gives the derivatives for some standard functions.

$f(x)$	$f'(x)$
C, where C is a constant	0
x^n, where n is a positive integer	nx^{n-1}
$\sin x$	$\cos x$
e^x	e^x

With these basic functions, it is possible to construct many other functions via function addition, subtraction, multiplication and division. The next theorem describes what happens to the resulting derivatives.

Theorem 4.2.1 (Rules for differentiation)**.** *Suppose that f and g are differentiable at x. Then $f+g$, $f-g$ and fg are differentiable at x. If $g(x) \neq 0$ then $\frac{f}{g}$ is also differentiable at x. Moreover,*

(i) $(f+g)'(x) = f'(x) + g'(x)$,

(ii) $(C.f)'(x) = C.f'(x)$, *where C is a constant,*

(iii) $(fg)'(x) = f'(x)g(x) + f(x)g'(x)$, *and*

(iv) $\left(\dfrac{f}{g}\right)'(x) = \dfrac{f'(x)g(x) - f(x)g'(x)}{g(x)^2}$, *provided that $g(x) \neq 0$.*

Rules (iii) and (iv) are usually called the product and quotient rules. A familiar way of expressing these rules is

$$\frac{d(uv)}{dx} = v\frac{du}{dx} + u\frac{dv}{dx}$$

and

$$\frac{d}{dx}\left(\frac{u}{v}\right) = \frac{v\frac{du}{dx} - u\frac{dv}{dx}}{v^2},$$

where u and v are both functions of x.

Given two functions f and g, a new function $f \circ g$ may be constructed by function composition. The next theorem describes the derivative of this new function.

Theorem 4.2.2 (The chain rule)**.** *Suppose that g is differentiable at the point x and f is differentiable at the point $g(x)$. Then $f \circ g$ is differentiable at x and*

$$(f \circ g)'(x) = f'(g(x))g'(x). \tag{4.2}$$

You are probably more familiar with the chain rule expressed in the following way: if $y = f(u)$ and $u = g(x)$ then

$$\frac{dy}{dx} = \frac{dy}{du}\frac{du}{dx}. \tag{4.3}$$

While formula (4.3) looks simpler than (4.2), it does not specify the relationship between the variables or the points at which each derivative is to be evaluated. It should therefore be used with care.

The next example illustrates how the chain rule formula (4.2) is applied.

Example 4.2.3. Suppose that $y = (\sin x + 1)^7$. Find the derivative of y at the point x.

Solution. Suppose that $f(x) = x^7$ and $g(x) = \sin x + 1$. Then $f'(x) = 7x^6$ and $g'(x) = \cos x$. Then $y = (f \circ g)(x)$. By the chain rule, the derivative of y at x is given by

$$\begin{aligned}(f \circ g)'(x) &= f'(g(x))g'(x)\\ &= 7(g(x))^6 \times \cos x\\ &= 7(\sin x + 1)^6 \cos x.\end{aligned}$$

□

As mentioned at the beginning of this section, derivatives of many functions can be obtained from those of a few basic functions by the application of simple rules. The next example illustrates this point.

Example 4.2.4. Show that

(a) $\dfrac{d}{dx}\big(x^m\big) = mx^{m-1}$, where m is an integer;

(b) $\dfrac{d}{dx}\big(\cos x\big) = -\sin x$; and

(a) $\dfrac{d}{dx}\big(\tan x\big) = \sec^2 x$.

4.3 Proofs of results in Section 4.2

In Section 4.2 we stated two theorems and gave a table of derivatives. In this section we prove some of these results. We begin with the derivative of the sine function, using the standard result that

$$\lim_{\theta\to 0}\frac{\sin\theta}{\theta} = 1. \tag{4.4}$$

For a geometric argument for this limit, see Question 14 from the problems for Chapter 2.

Proposition 4.3.1. *If $f(x) = \sin x$ then $f'(x) = \cos x$ for all real x.*

Proof. The difference quotient for the sine function and the angle sum formula for sine gives

$$\begin{aligned}\frac{\sin(x+h)-\sin x}{h} &= \frac{\sin x\cos h + \cos x\sin h - \sin x}{h}\\ &= \sin x\,\frac{\cos h - 1}{h} + \cos x\,\frac{\sin h}{h}.\end{aligned} \tag{4.5}$$

Now $\cos^2 h + \sin^2 h = 1$ so

$$\begin{aligned}\frac{\cos h - 1}{h} &= \frac{\cos^2 h - 1}{h(\cos h + 1)}\\ &= \frac{-\sin^2 h}{h(\cos h + 1)}\\ &= -\frac{\sin h}{h}\,\frac{\sin h}{\cos h + 1},\end{aligned}$$

and by (4.4) we have

$$-\frac{\sin h}{h}\,\frac{\sin h}{\cos h + 1} \to -1\times 0 \qquad \text{as} \qquad h\to 0.$$

If we combine this with (4.5) then

$$\begin{aligned}\lim_{h\to 0}\frac{\sin(x+h)-\sin x}{h} &= \sin x\lim_{h\to 0}\frac{\cos h - 1}{h} + \cos x\lim_{h\to 0}\frac{\sin h}{h}\\ &= \sin x\times 0 + \cos x\times 1\\ &= \cos x.\end{aligned}$$

Thus $f'(x) = \cos x$. ☐

The derivative of e^x will be dealt with in Chapter 9. The derivative of constant functions can be easily computed from Definition 4.1.2.

Suppose that f is given by $f(x) = x^n$, where n is a positive integer. The derivative of f can be computed using the same method as Example 4.1.1, with the assistance of the binomial theorem. Alternatively, one can use mathematical induction and the product rule for differentiation.

Proof of the product rule for differentiation. Suppose that f and g are differentiable at the point x. The difference quotient of fg at x gives

$$\begin{aligned}
&\frac{(fg)(x+h)-(fg)(x)}{h}\\
&\quad= \frac{f(x+h)g(x+h)-f(x)g(x)}{h}\\
&\quad= \frac{f(x+h)g(x+h)-f(x)g(x+h)+f(x)g(x+h)-f(x)g(x)}{h}\\
&\quad= g(x+h)\frac{f(x+h)-f(x)}{h}+f(x)\frac{g(x+h)-g(x)}{h}. \qquad (4.6)
\end{aligned}$$

Now g is differentiable at x, so g is also continuous at x (see Theorem 4.5.1). Hence $g(x+h) \to g(x)$ as $h \to 0$. Therefore (4.6) implies that

$$\frac{(fg)(x+h)-(fg)(x)}{h} \to g(x)f'(x)+f(x)g'(x)$$

as $h \to 0$. □

Proofs of other differentiation rules are found in most undergraduate calculus textbooks.

4.4 Implicit differentiation

(Ref: SH10 §3.7)

In Section 1.8 we introduced functions that are defined by an equation relating the variables x and y. Using the chain rule, one can calculate the derivative of such functions without first expressing y explicitly as a function of x. We begin with a simple example to illustrate the principle. The point to keep in mind is that y is a function of x.

Example 4.4.1. Suppose that y is a function of x, and that y and x are related by the formula

$$x^2 + y^2 = 1.$$

Calculate $\frac{dy}{dx}$.

Solution. We begin with the equation

$$x^2 + y^2 = 1$$

and differentiate both sides with respect to x to obtain

$$\frac{d}{dx}(x^2) + \frac{d}{dx}(y^2) = \frac{d}{dx}(1). \qquad (4.7)$$

Now clearly, $\frac{d}{dx}(x^2) = 2x$ and $\frac{d}{dx}(1) = 0$. To simplify the central term of (4.7), we regard y as a function of x and suppose that $g(y) = y^2$. The chain rule gives

$$\begin{aligned}\frac{d}{dx}(y^2) &= \frac{dg}{dx}\\ &= \frac{dg}{dy}\,\frac{dy}{dx}\\ &= 2y\,\frac{dy}{dx}.\end{aligned}$$

So continuing from (4.7) we have

$$2x + 2y\frac{dy}{dx} = 0.$$

By isolating $\frac{dy}{dx}$ we obtain

$$\begin{aligned}\frac{dy}{dx} &= \frac{-2x}{2y}\\ &= -\frac{x}{y}\end{aligned}$$

as required.

(We note in passing that this answer coincides with the derivative obtained by first expressing y as a function of x and then differentiating. In particular, if

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

then

$$\frac{dy}{dx} = \frac{1}{2}(1 - x^2)^{-1/2}(-2x) = -\frac{x}{(1 - x^2)^{1/2}} = -\frac{x}{y}$$

as above.) □

In general, if g is a function of y and y is a function of x then

$$\frac{d}{dx}g(y) = g'(y)\frac{dy}{dx} \tag{4.8}$$

by the chain rule. The use of this formula is illustrated below.

Example 4.4.2. Suppose that y is a function of x, implicitly related by the equation

$$y^4 + x^3 - x^2 e^{3y} = 4. \tag{4.9}$$

Find the equation of the tangent to the corresponding curve at the point where $(x, y) = (2, 0)$.

In Section 4.2 we saw that $\frac{d}{dx}x^n = nx^{n-1}$ whenever n is an integer. A similar formula holds when n is any rational number. This can be easily proved using implicit differentiation.

Proposition 4.4.3. *Suppose that q is a rational number. Then*

$$\frac{d}{dx}x^q = qx^{q-1}.$$

Proof. Since q is a rational number, there are integers m and n such that $q = \frac{m}{n}$ and $n \neq 0$. Now if $y = x^q$ then $y = x^{m/n}$ and taking the nth power of both sides gives

$$y^n = x^m.$$

Differentiating both sides with respect to x yields

$$ny^{n-1}\frac{dy}{dx} = mx^{m-1}.$$

Hence

$$\begin{aligned}\frac{dy}{dx} &= \frac{mx^{m-1}}{ny^{n-1}} \\ &= \frac{m}{n}\frac{x^{m-1}}{x^{q(n-1)}} \\ &= qx^{(m-1)-qn+q} \\ &= qx^{q-1}\end{aligned}$$

as required. □

4.5 Differentiation, continuity and split functions

(Ref: SH10 §3.1)

Not every function that is continuous at a point a is differentiable at a. (Consider, for example, the function f, given by $f(x) = |x|$, at the point 0.) However, we do have the following result.

Theorem 4.5.1. *If f is differentiable at a then f is continuous at a.*

Proof. Suppose that f is differentiable at the point a. Then

$$\lim_{h\to 0}\frac{f(a+h)-f(a)}{h}$$

exists. To show that f is continuous at a, we need to show that

$$\lim_{x\to a} f(x) = f(a),$$

or equivalently, that

$$\lim_{h\to 0} f(a+h) = f(a).$$

Now

$$\begin{aligned}\lim_{h\to 0} f(a+h) - f(a) &= \lim_{h\to 0}\left(\frac{f(a+h)-f(a)}{h}\times h\right) \\ &= \left(\lim_{h\to 0}\frac{f(a+h)-f(a)}{h}\right)\times \lim_{h\to 0} h \qquad (4.10) \\ &= f'(a)\times 0 \\ &= 0.\end{aligned}$$

(Note that we can separate the limits in (4.10) because we know that the limit of the difference quotient exists.) Hence

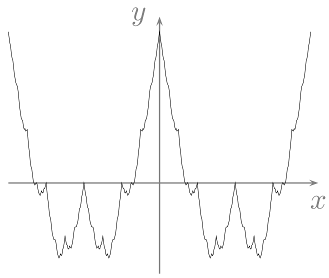
$$\lim_{h\to 0} f(a+h) = f(a),$$

proving the theorem. □

The *contrapositive* form of the theorem is stated below.

Corollary 4.5.2. *If f is not continuous at a then it is not differentiable at a.*

Remark 4.5.3. Differentiability is a much stronger property than continuity. While every differentiable function is continuous, there exist functions that are continuous everywhere but differentiable nowhere. In fact, functions that are differentiable everywhere are a very rare breed, even among the continuous functions. An example of a function that is continuous everywhere but differentiable nowhere is the *Weierstrass function* whose graph is shown below.



The next function studied is an example of a 'split function'. Whether or not the function is differentiable at the 'split point' can be determined by calculating left- and right-hand limits of the difference quotient.

Example 4.5.4. Suppose that $f : \mathbb{R} \to \mathbb{R}$ is defined by

$$f(x) = \begin{cases} \sin x & \text{if } x \geq 0 \\ x^2 + bx + c & \text{if } x < 0, \end{cases}$$

where b and c are real numbers. Find all possible values of b and c such that f is (a) continuous at 0 and (b) differentiable at 0.

Solution. (a) For f to be continuous at 0, we require that

$$\lim_{x\to 0^-} f(x) = f(0) = \lim_{x\to 0^+} f(x).$$

Now $f(0) = \sin 0 = 0$ while

$$\lim_{x\to 0^-} f(x) = c \qquad \text{and} \qquad \lim_{x\to 0^+} f(x) = 0.$$

Hence f is continuous if and only if $c = 0$ and $b \in \mathbb{R}$.

(b) If f is differentiable at 0 then it is continuous at 0 and hence $c = 0$ by part (a). Also, for differentiability we require that

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

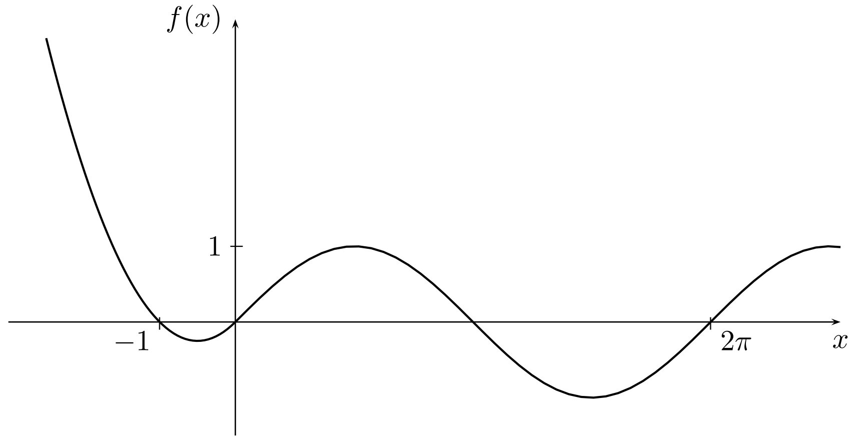
Now,

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

while

$$\lim_{h\to 0^-} \frac{f(0+h) - f(0)}{h} = \lim_{h\to 0^-} \frac{h^2 + bh - 0}{h} = \lim_{h\to 0^-} h + b = b.$$

Hence we require that $b = 1$. So f is differentiable at 0 if and only if $b = 1$ and $c = 0$. The graph of the corresponding function is sketched below.



□

An important application of split functions is the theory of splines. A *spline* is a function defined piecewise by polynomials. They are frequently used in computer graphics and elsewhere because of their simplicity and their capacity to approximate complex curves. The following is an example of a spline.

Example 4.5.5.

$$f(x) = \begin{cases} (x-1)^2 - 1 & \text{if } -2 \le x < 0 \\ 1 - (x-1)^2 & \text{if } 0 \le x \le 2. \end{cases}$$

In this example, each of the functions $p(x) = (x-1)^2 - 1$ and $q(x) = 1 - (x-1)^2$ have had their domains restricted and then 'glued' together at the point 0.

If the constituent functions are continuous and differentiable in some interval containing the point a, then the following theorem tells us when the corresponding split function is differentiable.

Theorem 4.5.6. *Suppose a is a fixed real number and the function f is defined by*

$$f(x) = \begin{cases} p(x) & \text{if } x \ge a \\ q(x) & \text{if } x < a, \end{cases}$$

where $p(x)$ and $q(x)$ are continuous and differentiable in some open interval containing a. Then if f is continuous at a and $p'(a) = q'(a)$, then f is differentiable at $x = a$.

This theorem in particular can be applied to splines since the constituent functions are polynomials.

Example 4.5.7.

$$f(x) = \begin{cases} \sin x & \text{if } x \ge 0 \\ x^2 + x & \text{if } x < 0. \end{cases}$$

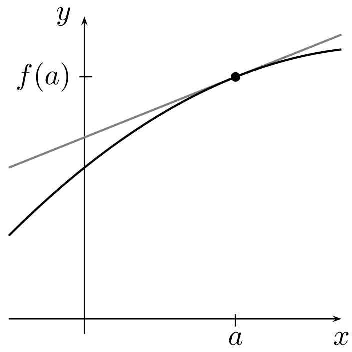
The functions $\sin x$ and $x^2 + x$ are continuous and differentiable everywhere. We saw in Example 4.5.4 that f is continuous at $x = 0$. Also, with $p(x) = \sin x$ and $q(x) = x^2 + x$, we have $p'(0) = \cos 0 = 1$ and $q'(0) = 1$ so f is differentiable at $x = 0$.

4.6 Derivatives and function approximation

(Ref: SH10 §4.11)

The last three sections of this chapter are devoted to some applications of the derivative.

Suppose that a function f is differentiable at a. Consider its graph and the tangent to f at a, shown below.



The equation of the tangent to f at a may be calculated using the point-gradient formula:

$$\begin{aligned} y - f(a) &= f'(a)(x-a) \\ y &= f'(a)(x-a) + f(a). \end{aligned}$$

Observe in the above diagram that the tangent line is close to the graph of f for all points that are close to a. In other words,

$$f(x) \approx f'(a)(x-a) + f(a) \tag{4.11}$$

when x is close to a.

Formula (4.11) reflects the fact that every differentiable function can be *locally* approximated by a linear function. We use this principle in the following example.

Example 4.6.1. Estimate $\sqrt{9.001}$ without using a calculator.

Solution. Suppose that $f(x) = \sqrt{x}$. Since 9.001 is close to 9 and f is a continuous function, we could use the approximation

$$\sqrt{9.001} = f(9.001) \approx f(9) = 3. \tag{4.12}$$

However, we can do better than this. We shall approximate f with a linear function at the point 9 using (4.11). Since $f'(9) = 1/6$ we have

$$f(x) \approx (x-9)/6 + 3$$

whenever x is close to 9. When $x = 9.001$ this gives

$$\sqrt{9.001} = f(9.001) \approx (9.001-9)/6 + 3 = 3 + 1/6000. \tag{4.13}$$

The table below gives the error involved in each approximation.

Approximation	Error
Approximation (4.12)	1.667^{-4}
Approximation (4.13)	4.637^{-9}

Clearly (4.13) is the superior approximation. $\square$

4.7 Derivatives and rates of change

(Ref: SH10 §3.4)

Many physical processes involve quantities (such as temperature, volume, concentration, velocity) that change with time. If Q is a quantity that varies with time, then the derivative $\frac{dQ}{dt}$ gives the rate of change of that quantity with respect to time. The chain rule is a major tool for answering questions about rates of quantities which are related.

Example 4.7.1. A spherical balloon is being inflated and its radius is increasing at a constant rate of 6 mm/sec. At what rate is its volume increasing when the radius of the balloon is 20 mm?

Solution. Suppose that $V(t)$ is the volume of the balloon and $r(t)$ is its radius at time t. We are told that $\frac{dr}{dt} = 6$ and we need to find $\frac{dV}{dt}$ when $r = 20$. By the chain rule,

$$\frac{dV}{dt} = \frac{dV}{dr}\frac{dr}{dt}.$$

Now $V = \frac{4}{3}\pi r^3$ so $\frac{dV}{dr} = 4\pi r^2$. Hence

$$\frac{dV}{dt} = 4\pi r^2 \times 6.$$

When $r = 20$,

$$\frac{dV}{dt} = 24\pi(20)^2 = 9600\pi.$$

So the volume is increasing at a rate of 9600π mm^3/sec when the radius is 20 mm. □

The above example illustrates an approach to solving such problems.

1. Define variables for the quantities involved.
2. Write down what is known in terms of these variables and their derivatives.
3. Write down what you need to find in terms of the these variables and their derivatives.
4. Write down anything else you know that relates the variables (for example, a volume or area formula).
5. Use the chain rule (or implicit differentiation) to find the relevant derivative.

Sometimes drawing a diagram helps in the first few steps.

The next example follows this procedure.

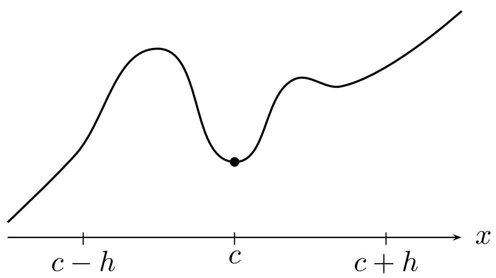
Example 4.7.2. A point P moves to the right along the positive x-axis at a constant rate of 5 cm per second, and a point Q moves up the positive y-axis at a constant rate of 10 cm per second. How fast is the distance between them changing when $OP = 30$ cm and $OQ = 40$ cm?

4.8 Local maximum, local minimum and stationary points

(Ref: SH10 §4.3)

In Example 3.4.4, we determined the absolute maximum and minimum points of a continuous function over a closed interval. While the existence of these extreme points is given by the max-min theorem, the theorem provides no systematic way of locating these points. In this section we begin to develop a systematic approach to locating maxima and minima. A complete approach will be presented in the next chapter.

We begin by defining what is meant by a *local minimum point* c. This is a point such that for all x sufficiently close to c, $f(c) \leq f(x)$. A more precise statement is below.



Local minimum point

Definition 4.8.1. Suppose that f is defined on some interval I. We say that a point c in I is a *local mininum point* if there is a positive number h such that $f(c) \leq f(x)$ whenever $x \in (c-h, c+h)$ and $x \in I$. We say that a point d in (a,b) is a *local maximum point* if there is a positive number h such that $f(x) \leq f(d)$ whenever $x \in (d-h, d+h)$ and $x \in I$.

The following theorem will be familiar from high school.

Theorem 4.8.2. *Suppose that f is defined on (a,b) and has a local maximum or minimum point at c for some c in (a,b). If f is differentiable at c then $f'(c) = 0$.*

Proof. Suppose that f is a local maximum at c. To show that $f'(c) = 0$ we consider the difference quotient

$$\frac{f(c+h) - f(c)}{h}.$$

Since f is a maximum at c, $f(c+h) \leq f(c)$ for h sufficiently close to 0. So if h is sufficiently small and positive,

$$\frac{f(c+h) - f(c)}{h} \leq 0,$$

while if h is sufficiently small and negative,

$$\frac{f(c+h) - f(c)}{h} \geq 0.$$

That is,

$$\lim_{h\to 0^+}\frac{f(c+h)-f(c)}{h}\le 0$$

while

$$\lim_{h\to 0^-}\frac{f(c+h)-f(c)}{h}\ge 0.$$

Hence

$$\lim_{h\to 0}\frac{f(c+h)-f(c)}{h}=0$$

which is equivalent to saying that $f'(c)=0$.

The proof when c is a local minimum is similar and will be omitted. ☐

The set of points where the derivative of a function is zero is important. In particular, tangents to the graph of the function at these points are horizontal lines. This motivates the following definition.

Definition 4.8.3. If a function f is differentiable at a point c and $f'(c)=0$ then c is called a *stationary point* of f.

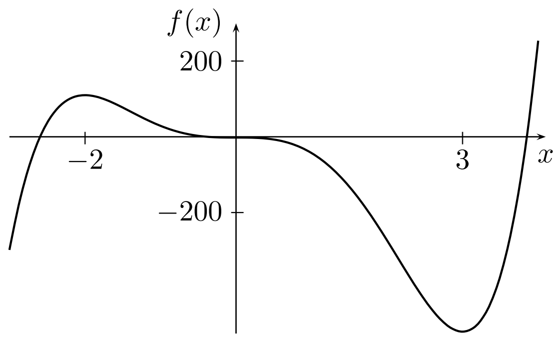
Example 4.8.4. Suppose that $f:\mathbb{R}\to\mathbb{R}$ is given by $f(x)=4x^5-5x^4-40x^3-2$. Find all the stationary points of f.

Solution. We begin by differentiating f and factorising the derivative:

$$\begin{aligned} f'(x) &= 20x^4-20x^3-120x^2 \\ &= 20x^2(x^2-x-6) \\ &= 20x^2(x+2)(x-3). \end{aligned}$$

Hence the solutions to $f'(x)=0$ are $x=-2,0,3$. So the points -2, 0 and 3 are the stationary points of f. ☐

The graph of the function f appearing in Example 4.8.4 is shown below.



It is clear that -2 is a local maximum point of f and that 3 is a local minimum point of f. However, 0 is neither a local minimum point nor local maximum point, and none of these points are global maximum points or global minimum points. So while the local maxima and minima of a function occur where f' is zero, it does not follow that any point c satisfying $f'(c) = 0$ is a local maximum point or local minimum point of f.

Tools for identifying whether a stationary point is a local maximum, local minimum or neither are developed in the next chapter using an important result called the *mean value theorem*.

4.9 Maple notes

The Maple `diff` command is used to compute derivatives: `diff(f(x),x);` calculates $\dfrac{df}{dx}$ and `diff(f(x),x$n);` calculates $\dfrac{d^n f}{dx^n}$. For example,

```
> # Note Maple use of csc to denote function we usually call cosec
> diff(sqrt(csc(x^12-53*x^5-1)),x);
```

$$-\tfrac{1}{2}\sqrt{\csc\left(x^{12}-53\,x^5-1\right)}\cot\left(x^{12}-53\,x^5-1\right)\left(12\,x^{11}-265\,x^4\right)$$

```
> diff(x^3*exp(5*x),x);
```

$$3\,x^2e^{5\,x}+5\,x^3e^{5\,x}$$

```
> diff(%,x);
```

$$6\,xe^{5\,x}+30\,x^2e^{5\,x}+25\,x^3e^{5\,x}$$

```
> f2ndDeriv:=diff(x^3*exp(5*x),x$2);
```

$$\mathit{f2ndDeriv} := 6\,xe^{5\,x}+30\,x^2e^{5\,x}+25\,x^3e^{5\,x}$$

Problems for Chapter 4

Problems 4.1 : Gradients of tangents and derivatives

1. [**R**] Using the definition of the derivative, show that:

   a) if $f(x) = x^2$ then $f'(x) = 2x$;

   b) if $f(x) = x^3$ then $f'(x) = 3x^2$;

   c) if $f(x) = \frac{1}{x}$ then $f'(x) = \frac{-1}{x^2}$;

   d) [**H**] if $f(x) = \sqrt{x}$ then $f'(x) = \frac{1}{2\sqrt{x}}$.

Problems 4.2 : Rules for differentiation

2. [**R**] Find the derivative in each case.

   a) $f(x) = 5(x^4 + 3x^7)$ b) $g(x) = (x^4 - 2x)(4x^2 + 2x + 5)$

   c) $h(y) = \dfrac{y^2}{y^3 + 8}$ d) $f(x) = x(x^2 - 4)^{1/2}$

   e) $f(t) = t/\sqrt{t^2 - 4}$ f) $g(y) = \sin 3y - 3\cos^2 2y$

   g) $g(x) = x^4 e^{-x}$ h) $f(x) = (x^2 + 1)\ \ln\sqrt{x^3 + 1}$

   i) $f(x) = \ln(e^{\tan x})$ j) $f(x) = \ln(\cos x)$

3. [**H**] Suppose that $f : \mathbb{R} \to \mathbb{R}$ is defined by $f(x) = x|x|$ for all x in $\mathbb{R}$.

   a) If it exists, evaluate $\displaystyle\lim_{h\to 0^+} \frac{f(0+h) - f(0)}{h}$.

   b) If it exists, evaluate $\displaystyle\lim_{h\to 0^-} \frac{f(0+h) - f(0)}{h}$.

   c) State the value of $f'(0)$ or explain why f is not differentiable at 0.

4. [**R**] [**V**] Determine at which points each function f is (i) differentiable; (ii) continuous.

   a) $f(x) = |x|$ b) $f(x) = \begin{cases} \sin x & \text{if } x \le 0 \\ x & \text{if } x > 0 \end{cases}$ c) $f(x) = \dfrac{x^3 - 6x + 4}{x^2 + 4x + 4}$

5. [**R**] Sketch the graph of f, where $f(x) = x^{1/3}$. Is f differentiable at 0? Give reasons.

6. [**X**] Prove that the function $f : \mathbb{R} \to \mathbb{R}$, given by

$$f(x) = \begin{cases} x^2 \sin \frac{1}{x} & \text{if } x \neq 0 \\ 0 & \text{if } x = 0, \end{cases}$$

is continuous and differentiable everywhere, but that f' is not continuous at 0.

7. **[X]** The function f is differentiable at a. Find

$$\lim_{h\to 0} \frac{f(a+ph)-f(a-ph)}{h}.$$

8. **[R]** (*An exercise on notation.*) Suppose that $f(x) = x + \cos 2x$. Write down

   a) $f(x+17\pi)$ b) $f'(x+n\pi)$ c) $f(2-x^2)$

   d) $f'(2-x^2)$ e) $\frac{d}{dx} f(2-x^2)$.

Problems 4.4 : Implicit differentiation

9. **[R]** Find $\dfrac{dy}{dx}$ in terms of x and y if

   a) $x^3 + y^3 = xy$ b) $x^2 - \sqrt{xy} + y^2 = 6$.

10. **[R]** Find $\dfrac{dy}{dx}$ for the curve $x^4 + y^4 = 16$. Sketch the graph of the curve.

11. **[R]** **[V]** Find the equation of the line tangent to the curve $x^3+y^3 = 3(x+y)$ at the point $(1, 2)$.

Problems 4.5 : Differentiation, continuity and split functions

12. **[R]** Suppose that a and b are real numbers. Find all values of a and b (if any) such that the functions f and g, given by

    a) $f(x) = \begin{cases} ax+b & \text{if } x<0 \\ \sin x & \text{if } x \ge 0 \end{cases}$ and b) $g(x) = \begin{cases} ax+b & \text{if } x<0 \\ e^{2x} & \text{if } x \ge 0 \end{cases}$,

    are (i) continuous at 0 and (ii) differentiable at 0.

13. **[H]** The function $f : \mathbb{R} \to \mathbb{R}$ is is defined by

$$f(x) = \begin{cases} \sqrt{x}\sin\sqrt{x} & \text{if } x \ge 0 \\ ax+b & \text{if } x < 0, \end{cases}$$

    where a and b are real numbers. Find all values of a and b (if any) such that f is differentiable at 0.

Problems 4.6 : Derivatives and function approximation

14. **[R]** **[V]** Suppose that $f(x) = \sqrt[3]{x}$.

    a) Without using a calculator, give a rough estimate of $f(8.01)$.

    b) i) Find the equation of the tangent to f at the point $(8, 2)$.

       ii) Use your answer to part (i) to find a different approximation for $f(8.01)$.

    c) Using a calculator, determine the error for the approximation in (a) and in (b). Which approximation is better?

Problems 4.7 : Derivatives and rates of change

15. [**R**] At a certain instant the side length of an equilateral triangle is a cm and this length is increasing at r cm/s. How fast, in cm^2/s, is the area increasing?

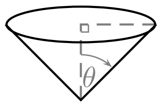
16. [**R**] [**V**] A 5 m ladder is leaning against a vertical wall. Suppose that the bottom of the ladder is being pulled away from the wall at a rate of 1 m/s. How fast is the area of the triangle underneath the ladder changing at the instant that the top of the ladder is 4 m from the floor?

17. [**R**] A spherical balloon is to be filled with water so that there is a constant increase in the rate of its surface area of $3\,\mathrm{cm}^2$/s.
(The surface area A and volume V of a sphere of radius r is given by $A = 4\pi r^2$ and $V = \frac{4}{3}\pi r^3$.)

    a) Find the rate of increase in the radius when the radius is 3 cm.
    b) Find the volume when the volume is increasing at a rate of $10\,\mathrm{cm}^3$/s.

18. [**R**]

    a) A container in the shape of a right circular cone, of semi-vertical angle $\tan^{-1}(\frac{1}{2})$, is placed vertex downwards with its axis vertical.

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

    Water is poured in at the rate of $10\,\mathrm{mm}^3$/s. Find the rate at which the depth, h mm, is increasing when the depth of water in the cone is 50 mm.

    b) [**H**] The cone is filled to a depth of 100 mm and pouring is then stopped. A hole is then opened at the vertex of the cone and water flows out of the hole at the rate of $50\pi\sqrt{h}\,\mathrm{mm}^3$ per second, where h is the depth at time t. Show that it takes 200 seconds to empty the cone.

Chapter 5

The mean value theorem and applications

The mean value theorem is one of the most important results for establishing the theoretical structure of calculus. Several results that you are familiar with from high school calculus are based on the mean value theorem. Applications of the mean value theorem include

- identifying where a function is increasing or decreasing,
- identifying different types of stationary points,
- determining how many zeros a polynomial has,
- evaluating limits which are indeterminate forms of type $\frac{\infty}{\infty}$ and $\frac{0}{0}$,
- proving useful inequalities and
- estimating errors in approximations.

We begin the chapter by introducing the theorem.

5.1 The mean value theorem

(Ref: SH10 §4.1)

The basic idea behind the mean value theorem is straightforward. Consider a differentiable function f and choose an interval $[a, b]$. We construct a chord AB joining the points $(a, f(a))$ and $(b, f(b))$ as shown in Figure 5.1. Because the function is continuous and differentiable, there is a point c in (a, b) such that tangent to the graph of f at c lies parallel to the chord AB.

Since the tangent and the chord are parallel, their gradients are equal. The gradient of the tangent at c is $f'(c)$ while the gradient of the chord is given by

$$\frac{f(b) - f(a)}{b - a}$$

(using the 'rise over run' formula). Hence

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

A precise statement of the theorem is given below.

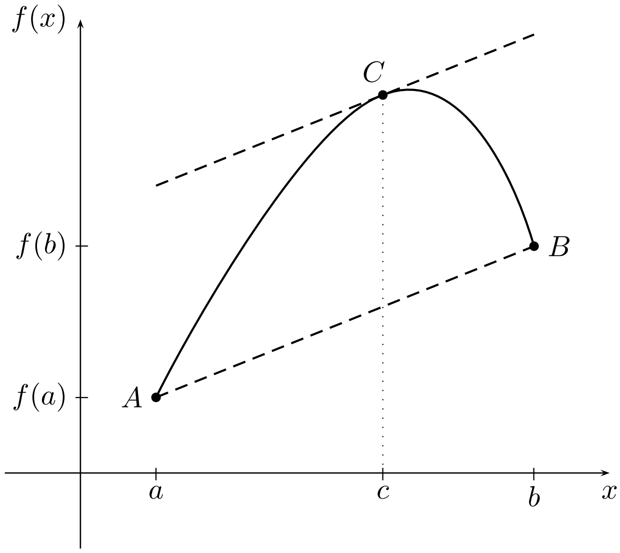


Figure 5.1: The mean value theorem – AB is parallel to the tangent at C.

Theorem 5.1.1 (The mean value theorem). *Suppose that f is continuous on $[a, b]$ and differentiable on (a, b). Then there is at least one real number c in (a, b) such that*

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

The proof of the theorem will be presented in the next section. For now we illustrate the theorem with a simple example.

Example 5.1.2. Suppose that $f : [1, 8] \to \mathbb{R}$ is given by $f(x) = \sqrt[3]{x}$. Find a number c in $(1, 8)$ that satisfies the conclusions of the mean value theorem for f on $[1, 8]$.

Solution. Note that f is continuous on $[1, 8]$ and differentiable on $(1, 8)$. By the mean value theorem there is a real number c in $(1, 8)$ such that

$$\frac{f(8) - f(1)}{8 - 1} = f'(c).$$

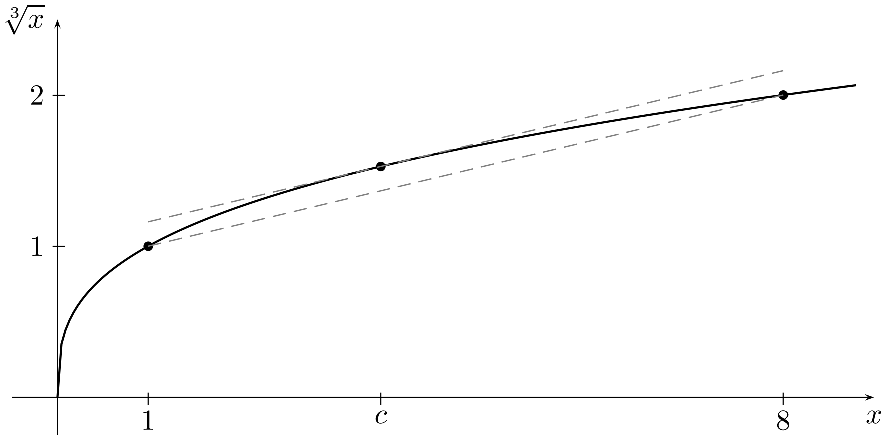
Now $f'(x) = \frac{1}{3}x^{-2/3}$ so the above equation becomes

$$\frac{2 - 1}{7} = \frac{1}{3\sqrt[3]{c^2}}.$$

Rearranging to find c gives

$$\begin{aligned} \sqrt[3]{c^2} &= \frac{7}{3} \\ c &= \sqrt{\frac{343}{27}}. \end{aligned}$$

Since $c \approx 3.56$ it is clear that $c \in (1, 8)$. This is illustrated graphically below.



□

5.2 Proof of the mean value theorem

(Ref: SH10 §4.1)

In this section we prove the mean value theorem. The proof is part of MATH1141 only, so MATH1131 students may want to jump straight to Section 5.3.

To prove the mean value theorem, we begin by considering the special case when $f(a) = f(b)$. This case is known as *Rolle's theorem* and is illustrated in Figure 5.2.

Theorem 5.2.1. *Suppose that g is continuous on $[a, b]$ and differentiable on (a, b). Suppose also that $g(a) = g(b) = 0$. Then there is a real number c in (a, b) such that $g'(c) = 0$.*

Proof. The proof considers three cases.

Case 1: Suppose that g is the constant function given by $g(x) = 0$. Then $g'(c) = 0$ for every c in (a, b) (see Section 4.2).

Case 2: Suppose that there is a point d in (a, b) such that $g(d) > 0$. By the max-min theorem (Theorem 3.4.3), g attains a maximum value at some point c in $[a, b]$. Moreover, c cannot be a or b since $g(d) > g(a) = g(b) = 0$. Hence c lies in (a, b), and since g is differentiable on (a, b), we have $g'(c) = 0$ (see Theorem 4.8.2).

Case 3: Suppose that $g(x) \leq 0$ for all x in $[a, b]$ and that g is not constant on $[a, b]$. Then g attains a minimum at a point c in (a, b), and similarly to Case 2 one can show that $g'(c) = 0$. □

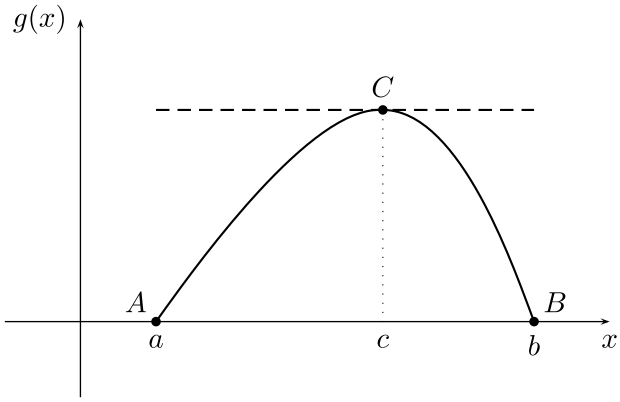


Figure 5.2: Rolle's theorem – the tangent at C is horizontal.

From here it is not hard to see how to prove the mean value theorem. Consider the graph of the function f in Figure 5.1 and 'subtract' the chord AB. The result will look something like the graph of g in Figure 5.2, to which we can apply Rolle's theorem and obtain a horizontal tangent at C. We then 'add' the chord back again to obtain the tangent to f of Figure 5.1. This rough geometric argument is made rigorous below.

Proof of the mean value theorem. Suppose that f is continuous on $[a, b]$ and differentiable on (a, b). We consider the function g given by

$$g(x) = f(x) - \left[\frac{f(b) - f(a)}{b - a}(x - a) + f(a)\right].$$

(The part in square brackets is taken from the equation of the chord AB in Figure 5.1.) One can check that g is continuous on $[a, b]$, differentiable on (a, b) and that $g(a) = g(b) = 0$. By Rolle's theorem there is a c in (a, b) such that $g'(c) = 0$; that is, such that

$$f'(c) - \left[\frac{f(b) - f(a)}{b - a}\right] = 0.$$

If we rearrange this equation then we obtain

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

as required. $\square$

5.3 Proving inequalities using the mean value theorem

Our first application of the mean value theorem will be to prove some useful inequalities. We begin with an example.

Example 5.3.1. By using the mean value theorem, show that $\ln x < x - 1$ whenever $x > 1$.

Solution. Suppose that $x > 1$ and consider the closed interval $[1, x]$. We define a function $f : [1, x] \to \mathbb{R}$ by $f(t) = \ln t$. Now f is continuous on $[1, x]$ and differentiable on $(1, x)$. So we may apply the mean value theorem to f on the interval $[1, x]$. Now $f'(t) = 1/t$ so

$$\frac{f(x) - f(1)}{x - 1} = \frac{1}{c}$$

for some c in $(1, x)$ by the mean value theorem. That is,

$$\frac{\ln x}{x-1} = \frac{1}{c}$$

for some c between 1 and x. Since $c > 1$ we have $1/c < 1$ and hence

$$\frac{\ln x}{x-1} < 1.$$

Rearranging the inequality gives $\ln x < x - 1$ as desired. $\square$

The technique used in the above example is as follows. Consider a function f that is continuous on the interval $[a, x]$ and differentable on the interval (a, x). The mean value theorem gives

$$\frac{f(x) - f(a)}{x - a} = f'(c) \tag{5.1}$$

for some c between a and x. If $f'(c) < M$ then

$$\frac{f(x) - f(a)}{x - a} < M$$

and hence

$$f(x) < M(x - a) + f(a). \tag{5.2}$$

whenever $x > a$.

Sometimes it is desirable for the inequality to go the other way. In this case we find m such that $f'(c) > m$. Then, instead of (5.2), we obtain

$$f(x) > m(x - a) + f(a).$$

Often, the tricky part of the process is determining which function f will give the desired inequality.

Example 5.3.2. Use the mean value theorem to show that

$$\sqrt{x+4} - 2 < \frac{x}{4}$$

whenever $x > 0$.

Solution. Fix positive x. The first step is to identify a good choice of f. If we rewrite the inequality as

$$\sqrt{x+4} < \frac{x}{4} + 2$$

then we have a similar form to inequality (5.2). This suggests that we should define $f : [0, x] \to \mathbb{R}$ by the formula $f(t) = \sqrt{t+4}$. Then f is continuous on $[0, x]$ and differentiable on $(0, x)$. An application of the mean value theorem gives

$$\frac{f(x) - f(0)}{x - 0} = \frac{1}{2\sqrt{c+4}}$$

for some c in $(0, x)$. Now the right-hand side is bounded above by $\frac{1}{2\sqrt{4}}$, so we have

$$\frac{\sqrt{x+4} - \sqrt{0+4}}{x - 0} < \frac{1}{2\sqrt{4}}.$$

Simplifying and rearranging gives

$$\sqrt{x+4} - 2 < \frac{x}{4}$$

as required. $\square$

5.4 Error bounds

A second application of the mean value theorem is in calculating error bounds. Suppose that a calculation involves $\sqrt{27}$. Clearly $\sqrt{27}$ is close to $\sqrt{25}$, but the latter is easier to work with because it is equal to 5. If we use 5 instead of $\sqrt{27}$ then calculations are made easier with the tradeoff that we introduce an error. How bad is this error?

Example 5.4.1. Use the mean value theorem to find an upper bound for the error involved if we approximate $\sqrt{27}$ by $\sqrt{25}$.

Solution. The precise error in the approximation is given by

$$\sqrt{27} - \sqrt{25}.$$

We apply the mean value theorem to the function f, given by $f(x) = \sqrt{x}$, on the interval $[25, 27]$. This gives

$$\frac{\sqrt{27} - \sqrt{25}}{27 - 25} = \frac{1}{2\sqrt{c}}$$

for some c in $(25, 27)$. Rearranging this gives us

$$\sqrt{27} - \sqrt{25} = \frac{1}{\sqrt{c}}$$

where $25 < c < 27$. Hence

$$\begin{aligned} \text{error} &= \sqrt{27} - \sqrt{25} \\ &= \frac{1}{\sqrt{c}} \\ &< \frac{1}{\sqrt{25}} \\ &= \frac{1}{5}, \end{aligned}$$

since $c > 25$. Hence an upper bound for the error is $1/5$. □

We illustrate the process with one more example.

Example 5.4.2. Use the mean value theorem to find an upper bound for the error involved if we approximate $\sin \frac{5\pi}{7}$ by $\sin \frac{2\pi}{3}$.

Solution. Note that since

$$\frac{2\pi}{3} = \frac{14\pi}{21} \approx \frac{15\pi}{21} = \frac{5\pi}{7}$$

and the sine function is continuous,

$$\sin \frac{2\pi}{3} \approx \sin \frac{5\pi}{7}.$$

The (absolute) error of the approximation is exactly

$$\left| \sin \frac{5\pi}{7} - \sin \frac{2\pi}{3} \right|.$$

If we apply the mean value theorem to the sine function on the interval $[2\pi/3, 5\pi/7]$ then

$$\frac{\sin\frac{5\pi}{7} - \sin\frac{2\pi}{3}}{\frac{5\pi}{7} - \frac{2\pi}{3}} = \cos c$$

for some c in $(2\pi/3, 5\pi/7)$. Taking absolute values of both sides and rearranging gives

$$\left|\sin\frac{5\pi}{7} - \sin\frac{2\pi}{3}\right| = \left|\frac{5\pi}{7} - \frac{2\pi}{3}\right| |\cos c|$$

where $2\pi/3 < c < 5\pi/7$. Now $|\cos c| < |\cos\frac{5\pi}{7}|$ (because the absolute value of the cosine is increasing over the interval $(\pi/2, \pi)$) but we don't have an exact value for $\cos\frac{5\pi}{7}$. A simple way out of this situation is to just use the fact that $|\cos c| < 1$ for all c in $(2\pi/3, 5\pi/7)$. Hence we have shown that

$$\text{absolute error} = \left|\sin\frac{15\pi}{21} - \sin\frac{2\pi}{3}\right| < \left|\frac{5\pi}{7} - \frac{2\pi}{3}\right| \times 1.$$

This proves that

$$\text{absolute error} < \frac{\pi}{21} < 0.15.$$

If we aren't so lazy with our estimates we can quickly spot that

$$\left|\cos\frac{5\pi}{7}\right| = \left|\cos\frac{15\pi}{21}\right| < \left|\cos\frac{15\pi}{20}\right| = \frac{1}{\sqrt{2}}.$$

This allows us to prove the more accurate statement that

$$\text{absolute error} < \frac{\pi}{21\sqrt{2}} < 0.11.$$

□

5.5 The sign of a derivative

(Ref: SH10 §§4.2, 4.3)

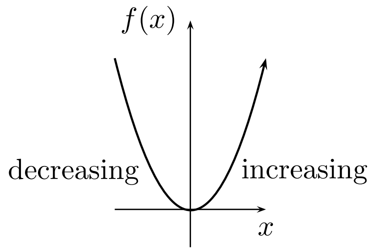
Suppose that a function f is differentiable at 4 and that $f'(4) = -2$. In high school, it was taught that, since the derivative of f at 4 is negative, the function f is decreasing at the point 4. This makes intuitive sense since the gradient of the tangent to f at the point 4 is -2 and so the tangent slopes downward. In this section, these ideas are given a rigorous foundation by using the mean value theorem.

Definition 5.5.1. Suppose that a function f is defined on an interval I. We say that

a) f *is increasing* on I if for every two points x_1 and x_2 in I,
$$x_1 < x_2 \text{ implies that } f(x_1) < f(x_2);$$

b) f *is decreasing* on I if for every two points x_1 and x_2 in I,
$$x_1 < x_2 \text{ implies that } f(x_1) > f(x_2).$$

Example 5.5.2. The function $f : \mathbb{R} \to \mathbb{R}$ given by $f(x) = x^2$ is increasing on $[0, \infty)$ and decreasing on $(-\infty, 0]$.



The following theorem should be familiar to you. It is proved using the mean value theorem.

Theorem 5.5.3. *Suppose that f is continuous on $[a, b]$ and differentiable on (a, b).*

(i) If $f'(x) > 0$ for all x in (a, b) then f is increasing on $[a, b]$.

(ii) If $f'(x) < 0$ for all x in (a, b) then f is decreasing on $[a, b]$.

(iii) If $f'(x) = 0$ for all x in (a, b) then f is constant on $[a, b]$.

Proof. We only prove statement (i) since the proofs of statements (ii) and (iii) are similar. Suppose that $f'(x) > 0$ for all x in (a, b) and choose two points x_1 and x_2 in $[a, b]$ such that $x_1 < x_2$. By Definition 5.5.1, it suffices to show that $f(x_1) < f(x_2)$.

Since f is differentiable on I, it is continuous on $[x_1, x_2]$ (see Theorem 4.5.1) and differentiable on (x_1, x_2). Hence, by the mean value theorem, there is a c in (x_1, x_2) such that

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

But $f'(c) > 0$ and $x_2 - x_1 > 0$. This means that $f(x_2) - f(x_1) > 0$ and hence $f(x_1) < f(x_2)$ as required. □

Theorem 5.5.3 has several important applications. The first of these is classifying different types of stationary points. The basic idea is as follows. Suppose that f is differentiable on the open interval $(1, 3)$ and that for some small positive number h we have the following:

- $f'(x) > 0$ for all x in $(2 - h, 2)$,
- $f'(2) = 0$, and
- $f'(x) < 0$ for all x in $(2, 2 + h)$.

Now f is differentiable on $(2 - h, 2 + h)$ so it is also continuous on this interval. Moreover, f increases on the interval $(2 - h, 2)$, is stationary at 2 and then decreases on the interval $(2, 2 + h)$. From these facts we conclude that the stationary point 2 is a local maximum point for f.

Example 5.5.4. Find and classify all stationary points of the function $f : \mathbb{R} \to \mathbb{R}$ whose derivative is given by

$$f'(x) = (x - 4)(x - 1)(x + 5)^2.$$

Solution. The stationary points, which are found by solving the equation $f'(x) = 0$, are 4, 1 and -5. To classify each stationary point, we examine the sign of f' on small intervals either side of each stationary point. This process is concisely documented in the following table.

	-5^-	-5	-5^+	1^-	1	1^+	4^-	4	4^+
$x-4$	$-$	$-$	$-$	$-$	$-$	$-$	$-$	0	$+$
$x-1$	$-$	$-$	$-$	$-$	0	$+$	$+$	$+$	$+$
$(x+5)^2$	$+$	0	$+$	$+$	$+$	$+$	$+$	$+$	$+$
$f'(x)$	$+$	0	$+$	$+$	0	$-$	$-$	0	$+$
Gradient	$/$	$-$	$/$	$/$	$-$	$\backslash$	$\backslash$	$-$	$/$

For example, the notation 4^- is shorthand for those points x lying in a small interval to the immediate left of 4. For such a point x, the factor $x-4$ is negative, while the factors $x-1$ and $(x+5)^2$ are positive. Hence $f'(x)$ is negative when x is slightly to the left of 4. Other columns of the table are filled by a similar process.

In conclusion, 4 is a local minimum point, 1 is a local maximum point and -5 is neither a maximum nor minimum point. (In fact, -5 is called a *horizontal point of inflexion* since the tangent to f at -5 is horizontal and the concavity of the function changes about this point. See any standard undergraduate calculus text or any advanced high school calculus text for further details.) □

5.6 The second derivative and applications

(Ref: SH10 §4.3)

In this section we give another method for classifying the stationary points of a function f. To do so, we use the *second derivative of* f, which is denoted by f''.

Example 5.6.1. Find the second derivative of the function $f : \mathbb{R} \to \mathbb{R}$, given by the rule

$$f(x) = 3x^4 + 2x - \sin x.$$

Solution. The *first derivative* of f is given by

$$f'(x) = 12x^3 + 2 - \cos x.$$

To find the second derivative f'' of f, we simply differentiate f'. Hence

$$f''(x) = 36x^2 + \sin x.$$

□

Remark 5.6.2. If $y = f(x)$ then the first derivative of f is often written as $\frac{dy}{dx}$ or y' while the second derivative of f is often written as $\frac{d^2y}{dx^2}$ or y''.

Remark 5.6.3. Even if a function is differentiable at a point a, the function may not have a second derivative at a. (Consider, for example, the function $f : \mathbb{R} \to \mathbb{R}$ given by

$$f(x) = x^{4/3},$$

at the point 0.) A function which has a second derivative at a is called *twice differentiable* at a.

If f is twice differentiable, the sign of f'' can often be used to determine whether a stationary point is a local maximum point or a local minimum point.

Theorem 5.6.4 (The second derivative test). *Suppose that a function f is twice differentiable on (a, b) and that $c \in (a, b)$.*

(i) If $f'(c) = 0$ and $f''(c) > 0$ then c is a local minimum point of f;

(ii) If $f'(c) = 0$ and $f''(c) < 0$ then c is a local maximum point of f.

The proof of the theorem uses the following result.

Lemma 5.6.5. *Suppose that g is differentiable at a point c.*

(i) If $g'(c) > 0$ then $g(c - h) < g(c) < g(c + h)$ for all positive h sufficiently small.

(ii) If $g'(c) < 0$ then $g(c + h) < g(c) < g(c - h)$ for all positive h sufficiently small.

Sketch proof. We give an intuitive argument for statement (i) and leave (ii) as an exercise. Suppose that $g'(c) > 0$. By the definition of the derivative,

$$\lim_{h \to 0} \frac{g(c + h) - g(c)}{h} = g'(c). \tag{5.3}$$

That is,

$$\frac{g(c + h) - g(c)}{h} \approx g'(c)$$

when h is close to 0. Now, if $h > 0$ then

$$g(c + h) - g(c) \approx hg'(c) > 0$$

Hence

$$g(c + h) - g(c) > 0 \tag{5.4}$$

provided that $h > 0$ and that h is sufficiently small.

On the other hand, the definition of the derivative also implies that

$$\lim_{h \to 0} \frac{g(c - h) - g(c)}{-h} = g'(c)$$

(to see this, simply replace h with $-h$ in (5.3)). Now, if $h > 0$ and h is close to 0 then

$$g(c) - g(c - h) \approx hg'(c) > 0.$$

Hence

$$g(c) - g(c - h) > 0 \tag{5.5}$$

provided that $h > 0$ and that h is sufficiently small.

If we combine (5.4) and (5.5) then we have

$$g(c - h) < g(c) < g(c + h)$$

for all positive h sufficiently small. □

Proof of Theorem 5.6.4. We prove statement (i) and leave the proof of (ii) as an exercise. Suppose that $f'(c) = 0$, $f''(c) > 0$ and that $g = f'$. Then $g'(c) > 0$ and

$$g(c-h) < g(c) < g(c+h)$$

for all positive h sufficiently small, by Lemma 5.6.5. Since $f'(c) = g(c) = 0$ we conclude that

$$f'(c-h) < 0 < f'(c+h)$$

for all positive h sufficiently small. So f' is positive on a small interval immediately to the right of c, zero at x and negative on a small interval immediately to the left of c. By Theorem 5.5.3 we now conclude that f has a local minimum at the point c. □

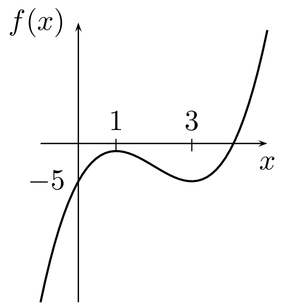
Example 5.6.6. Find and classify the stationary points of the function $f : \mathbb{R} \to \mathbb{R}$ given by

$$f(x) = x^3 - 6x^2 + 9x - 5.$$

Solution. First we calculate the first and second derivatives of f:

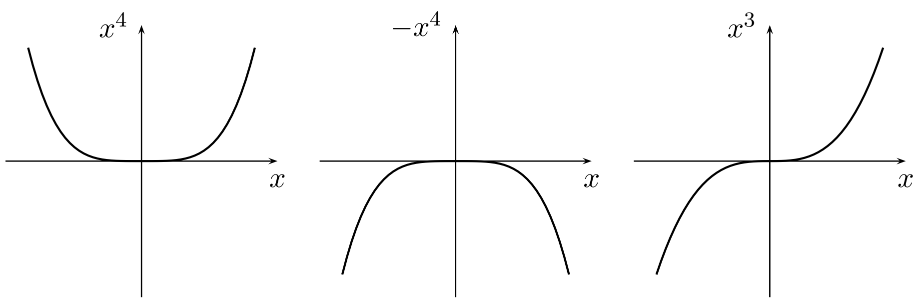
$$\begin{aligned} f'(x) &= 3x^2 - 12x + 9 = 3(x-1)(x-3) \\ f''(x) &= 6x - 12. \end{aligned}$$

Hence f has stationary points at 1 and 3. Now $f''(1) = -6 < 0$, so f has a local maximum point at 1. For second stationary point, $f''(3) = 6 > 0$, so f has a local minimum stationary point at 3. (To illustrate these conclusions, we have included the graph of f, shown below.)



□

Remark 5.6.7. Suppose that $f'(c) = 0$ and that $f''(c) = 0$. Then we cannot apply the second derivative test. In fact, the point c may be a local maximum, a local minimum or neither of these. The examples below illustrate various possibilities.



- If $f(x) = x^4$ then $f'(0) = f''(0) = 0$ and there is a local *minimum* at 0.
- If $f(x) = -x^4$ then $f'(0) = f''(0) = 0$ and there is a local *maximum* at 0.
- If $f(x) = x^3$ then $f'(0) = f''(0) = 0$ and there is neither a a local maximum and minimum at 0 (in fact, we have a horizontal point of inflexion at 0).

Hence if $f'(c) = f''(c) = 0$ then it is best to classify the stationary point c by examining the sign of the derivative on either side of c, as illustrated in Section 5.5.

5.7 Critical points, maxima and minima

(Ref: SH10 §§4.3–4.5)

By combining results from Chapters 3, 4 and 5, we now give a comprehensive presentation of how to locate global maxima and minima for any real-valued continuous function defined on a closed interval.

Suppose that a function $f : [a, b] \to \mathbb{R}$ is continuous on $[a, b]$. By the max-min theorem (Theorem 3.4.3), f attains a global maximum and a global minimum on $[a, b]$. Suppose that the global maximum occurs at the point c. Now either c is one of the endpoints a or b, or c lies in the open interval (a, b). In the latter case, f is either differentiable at c or it is not. If it is differentiable at c, then $f'(c) = 0$ by Theorem 4.8.2.

Thus we have identified that a global maximum point c must satisfy one of the three following properties:

- c is one of the endpoints a or b of the closed interval $[a, b]$,
- f is differentiable at c and $f'(c) = 0$, or
- f is not differentiable at c.

The set of points satisfying these properties is useful so we give it a special name.

Definition 5.7.1. Suppose that f is defined on $[a, b]$. We say that a point c in $[a, b]$ is a *critical point* for f on $[a, b]$ if c satisfies one of the following properties:

(a) c is an endpoint a or b of the interval $[a, b]$,

(b) f is not differentiable at c, or

(c) f is differentiable at c and $f'(c) = 0$.

Theorem 5.7.2. *Suppose that f is continuous on $[a, b]$. Then f has a global maximum and global minimum on $[a, b]$. Moreover, the global maximum point and the global minimum point are both critical points for f on $[a, b]$.*

The proof of the theorem follows the same reasoning given at the beginning of this section. The theorem provides a systematic way of finding global minima and maxima of functions.

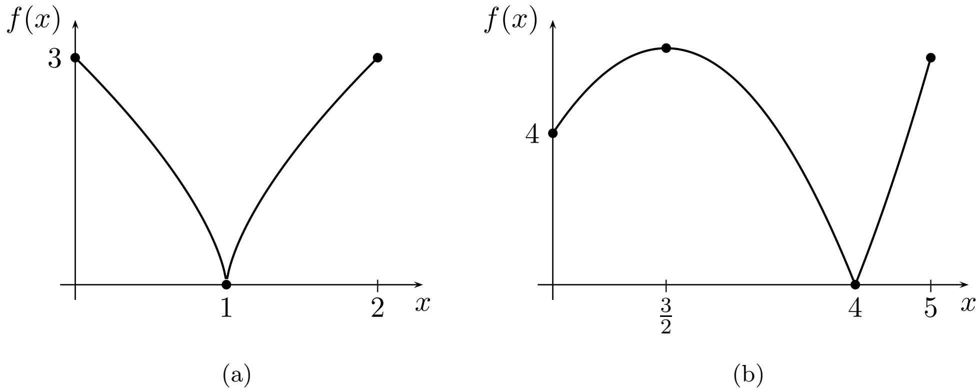


Figure 5.3: Graphs of functions and their critical points.

Example 5.7.3. Suppose that $f : \mathbb{R} \to \mathbb{R}$ is given by

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

Find the absolute maximum and minimum values of f on $[0, 2]$.

Solution. The derivative of f is given by

$$f'(x) = 2(x-1)^{-1/3} = \frac{2}{\sqrt[3]{x-1}}.$$

Clearly f is not differentiable at 1. Morover, its derivative is never zero. So the only critical points of f on $[0, 2]$ are 1 and the endpoints 0 and 2. By Theorem 5.7.2, the global maximum and minimum points must be among the points 0, 1 and 2. But

$$f(0) = 3, \quad f(1) = 0 \quad \text{and} \quad f(2) = 3.$$

Hence the global minimum value of f on $[0, 2]$ is 0 while the global maximum value of f on $[0, 2]$ is 3. (See Figure 5.3 (a) for a sketch of f that shows its critical points.) □

Example 5.7.4. Suppose that the function $f : \mathbb{R} \to \mathbb{R}$ is given by the rule

$$f(x) = |x^2 - 3x - 4|.$$

Find the absolute maximum and absolute minimum values of f on the interval $[0, 5]$.

Example 5.7.5. Consider the parabolic arc

$$\{(x, y) \in \mathbb{R}^2 : y = x^2 - x - 1, -2 \leq x \leq 3\}$$

(see Figure 5.4 (a)). Find where the arc is (i) closest to and (ii) furthest from the origin.

Solution. If (x, y) is a point on the arc then its distance from the origin is given by

$$\sqrt{x^2 + y^2}.$$

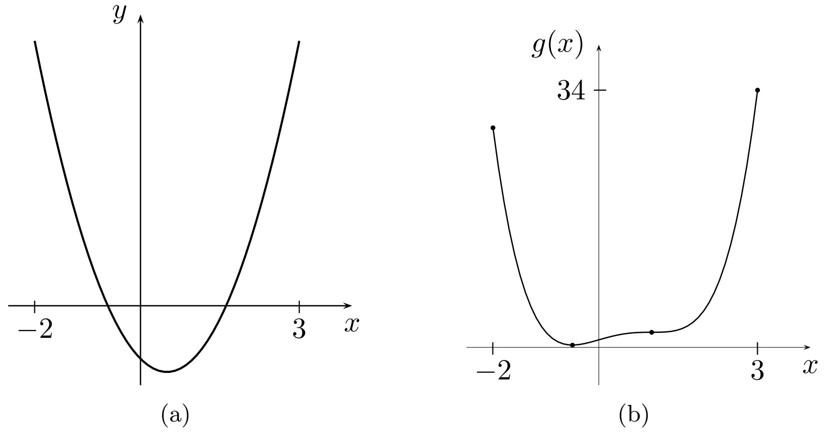


Figure 5.4: Graphs for Example 5.7.5.

To find where the arc is closest to the origin, we want to minimise $\sqrt{x^2+y^2}$ subject to the constraints $y = x^2 - x - 1$ and $-2 \le x \le 3$. In other words, we want to find a global minimum for the function

$$f(x) = \sqrt{x^2 + (x^2 - x - 1)^2}$$

on the interval $[-2, 3]$. Because calculations involving square roots can get messy, we instead consider the function g, given by

$$g(x) = x^2 + (x^2 - x - 1)^2$$

on $[-2, 3]$. (Note that $g(x)$ may be interpreted as the square of the distance between the origin and a point on the arc with coordinate $(x, x^2 - x - 1)$. The global minimum of both f and g will occur at the same point x.)

We seek the critical points of g, so we compute its derivative:

$$\begin{aligned} g'(x) &= 2x + 2(x^2 - x - 1)(2x - 1) \\ &= 2(x + 2x^3 - 2x^2 - 2x - x^2 + x + 1) \\ &= 2(2x^3 - 3x^2 - x^2 + 1) \\ &= 2(x - 1)(2x^2 - x - 1) \\ &= 2(x - 1)(x - 1)(2x + 1). \end{aligned}$$

Hence the critical points of g on $[-2, 3]$ are -2, $-1/2$, 1 and 3. Now

$$g(-2) = 29, \quad g(-1/2) = 5/16, \quad g(1) = 2, \quad g(3) = 34$$

(see Figure 5.4 (b)). Therefore the arc is closest to the origin when $x = -1/2$ and furthest from the origin when $x = 3$. □

5.8 Counting zeros

In this section we combine the intermediate value theorem with results about the sign of derivatives to determine how many solutions there are to various equations. The main ideas are illustrated in the next example.

Example 5.8.1. Determine how many real numbers satisfy the equation

$$x^3 - 6x^2 - 15x + 8 = 0. \tag{5.6}$$

Solution. Suppose that $f(x) = x^3 - 6x^2 - 15x + 8$ for all real x. Ideally, we would look for one root of f, factor this out and solve a quadratic equation to determine the remaining roots (if any). However, finding one root by inspection is not straightforward, so we need another approach.

Note that f is differentiable (and hence continuous) everywhere. We begin by identifying intervals where f is monotonically decreasing and intervals where f is monotonically increasing. The derivative of f is given by

$$f'(x) = 3x^2 - 12x - 15 = 3(x-5)(x+1).$$

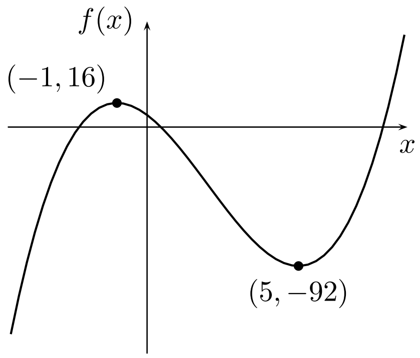
So the stationary points of f are 5 and -1. Morover, the table below shows that f is increasing on the intervals $(-\infty, -1]$ and $[5, \infty)$, while it is decreasing on $[-1, 5]$.

Interval	$(-\infty, -1)$	$(-1, 5)$	$(5, \infty)$
$(x-5)$	$-$	$-$	$+$
$(x+1)$	$-$	$+$	$+$
$f'(x)$	$+$	$-$	$+$

We evaluate f at the (finite) endpoints of these intervals and see that

$$f(-1) = 16 \quad \text{and} \quad f(5) = -92.$$

Since f is a cubic, we can now give a rough sketch of the graph of f.



We now argue that f has *exactly* one root on each of the intervals mentioned above.

- On the interval $[-1, 5]$, the function f is decreasing. Hence f cannot have more than one root on $[-1, 5]$. On the other hand, since $f(-1) > 0$ and $f(5) < 0$ there is at least one c in $[-1, 5]$ such that $f(c) = 0$ by the intermediate value theorem. Hence f has exactly one root on $[-1, 5]$.

- On the interval $(-\infty, -1]$, the function f is increasing. Hence f cannot have more than one root on $(-\infty, -1]$. On the other hand, since $f(-1) > 0$ and $f(-5) < 0$ there is at least one c in $[-5, -1]$ such that $f(c) = 0$ by the intermediate value theorem. Hence f has exactly one root on $(-\infty, -1]$.

- On the interval $[5, \infty)$, the function f is increasing. Hence f cannot have more than one root on $[5, \infty)$. On the other hand, since $f(5) < 0$ and $f(10) > 0$ there is at least one c in $[5, 10]$ such that $f(c) = 0$ by the intermediate value theorem. Hence f has exactly one root on $[5, \infty)$.

In conclusion, equation (5.6) has exactly three real solutions, one in each of the intervals $[-5, -1]$, $[-1, 5]$ and $[5, 10]$. $\square$

Guided by our solution to Example 5.8.1, we now outline a general approach for solving problems of this type.

Suppose that f is continuous and differentiable everywhere and that its derivative is also continuous. To determine how many real solutions the equation $f(x) = 0$ has, one may follow the procedure below.

1. Calculate f' and solve $f'(x) = 0$.

2. Determine the intervals where the derivative is positive and the intervals where it is negative.

3. By step 2 we know the intervals where the function f is monotonically increasing, and the intervals where f is monotonically decreasing.

4. Evaluate the function f at the endpoints of each interval. If f changes sign on the interval, there is exactly one root on that interval. If it does not change sign, there are no roots on that interval. (This step must be modified slightly for intervals of the form $(-\infty, a]$ or $[b, \infty)$.)

Example 5.8.2. Determine how many real numbers satisfy the equation

$$2x^3 - 9x^2 + 12x - 1 = 0. \tag{5.7}$$

Solution. We follow the steps outlined above. Suppose that

$$f(x) = 2x^3 - 9x^2 + 12x - 1.$$

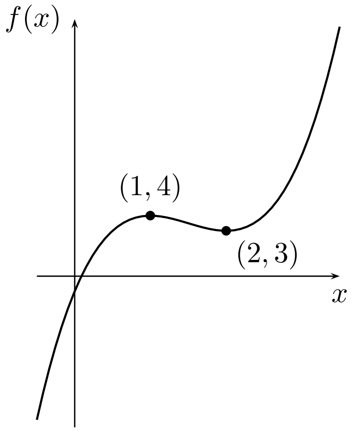
By differentiating f we obtain

$$f'(x) = 6x^2 - 18x + 12 = 6(x - 1)(x - 2).$$

The stationary points of f are 1 and 2. By studying the sign of f', it is clear that f is increasing on the intervals $(-\infty, 1)$ and $(2, \infty)$ while it is decreasing on $(1, 2)$. Now

$$f(1) = 4 \quad \text{and} \quad f(2) = 3,$$

as illustrated in the sketch below.



So it is clear that f has no zeros on $[1, 2]$ or on $[2, \infty)$. On $(-\infty, 1)$, the function f has no more than one zero (since it is increasing on this interval). Moreover, f changes sign on the interval $[0, 1]$ (since $f(0) = -1 < 0$ and $f(1) > 0$). Hence f has exactly one root on $[0, 1]$ by the intermediate value theorem.

In conclusion, equation (5.7) has exactly one real solution and this solution lies somewhere in the interval $[0, 1]$. $\square$

5.9 Antiderivatives

We open with an example.

Example 5.9.1. While filming a movie, a stunt man jumps out of a stationary helicopter 900 meters above the ground. The scriptwriter wants the man to make a 20 second mobile phone call before opening the parachute. Physical considerations (gravity and air resistance) suggest that the man's velocity $f(t)$ (in metres per second) without an open parachute is given by

$$f(t) = 50(1 - e^{-t/5}),$$

where t is the number of seconds after jumping from the helicopter. How far will the man fall 20 seconds?

Solution. Write $F(t)$ for the distance fallen (in metres) after t seconds. We note that $F(0) = 0$. Since velocity is the rate of change of displacement, F satisfies the equation

$$F'(t) = f(t).$$

(Any equation like this involving differentiation is called a *differential equation*). The differential equation may be rewritten as

$$\frac{dF}{dt} = 50 - 50e^{-t/5}. \tag{5.8}$$

One can check by differentiation that if

$$F(t) = 50t + 250e^{-t/5} + C, \tag{5.9}$$

where C is a real constant, then F satisfies (5.8). Let's *assume* (for the moment) that all possible solutions to the differential equation (5.8) are of the form (5.9). Imposing the condition that $F(0) = 0$ allows us to evaluate the unknown constant C:

$$0 = F(0) = 50 \times 0 + 250e^0 + C,$$

whence $C = -250$. Therefore $F(t) = 50t + 250e^{-t/5} - 250$. To complete the solution,

$$F(20) = 50 \times 20 + 250e^{-4} - 250 \approx 754.6.$$

Hence the stunt man will fall about 755 metres in the first 20 seconds. $\square$

To be certain of this solution, we need to show that *every* function F satisfying the differential equation (5.8) is of the form (5.9). This can be done using the mean value theorem. We begin by introducing the notion of an *antiderivative*.

Definition 5.9.2. Suppose that f is continuous on an open interval I. A function F is said to be an *antiderivative* (or a *primitive*) of f on I if $F'(x) = f(x)$ for all x in I. The process of finding an antiderivative of a function is called *antidifferentiation*.

Example 5.9.3. (a) Suppose that

$$f(t) = 50 - 50e^{-t/5} \qquad \forall t \in \mathbb{R}$$

Then the function F, given by

$$F(t) = 50t + 250e^{-2t} - 250$$

is an antiderivative of f on $(-\infty, \infty)$.

(b) Suppose that n is a positive integer and $f(x) = x^n$. If

$$F(x) = \frac{x^{n+1}}{n+1} \qquad \text{and} \qquad G(x) = \frac{x^{n+1}}{n+1} + 5$$

then F and G are both antiderivatives of f on the interval $(-\infty, \infty)$. So f has more than one antiderivative on $(-\infty, \infty)$. In fact, if C is any real number and

$$H(x) = x^{n+1}/(n+1) + C$$

then H is an antiderivative of f on $(-\infty, \infty)$. Thus the function f has infinitely many antiderivatives on $(-\infty, \infty)$.

Example 5.9.3 (b) illustrates the general principle that if F is an antiderivative of f on I then

$$F + C$$

is also an antiderivative of f on I, for any real constant C. The next theorem says that all antiderivatives are of this form.

Theorem 5.9.4. *Suppose that f is a continuous function on an open interval I and that F and G are two antiderivatives of f on I. Then there is a real constant C such that*

$$G(x) = F(x) + C$$

for all x in I.

Proof. Suppose that F and G are two antiderivatives of f on I and let H denote the function given by

$$H(x) = G(x) - F(x)$$

for all x in I. Then H is differentiable on I and

$$\begin{aligned} H'(x) &= G'(x) - F'(x) \\ &= f(x) - f(x) \\ &= 0 \end{aligned}$$

for all x in I. Hence there is a constant C such that $H(x) = C$ for all x in I (by Theorem 5.5.3 (c)). Therefore

$$G(x) = F(x) + H(x) = F(x) + C$$

for all x in I. □

Theorem 5.9.4 justifies the assumption made in the solution to Example 5.9.1. It also enables us to write down all possible antiderivatives of some well-known functions. In the table below, C is any real constant.

Function	Antiderivative
x^r, where r is rational and $r \neq -1$	$\frac{1}{r+1}x^{r+1} + C$
$\sin x$	$-\cos x + C$
$\cos x$	$\sin x + C$
e^{ax}	$\frac{1}{a}e^{ax} + C$
$\dfrac{f'(x)}{f(x)}$	$\ln\lvert f(x)\rvert + C$

5.10 L'Hôpital's rule

(Ref: SH10 §§11.5, 11.6)

Consider the limit

$$\lim_{x \to 0} \frac{e^x - 1}{5x + x^2}.$$

Since both the numerator and denominator approach 0 as $x \to 0^+$, the limit is an example of an indeterminate form of type $\frac{0}{0}$. None of the limit rules of Chapter 2 can be easily applied to evaluate the limit. However, the mean value theorem gives another rule which helps in this situation.

Suppose that $f(x) = e^x - 1$ and $g(x) = 5x + x^2$. Instead of considering the quotient $f(x)/g(x)$ (whose limit is difficult to calculate), we consider the quotient $f'(x)/g'(x)$ of derivatives (whose limit is easier to calculate):

$$\lim_{x \to 0} \frac{f'(x)}{g'(x)} = \lim_{x \to 0} \frac{e^x}{5 + 2x} = \frac{1}{5}.$$

Since this second limit exists, *l'Hôpital's rule* (which is stated below) implies that

$$\lim_{x\to 0}\frac{f(x)}{g(x)} = \lim_{x\to 0}\frac{f'(x)}{g'(x)}.$$

Hence

$$\lim_{x\to 0}\frac{e^x-1}{5x+x^2} = \frac{1}{5}.$$

Theorem 5.10.1 (l'Hôpital's rule). *Suppose that f and g are both differentiable functions and a is a real number. Suppose also that either one of the two following conditions hold:*

- $f(x)\to 0$ *and* $g(x)\to 0$ *as* $x\to a$*;*
- $f(x)\to\infty$ *and* $g(x)\to\infty$ *as* $x\to a$*.*

If

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

exists then

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

Remark 5.10.2. The theorem also holds for

- limits as $x\to\infty$ or $x\to-\infty$;
- one-sided limits (that is, as $x\to a^+$ or as $x\to a^-$).

L'Hôpital's rule is proved using the mean value theorem (see the sketch proof at the end of this section).

Example 5.10.3. Calculate $\displaystyle\lim_{x\to 1}\frac{\ln x}{1-x}$.

Solution. If $f(x)=\ln x$ and $g(x)=1-x$ then both $f(x)$ and $g(x)$ approach 0 as $x\to 1$. Moreover, f and g are both differentiable. So we look at the quotient of the derivatives:

$$\frac{f'(x)}{g'(x)} = \frac{1/x}{-1} = -\frac{1}{x} \to -1$$

as $x\to 1$. We conclude by l'Hôpital's rule that

$$\lim_{x\to 1}\frac{f(x)}{g(x)} = -1.$$

(*Note:* for short we may write

$$\begin{aligned}\lim_{x\to 1}\frac{\ln x}{1-x} &= \lim_{x\to 1}\frac{1/x}{-1} \qquad (5.10)\\ &= \lim_{x\to 1}-\frac{1}{x}\\ &= -1\end{aligned}$$

provided that we check that our application of l'Hôpital's rule in (5.10) is valid.) □

Sometimes l'Hôpital's rule must be applied more than once to calculate a limit. The next example illustrates this.

Example 5.10.4. Determine the limiting behaviour of $\dfrac{x^2}{e^{2x}}$ as $x \to \infty$.

Solution. Observe that both the numerator and denominator approach ∞ as $x \to \infty$. Differentiating the numerator and denominator gives

$$\frac{2x}{2e^{2x}}.$$

Again, as $x \to \infty$ we have an indeterminate form of type $\frac{\infty}{\infty}$. If we differentiate the numerator and denominator once more we have

$$\frac{2}{4e^{2x}}$$

and this limit can be evaluated. Piecing this together, we have

$$\lim_{x\to\infty} \frac{x^2}{e^{2x}} = \lim_{x\to\infty} \frac{2x}{2e^{2x}} = \lim_{x\to\infty} \frac{2}{4e^{2x}} = 0$$

by l'Hôpital's rule. □

As the next example shows, l'Hôpital's rule can be used to solve limits that are indefinite forms of type $0 \times \infty$.

Example 5.10.5. Determine the limiting behaviour of $x \ln x$ as $x \to 0^+$.

The next example warns against the improper application of l'Hôpital's rule.

Example 5.10.6. Determine the limiting behaviour of $\dfrac{2x - \cos x}{3x + \cos x}$ as $x \to \infty$.

We shall now sketch the proof of l'Hôpital's rule for the case when $x \to a^+$. Other cases are done similarly. (A complete proof uses a generalisation of the mean value theorem known as *Cauchy's mean value theorem.*)

Sketch proof of Theorem 5.10.1. Suppose that f and g are differentiable (and hence continuous) everywhere, that $f(x) \to 0$ and $g(x) \to 0$ as $x \to a^+$, and that

$$\lim_{x\to a^+} \frac{f'(x)}{g'(x)}$$

exists. We need to show that

$$\lim_{x\to a^+} \frac{f(x)}{g(x)} = \lim_{x\to a^+} \frac{f'(x)}{g'(x)}. \tag{5.11}$$

We begin by applying the mean value theorem to both functions f and g on the interval $[a, x]$. So there are real numbers c and d in (a, x) such that

$$\frac{f(x) - f(a)}{x - a} = f'(c) \qquad \text{and} \qquad \frac{g(x) - g(a)}{x - a} = g'(d). \tag{5.12}$$

Since f is continuous and $f(x) \to 0$ as $x \to a^+$, we have $f(a) = 0$. Similarly $g(a) = 0$. So by (5.12) we have

$$f(x) = (x - a)f'(c) \qquad \text{and} \qquad g(x) = (x - a)g'(d)$$

for some c and d in (a, x). Hence

$$\frac{f(x)}{g(x)} = \frac{(x-a)f'(c)}{(x-a)g(d)} = \frac{f'(c)}{g'(d)}. \tag{5.13}$$

Now $c \to a^+$ and $d \to a^+$ as $x \to a^+$. Hence (5.13) becomes (5.11) as required. $\square$

Problems for Chapter 5

Problems 5.1 : The mean value theorem

1. [**R**] Find a real number c which satisfies the conclusions of the mean value theorem for each function f on the given interval.

   a) $f(x) = x^3$ on $[1, 2]$ b) $f(x) = \sqrt{x}$ on $[0, 2]$.

2. [**R**] Suppose that $f(x) = 1/x$. Show that there is no real number c in $[-1, 2]$ such that
$$f'(c) = \frac{f(2) - f(-1)}{2 - (-1)}.$$
   Why does this not contradict the mean value theorem?

3. [**R**] Consider the function f given by $f(x) = (x-2)^4 \cos(x^2 - 4x + 4)$. Use the mean value theorem to show that f' has a zero on the interval $[1, 3]$.

Problems 5.3 : Proving inequalities using the mean value theorem

4. [**R**] [**V**] By using the mean value theorem, show that

   a) $\ln(1 + x) < x$ whenever $x > 0$;
   b) $-\ln(1 - x) < x/(1 - x)$ whenever $0 < x < 1$;
   c) $1 + x < e^x$ whenever $x > 0$.

5. [**R**]

   a) Use the mean value theorem to show that $\sin t \le t$ whenever $t > 0$.
   b) Hence show that $\sin t < t$ whenever $t > 0$.
   c) Using the pinching theorem and part (a), evaluate the limit $\lim\limits_{x\to\infty} \sin \frac{1}{x}$.

6. [**H**] Prove that
$$1 + \frac{x}{2\sqrt{1+x}} < \sqrt{1+x} < 1 + \frac{x}{2} \qquad \text{whenever } x > 0.$$

Problems 5.4 : Error bounds

7. [**R**] [**V**] Use the mean value theorem to find an upper bound for the error involved if we approximate

   a) $\sqrt{17}$ by $\sqrt{16} = 4$;
   b) $\left(\frac{1998}{1000}\right)^2$ by $2^2 = 4$;
   c) $\frac{1}{1002}$ by $\frac{1}{1000}$.

Problems 5.5 : The sign of a derivative, 5.6 : The second derivative and applications and 5.7 : Critical points, maxima and minima

8. [**R**] The derivative of a function $f : \mathbb{R} \to \mathbb{R}$ is given by $f'(x) = 3(x+1)(x-1)^2(x-4)^3$. Locate all stationary points of f and identify any local maximum or minimum points of f.

9. [**X**] The function $f : \mathbb{R} \to \mathbb{R}$, given by
$$f(x) = \begin{cases} x \sin \frac{1}{x} & \text{if } x \neq 0 \\ 0 & \text{if } x = 0, \end{cases}$$
is continuous but not differentiable at 0. Does f have a local maximum or a local minimum at 0? Prove your answer.

10. [**R**] Find the maximum and minimum values for each function f over the given interval.

    a) $f(x) = 3 - x^3$ over $[-2, 4]$ b) $f(x) = 3 - x^4$ over $[-2, 4]$

    c) $f(x) = x^3 - x^4$ over $[-5, 5]$ d) $f(x) = 2x(x+4)^3$ over $[-2, 1]$

    e) $f(x) = |x^2 - 3x + 2|$ over $[0, 3]$

11. [**R**] Find the point on the straight line $2x + 3y = 6$ which is closest to the origin.

12. a) i) [**R**] Show that the polynomial p_3, where $p_3(x) = 1 + x + \dfrac{x^2}{2!} + \dfrac{x^3}{3!}$, has *at least* one real root.

    ii) [**H**] Show that the polynomial p_2, where $p_2(x) = 1 + x + \dfrac{x^2}{2!}$, has no real roots and deduce that p_3 has *exactly* one real root.

    iii) [**X**] Deduce that $p_4(x) = 1 + x + \dfrac{x^2}{2!} + \dfrac{x^3}{3!} + \dfrac{x^4}{4!} > 0$ for all real numbers x.

    b) [**X**] Suppose that $p_n(x) = \displaystyle\sum_{k=0}^{n} \frac{x^k}{k!}$ whenever $n = 1, 2, 3, \ldots$. Use induction to prove that

    i) if n is even then $p_n(x) > 0$ for all real numbers x, and
    ii) if n is odd then $p_n(x)$ has exactly one real root and this root is negative.

13. [**R**] A wire of length 100 cm is cut into two pieces of length x cm and y cm. The piece of length x cm is bent into the shape of a square and the piece of length y cm into the shape of a circle. Find x and y so that the sum of the areas enclosed by the shapes will be

    a) a minimum b) a maximum.

14. [**X**] Suppose that $a \geq 0$. Find the greatest and least distances from the point $(a, 0)$ to the ellipse
$$\frac{x^2}{4} + \frac{y^2}{1} = 1.$$
(Have a precise answer before comparing with the given answer.)

15. [**X**] Find all the values of a and x, both in $[0, 2\pi]$, where

$$f(x) = \cos a + 2\cos(2x) + \cos(4x - a)$$

has a horizontal point of inflexion.

Problems 5.8 : Counting zeros

16. [**R**] Show that $x^3 + x - 9 = 0$ has only one real solution.

17. [**R**] [**V**] Suppose that $p(x) = x^3 - 12x^2 + 45x - 51$ whenever $x \in \mathbb{R}$. How many real zeros does p have?

Problems 5.9 : Antiderivatives

18. [**R**]

a) Find a function f that has the following properties:

$$\begin{aligned} &f'(t) = \sin t + t \qquad \text{whenever } t \in \mathbb{R}, \\ &f(0) = 2. \end{aligned}$$

b) Are there any other functions with these properties? Explain your answer.

19. [**R**] [**V**] A particle moving along the x-axis has velocity $2t - t^2$ units per second after t seconds. Find

a) the distance from the starting point after three seconds;

b) the total distance travelled after three seconds.

Problems 5.10 : L'Hôpital's rule

20. [**R**] Calculate the following limits.

a) $\displaystyle\lim_{x\to 0} \frac{e^x - 1}{x(3+x)}$ b) $\displaystyle\lim_{x\to 1} \frac{x^m - 1}{x^n - 1}, \quad n \neq 0$ c) $\displaystyle\lim_{x\to \pi/2} \frac{x - \pi/2}{\cos x}$

d) $\displaystyle\lim_{x\to 0} \frac{\ln(1+x) - x}{x^2}$ e) $\displaystyle\lim_{x\to \pi/2} \left(\frac{1 - \sin x}{1 + \cos 2x}\right)$ f) $\displaystyle\lim_{x\to 0} \frac{\tan x - x}{x^3}$

21. [**R**] Determine the limiting behaviour in the following cases.

a) $\dfrac{x^3+1}{x^4+1}$ as $x \to \infty$ b) $\dfrac{e^{5x}}{x^3}$ as $x \to \infty$

c) $\dfrac{e^{5x}}{x^3}$ as $x \to -\infty$ d) $x\sin(1/x)$ as $x \to \infty$

e) $\dfrac{\sqrt{x^4+1}}{\sqrt[3]{x^6+1}}$ as $x \to \infty$ f) $\dfrac{\ln(x^3+1)}{\ln(x^2+1)}$ as $x \to \infty$

22. [**H**] Find the value of $\displaystyle\lim_{t\to 0}\left(\frac{1}{\ln(1+t)}+\frac{1}{\ln(1-t)}\right)$.

23. [**H**] Find (a,b) such that $\displaystyle\lim_{x\to 0}\frac{ax-1+e^{bx}}{x^2}=1$.

24. [**R**] Explain why l'Hôpital's Rule cannot be used to find $\displaystyle\lim_{x\to\infty}\frac{4x+\sin x}{2x-\sin x}$. Use another method to find this limit.

25. [**R**] [**V**] Show that the function f, given by

$$f(x)=\begin{cases} e^{2x} & \text{if } x\geq 0\\ 2x+1 & \text{if } x<0,\end{cases}$$

is differentiable at 0.

26. [**R**]

a) Evaluate $\displaystyle\lim_{h\to 0^+}\frac{\cos\sqrt{h}-1}{h}$.

b) A function f is defined by

$$f(x)=\begin{cases} \cos\sqrt{x} & \text{if } x\geq 0\\ ax+b & \text{if } x<0,\end{cases}$$

where a and b are real numbers. By using the limit calculated in (a), find all possible values of a and b such that f is differentiable at 0.

27. [**H**] [**V**]

a) Use l'Hôpital's rule to show that $\displaystyle\lim_{x\to 0+} x\ln x=0$.

b) By using part (a), or otherwise, show that $\displaystyle\lim_{x\to 0^+} x^2\ln x=0$.

c) A function f is defined by

$$f(x)=\begin{cases} x^2\ln x & \text{if } x>0\\ ax+b & \text{if } x\leq 0,\end{cases}$$

where a and b are real numbers. Find all possible values of a and b such that f is differentiable at 0.

Chapter 6

Inverse functions

We return again to the pollution example (see the introductions to Chapter 2). According to our model, the volume of pollution $P(t)$ in the lake after t days of the factory's operation is given by

$$P(t) = \frac{10^9}{101}\big(1 - e^{-101t/10^5}\big).$$

Environmental authorities, after hearing that the eventual amount of pollution in the lake will be $10^9/101$ litres, determine that this level would be environmentally devastating. Recommendations are made that the amount of pollution should not exceed 8×10^6 litres. If factory operations do not change, in how many days will pollution levels in the lake exceed this amount?

Note that the function P expresses pollution as a function of time. One approach to solving this problem is to find a function T that expresses time as a function of pollution. Such a function is called an *inverse function* of P. This chapter is devoted to a study of inverse functions.

6.1 Some preliminary examples

The examples of this section give an introduction to some of the key concepts appearing in Chapter 6. (A rigorous treatment of these concepts will be given in later sections.)

Example 6.1.1. Jack and Jill are playing a number game. Jack says, 'A positive number, when squared, is equal to 16. What was the original number?' Jill correctly answers that the number was 4.

Jack goes again: 'A positive number, when squared, is equal to 49. What was the original number?' This time Jill answers '7.'

This game may be modelled using two functions. Jack's function f takes a positive number and squares it. That is, f is given by

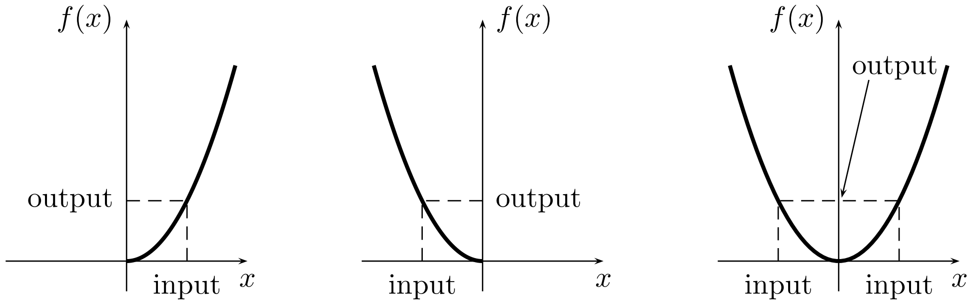
$$f : (0, \infty) \to (0, \infty), \qquad f(x) = x^2.$$

The function g that Jill uses to answer Jack's question is given by

$$g : (0, \infty) \to (0, \infty), \qquad g(x) = \sqrt{x}.$$

So f takes an initial number and squares it, while g takes the square root and recovers the original number. A way of writing down the fact that g undoes (or reverses) what f does to any positive number x is:

$$g(f(x)) = x \qquad \forall x \in \mathrm{Dom}(f). \tag{6.1}$$



(a) Example 6.1.1 (b) Example 6.1.2 (c) Example 6.1.3

Figure 6.1: Correspondence between inputs and outputs

In fact, f also undoes what g does to any positive number. This is easily verified by checking that

$$f(g(x)) = x \qquad \forall x \in \mathrm{Dom}(g). \tag{6.2}$$

Because f and g have this special relationship, they are called *inverses* of each other.

The relationship between any function f and its inverse function g can also be viewed in the following way. If $y = f(x)$ and we want to express x as a function of y, then we find that $x = g(y)$. For the functions of Example 6.1.1, this is easily verified by the following calculation:

$$y = f(x) \quad \Longleftrightarrow \quad y = x^2 \quad \Longleftrightarrow \quad \sqrt{y} = x \quad \Longleftrightarrow \quad x = g(y).$$

The next example shows that changing the domain of f gives a different inverse function g.

Example 6.1.2. Jack changes the game slightly. He says, 'A negative number, when squared, is equal to 36. What was the original number?' Jill correctly answers that the number was -6.

Again, we model the game with two functions. Jack's function f is given by

$$f : (-\infty, 0) \to (0, \infty), \qquad f(x) = x^2,$$

while Jill's function g is given by

$$g : (0, \infty) \to (-\infty, 0), \qquad g(x) = -\sqrt{x}.$$

Note that g undoes what f does to any negative number, and vice-versa. That is, f and g satisfy equations (6.1) and (6.2) and are therefore inverse functions of each other.

Each function f in Example 6.1.1 and Example 6.1.2 had the property that for every output there is a unique corresponding input (See Figure 6.1 (a) and (b)). However, the function f graphed in Figure 6.1 (c) does not have a one-to-one correspondence between its outputs and inputs. As we see in the next example, this function f has no inverse function.

Example 6.1.3. Jack changes the game one more time. This time he says, 'A real number, when squared, is equal to 64. What was the original number?' This time Jill cannot give a definite answer to his question. The answer is either 8 or -8, but she has no way of telling which it is.

The model for this situation is the following. Jack's function f is now given by

$$f : \mathbb{R} \to [0, \infty), \qquad f(x) = x^2;$$

but there is no function g that undoes what f does to a real number x. That is, there is no function g that satisfies the equation (6.1). Therefore f has no inverse function. The reason why this is the case is clear: the domain of f now includes the positive *and* negative real numbers, and therefore some (in fact most) outputs of f have two possible corresponding inputs. That is, there is no *one-to-one correspondence* between inputs and outputs of f.

The ideas introduced through these examples are summarised below.

- If f has an inverse function g, then g undoes (or reverses) what f does to elements of $\mathrm{Dom}(f)$.
- Suppose f and g are inverse functions. Then $y = f(x)$ if and only if $x = g(y)$.
- A function f has an inverse function only if there is a one-to-one correspondence between inputs and outputs of f.
- The domain of a function f plays an important role in determining whether f has an inverse function g, and if so, exactly what the inverse function is.
- If a function f does not have an inverse function, then we can sometimes modify the domain of f so that it does have an inverse function. (For example, if $f(x) = x^2$ then $f : \mathbb{R} \to [0, \infty)$ does not have an inverse function while the function $f : (0, \infty) \to (0, \infty)$, with domain *restricted* to the positive real numbers, does have an inverse function.)

In the following sections we treat these ideas more rigorously.

6.2 One-to-one functions

(Ref: SH10 §7.1)

As was demonstrated in Section 6.1, the property that a function has a one-to-one correspondence between inputs and outputs is critical to our discussion.

Definition 6.2.1. A function f is said to be *one-to-one* if
$$f(x_1) = f(x_2) \quad \text{implies that} \quad x_1 = x_2$$
whenever $x_1, x_2 \in \mathrm{Dom}(f)$.

In other words, a function is one-to-one if every output has a unique input. Equivalently, f is one-to-one if
$$f(x_1) \neq f(x_2) \qquad \text{whenever } x_1 \neq x_2,$$
provided that both x_1 and x_2 belong to $\mathrm{Dom}(f)$. One-to-one functions are sometimes called *injective* functions.

Example 6.2.2. Show that the function $f : \mathbb{R} \to \mathbb{R}$, given by $f(x) = 2x - 3$, is one-to-one.

Solution. Suppose that $f(x_1) = f(x_2)$. Then

$$2x_1 - 3 = 2x_2 - 3.$$

If we add 3 to both sides and afterwards divide by 2 then we obtain

$$x_1 = x_2.$$

Hence f is one-to-one. □

Example 6.2.3. Show that the function $f : \mathbb{R} \to \mathbb{R}$, given by $f(x) = x^2$, is not one-to-one.

Proof. Clearly $f(-2) = f(2)$. However, $-2 \neq 2$ and so f is not one-to-one. □

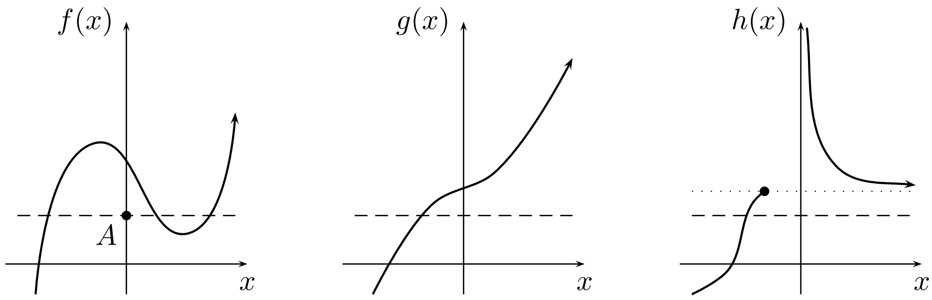
There are other ways of identifying one-to-one functions. The following method is a simple geometric test.

Proposition 6.2.4 (The horizontal line test)**.** *Suppose that f is a real-valued function defined on some subset of $\mathbb{R}$. Then the following statements are equivalent:*

(a) f is one-to-one;

(b) every horizontal line in the Cartesian plane intersects the graph of f at most once.

Example 6.2.5. Consider the functions graphed below.



- f is not one-to-one because the horizontal line passing through the point A cuts the graph of f more than once;
- g is one-to-one (in fact, since g is increasing, every horizontal line can cut the graph of g graph no more than once);
- h is also one-to-one (even though it is not always increasing).

Although not every one-to-one function is increasing (or decreasing), it is true that every increasing function is one-to-one.

Proposition 6.2.6. *If a function f is either increasing or decreasing, then f is one-to-one.*

Sketch proof. If f is increasing and $x_1, x_2 \in \mathrm{Dom}(f)$, then

$$x_1 < x_2 \quad \Longleftrightarrow \quad f(x_1) < f(x_2). \tag{6.3}$$

Now compare (6.3) with Definition 6.2.1. □

Example 6.2.7. Show that the function $f : \mathbb{R} \to \mathbb{R}$, given by

$$f(x) = 2x^5 + x^3 + x - 10,$$

is one-to-one.

Solution. The derivative of f is given by $f'(x) = 10x^4 + 3x^2 + 1$. Since $f'(x) \geq 1$ for all real x, the function f is increasing and hence one-to-one. $\square$

Remark 6.2.8. Not every function whose derivative is only positive (or only negative) on its domain is one-to-one. For example, on Dom(tan)

$$\frac{d}{dx} \tan x = \sec^2 x \geq 1,$$

but tan is *not* one-to-one, nor is it increasing on its domain. The problem here is that the domain of tan has gaps.

6.3 Inverse functions

(Ref: SH10 §7.1)

We have already mentioned the term *inverse function* in Section 6.1. In this section we define precisely what is meant by an inverse function and explore some of the relationships between a function and its inverse. Our starting point is the following theorem. It says that if f is one-to-one, then there is a function g which undoes what f does to any input x (c.f. equation (6.1)).

Theorem 6.3.1. *Suppose that f is a one-to-one function. Then there is a unique function g with* $\mathrm{Dom}(g) = \mathrm{Range}(f)$ *satisfying*

$$g(f(x)) = x \qquad \forall x \in \mathrm{Dom}(f) \tag{6.4}$$

and

$$f(g(x)) = x \qquad \forall x \in \mathrm{Range}(f). \tag{6.5}$$

Moreover,

$$\mathrm{Range}(g) = \mathrm{Dom}(f)$$

and g is one-to-one.

Proof. Suppose that f is one-to-one and let D and R denote the domain and range of f respectively. Since f is one-to-one, for every y in R there is a unique x in D such that $y = f(x)$.

Now define a function $g : R \to D$ by the following rule: if $y \in R$ then $g(y)$ is the unique x in D such that $f(x) = y$.

It is left as an exercise to show that g is one-to-one and that equations (6.4) and (6.5) hold. $\square$

The theorem allows us to define the term *inverse function.*

Definition 6.3.2. Suppose that f is a one-to-one function. Then the *inverse function* of f is the unique function g given by Theorem 6.3.1. The inverse function for f is often written as f^{-1}.

Remark 6.3.3. If f^{-1} denotes the inverse function of a one-to-one function f, then equations (6.4) and (6.5) can be expressed as

$$f^{-1}(f(x)) = x \qquad \forall x \in \mathrm{Dom}(f)$$

and

$$f(f^{-1}(x)) = x \qquad \forall x \in \mathrm{Range}(f).$$

(*Warning:* This notation is potentially confusing if not understood correctly. The function f^{-1} is *not* the reciprocal function of f (that is, $1/f(x)$). To denote the reciprocal of $f(x)$ using index notation we write $[f(x)]^{-1}$ rather than $f^{-1}(x)$.)

Remark 6.3.4. If f is a bijection and g is the inverse of a function f, then f is the inverse of g. This is due to the symmetry between f and g in the statements of Theorem 6.3.1.

If f is a one-to-one function then equation (6.5) can often be used to find an explicit formula for its inverse g.

Example 6.3.5. Suppose that $f : \mathbb{R} \to \mathbb{R}$ is given by $f(x) = 4 - \frac{1}{3}x^3$.

(a) Explain why f has an inverse function.

(b) Find the inverse function g of f.

(c) Sketch f and g on the Cartesian plane.

(d) Use the inverse function g to solve the equation $f(x) = 13$.

Solution. (a) Since $f'(x) = -x^2 \le 0$ for all x in $\mathbb{R}$ (with $f'(x) = 0$ only when $x = 0$), the function f is decreasing and is therefore one-to-one. Hence f has an inverse function.

(b) Let g denote the inverse function of f. Then $\mathrm{Dom}(g) = \mathrm{Range}(f) = \mathbb{R}$ and $\mathrm{Range}(g) = \mathrm{Dom}(f) = \mathbb{R}$. To find an explicit formula for $g(x)$, suppose that

$$y = g(x).$$

If we apply f to both sides of the equation, and use (6.5), then

$$f(y) = f(g(x)) = x.$$

Using the definition of f, this simplifies to

$$4 - \frac{1}{3}y^3 = x,$$

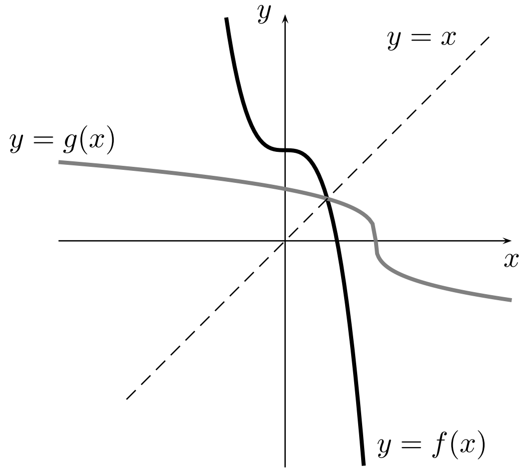
whereupon rearrangement gives

$$y = \sqrt[3]{12 - 3x}.$$

Therefore $g : \mathbb{R} \to \mathbb{R}$ is given by

$$g(x) = \sqrt[3]{12 - 3x}. \tag{6.6}$$

(c) The graphs of f and g are shown below. Note that each graph is the reflection the other in the line $y = x$.



(d) To solve the equation $f(x) = 13$, apply g to both sides. This gives

$$g(f(x)) = g(13).$$

By (6.4) and (6.6) this simplifies to

$$\begin{aligned} x &= \sqrt[3]{12 - 3 \times 13} \\ &= -3. \end{aligned}$$

□

Remark 6.3.6. As was illustrated in Example 6.3.5 (c), the graph of a one-to-one function f and its inverse g are reflections of each other in the line $y = x$. We give a proof of this fact below:

$$\begin{aligned} (x, y) \text{ lies on the graph of } f \quad &\Longleftrightarrow \quad y = f(x) \\ &\Longleftrightarrow \quad g(y) = g(f(x)) \\ &\Longleftrightarrow \quad g(y) = x \\ &\Longleftrightarrow \quad (y, x) \text{ lies on the graph of } g. \end{aligned}$$

If a function f is not one-to-one, then we can sometimes *restrict* the domain of f so that it becomes one-to-one. The function f with restricted domain then has an inverse function.

Example 6.3.7. Consider the function $f : \mathbb{R} \to \mathbb{R}$ given by

$$f(x) = 4 - x^2.$$

Find a restriction of f such that f becomes one-to-one. Find the inverse function g of this restriction.

Solution. A quick sketch of the graph of f (see Figure 6.2 (a)) shows that f is not one-to-one. If we restrict the domain of f to $[0, \infty)$, then the restricted function $f : [0, \infty) \to \mathbb{R}$ passes the horizontal line test and is therefore one-to-one. Let g denote the inverse of $f : [0, \infty) \to \mathbb{R}$. Then

$$g : (-\infty, 4] \to [0, \infty)$$

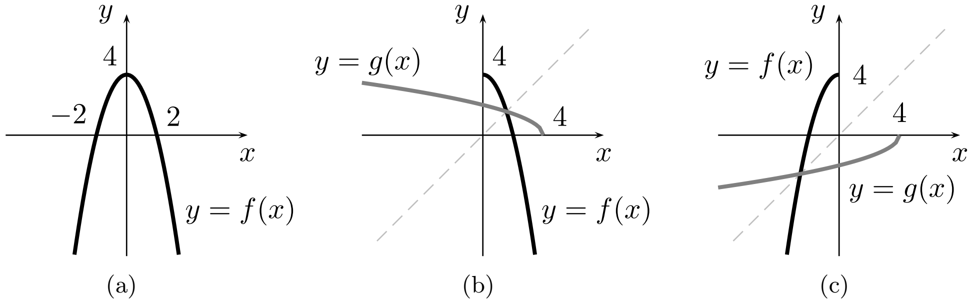


Figure 6.2: Two different restrictions of the function f, and their corresponding inverses (see Example 6.3.7).

since $\mathrm{Dom}(g) = \mathrm{Range}(f) = (-\infty, 4]$ and $\mathrm{Range}(g) = \mathrm{Dom}(f) = [0, \infty)$. We find an explicit formula for g:

$$\begin{aligned} y &= g(x) \\ f(y) &= x && \text{(by equation 6.5)} \\ 4 - y^2 &= x \\ y &= \sqrt{4-x} && \text{(taking the positive root since } \mathrm{Range}(g) = [0,\infty)). \end{aligned}$$

Hence $g(x) = \sqrt{4-x}$. The graphs of f and g are sketched in Figure 6.2 (b). $\square$

Remark 6.3.8. In the solution to Example 6.3.7, one could have instead chosen the restriction $f : (-\infty, 0] \to \mathbb{R}$. The corresponding inverse g is given by

$$g : (-\infty, 4] \to (-\infty, 0], \qquad g(x) = -\sqrt{4 - x^2}.$$

The graphs of f and g are given in Figure 6.2 (c).

Example 6.3.9. Consider the function $f : \mathbb{R} \to \mathbb{R}$, given by

$$f(x) = -x^3 + 3x^2 + 24x - 13.$$

Find all intervals I, each as large as possible, such that the function $f : I \to \mathbb{R}$ (with domain restricted to I) has an inverse function.

Remark 6.3.10. If you study MATH1081 you will see that an abstract function $f : A \to B$ is said to be invertible if it is a bijection from the set A onto the set B. That is, it is both one-to-one and its range is all of its codomain B and so the problem of solving $f(x) = b$ has a unique solution for all $b \in B$. In the algebra strand of this course you will see that an $m \times n$ matrix A is invertible if the matrix equation $A\mathbf{x} = \mathbf{b}$ has a unique solution for all $\mathbf{b} \in \mathbb{R}^m$, and we'll see the idea again with linear transformations in Mathematics 1B.

In calculus we often initially define a function without precisely knowing its range, and so we usually say,

"Define $f : (a, b) \to \mathbb{R}$ by ...".

For example, a very important function is $f : \mathbb{R} \to \mathbb{R}$, $f(x) = \int_0^x e^{-t^2} dt$. It is easy to check that this function f is one-to-one, but it is hard to see that the range of f is $(-\sqrt{\pi}/2, \sqrt{\pi}/2)$. In calculus, when we say that this function f is invertible, we are *not* saying that f is a bijection from its domain onto it's codomain $\mathbb{R}$, but rather that f is a bijection from it domain onto its *range*.

6.4 The inverse function theorem

(Ref: SH10 §§7.1, B.3)

If a function f is one-to-one then its inverse function g is also one-to-one (see Theorem 6.3.1). What other properties of f does its inverse function g inherit?

First, if f is continuous, then so is g. (A rough argument for this fact would go like this: if f is continuous then its graph can be drawn as an unbroken line. Now reflect this graph in the line $y = x$. The reflected graph is also an unbroken line. Moreover it corresponds to the graph of g. Hence g is continuous.)

What about differentiability? If a one-to-one function f is differentiable then its inverse g is not always differentiable. For example, if $f : \mathbb{R} \to \mathbb{R}$ is given by $f(x) = x^3$, then f is differentiable everywhere. Its inverse function g is given by

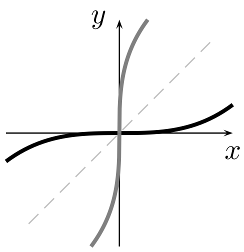
$$g : \mathbb{R} \to \mathbb{R}, \qquad g(x) = \sqrt[3]{x}.$$

Since

$$g'(x) = \frac{1}{\sqrt[3]{x^2}},$$

g is not differentiable at 0. So while f is differentiable everywhere, its inverse g is not.

The problem with the above example occurs where the derivative of f is zero. When a horizontal tangent to f is reflected in the line $y = x$, it becomes a vertical tangent to g and hence g is not differentiable at this point.



However, if f is a differentiable function whose derivative is never zero, its inverse g will also be differentiable.

Theorem 6.4.1 (The inverse function theorem)**.** *Suppose that I is an open interval, $f : I \to \mathbb{R}$ is differentiable and $f'(x) \neq 0$ for all x in I. Then*

(i) f is one-to-one and has an inverse function $g : \mathrm{Range}(f) \to \mathrm{Dom}(f)$,

(ii) g is differentiable at all points in $\mathrm{Range}(f)$, and

(iii) the derivative of g is given by the formula

$$g'(x) = \frac{1}{f'(g(x))}$$

for all x in $\mathrm{Range}(f)$.

Some parts of the inverse function theorem are easy to prove while others are hard. Statement (i) is true since $f'(x) \neq 0$ for x in I implies that f is either increasing on I or decreasing on I (by the mean value theorem). Hence f is one-to-one. Statement (ii) is difficult to prove and involves a delicate limiting argument with the difference quotients of f and g.

It is important that students understand the proof of statement (iii). Suppose (by statement (ii)) that g is differentiable. Beginning with the equation

$$f(g(x)) = x$$

(see 6.5), differentiation with respect to x gives

$$f'(g(x)).g'(x) = 1$$

by the chain rule. Since f' is never zero on I, we can divide by $f'(g(x)))$ to obtain

$$g'(x) = \frac{1}{f'(g(x))}$$

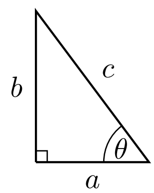
as required.

In the next section the inverse function theorem will be applied to the trigonometric functions.

6.5 Applications to the trigonometric functions

(Ref: SH10 §7.7)

In elementary trigonometry, the sine function gives the ratio $\frac{b}{a}$ of two side lengths of a right-angle triangle corresponding to the angle θ.



However, if we know the ratio and instead want to find the corresponding angle, we use the inverse sine function. Thus relationship between the sine and inverse sine functions is given by

$$\sin\theta = \frac{b}{c}, \qquad \sin^{-1}\frac{b}{c} = \theta. \tag{6.7}$$

This relationship works in elementary trigonometry where the angle θ is acute. However, if θ is allowed to be any real number then the relationship between sin and $\sin^{-1}$ is not so straightforward. This is because the function $\sin : \mathbb{R} \to \mathbb{R}$ is not one-to-one and therefore has no inverse (see Figure 6.3).

In the next example, we define exactly what we mean by $\sin^{-1}$ and compute its derivative.

Example 6.5.1 (The inverse sine function)**.** Since the usual sine function is not a one-to-one function, we consider instead a *restricted* sine function

$$\sin : [-\tfrac{\pi}{2}, \tfrac{\pi}{2}] \to [-1, 1]$$

which is one-to-one (see Figure 6.3 (a)). This restricted function has an inverse

$$\sin^{-1} : [-1, 1] \to [-\tfrac{\pi}{2}, \tfrac{\pi}{2}],$$

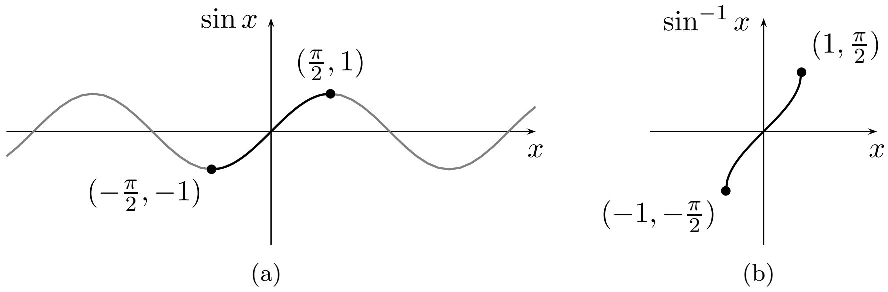


Figure 6.3: Graphs of sin and $\sin^{-1}$

which is graphed in Figure 6.3 (b).

We note that $\sin^{-1}$ is an odd function, that is,

$$\sin^{-1}(-x) = -\sin^{-1} x \qquad \forall x \in [-\tfrac{\pi}{2}, \tfrac{\pi}{2}].$$

We now examine the differentiability of $\sin^{-1}$. If $I = (-\frac{\pi}{2}, \frac{\pi}{2})$ then the sine function is differentiable on I and

$$\frac{d}{dx}(\sin x) = \cos x \neq 0$$

for all x in I. So by the inverse function theorem (Theorem 6.4.1), $\sin^{-1}$ is differentiable on $(-1, 1)$. While the derivative of $\sin^{-1}$ could be computed using the formula of Theorem 6.4.1 (iii), we prefer to compute it directly by implicit differentiation.

If $y = \sin^{-1} x$ then

$$\begin{aligned} \sin y &= x && \text{(by applying sin to both sides)} \\ \cos y \frac{dy}{dx} &= 1 && \text{(by implicit differentiation)} \\ \frac{dy}{dx} &= \frac{1}{\cos y} && \text{(since } \cos y > 0 \text{ when } -\tfrac{\pi}{2} < y < \tfrac{\pi}{2}). \end{aligned} \tag{6.8}$$

We seek a simple expression for $\cos y$ in terms of x. By the trigonometric identity $\sin^2 y + \cos^2 y = 1$, and the fact that $y = \sin^{-1} x$, we have

$$\cos y = \sqrt{1 - \sin^2 y} = \sqrt{1 - (\sin(\sin^{-1} x))^2} = \sqrt{1 - x^2}, \tag{6.9}$$

where the positive square root was taken since $\cos y > 0$. Substituting (6.9) into (6.8) gives

$$\frac{dy}{dx} = \frac{1}{\sqrt{1 - x^2}}$$

and hence

$$\frac{d}{dx}(\sin^{-1}) = \frac{1}{\sqrt{1 - x^2}}$$

whenever $-1 < x < 1$.

As a corollary, note that the derivative of $\sin^{-1}$ is positive, so $\sin^{-1}$ is an increasing function.

The following table summarises the domain, range and derivatives of the restricted (and hence one-to-one) trigonometric functions and their inverses. The derivatives exist on the largest open

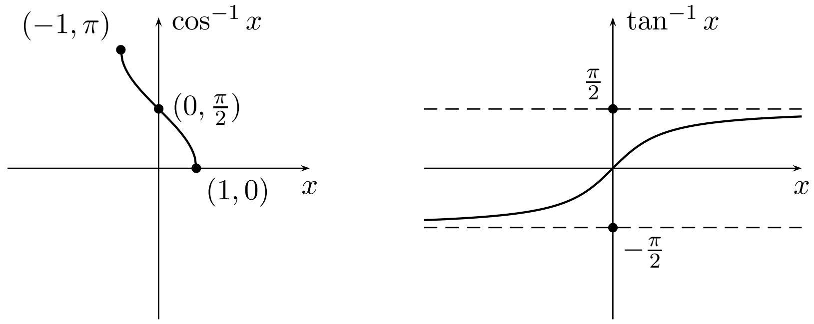


Figure 6.4: Graphs of $\cos^{-1}$ and $\tan^{-1}$

intervals contained in the domain of each function. Proofs of the derivatives for the inverse trigonometric functions are done in much the same way as illustrated in Example 6.5.1. It is expected that students will be able to derive these derivatives, using the method of Example 6.5.1, if required. The graphs of $\cos^{-1}$ and $\tan^{-1}$ are shown in Figure 6.4.

Function	Domain	Range	Derivative
sin	$[-\frac{\pi}{2}, \frac{\pi}{2}]$	$[-1, 1]$	$\frac{d}{dx}(\sin x) = \cos x$
$\sin^{-1}$	$[-1, 1]$	$[-\frac{\pi}{2}, \frac{\pi}{2}]$	$\frac{d}{dx}(\sin^{-1} x) = \frac{1}{\sqrt{1-x^2}}$
cos	$[0, \pi]$	$[-1, 1]$	$\frac{d}{dx}(\cos x) = -\sin x$
$\cos^{-1}$	$[-1, 1]$	$[0, \pi]$	$\frac{d}{dx}(\cos^{-1} x) = -\frac{1}{\sqrt{1-x^2}}$
tan	$(-\frac{\pi}{2}, \frac{\pi}{2})$	$(-\infty, \infty)$	$\frac{d}{dx}(\tan x) = \sec^2 x$
$\tan^{-1}$	$(-\infty, \infty)$	$(-\frac{\pi}{2}, \frac{\pi}{2})$	$\frac{d}{dx}(\tan^{-1} x) = \frac{1}{1+x^2}$

We note also that $\sin^{-1}$ and $\tan^{-1}$ are both odd functions, but that

$$\cos^{-1}(-x) = \pi - \cos^{-1} x \qquad \forall x \in [-1, 1].$$

The next example illustrates that one must be careful with the notation sin and $\sin^{-1}$. It is not always true that $\sin^{-1}(\sin x) = x$. This is because $\sin^{-1}$ is *not* the inverse of $\sin : \mathbb{R} \to \mathbb{R}$. It is only the inverse of $\sin : [-\frac{\pi}{2}, \frac{\pi}{2}] \to \mathbb{R}$. In particular,

$$\sin^{-1}(\sin x) = x$$

only when $x \in [-\frac{\pi}{2}, \frac{\pi}{2}]$. Similar comments apply to the other trigonometric functions.

Example 6.5.2. Evaluate the following numbers without using a calculator:

(a) $\sin^{-1}(\sin \frac{5\pi}{4})$,

(b) $\sin(\cos^{-1} \frac{3}{5})$.

Solution. (a) We cannot write $\sin^{-1}(\sin\frac{5\pi}{4}) = \frac{5\pi}{4}$, since $\frac{5\pi}{4}$ is not in the interval $[-\frac{\pi}{2}, \frac{\pi}{2}]$. However, since $\sin^{-1}$ is an odd function,

$$\sin^{-1}\left(\sin\frac{5\pi}{4}\right) = \sin^{-1}\left(-\frac{1}{\sqrt{2}}\right) = -\sin^{-1}\left(\frac{1}{\sqrt{2}}\right) = -\frac{\pi}{4}.$$

(b) Let θ denote $\cos^{-1}\frac{3}{5}$. Then we seek $\sin\theta$. By the trigonometric identity

$$\sin^2\theta + \cos^2\theta = 1,$$

we have

$$\sin\theta = \pm\sqrt{1-\cos^2\theta},$$

which simplifies to

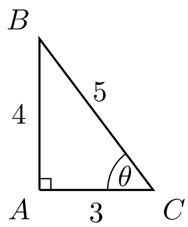
$$\sin\theta = \pm\sqrt{1-\left(\frac{3}{5}\right)^2} = \pm\frac{4}{5}.$$

Since $\theta = \cos^{-1}\frac{3}{5} > 0$ we have $\sin\theta > 0$ and hence

$$\sin\theta = +\frac{4}{5},$$

completing the question.

An alternate solution to (b) is the following. Once again, let θ denote the angle $\cos^{-1}\frac{3}{5}$. This means that $\cos\theta = \frac{3}{5}$, whereupon we construct the right-angled triangle shown below.

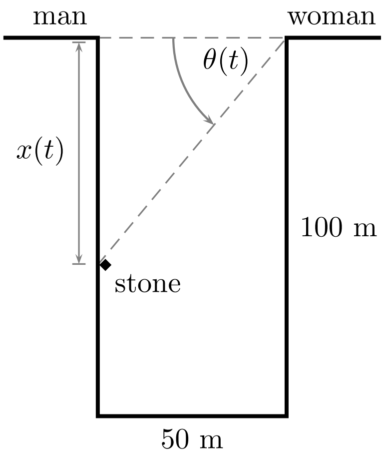


By Pythagoras' theorem, the length of side AB is 4. Hence $\sin\theta = \frac{4}{5}$. □

The final example uses theory from all chapters covered in the course so far.

Example 6.5.3. A man and a woman sit on opposite edges of a canyon which is 50 metres wide and 100 metres deep. The man drops a small stone from rest into the canyon. The woman watches the stone fall to the bottom of the canyon through a telescope. At what time during the stone's fall will the angle that the telescope makes with the horizontal change the fastest? (That is, when will the angular velocity of the telescope be greatest?) Assume that the air resistance experienced by the stone is negligible.

Solution. Let $x(t)$ denote the distance (in metres) that the stone has travelled exactly t seconds after the stone was dropped. Let $\theta(t)$ denote the angle that the telescope makes with the horizontal at t seconds. Let g denote acceleration due to gravity ($g \approx 9.8\ \mathrm{ms}^{-2}$).



If $v(t)$ denotes the velocity of the stone at time t, then we have

$$\frac{dx}{dt} = v, \qquad x(0) = 0, \tag{6.10}$$

$$\frac{dv}{dt} = g, \qquad v(0) = 0. \tag{6.11}$$

From (6.11) we obtain $v(t) = gt + C_1$. By imposing the initial condition that $v(0) = 0$ we see that $C_1 = 0$. Hence $v(t) = gt$ and substituting this into (6.10) gives

$$\frac{dx}{dt} = gt, \qquad x(0) = 0.$$

It follows that $x(t) = \frac{1}{2}gt^2 + C_2$ and one easily shows that $C_2 = 0$.

To find out when the stone hits the canyon floor, we solve the equation

$$100 = \frac{1}{2}gt^2$$

for t. Hence $t = \sqrt{200/g} \approx 4.5$. So it takes about 4.5 seconds for the stone to hit the canyon floor.

By examining the triangle in the above diagram, we see that

$$\tan\theta = \frac{x(t)}{50} = \frac{gt^2}{100}.$$

To solve the problem, we need to determine the value of t when $\frac{d\theta}{dt}$ is a maximum. Now

$$\theta = \tan^{-1}\left(\frac{gt^2}{100}\right)$$

and so

$$\frac{d\theta}{dt} = \frac{1}{1 + \left(\frac{gt^2}{100}\right)^2} \times \frac{gt}{50}$$

by the derivative of $\tan^{-1}$ and the chain rule. This simplifies to

$$\frac{d\theta}{dt} = \frac{200gt}{10^4 + g^2t^4}.$$

For simplicity, denote $\frac{d\theta}{dt}$ by ω. We need to find when ω attains its maximum on the interval $[0, \sqrt{200/g}]$ (that is, over the time interval when the stone is falling). To locate the critical points of ω, we calculate $\frac{d\omega}{dt}$:

$$\begin{aligned}\frac{d\omega}{dt} &= \frac{d^2\theta}{dt^2}\\ &= \frac{(10^4+g^2t^4)200g - 200gt(4g^2t^3)}{(10^4+g^2t^4)^2}\\ &= \frac{(10^4)200g - 600g^3t^4}{(10^4+g^2t^4)^2}.\end{aligned}$$

The stationary points of ω occur when

$$(10^4)200g - 600g^3t^4 = 0,$$

that is, when

$$t^4 = \frac{10^4}{3g^2}.$$

Since $t \geq 0$, we obtain $t = 10/\sqrt[4]{3g^2} \approx 2.42$. So the critical points of ω on $[0, \sqrt{200/g}]$ are

$$0, \qquad 10/\sqrt[4]{3g^2} \qquad \text{and} \qquad \sqrt{200/g}.$$

The value of ω at each of these points is calculated below:

$$\omega(0) = 0, \qquad \omega(10/\sqrt[4]{3g^2}) \approx 1.10, \qquad \omega(\sqrt{200/g}) \approx 0.26.$$

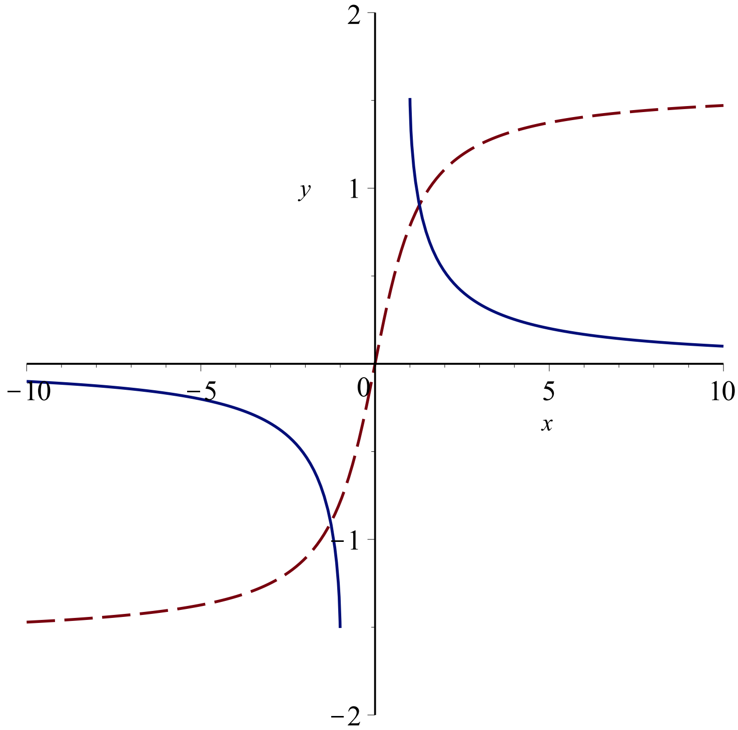
Clearly $\omega(t)$ attains its maximum when $t = 10/\sqrt[4]{3g^2}$.

In conclusion, the angular velocity of the telescope is greatest $10/\sqrt[4]{3g^2}$ (or about 2.26) seconds after the stone is dropped. At this time, the angular velocity is about 1.10 radians per second (that is, about 62.9 degrees per second). □

6.6 Maple notes

Maple knows about the inverse trigonometric functions. The functions $\sin^{-1}$, $\cos^{-1}$, $\tan^{-1}$, cosec^{-1}, $\sec^{-1}$, $\cot^{-1}$ are written respectively as: `arcsin`, `arccos`, `arctan`, `arccsc`, `arcsec` and `arccot`. For example,

```
> plot([arctan(x), arccsc(x)], x=-10..10, y=-2..2, linestyle=[dash,solid]);
```




```
> # Note the gap around zero in the graph of arccsc(x) --- please explain!
```

Problems for Chapter 6

Problems 6.1 : Some preliminary examples

1. [**R**] [**V**] Suppose that the functions $f : [0, \infty) \to [1, \infty)$ and $g : [1, \infty) \to [0, \infty)$ are given by $f(x) = \sqrt{1 + x^2}$ and $g(x) = \sqrt{x^2 - 1}$.

   a) By calculating $(f \circ g)(x)$ and $(g \circ f)(x)$, verify that g is the inverse function to f.

   b) What are the domains of $f \circ g$ and $g \circ f$?

2. [**R**]

   a) Suppose that $f : \mathbb{R} \to \mathbb{R}$ is given by $f(x) = 3x + 1$. Find $f^{-1}(x)$. Sketch the graph of f and the graph of its inverse function, f^{-1}, on the same diagram.

   b) The function $g : (-\infty, 0] \to \mathbb{R}$ is defined by $g(x) = x^2+1$. Write down the domain and range of the inverse function g^{-1} and find a formula for $g^{-1}(x)$. Find the derivative of g^{-1}.

Problems 6.2 : One-to-one functions and 6.3 : Inverse functions

3. [**R**] Show that the function $f : \mathbb{R} \to \mathbb{R}$, given by $f(x) = x^3+3x+1$, has an inverse function whose domain is $\mathbb{R}$.

4. [**R**] [**V**] Suppose that $f : \mathbb{R} \to \mathbb{R}$ is given by $f(x) = 4x + \cos x$.

   a) Show that f has an inverse function g.

   b) By using the inverse function theorem, find $g'(2\pi)$.

5. [**R**] Suppose that $f : \mathbb{R} \to \mathbb{R}$ is defined by $f(x) = x^3 - 3x + 1$.

   a) Show that $f : \mathbb{R} \to \mathbb{R}$ is not a one-to-one function.

   b) Find all possible intervals I of $\mathbb{R}$, each as large as possible, such that the restricted function $f : I \to \mathbb{R}$ has an inverse. What is the domain of each of corresponding inverse function?

6. [**H**]

   a) Can you find a quadratic function from $\mathbb{R}$ to $\mathbb{R}$ which is one-to-one?

   b) Can you find a cubic function from $\mathbb{R}$ to $\mathbb{R}$ which is not one-to-one?

Problems 6.4 : The inverse function theorem

7. [**H**] For each function $f : \mathbb{R} \to \mathbb{R}$ given below, find all possible intervals I of $\mathbb{R}$, each as large as possible, such that the restricted function $f : I \to \mathbb{R}$ is one-to-one. State the range of each restricted function $f : I \to \mathbb{R}$. What can you say about existence, domain of definition, continuity and differentiability of the corresponding inverse functions?

a) $f(x) = x(x^2 - 1)(x + 2)$
b) $f(x) = (x + 1)^{17}$
c) $f(x) = |x|\ |x + 1|$

Problems 6.5 : Applications to the trigonometric functions

8. **[R]** Simplify each expression without using a calculator.
   a) $\sin^{-1}(\sqrt{3}/2)$ b) $\cos(\cos^{-1}(2/5))$ c) $\sin^{-1}(\sin(5\pi/3))$
   d) $\cos^{-1}(\cos(-\pi/3))$ e) $\cos(\sin^{-1}(3/5))$ f) $\sin(\tan^{-1}(3/5))$
   g) $\sec^{-1}(2)$ h) $\sin^{-1}(\sin x)$ when $\frac{\pi}{2} \le x \le \frac{3\pi}{2}$

9. **[R]** Sketch the graph of $f : [1, 3] \to \mathbb{R}$, where $f(x) = \cos^{-1}(x - 2)$.

10. **[R]** Show that

    a) $\dfrac{d}{dx}\big(\cos^{-1} x\big) = \dfrac{-1}{\sqrt{1 - x^2}}$ b) $\dfrac{d}{dx}\big(\tan^{-1} x\big) = \dfrac{1}{1 + x^2}.$

11. **[R]** Differentiate

    a) $\cos^{-1}(2x)$ b) $\sin^{-1}\sqrt{x}$ c) $\tan^{-1}(2x - 3)$.

12. **[R]** Prove that $\sin^{-1} x + \cos^{-1} x$ is constant. For what values of x is this valid and what is the constant?

13. **[H]** **[V]** Suppose that $f(x) = \tan^{-1} x + \tan^{-1}(1/x)$ whenever $x \neq 0$.

    a) Show that $f'(x) = 0$ whenever $x \neq 0$.
    b) Hence evaluate f on the intervals $(0, \infty)$ and $(-\infty, 0)$.
    c) How do you account for this result geometrically?

14. **[H]**

    a) Draw the graph of $\operatorname{cosec} x$.
    b) Show that cosec restricted to the interval $(0, \frac{\pi}{2}]$ has an inverse function. Sketch the graph of the inverse and calculate its derivative.

15. **[X]**

    a) Show that $2\tan^{-1} 2 = \pi - \cos^{-1}(3/5)$.
    b) Show that $\cos^{-1}(1 - 2x^2) = 2\sin^{-1} x$ whenever $0 \le x \le 1$.
    c) Suppose that $q(x) = \cos^{-1}(1 - 2x^2)$. Is q differentiable at 0?

16. **[H]** A function $f : \mathbb{R} \to \mathbb{R}$ is defined by

$$f(x) = \begin{cases} x\tan^{-1}\left(\frac{1}{\sqrt{x}}\right) & \text{if } x > 0 \\ ax + b & \text{if } x \le 0, \end{cases}$$

where a and b are real numbers. Find all values of a and b such that f is differentiable at 0.

17. [**H**] A lighthouse containing a revolving beacon is located 3 km from P, the nearest point on a straight shoreline. The beacon revolves with a constant rotation rate of 4 revolutions per minute and throws a spot of light onto the shoreline. How fast is the spot of light moving when it is (a) at P and (b) at a point on the shoreline 2 km from P?

18. [**H**] [**V**] A picture 2 metres high is hung on a wall with its bottom edge 6 metres above the eye of the viewer. How far from the wall should the viewer stand for the picture to subtend the largest possible vertical angle with her eye?

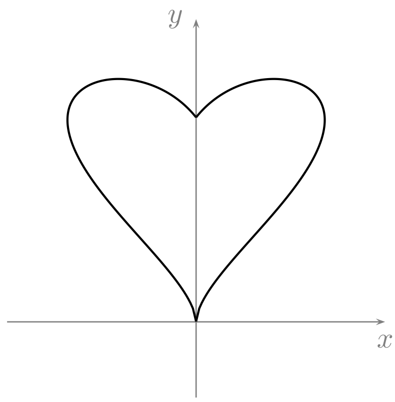
Chapter 7

Curve sketching

There are a variety of ways to describe curves that lie in the plane. In this chapter we study curves which are described by using

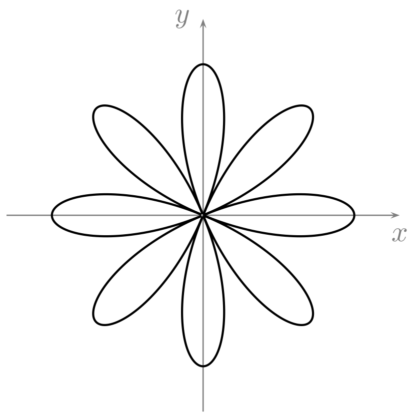
- a Cartesian equation (for example, $y = x^2$ or $y = \sqrt{1 - x^2}$),
- parametric equations, and
- polar coordinates.

Each section of this chapter is devoted to one of these methods. The use of Cartesian equations will already be familiar. However, many curves cannot be easily described by Cartesian equations. The following diagrams illustrate two curves that are better described by a parameter or with polar coordinates.



An example of a parametric curve:

$$\begin{cases} x(t) = \sin t \cos t \ln |t| \\ y(t) = \sqrt{|t|} \cos t \\ t \in [-1, 1], \quad t \neq 0 \end{cases}$$



An example of a polar curve:

$$r = |\cos(4\theta)|$$

7.1 Curves defined by a Cartesian equation

(Ref: SH10 §§4.7, 4.8)

In this section we survey techniques for sketching curves that are described by a Cartesian equation of the form $y = f(x)$.

7.1.1 A checklist for sketching curves

Students will already be familiar with techniques for sketching the graph of a function f. For convenience we summarise some of the most useful here.

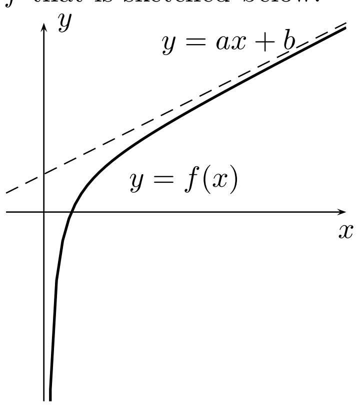
- Identify the domain of f. (In particular, $\mathrm{Dom}(f)$ should not include any points that lead to division by 0 or the square root of a negative number.)
- Identify any symmetries:

$$\begin{array}{lll} f \text{ is } \textit{odd} \text{ if} & f(-x) = -f(x) & \forall x \in \mathrm{Dom}(f) \\ f \text{ is } \textit{even} \text{ if} & f(-x) = f(x) & \forall x \in \mathrm{Dom}(f). \end{array}$$

- Find x- and y-axis intercepts.
- Identify vertical asymptotes.
- Examine the behaviour of $f(x)$ as $x \to \pm\infty$. (This will identify any other asymptotes that exist).
- Use calculus to identify stationary points and other features, if necessary.

7.1.2 Oblique asymptotes

Consider the graph of the function f that is sketched below.



It appears that the graph of f approaches the line $y = ax + b$ as $x \to \infty$. If this is indeed the case, we say that f is *asymptotic* to the line. Since this line is neither vertical, nor horizontal, we call it an *oblique asymptote* to the function f.

Definition 7.1.1. Suppose that a and b are real numbers and that $a \neq 0$. We say that a straight line, given by the equation

$$y = ax + b,$$

is an *oblique asymptote* for a function f if

$$\lim_{x\to\infty} \big(f(x) - (ax + b)\big) = 0.$$

Example 7.1.2. Suppose that f is defined by the equation

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

whenever $x \neq 2$. Identify all asymptotes to f and hence sketch its graph.

Solution. Clearly $f(x) \to \infty$ as $x \to 2^+$, therefore there is a vertical asymptote when $x = 2$. To identify the other asymptotes (if any), we examine the behaviour of $f(x)$ as $x \to \infty$ and as $x \to -\infty$.

Now as $x \to \infty$, $\frac{2}{x-2} \to 0$ from above. Hence we conclude that $f(x)$ approaches $x + 1$ from above. Therefore the line $y = x + 1$ is an oblique asymptote to f. (The calculation

$$\lim_{x\to\infty} \big(f(x) - (x+1)\big) = \lim_{x\to\infty} \frac{2}{x-2} = 0$$

demonstrates from Definition 7.1.1 that this is the case.)

On the other hand, as $x \to -\infty$, $\frac{2}{x-2} \to 0$ from below. Therefore $f(x)$ approaches $x + 1$ from below.

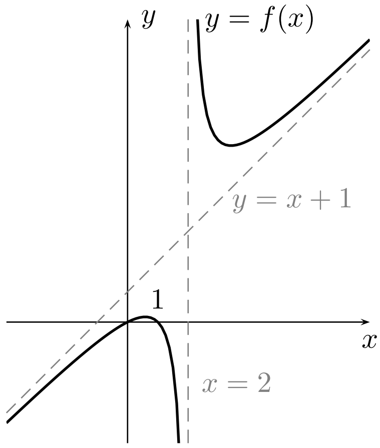
The asymptotes for f help us construct the sketch (which is pictured below). To complete the sketch, we identify the axes intercepts. For the y-axis, we have

$$f(0) = 0.$$

For the x-axis,

$$f(x) = 0 \quad \Rightarrow \quad (x+1)(x-2) + 2 = 0 \quad \Rightarrow x^2 - x = 0 \quad \Rightarrow \quad x = 0, 1.$$

Thus we have the following sketch:



(Note that we have obtained sufficient information to make a useful sketch of the graph without resorting to calculus. However, calculus would help to refine this by locating the turning points and confirming where the graph is increasing or decreasing.) ☐

If f is a rational function of the form p/q, where p and q are polynomials and $\deg(p) > \deg(q)$, then the asymptotic behaviour of f may be determined by polynomial division.

Example 7.1.3. Suppose that f is defined by the equation

$$f(x) = \frac{x^2 - 4}{x+1}$$

whenever $x \neq -1$. Identify all asymptotes to f and hence sketch its graph.

Solution. As in the previous example, clearly there is a vertical asymptote when $x = -1$. It is also clear that $f(x) \to \infty$ as $x \to \infty$, but the precise asymptotic behaviour of $f(x)$ as $x \to \infty$ is still not clear. However, polynomial division gives

$$\begin{array}{r|l} & x - 1 \\ \hline x+1 & x^2 \qquad - 4 \\ & \underline{x^2 + x} \\ & -x - 4 \\ & \underline{-x - 1} \\ & -3. \end{array}$$

Hence

$$f(x) = x - 1 - \frac{3}{x+1}. \tag{7.1}$$

(See Remark 7.1.4 for an alternative method of establishing (7.1).)

Since $\frac{3}{x+1} \to 0$ from above as $x \to \infty$, we conclude that $f(x)$ approaches $x - 1$ from below as $x \to \infty$. A similar observation shows that $f(x)$ approaches $x - 1$ from above as $x \to -\infty$. So the line $y = x - 1$ is an oblique asymptote to the graph of f at both ∞ and $-\infty$.

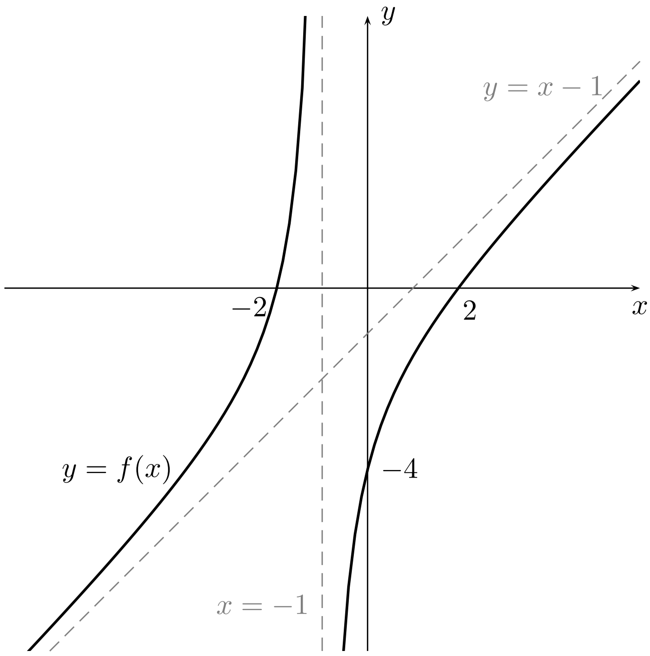
Finally, we calculate the axes intercepts:

$$f(0) = -4$$

and

$$f(x) = 0 \quad \Rightarrow \quad x^2 - 4 = 0 \quad \Rightarrow \quad x = \pm 2.$$

Hence we obtain the following sketch.



(Again, we have obtained sufficient information for a useful sketch of the graph without the use of calculus.) ☐

Remark 7.1.4. We give a method for establishing equation (7.1) that uses an algebraic 'trick' rather than polynomial 'long division':

$$\begin{aligned} f(x) &= \frac{x^2-4}{x+1} \\ &= \frac{x^2-1-3}{x+1} \\ &= \frac{(x-1)(x+1)-3}{x+1} \\ &= (x-1) - \frac{3}{x+1}. \end{aligned}$$

7.1.3 Examples

We complete this section with three examples.

Example 7.1.5. Sketch the graph of f, where f is defined by

$$f(x) = xe^{-x^2}.$$

Solution. It is clear that $\mathrm{Dom}(f) = \mathbb{R}$. Moreover,

$$f(-x) = -xe^{-x^2} = -f(x) \qquad \forall x \in \mathbb{R},$$

so f is an odd function. Therefore, it suffices to consider $x \geq 0$; the rest of the graph can be obtained by symmetry.

Since $e^{-x^2} > 0$, it is easy to see that the only axis intercept occurs at $(0, 0)$. We look now at the behaviour of $f(x)$ as $x \to \infty$. By l'Hôpital's rule,

$$\lim_{x\to\infty} f(x) = \lim_{x\to\infty} \frac{x}{e^{x^2}} = \lim_{x\to\infty} \frac{1}{2xe^{x^2}} = 0.$$

So $y = 0$ is an horizontal asymptote.

Finally, we use calculus to determine the shape of the curve. By the product rule,

$$\begin{aligned} f'(x) &= x\big(-2xe^{-x^2}\big) + e^{-x^2}(1) \\ &= (1-2x^2)e^{-x^2}. \end{aligned}$$

Since $e^{-x^2} > 0$ for all x in $\mathbb{R}$, we conclude that

$$f'(x) = 0 \text{ when } x = \frac{1}{\sqrt{2}},$$

while

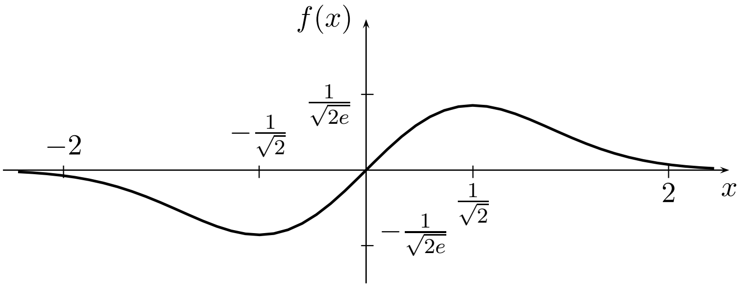
$$f'(x) > 0 \text{ when } 0 \leq x < \frac{1}{\sqrt{2}} \qquad \text{and} \qquad f'(x) < 0 \text{ when } x > \frac{1}{\sqrt{2}}.$$

So f is increasing on the interval $(0, \frac{1}{\sqrt{2}})$, stationary at $\frac{1}{\sqrt{2}}$ and decreasing on $(\frac{1}{\sqrt{2}}, \infty)$. Now

$$f(1/\sqrt{2}) = \frac{1}{\sqrt{2}}e^{-1/2} = \frac{1}{\sqrt{2e}} \approx 0.429.$$

Moreover, $f'(0) = 1$, implying that the tangent line to the graph at the origin has a gradient of 1.

This information can be used to sketch the graph when $x \geq 0$, and the fact that f is odd allows us to deduce the shape of the graph when $x < 0$.



□

Example 7.1.6. Sketch the graph of f, where f is defined by

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

The next function is important since it is later used to define the Si function (see Example 8.12.1). The Si function is used by electrical engineers for digital signal processing and by surveyors using GPS (Global Positioning System).

Example 7.1.7. Sketch the graph of f, where f is defined by

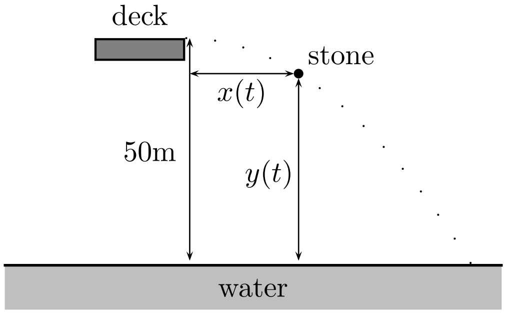
$$f(x) = \begin{cases} \dfrac{\sin x}{x} & \text{if } x \neq 0 \\ 1 & \text{if } x = 0. \end{cases}$$

7.2 Parametrically defined curves

(Ref: SH10 §§10.5, 10.6)

Many moving bodies trace out a path that lies in a plane. An important problem is to describe the curve of the trajectory. We give an example.

Example 7.2.1. Suppose that a stone is thrown horizontally from the deck of Sydney Harbour Bridge at 20 metres per second, and that the air resistance experienced by the stone is negligible. Let $x(t)$ denote the horizontal distance (in metres) travelled by the stone t seconds after being thrown and let $y(t)$ denote its height (in metres) above the water. The deck is 50 metres above sea level.



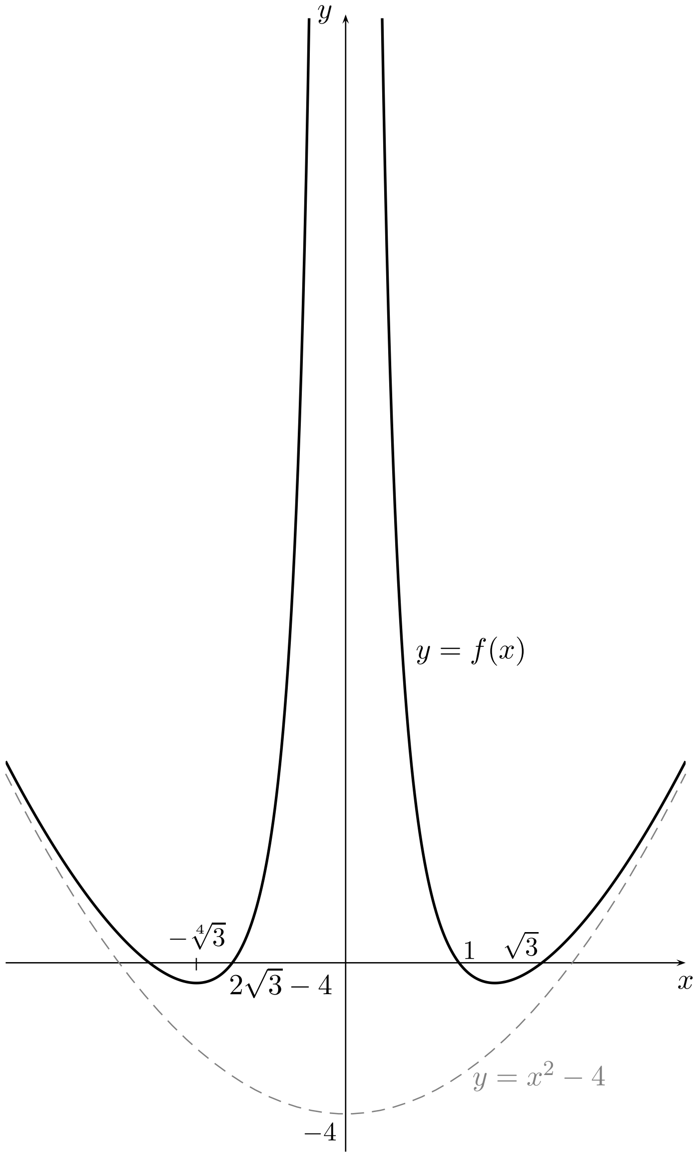


Figure 7.1: Sketch for Example 7.1.6.

If we take the acceleration due to gravity to be 10 m/s^2, then Newton's laws of motion show that

$$x(t) = 20t \qquad \text{and} \qquad y(t) = 50 - 5t^2 \tag{7.2}$$

when $0 < t < \sqrt{10}$ (that is, after the stone is thrown and just before it hits the water). Sketch the trajectory of the stone when $0 < t < \sqrt{10}$.

Solution. We want to express y as a function of x. By (7.2) we have

$$t = \frac{x(t)}{20} \qquad \text{and} \qquad y(t) = 50 - 5t^2.$$

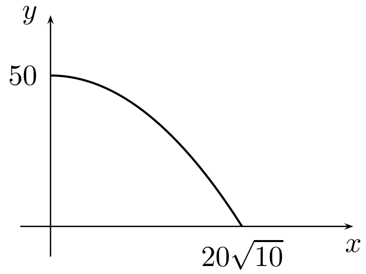
If we substitute the equation on the left into the equation on the right then

$$y(t) = 50 - 5\left(\frac{x(t)}{20}\right)^2$$

which becomes

$$y = 50 - \frac{x^2}{80}.$$

So the trajectory is a parabola. It is easy to see that $y = 50$ when $x = 0$ and that $x = 20\sqrt{10}$ when $y = 0$. We therefore obtain the following sketch.



□

Remark 7.2.2. The equation

$$\big(x(t), y(t)\big) = \big(20t, 50 - 5t^2\big), \qquad t \in \mathbb{R}$$

is called a *parametrisation* of the parabola

$$y = 50 - \frac{x^2}{80}.$$

The variable t is called a *parameter.*

Parametrisations of curves are important to study for several reasons.

- They arise naturally with moving bodies, as illustrated in Example 7.2.1. In this context they also contain more information than a corresponding Cartesian equation (since a Cartesian equation only records the trajectory of the body, not the position of the body at any given time).

- Many curves that cannot be described by an equation of the form $y = f(x)$ can be described parametrically.
- Even if a curve is the graph of a function, it may be much easier to describe the curve parametrically than to write down an equation describing the function. (This is the case with the cycloid, which will be studied in Example 7.2.7.)
- Curves in three (or higher) dimensional space are most easily described using parameters.

7.2.1 Parametrisation of conic sections

The trajectories of many heavenly bodies (such as planets and comets) can be modelled as conic sections. In this section we will give some standard parametrisations of conic sections.

Example 7.2.3. Show that the curve given by the functions

$$x(t) = a\cos t, \qquad y(t) = b\sin t, \qquad 0 \le t \le 2\pi \tag{7.3}$$

is an ellipse. Sketch the curve, showing how the point $(x(t), y(t))$ moves as t varies from 0 to 2π.

Solution. Suppose that $t \in [0, 2\pi]$. From (7.3) we see that

$$\frac{x(t)^2}{a^2} = \cos^2 t, \qquad \frac{y(t)^2}{b^2} = \sin^2 t.$$

Since $\cos^2 t + \sin^2 t = 1$,

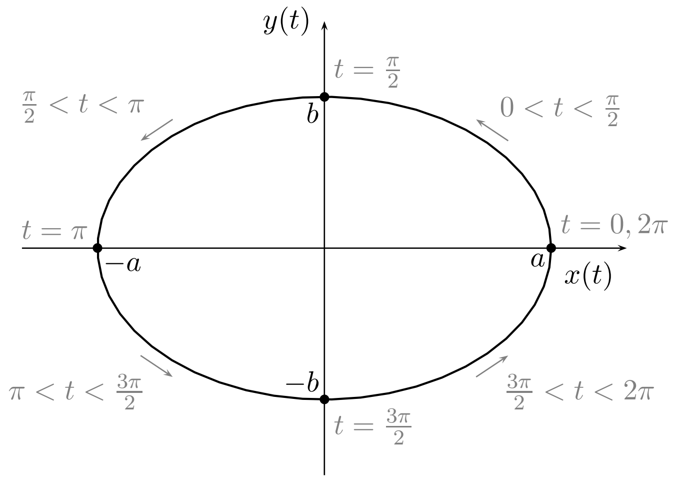
$$\frac{x(t)^2}{a^2} + \frac{y(t)^2}{b^2} = \cos^2 t + \sin^2 t = 1.$$

Hence the point $(x(t), y(t))$ lies on the ellipse

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1.$$

(Conversely, any point (x, y) that satisfies this equation can be expressed as $x = a\cos t$, $y = b\sin t$ for some t in $\mathbb{R}$.)

In fact, it is easy to see that, as t moves from 0 to 2π, $(x(t), y(t))$ traces out the entire ellipse. This is illustrated in the diagram below.



$\square$

The table below lists some commonly used parametrisations of conic sections.

Conic section	Cartesian equation	Parametric equation
Parabola	$4ay = x^2$	$x(t) = 2at$ $y(t) = at^2$
Circle	$x^2 + y^2 = a^2$	$x(t) = a\cos t$ $y(t) = a\sin t$
Ellipse	$\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} = 1$	$x(t) = a\cos t$ $y(t) = b\sin t$
Hyperbola	$\dfrac{x^2}{a^2} - \dfrac{y^2}{b^2} = 1$	$x(t) = a\sec t$ $y(t) = b\tan t$

There are also other parametrisations for each of these curves. (We give a second parametrisation of the hyperbola in Chapter 10.)

In Examples 7.2.1 and 7.2.3, we sketched a parametrically defined curve by first eliminating the parameter t. It is not always possible to do this, as the next example illustrates. Such curves can be sketched with the aid of a table of values.

Example 7.2.4. Sketch the curve defined parametrically by

$$\big(x(t), y(t)\big) = (t^2\cos t, t^2\sin t), \qquad 0 \le t \le 3\pi. \tag{7.4}$$

Solution. We begin by noting that the above parametric equation is similar to the parametrisation

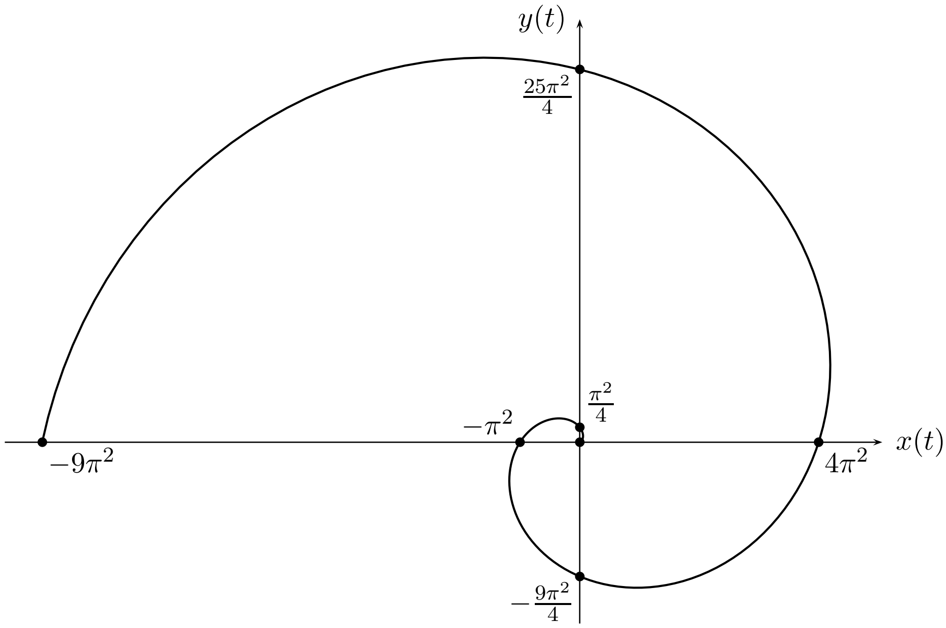
$$\big(x(t), y(t)\big) = (a\cos t, a\sin t) \tag{7.5}$$

of the circle. The only difference is that the points in (7.5) have a fixed distance a from the origin, whereas the points in (7.4) have a variable distance t^2 from the origin. So intuitively, as t increases, the point $(x(t), y(t))$ in (7.4) rotates about origin, while at the same time getting further away from the origin. Therefore it is likely that the curve is a spiral.

The following table of values helps verify this reasoning.

t	0	$\frac{\pi}{2}$	π	$\frac{3\pi}{2}$	2π	$\frac{5\pi}{2}$	3π
x	0	0	$-\pi^2$	0	$4\pi^2$	0	$-9\pi^2$
y	0	$\frac{\pi^2}{4}$	0	$\frac{9\pi^2}{4}$	0	$\frac{25\pi^2}{4}$	0

We plot these points on the Cartesian plane and interpolate between them to produce the following sketch.



□

7.2.2 Calculus and parametric curves

Given a smooth curve defined using a parameter t, one would like to calculate its gradient at any point along the curve. The following proposition enables us to do this without having to eliminate t.

Proposition 7.2.5. *Suppose that x and y are both differentiable functions of t and that y is a function of x. If $x'(t) \neq 0$ then*

$$\frac{dy}{dx} = \frac{\frac{dy}{dt}}{\frac{dx}{dt}} = \frac{y'(t)}{x'(t)}. \tag{7.6}$$

Proof. Since y is a function of x, we write $y = f(x)$. Hence

$$y(t) = f(x(t)).$$

Differentiating both sides with respect to t gives

$$y'(t) = \frac{df}{dx}(x(t))\, x'(t)$$

by the chain rule. Since $x'(t) \neq 0$,

$$\frac{df}{dx}(x(t)) = \frac{y'(t)}{x'(t)}$$

which is equivalent to (7.6). □

Example 7.2.6. A curve is described parametrically by

$$\big(x(t), y(t)\big) = (t^2 - 1, t^3 + 2t), \qquad t \in \mathbb{R}.$$

Find the equation of the tangent to the curve at the point $(0, 3)$.

Solution. By Proposition 7.2.5,

$$\begin{aligned}\frac{dy}{dx} &= \frac{y'(t)}{x'(t)} \\ &= \frac{3t^2+2}{2t} \qquad (7.7)\end{aligned}$$

provided that $t \neq 0$.

It is easy to see that

$$t^2 - 1 = 0 \qquad \text{and} \qquad t^3 + 2t = 3$$

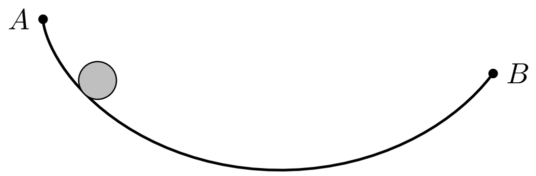
only when $t = 1$. By substituting this into (7.7), we find that the gradient of the curve at $(0, 3)$ is $5/2$. Hence the equation of the tangent at $(0, 3)$ is given by

$$y = \frac{5x}{2} + 3.$$

□

7.2.3 The cycloid and curve of fastest descent

Suppose that a point A is higher than a point B. What curve, starting at A and ending at B, would enable a particle moving under the influence of gravity (and ignoring friction) to move from A to B in the shortest time?

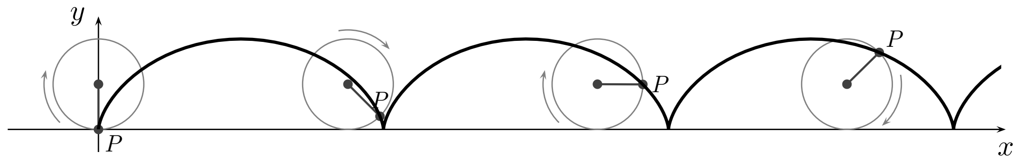


Curve of fastest descent.

Such a curve is known as a *curve of fastest descent* or a *brachistochrone* (which, in Greek, means 'shortest time').

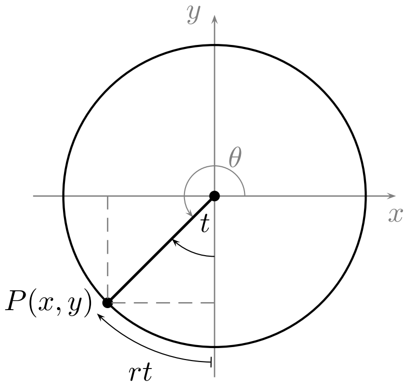
In 1638 Galileo stated that the curve of fastest descent is the arc of a circle. This was shown to be false by Johann Bernoulli who found the correct description of the curve. In 1696 he issued a public challenge for others to find this description. Four mathematicians responded with a correct answer: Jakob Bernoulli (Johann's brother), Gottfried Leibniz, Isaac Newton and Guillaume de l'Hôpital. The curve is the arc of a *cycloid.*

Example 7.2.7 (The cycloid)**.** A circle of radius r rolls along the x-axis, starting from the origin as shown below. Describe the locus $(x(t), y(t))$ of the point P on the edge of the circle that satisfies $(x(0), y(0)) = (0, 0)$

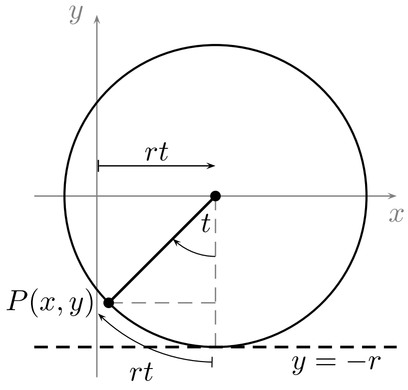


The cycloid.

Solution. The solution will be built up over three steps, of which the first two correspond to the diagrams below.



Step 1.



Step 2.

Step 1. First we consider a circle of radius r centred at the origin. If the circle rotates about its centre by t radians in a clockwise direction (as shown) then

$$\begin{aligned} x(t) &= r\cos\theta = r\cos(\tfrac{3\pi}{2} - t) = r\big(\cos\tfrac{3\pi}{2}\cos t + \sin\tfrac{3\pi}{2}\sin t\big) = -r\sin t, \\ y(t) &= r\sin\theta = r\sin(\tfrac{3\pi}{2} - t) = r\big(\sin\tfrac{3\pi}{2}\cos t - \cos\tfrac{3\pi}{2}\sin t\big) = -r\cos t \end{aligned}$$

and the point P sweeps out a distance of rt.

Step 2. We now consider what happens when the circle rolls along the line $y = -r$. If the circle rotates by t radians, then the centre of the circle moves to the right by rt (corresponding to the distance that P sweeps out – see Step 1). Hence

$$x(t) = -r\sin t + rt.$$

By Step 1 we still have

$$y(t) = -r\cos t.$$

Step 3. We retain the same situation as in Step 2, except that we shift the circle up by r (so that the circle rolls along the line $y = 0$ rather than along the line $y = -r$. Thus

$$y(t) = -r\cos t + r.$$

By Step 2 we still have

$$x(t) = -r\sin t + rt.$$

Final solution. The locus of the curve is given by

$$\begin{aligned} x(t) &= r(t - \sin t), \\ y(t) &= r(1 - \cos t) \end{aligned}$$

where $t \geq 0$. □

Remark 7.2.8. The curve of fastest descent from A to B is the unique arc of an (inverted) cycloid whose tangent at A is vertical.

7.3 Curves defined by polar coordinates

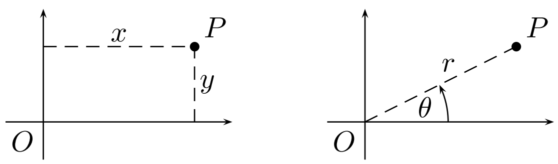
(Ref: SH10 §§10.2, 10.3)

In this final section, we introduce a different coordinate system for the plane. Many equations in this system correspond to beautiful curves.

7.3.1 Polar coordinates

Every point P in plane can be specified by

- (x, y), where x is the horizontal signed distance of P from the origin, and y is the vertical distance, or
- (r, θ), where r is the distance of P from the origin and θ is the angle (taken in the anticlockwise direction) between OP and the positive horizontal axis.

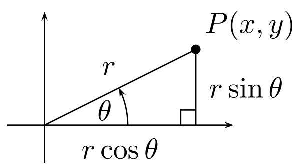


The pair (x, y) is called the *Cartesian coordinates* of P and the pair (r, θ) are *polar coordinates* of P. (If P is the origin then $r = 0$ and θ is not defined.)

The polar coordinates (r, θ) of P are related to the Cartesian coordinates (x, y) of P by the formulae

$$\begin{aligned} x &= r\cos\theta & \qquad r &= \sqrt{x^2+y^2} \\ y &= r\sin\theta & \qquad \tan\theta &= \frac{y}{x}, \quad \text{provided } x \neq 0. \end{aligned} \tag{7.8}$$

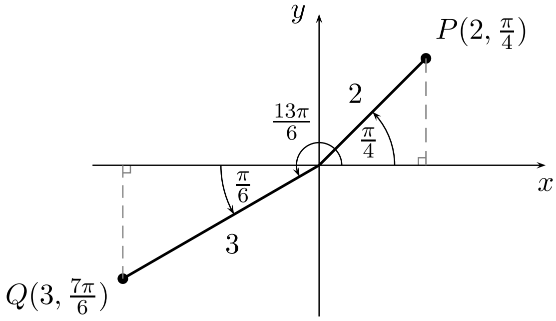
This is easily seen using trigonometry, Pythagoras' theorem and the diagram below.



When converting between the polar coordinates and Cartesian coordinates, it is best to draw a diagram.

Example 7.3.1. Suppose that the polar coordinates of P and Q are given by $(2, \frac{\pi}{4})$ and $(3, \frac{7\pi}{6})$ respectively. Find the Cartesian coordinates for P and Q.

Solution. First we plot these points on the Cartesian plane.



From the diagram, it is clear that the Cartesian coordinate of P is given by

$$\left(2\cos\frac{\pi}{4}, 2\sin\frac{\pi}{4}\right) = \left(\frac{2}{\sqrt{2}}, \frac{2}{\sqrt{2}}\right) = (\sqrt{2}, \sqrt{2}),$$

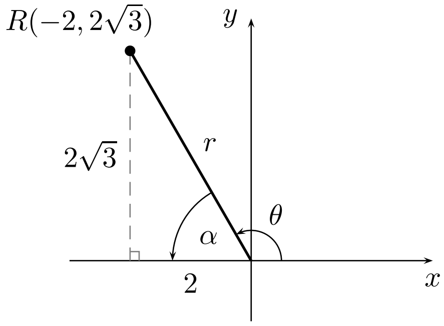
while the Cartesian coordinate of Q is given by

$$\left(-3\cos\frac{\pi}{6}, -3\sin\frac{\pi}{6}\right) = \left(-\sqrt{3}, -\frac{3}{2}\right).$$

□

Example 7.3.2. Suppose that the Cartesian coordinate of R is given by $(-2, 2\sqrt{3})$. Find polar coordinates for R.

Solution. We begin with the following diagram.



In the diagram, $\tan\alpha = \frac{2\sqrt{3}}{2} = \sqrt{3}$, so $\alpha = \frac{\pi}{3}$. Hence

$$\theta = \pi - \alpha = \frac{2\pi}{3}.$$

By Pythagoras' theorem,

$$r^2 = 2^2 + (2\sqrt{3})^2 = 16.$$

Hence in polar coordinates, R can be written as $(4, \frac{2\pi}{3})$. □

Remark 7.3.3. Note that the point R does not have unique polar coordinates. For example

$$(4, \tfrac{2\pi}{3} + 2\pi) \qquad \text{and} \qquad (4, \tfrac{2\pi}{3} - 2\pi)$$

are also polar coordinates for R. In practice, we often choose the value for θ such that $-\pi < \theta \le \pi$.

7.3.2 Basic sketches of polar curves

Many curves can be described by an equation relating the polar variables r and θ. We begin with the simplest case: when either r or θ are constant.

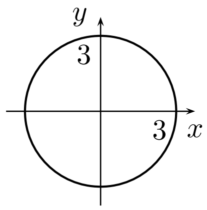
Example 7.3.4. Sketch the curves corresponding to the the equations

(a) $r = 3$

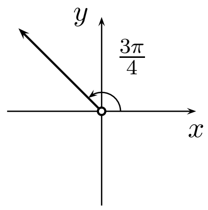
(b) $\theta = \frac{3\pi}{4}$.

Solution. (a) Since $r = 3$, we seek all points in the plane whose distance from the origin is 3. This is a circle of radius three centred at $(0, 0)$.

(b) This is the collection of all points P such that the angle between OP and the positive x-axis is $\frac{3\pi}{4}$. Hence the curve is the ray sketched below with the origin excluded.



$r = 3$



$\theta = \frac{3\pi}{4}$

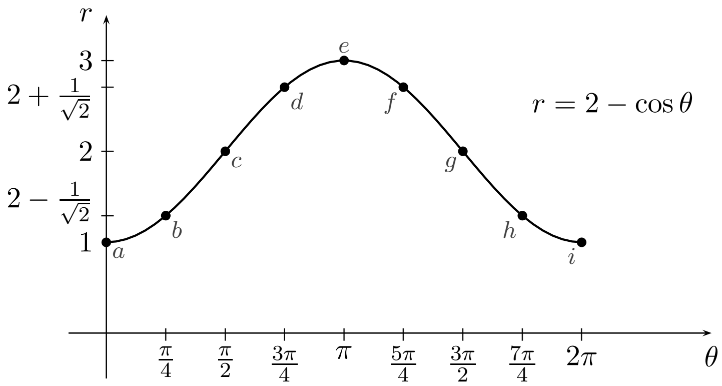
□

To sketch a curve, described by the polar equation $r = f(\theta)$, in the xy-plane, it may be helpful to first sketch $r = f(\theta)$ in the $r\theta$-plane.

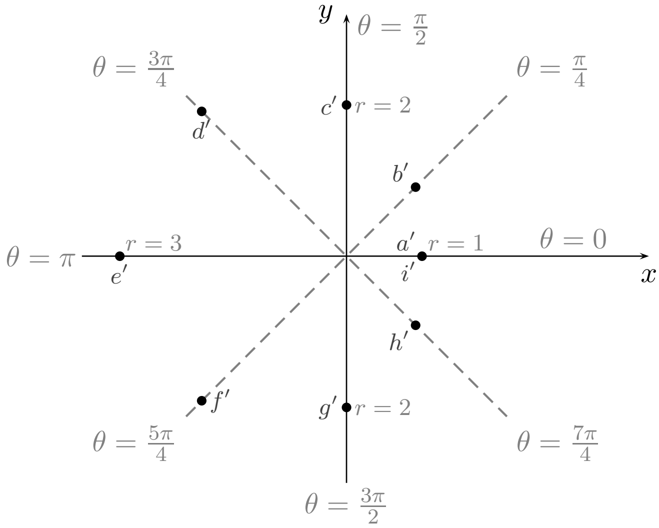
Example 7.3.5. Sketch the curve whose polar equation is

$$r = 2 - \cos\theta.$$

Solution. First we graph r against θ.



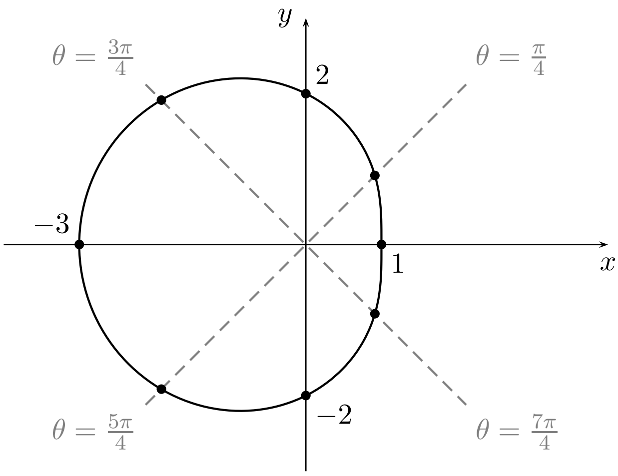
The r-value of each point $a, b, c, ..., i$ in the $r\theta$-plane corresponds to the distance from the origin to each point $a', b', ..., i'$ in the xy-plane.



Our task now is to interpolate between the points $a', b', ..., i'$ to obtain a sketch of the curve. Consider what happens to r as θ moves from 0 to 2π.

- On the $r\theta$-plane, as θ moves from 0 to π, r increases from 1 to 3. So on the xy-plane, if we travel along the curve in an anticlockwise direction from the positive to the negative x-axis, then the distance r from the origin to points on the curve increases from 1 to 3.
- On the $r\theta$-plane, as θ moves from π to 2π, r decreases from 3 to 1. So on the xy-plane, if we travel along the curve in an anticlockwise direction from the negative to the positive x-axis, then the distance r from the origin to points on the curve decreases from 3 to 1.

These considerations lead to the final sketch.



$r = 2 - \cos\theta$

□

We list a few more helpful tips for sketching a curve described by the polar equation $r = f(\theta)$.

- If f is an even function (that is, $f(-\theta) = f(\theta)$) then the polar curve is symmetric about the x-axis.

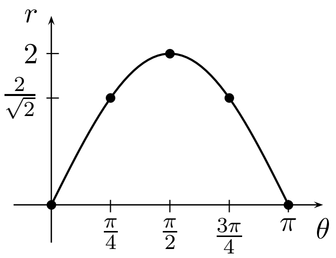
- If $f(\pi - \theta) = f(\theta)$ then the polar curve is symmetric about the y-axis.
- If f is 2π–periodic then it suffices to consider θ in the range $0 \le \theta < 2\pi$.
- Sometimes it is possible to rewrite the equation $r = f(\theta)$ in terms of x and y only.

The use of calculus will be discussed in the next subsection.

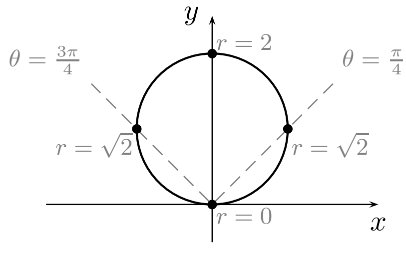
Example 7.3.6. Sketch the curve that is described by the polar equation

$$r = 2\sin\theta, \qquad 0 \le \theta \le \pi.$$

Solution. Graphing r against θ, followed by y against x leads to the following two diagrams.



$r = 2\sin\theta$



$r = 2\sin\theta$

The curve looks as though it is a circle, an observation that can be confirmed by rewriting $r = 2\sin\theta$ in terms of x and y:

$$\begin{aligned} r &= 2\sin\theta \\ r^2 &= 2r\sin\theta && \text{(multiply by } r\text{)} \\ x^2 + y^2 &= 2y && \text{(by (7.8))} \\ x^2 + y^2 - 2y + 1 &= 1 && \text{(completing the square)} \\ x^2 + (y-1)^2 &= 1. \end{aligned}$$

Therefore the polar curve is a circle with Cartesian centre $(0, 1)$ and radius 1. □

Example 7.3.7. Sketch the curve that is described by the polar equation

$$r = 4|\sin(3\theta)|.$$

7.3.3 Sketching polar curves using calculus

Suppose that a curve can be expressed in polar form as $r = f(\theta)$. Since $x = r\cos\theta$ and $y = r\sin\theta$, the curve can also be written in parametric form (with parameter θ) as

$$x(\theta) = f(\theta)\cos\theta \qquad \text{and} \qquad y(\theta) = f(\theta)\sin\theta.$$

Hence,

$$\begin{aligned} \frac{dy}{dx} &= \frac{\frac{dy}{d\theta}}{\frac{dx}{d\theta}} = \frac{y'(\theta)}{x'(\theta)} \\ &= \frac{f(\theta)\cos\theta + f'(\theta)\sin\theta}{-f(\theta)\sin\theta + f'(\theta)\cos\theta} \end{aligned}$$

by the product rule. Since $r = f(\theta)$, this is usually written as

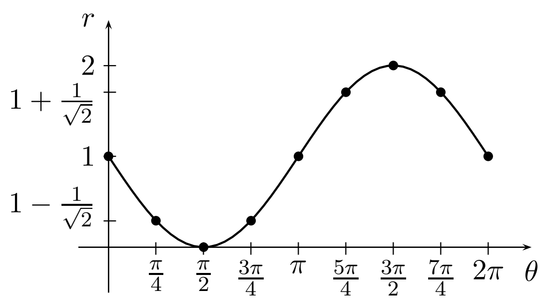
$$\frac{dy}{dx} = \frac{r\cos\theta + \frac{dr}{d\theta}\sin\theta}{-r\sin\theta + \frac{dr}{d\theta}\cos\theta}.$$

(Rather than memorising this formula, we recommend that students understand how it is obtained.)

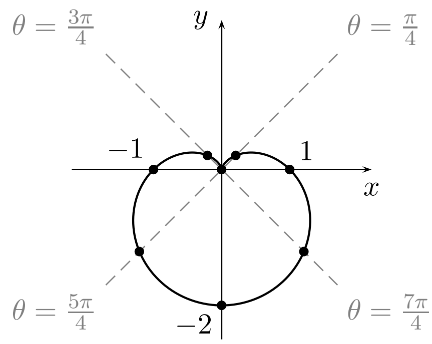
Knowing the value of derivative of the curve at various points allows for greater accuracy in sketching. We illustrate with one example.

Example 7.3.8. Sketch the curve described by the polar equation $r = 1 - \sin\theta$.

Solution. First we graph r against θ and then give a preliminary sketch of y against x.



$r = 1 - \sin\theta$



$r = 1 - \sin\theta$

To obtain a clearer idea of the behaviour of the curve, we calculate $\frac{dy}{dx}$. We have

$$x = r\cos\theta = (1 - \sin\theta)\cos\theta$$

and

$$y = r\sin\theta = (1 - \sin\theta)\sin\theta.$$

Hence

$$\begin{aligned} x'(\theta) &= \cos\theta(-\cos\theta) + (1 - \sin\theta)(-\sin\theta) \\ &= -\cos^2\theta - \sin\theta + \sin^2\theta \\ &= 2\sin^2\theta - \sin\theta - 1 \\ &= (2\sin\theta + 1)(\sin\theta - 1) \end{aligned}$$

and

$$\begin{aligned} y'(\theta) &= \cos\theta - 2\sin\theta\cos\theta \\ &= \cos\theta(1 - 2\sin\theta). \end{aligned}$$

Thus

$$\frac{dy}{dx} = \frac{y'(\theta)}{x'(\theta)} = \frac{\cos\theta(1 - 2\sin\theta)}{2\sin^2\theta - \sin\theta - 1}.$$

To locate horizontal and vertical tangents, we examine where

$$x'(\theta) = 0 \qquad \text{and} \qquad y'(\theta) = 0.$$

Solving $y'(\theta) = 0$ gives

$$\cos\theta = 0 \qquad \text{or} \qquad \sin\theta = \frac{1}{2},$$

that is,

$$\theta = \frac{\pi}{2}, \frac{3\pi}{2}, \frac{\pi}{6}, \frac{5\pi}{6}.$$

Solving $x'(\theta) = 0$ gives

$$\sin\theta = 1 \qquad \text{or} \qquad \sin\theta = -\frac{1}{2},$$

that is

$$\theta = \frac{\pi}{2}, \frac{7\pi}{6}, \frac{11\pi}{6}.$$

So horizontal tangents occur when

$$\theta = \frac{3\pi}{2}, \frac{\pi}{6}, \frac{5\pi}{6}$$

and vertical tangents occur when

$$\theta = \frac{7\pi}{6}, \frac{11\pi}{6}.$$

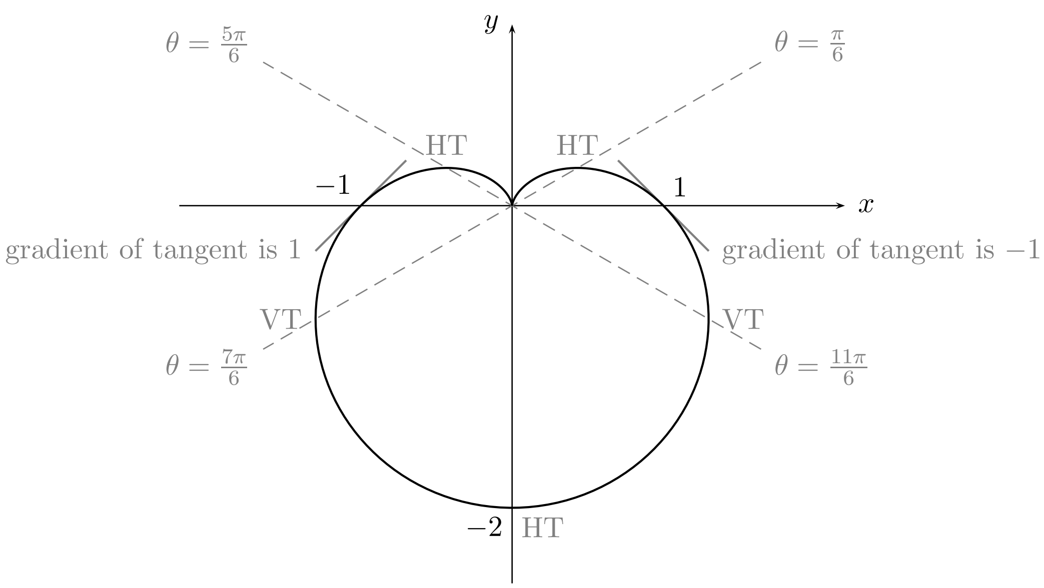
(Understanding how the derivative behaves when θ is close to $\frac{\pi}{2}$ is harder since

$$x'(\tfrac{\pi}{2}) = y'(\tfrac{\pi}{2}) = 0.$$

A full analysis is given in Remark 7.3.9.)

Finally, we observe that $\frac{dy}{dx} = -1$ when $\theta = 0$ and that $\frac{dy}{dx} = 1$ when $\theta = \pi$.

Hence we obtain the following sketch. (The abbreviations HT and VT, which appear in the diagram, stand for 'horizontal tangent' and 'vertical tangent.')



$r = 1 - \sin\theta$

□

Remark 7.3.9. (MATH1141 only) In the above example, the curve has a 'vertical cusp' at the origin. This is seen by analysing $\frac{dy}{dx}$ when θ is close to $\frac{\pi}{2}$. By l'Hôpital's rule we see that

$$\begin{aligned}\lim_{\theta\to\frac{\pi}{2}^-}\frac{x'(\theta)}{y'(\theta)} &= \lim_{\theta\to\frac{\pi}{2}^-}\frac{2\sin^2\theta-\sin\theta-1}{\cos\theta-2\sin\theta\cos\theta}\\ &= \lim_{\theta\to\frac{\pi}{2}^-}\frac{4\sin\theta\cos\theta-\cos\theta}{-\sin\theta-2(\cos\theta\cos\theta-\sin\theta\sin\theta)}\\ &= \lim_{\theta\to\frac{\pi}{2}^-}\frac{0}{1}\\ &= 0.\end{aligned}$$

Hence

$$\frac{dy}{dx}=\frac{y'(\theta)}{x'(\theta)}\to\infty \qquad \text{as} \qquad \theta\to\frac{\pi}{2}^-.$$

By symmetry,

$$\frac{dy}{dx}=\frac{y'(\theta)}{x'(\theta)}\to-\infty \qquad \text{as} \qquad \theta\to\frac{\pi}{2}^+.$$

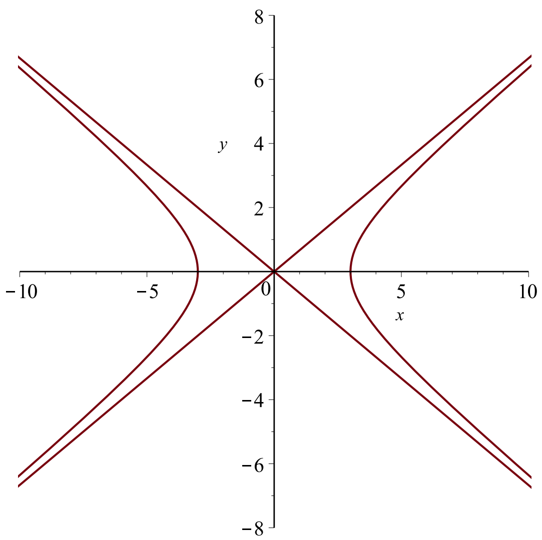
Therefore the derivative has an infinite discontinuity when $\theta=\frac{\pi}{2}$. Such a point is called a 'cusp point.'

Remark 7.3.10. In most cases, a reasonable sketch of a polar curve can be made without appealing to calculus. As seen in the example above, while employing calculus gives a more accurate sketch, it also dramatically increases the amount of work and time required.

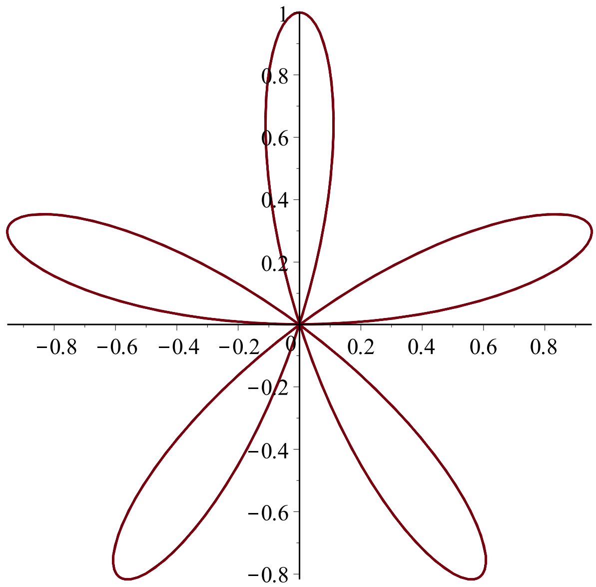
7.4 Maple notes

It is easy to draw parametric curves using Maple. Maple also knows about polar coordinates, but you should be aware that Maple allows $r<0$. For example,

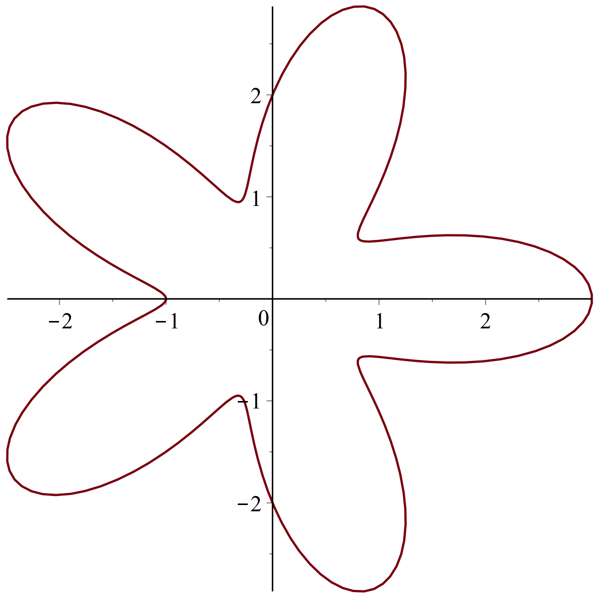
```
> # Draw a hyperbola via parametric specification
> plot([3*sec(t), 2*tan(t), t=-Pi/2..3*Pi/2], x=-10..10, y=-8..8, numpoints=100);
```



```
> # What is the reason for the straight lines through the origin in graph above?
> plot([sin(5*theta), theta, theta=-Pi..Pi], coords=polar);
```




```
> plot([2+cos(5*t), t, t=-Pi..Pi], coords=polar);
```



Problems for Chapter 7

Problems 7.1 : Curves defined by a Cartesian equation

1. [**R**] Find the maximal domain and range of the function f, given by $f(x) = \sqrt{5 + 4x - x^2}$, and sketch its graph.

2. [**R**] Write down the period of each of the following functions f (where possible). Determine which are odd or even. Sketch the graph of each function.

   a) $f(x) = \sin 3x$ b) $f(x) = 1 + \sin(2x/3)$ c) $f(x) = x \sin x$

   d) $f(x) = \tan 3x$ e) $f(x) = \cos^2 x$ f) $f(x) = \sin x + \cos x$

3. [**R**] Suppose that f is an odd function (not everywhere zero). Determine whether each function g below is odd, even or neither.

   a) $g(x) = x^2 f(x)$ b) $g(x) = x^3 f(x)$ c) $g(x) = x^2 + f(x)$

   d) $g(x) = x^3 + f(x)$ e) $g(x) = \sin(f(x))$ f) $g(x) = f(\cos x)$

4. [**R**] For each function f, identify any vertical and oblique asymptotes and hence sketch the graph. (Do *not* use calculus.)

   a) $f(x) = x + 2 + \dfrac{1}{x-3}$ b) $f(x) = \dfrac{x^2 - 2}{x + 1}$ c) [**H**] $f(x) = \dfrac{x^3 - 7x + 8}{x^2 + x - 6}$

5. [**R**] Sketch the following curves, showing their main features.

   a) $y = x^2 + \dfrac{1}{x^2}$ b) [**V**] $y = \dfrac{x-1}{x-2}$ c) $y = e^{-x^2/2}$

   d) $y = xe^{-x}$ e) $y^2 = x(x-4)^2$ f) $y = \dfrac{x^2}{x-2}$

   g) $y = \dfrac{x^2 - 1}{x^2 - 2x}$ h) $y = x \cos^{-1} x$

6. [**H**] (*Longer rather than difficult*) Suppose that $y = \dfrac{3x^2 - 10x + 3}{3x^2 + 10x + 3}$.

   a) Find the values of x for which $y \geq 0$. b) Find the asymptotes.

   c) Find the turning points. d) Find the domain and range.

   e) Sketch the graph.

Problems 7.2 : Parametrically defined curves

7. [**R**] Sketch the curves given by the following parametric equations. Also find, where possible, a Cartesian equation for the curve.

   a) $x = 4\cos t,\ y = 5\sin t$
   b) $x = 3\sec t,\ y = 2\tan t$
   c) $x = t^3,\ y = t^2$
   d) $x = e^t \cos t,\ y = e^t \sin t$

8. **[R]** For each of the curves given in parametric form by

a) $\begin{cases} x = 1 - t \\ y = 1 + t \end{cases}$ b) $\begin{cases} x = 3t + 2 \\ y = t^4 - 1 \end{cases}$ c) $\begin{cases} x = \cos t \\ y = \sin t, \end{cases}$

i) find the points on the curve corresponding to $t = -1, 0, 1,$ and 2;

ii) find any point on the curve where $y = 0$;

iii) find $\dfrac{dy}{dx}$ as a function of t.

9. **[R]** **[V]**

a) Find the equation of the normal to the curve $x = \dfrac{t}{t+1}$, $y = \dfrac{t}{t-1}$ at the point P when $t = 2$.

b) Eliminate t from the above equations and find the gradient of the normal at P using the Cartesian form.

10. **[X]** A curve is given in terms of the parameter t by $x = t^3$, $y = 3t^2$.

a) What is the equation of the curve? Can you sketch it?

b) Show that the equation of the chord joining the points with parameters t_1, t_2 is

$$(t_1^2 + t_1 t_2 + t_2^2)y = 3(t_1 + t_2)x + 3t_1^2 t_2^2.$$

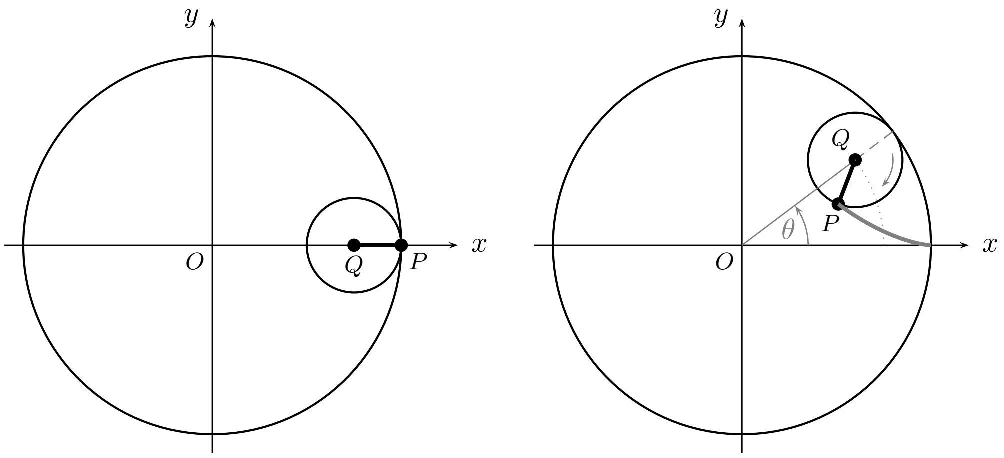
c) Show that the equation of the tangent at t is $ty = 2x + t^3$.

d) Suppose that P is a point with coordinates (a, b) and that P does not lie on on the curve or on the y-axis.

i) Show that either one or three tangents may be drawn from P to the given curve. Illustrate on a sketch the region in which P must lie so that there are three tangents from P to the curve.

ii) Assume that P lies in this region and let Q_1, Q_2, Q_3 denote the points of contact of the tangents from P to the curve. Show that the centroid of the triangle $Q_1Q_2Q_3$ is the point $(-2a, 2b)$.

11. **[H]** Consider a fixed circle of radius 1 centred at the origin and a smaller circle of radius $\frac{1}{4}$ initially centred at $(\frac{3}{4}, 0)$. The smaller circle rolls (without slipping) around the inside rim of the larger circle such that the centre Q of the smaller circle moves in an anticlockwise direction. A point P, fixed on the rim of the smaller circle and initially with coordinates $(1, 0)$, traces out a curve as the smaller circle moves inside the larger circle.



Initial configuration *Configuration after motion has begun*

The goal of this question is to find the Cartesian form of the trajectory of P. Let θ denote the angle (in radians) between OQ and the positive x-axis, as shown in the above diagram.

a) Explain why $\overrightarrow{OQ} = \frac{3}{4}(\cos\theta, \sin\theta)$.

b) [**X**] Explain why $\overrightarrow{QP} = \frac{1}{4}(\cos(-3\theta), \sin(-3\theta))$.

c) Show that $\overrightarrow{OP} = (\cos^3\theta, \sin^3\theta)$.
(You may find techniques from MATH1131 Algebra useful here.)

d) Hence the trajectory of P is given by

$$x = \cos^3\theta, \qquad y = \sin^3\theta, \qquad 0 \le \theta \le 2\pi.$$

By using an appropriate trigonometric identity, eliminate θ to find the cartesian equation of the trajectory of P.

e) Sketch the curve corresponding to this equation. (This curve is called an *astroid* after the Greek word for 'star'.)

12. [**R**] [**V**] In the 1960s two French car engineers, Paul de Casteljau and Pierre Bezier, independently discovered a remarkable new approach to parameterising curves. Let's introduce their approach by finding a quadratic curve determined by the control points $A(1,1), B(2,2)$ and $C(3,1)$ with respecitve coordinate vectors $\mathbf{a}, \mathbf{b}$ and $\mathbf{c}$.

a) The *linear de Casteljau Bezier* curve determined by A and B is given parametrically as

$$\mathbf{p}(t) = (1-t)\mathbf{a} + t\mathbf{b}.$$

This is called the *linear interpolation* of the vectors $\mathbf{a}$ and $\mathbf{b}$. Find a parametric vector equation for $\mathbf{p}(t)$ and evaluate $\mathbf{p}(0)$ and $\mathbf{p}(1)$.

b) Similarly, the linear de Casteljau Bezier determined by B and C is

$$\mathbf{q}(t) = (1-t)\mathbf{b} + t\mathbf{c}.$$

Find a Cartesian equation for this line, and evaluate $\mathbf{q}(1/2)$.

c) The *quadratic de Casteljau Bezier* curve $\mathbf{r}(t)$ determined A, B and C, is the linear interpolation of the vectors $\mathbf{p}(t)$ and $\mathbf{q}(t)$:

$$\mathbf{r}(t) = (1-t)\mathbf{p}(t) + t\mathbf{q}(t).$$

Find an explicit expression for $\mathbf{r}(t)$ and find the polynomials p_0, p_1 and p_2 such that

$$\mathbf{r}(t) = p_0(t)\mathbf{a} + p_1(t)\mathbf{b} + p_2(t)\mathbf{c}.$$

Problems 7.3 : Curves defined by polar coordinates

13. [**R**] The following points are given in polar coordinate form. Plot them on a diagram and find their Cartesian coordinates.

a) $(3, 0)$ b) $(6, 7\pi/6)$ c) $(2, 7\pi/4)$

14. [**R**] Convert these Cartesian coordinates into polar forms with $r \geq 0$ and $-\pi < \theta \leq \pi$.

a) $(-3, 0)$ b) $(-1, -1)$ c) $(-2, 2\sqrt{3})$

d) $(0, 1)$ e) $(-2\sqrt{3}, 2)$ f) $(-2\sqrt{3}, -2)$

15. [**R**] Sketch the graph corresponding to each polar equation.

a) $r = 4$ b) $\theta = 2$ c) $r = 3\theta$, for $\theta \geq 0$.

16. [**R**]

a) Express $r = 6\sin\theta$, where $0 \leq \theta \leq \pi$, in Cartesian form and hence draw its graph.

b) Repeat this for $r = 2\cos\theta$, where $-\pi/2 \leq \theta \leq \pi/2$.

17. [**R**] Sketch the graph corresponding to each polar equation.

a) $r = 2 + \sin\theta$ b) $r = 3 + \cos\theta$ c) [**V**] $r = 2 - 2\cos\theta$

d) $r = 2|\cos\theta|$ e) $r = 3|\sin 6\theta|$ f) $r = |\tan\frac{\theta}{2}|$ $(-\pi < \theta < \pi)$

18. [**H**] [**V**] The hyperbolic spiral is described by the equation $r\theta = a$ whenever $\theta > 0$, where a is a positive constant. Using the fact that $\lim\limits_{\theta \to 0} \dfrac{\sin\theta}{\theta} = 1$, show that the line $y = a$ is a horizontal asymptote to the spiral. Sketch the spiral.

19. [**H**] Show that $r = \dfrac{5}{3 - 2\cos\theta}$ is the polar equation of an ellipse by finding the Cartesian equation of the curve (and completing the square).

20. [**X**]

a) For what values of θ is $r^2 = 25\cos 2\theta$ defined?

b) Sketch the graph of this curve. What difference would it make if you allowed negative values of r?

21. [**H**] Find $\frac{dy}{dx}$ for the curves in Q16a,b,c.

Chapter 8

Integration

The problem of understanding and calculating the area of regions in the plane has a long history. Some of the main contributors to this field are Archimedes (the great Greek mathematician of antiquity), Isaac Barrow (Isaac Newton's mentor), Isaac Newton (one of the greatest mathematicians and physicists of all time), Gottfried Leibniz (Newton's contemporary), Bernhard Riemann (a German mathematician of the nineteenth century) and Henri Lebesgue (a French mathematician of the early twentieth century). In this chapter we see a method for calculating the area of regions with curved boundaries known as *Riemann integration*. We then give a startling result, known as *the fundamental theorem of calculus*, which connects the problem of calculating areas with antidifferentiation. This theorem was known to Barrow and its implications were developed by Newton, Leibniz and their disciples. A consequence of this theorem and other related results, is that the problem of calculating area is much the same as that of calculating mass, volume, work and probability. The unifying feature is a body of theory known as the *integral calculus*.

8.1 Area and the Riemann Integral

(Ref: SH10 §§5.1, 5.2, B.5)

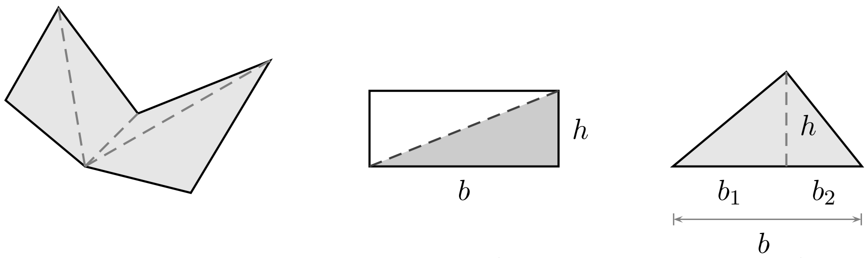
In this section we examine the problem of defining the area of regions with curved boundaries. The main technique introduced is called *integration*, which leads to defining 'the area under the graph of a function' via the *Riemann integral*.

8.1.1 Area of regions with curved boundaries

We all know how to calculate the area of a rectangle or a triangle. The area of more exotic polygons can be computed by partitioning the polygon into triangles and then summing the areas of these triangles (see Figure 8.1 (a)). Such calculations are based on an intuitive understanding of area that takes for granted that

- the area of a rectangle is the product of its length and height,
- areas of congruent regions are equal, and
- the area of a whole region is the sum of the areas of its parts.

When precisely formulated, these points could be taken together as a *definition* of area. With this definition, one can derive a formula for the area of a right-angled triangle (see Figure 8.1 (b)), then



(a) Partition of a polygon (b) $A = \frac{1}{2}bh$ (c) $A = \frac{1}{2}bh$

Figure 8.1: Calculating the area of polygons

for an arbitrary triangle (see Figure 8.1 (c)) and then calculate the area of an arbitrary polygon (as in Figure 8.1 (a)).

When it comes to calculating the area of a region with curved boundaries (see, for example, Figures 8.2 and 8.3), a more sophisticated definition of area is needed. Before asking what shape this definition might take, it is worthwhile asking the question, What *is* area?

First, we recognise that the area of a region is a non-negative real number. Also, our intuition of area informs us that it should satisfy the following properties (see Figure 8.2 for a diagram illustrating each property). We assume throughout that the regions considered are bounded.

(A1) If Ω is a region of the plane then
$$\text{area}(\Omega) \geq 0.$$

(A2) If one region Ω_1 is contained in another region Ω_2, then
$$\text{area}(\Omega_1) \leq \text{area}(\Omega_2).$$

(A3) If the area of a region Ω is partitioned into two smaller disjoint regions Ω_1 and Ω_2, then
$$\text{area}(\Omega) = \text{area}(\Omega_1) + \text{area}(\Omega_2).$$

(A4) If Ω_1 and Ω_2 are congruent regions then
$$\text{area}(\Omega_1) = \text{area}(\Omega_2).$$

(A5) If Ω is a rectangle of length a and height b then
$$\text{area}(\Omega) = ab.$$

These properties are *axioms* that any definition of area should satisfy.

Rather than give a definition for the area of any bounded region of the plane, we shall only give a definition for 'the area under the graph of a function.' The following example gives some justification for this decision.

Example 8.1.1. Consider the region Ω of Figure 8.3. By partitioning Ω with straight lines as shown, one can see by axiom (A3) that
$$\text{area}(\Omega) = \text{area}(\Omega_1) + \text{area}(\Omega_2) + \text{area}(\Omega_3).$$
By rotating and translating each subregion, we see by axiom (A4) that each of area(Ω_1), area(Ω_2) and area(Ω_3) is equal to the area under the graph of a function.

This procedure can be done for any region in the plane with a 'reasonable' boundary. Hence the problem of defining the area of a region with a curved boundary can be reduced to defining what is meant by 'the area under the graph of a function.'

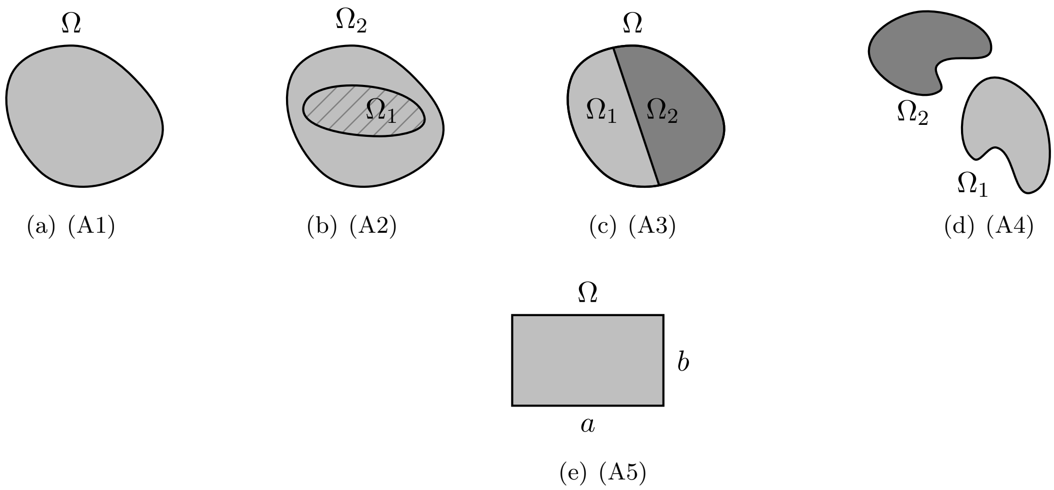


Figure 8.2: Properties of area

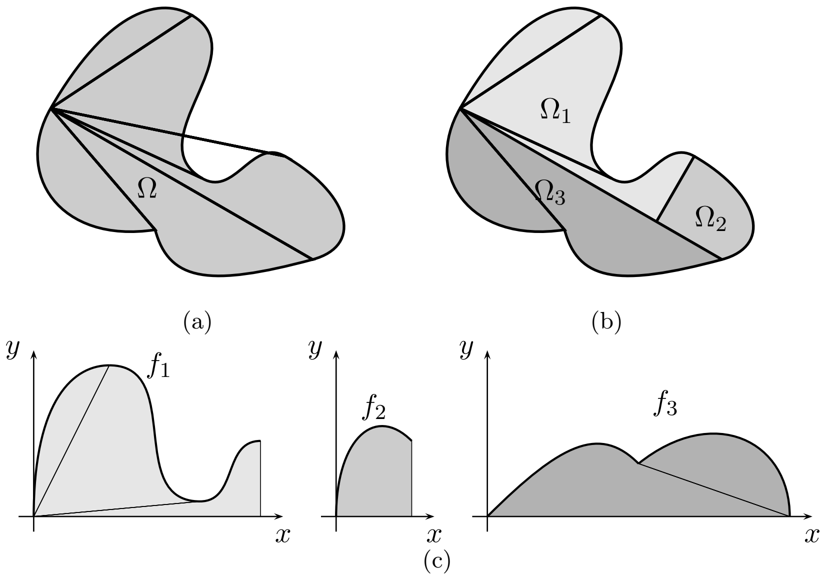


Figure 8.3: Partitioning of a region with curved boundary

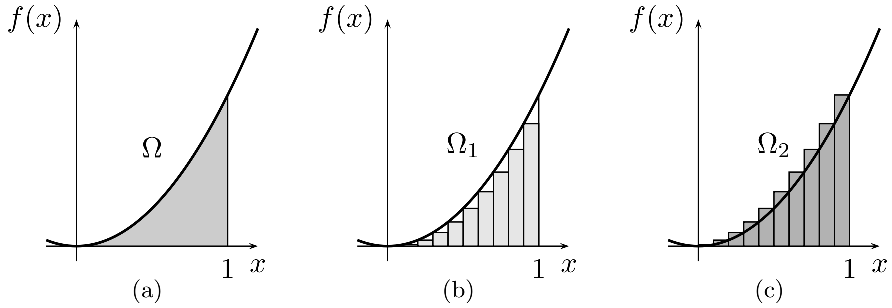


Figure 8.4: Area under the graph of $f(x) = x^2$.

8.1.2 Approximations of area using Riemann sums

To guide us towards a definition for 'the area under the graph of a function,' we study the following example.

Consider the function $f : \mathbb{R} \to \mathbb{R}$, given by the rule $f(x) = x^2$. Let Ω denote the region bounded by the graph of f, the x-axis and the lines $x = 0$ and $x = 1$. The region Ω is shaded in Figure 8.4 (a). The region Ω may be 'approximated' with rectangles as in Figure 8.4 (b) and (c). Since the regions Ω_1 and Ω_2 are composed of disjoint rectangles, their areas are easily calculated using axioms (A3) and (A5) (see Subsection 8.1.1). Moreover, however we eventually choose to define area(Ω), it is clear by axiom (A2) that we want

$$\text{area}(\Omega_1) \le \text{area}(\Omega) \le \text{area}(\Omega_2).$$

This gives a lower and upper bound for area(Ω).

The regions Ω_1 and Ω_2 each consist of ten rectangles. By changing the number of rectangles used, we obtain different upper and lower bounds for area(Ω).

Example 8.1.2. Find different upper and lower bounds for the area of the shaded region Ω of Figure 8.4 (a).

The following formula will help our solution:

$$\sum_{k=1}^{n} k^2 = 1^2 + 2^2 + 3^2 + \cdots + n^2 = \frac{1}{6}n(n+1)(2n+1). \tag{8.1}$$

(This may be proved using mathematical induction. If students need to use such a formula in an exam, they will be given it).

Solution to Example 8.1.2. Let A denote the (as yet undefined) area of the region Ω. Rather than calculating an upper bound for A by using 10 rectangles (as in Figure 8.4 (c)), we use n rectangles, where n is a positive integer (see Figure 8.5). The bases of the rectangles are constructed by subdividing the interval $[0, 1]$ into n subintervals

$$[0, \tfrac{1}{n}], [\tfrac{1}{n}, \tfrac{2}{n}], [\tfrac{2}{n}, \tfrac{3}{n}], \ldots, [\tfrac{n-1}{n}, 1].$$

The set $\mathcal{P}_n$, given by

$$\mathcal{P}_n = \{0, \tfrac{1}{n}, \tfrac{2}{n}, \tfrac{3}{n}, \ldots, \tfrac{n-1}{n}, 1\},$$

that divides the interval $[0, 1]$ into these subintervals is called a *partition* of $[0, 1]$.

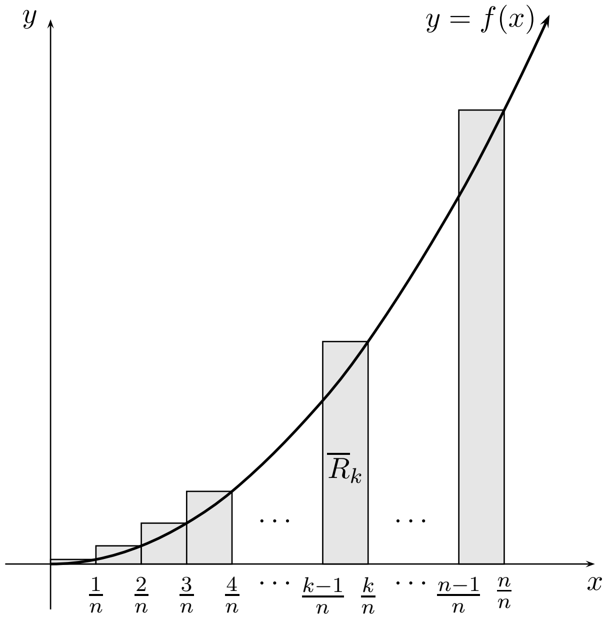


Figure 8.5: An upper Riemann sum for Example 8.1.2.

Let $\overline{R}_k$ denote the area of the kth rectangle in Figure 8.5. Then

$$\begin{aligned}
\overline{R}_k &= \text{width} \times \text{height} \\
&= \frac{1}{n} \times f(\tfrac{k}{n}) \\
&= \frac{1}{n} \times \left(\frac{k}{n}\right)^2 \\
&= \frac{1}{n^3} \times k^2.
\end{aligned}$$

If $\overline{S}_{\mathcal{P}_n}(f)$ denotes the total area of the shaded region in Figure 8.5, then

$$\begin{aligned}
\overline{S}_{\mathcal{P}_n}(f) &= \sum_{k=1}^{n} \overline{R}_k \\
&= \sum_{k=1}^{n} \frac{1}{n^3} \times k^2 \\
&= \frac{1}{n^3} \sum_{k=1}^{n} k^2 \\
&= \frac{1}{n^3}\bigl(1^2 + 2^2 + 3^2 + \cdots + (n-1)^2 + n^2\bigr) \\
&= \frac{1}{n^3} \times \frac{1}{6} n(n+1)(2n+1) \qquad \text{(by formula (8.1))} \\
&= \frac{1}{6}\left(1 + \frac{1}{n}\right)\left(2 + \frac{1}{n}\right) \\
&= \frac{1}{3} + \frac{1}{2n} + \frac{1}{6n^2}.
\end{aligned}$$

The quantity $\overline{S}_{\mathcal{P}_n}(f)$ is called the *upper Riemann sum* of f with respect to the partition $\mathcal{P}_n$. Axiom (A2) imples that

$$A \le \overline{S}_{\mathcal{P}_n}(f) = \frac{1}{3} + \frac{1}{2n} + \frac{1}{6n^2},$$

and by taking different values of n we obtain different upper bounds for A.

To compute a lower bound corresponding to n rectangles, consider Figure 8.6. The area $\underline{R}_k$ of the kth rectangle is given by

$$\underline{R}_k = \frac{1}{n} \times f(\tfrac{k-1}{n}) = \frac{1}{n^3}(k-1)^2.$$

The sum of all the areas of the rectangles is called the *lower Riemann sum* for the function f over the partition $\mathcal{P}_n$ and is denoted by $\underline{S}_{\mathcal{P}_n}(f)$. We see that

$$\begin{aligned}
\underline{S}_{\mathcal{P}_n}(f) &= \sum_{k=1}^{n} \underline{R}_k \\
&= \sum_{k=1}^{n} \frac{1}{n^3} \times (k-1)^2 \\
&= \frac{1}{n^3} \sum_{k=1}^{n} (k-1)^2 \\
&= \frac{1}{n^3}\big(1^2 + 2^2 + 3^2 + \cdots + (n-1)^2\big) \\
&= \frac{1}{n^3} \times \frac{1}{6}(n-1)n(2n-1) \qquad \text{(by formula (8.1))} \\
&= \frac{1}{6}\left(1 - \frac{1}{n}\right)\left(2 - \frac{1}{n}\right) \\
&= \frac{1}{3} - \frac{1}{2n} + \frac{1}{6n^2}.
\end{aligned}$$

By axiom (A2),

$$A \ge \underline{S}_{\mathcal{P}_n}(f) = \frac{1}{3} - \frac{1}{2n} + \frac{1}{6n^2}.$$

Hence for each positive integer n, the inequality

$$\frac{1}{3} - \frac{1}{2n} + \frac{1}{6n^2} \le A \le \frac{1}{3} + \frac{1}{2n} + \frac{1}{6n^2} \tag{8.2}$$

gives an upper and lower bound for A. $\square$

From inequality (8.2) we see that, however the area A of Ω is eventually defined, it must be squeezed between

$$\frac{1}{3} - \frac{1}{2n} + \frac{1}{6n^2} \qquad \text{and} \qquad \frac{1}{3} + \frac{1}{2n} + \frac{1}{6n^2}$$

for *every* value of n. But as $n \to \infty$,

$$\frac{1}{3} - \frac{1}{2n} + \frac{1}{6n^2} \to \frac{1}{3} \qquad \text{and} \qquad \frac{1}{3} + \frac{1}{2n} + \frac{1}{6n^2} \to \frac{1}{3}. \tag{8.3}$$

So there is only *one* number A that satisfies (8.2) for every positive integer n, namely $A = \frac{1}{3}$. Our definition of the area under the graph of a function must agree with this calculation.

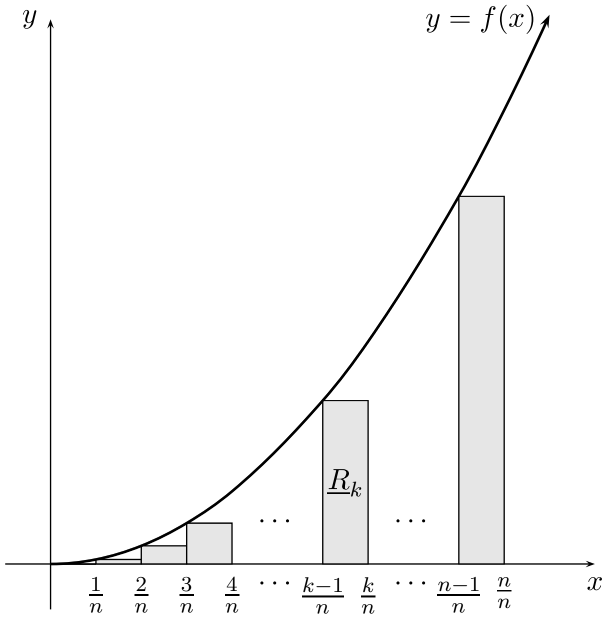


Figure 8.6: A lower Riemann sum for Example 8.1.2.

Remark 8.1.3. The process of calculating upper and lower Riemann sums and taking a limit as in (8.3) is called *integration*.

Guided by the ideas introduced in this subsection, we are now ready to give a definition for the area under the graph of a function.

8.1.3 The definition of area under the graph of a function and the Riemann integral

Suppose that f is a bounded function on $[a, b]$ and that $f(x) \geq 0$ for all x in $[a, b]$. In this subsection we define what is meant by 'the area under the graph of f from a to b'. This is done by constructing upper and lower Riemann sums with respect to partitions of $[a, b]$.

Definition 8.1.4. A finite set $\mathcal{P}$ of points in $\mathbb{R}$ is said to be a *partition* of $[a, b]$ if

$$\mathcal{P} = \{a_0, a_1, a_2, \ldots, a_n\}$$

and

$$a = a_0 < a_1 < a_2 < \ldots < a_n = b.$$

Suppose that $\mathcal{P}$ is a partition of $[a, b]$. In a manner similar to Example 8.1.2, $\mathcal{P}$ is used to construct rectangles that approximate the region under the graph of f (see Figure 8.7). As illustrated in the diagram, the points of $\mathcal{P}$ need not be evenly spaced.

The area of the kth rectangle in Figure 8.7 is

$$\text{width} \times \text{height} = (a_k - a_{k-1}) \times \overline{f}_k,$$

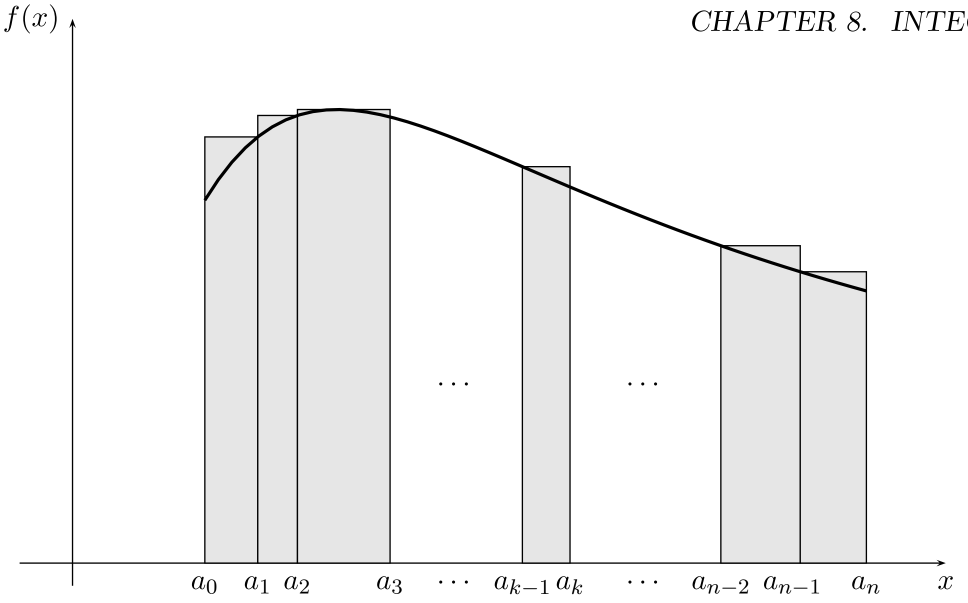


Figure 8.7: An example of an upper Riemann sum

where $\overline{f}_k$ is the maximum value of the function f on the subinterval $[a_{k-1}, a_k]$. The *upper Riemann sum* $\overline{S}_{\mathcal{P}}(f)$ for f with respect to the partition $\mathcal{P}$ is defined by the formula

$$\overline{S}_{\mathcal{P}}(f) = \sum_{k=1}^{n} (a_k - a_{k-1})\overline{f}_k. \tag{8.4}$$

Hence the upper Riemann sum corresponds to the total area of the rectangles in Figure 8.4. Likewise, the *lower Riemann sum* $\underline{S}_{\mathcal{P}}(f)$ for f with respect to the partition $\mathcal{P}$ is defined by the formula

$$\underline{S}_{\mathcal{P}}(f) = \sum_{k=1}^{n} (a_k - a_{k-1})\underline{f}_k, \tag{8.5}$$

where $\underline{f}_k$ is the minimum value of f on the subinterval $[a_{k-1}, a_k]$.

We are now ready to give a definition of the area under the graph of a function f.

Definition 8.1.5 (Definition of the area under the graph of a function)**.** Suppose that a function f is bounded on $[a, b]$ and that $f(x) \geq 0$ for all x in $[a, b]$. If there exists a *unique* real number A such that

$$\underline{S}_{\mathcal{P}}(f) \leq A \leq \overline{S}_{\mathcal{P}}(f) \qquad \text{for } \textit{every} \text{ partition } \mathcal{P} \text{ of } [a, b], \tag{8.6}$$

then we say that A is the *area under the graph of f from a to b.*

It may not seem obvious now but, as we shall see in Section 8.2, not every bounded function f has an area under its graph in the sense of Definition 8.1.5. The following definition gives a special name to those functions that do have a well-defined area under their graph. At this point, we lift the restriction that $f(x) \geq 0$ for all x in $[a, b]$.

Definition 8.1.6 (Definition of the Riemann integral)**.** Suppose that a function f is bounded on $[a, b]$. If there exists a *unique* real number I such that

$$\underline{S}_{\mathcal{P}}(f) \leq I \leq \overline{S}_{\mathcal{P}}(f) \qquad \text{for } \textit{every} \text{ partition } \mathcal{P} \text{ of } [a, b], \tag{8.7}$$

then we say that f is *Riemann integrable* on the interval $[a, b]$. If f is Riemann integrable, then the unique real number I satisfying (8.7) is called the *definite integral of f from a to b* and we write

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

The function f is called the *integrand* of the definite integral, while the points a and b are called the *limits* of the definite integral.

Remark 8.1.7. The notation $\int_a^b f(x)\,dx$ is due to Leibniz. It evolved from a slightly different way of writing down lower and upper Riemann sums. For example, $\overline{S}_{\mathcal{P}}(f)$ may be written as

$$\overline{S}_{\mathcal{P}}(f) = \sum_{k=1}^{n} f(\overline{x}_k)\Delta x_k,$$

where $\Delta x_k = a_k - a_{k-1}$ and f attains its maximum value on $[a_{k-1}, a_k]$ at the point $\overline{x}_k$. When taking a limit, as in (8.3), Δx_k was replaced with dx and the symbol $\sum$ was replaced with an elongated 'S' ('S' stands for 'sum').

Remark 8.1.8. If f is Riemann integrable on $[a, b]$, then the real number $\int_a^b f(x)\,dx$ is the area under the graph of f from a to b, provided that $f(x) \geq 0$ for all x in $[a, b]$. A geometric interpretation of the real number $\int_a^b f(x)\,dx$ when f takes negative values will be discussed in Section 8.3.

Remark 8.1.9. (MATH1141 only.) If f is bounded but not continuous, then f may not attain a maximum or minimum value on the closed interval $[a_k, a_{k-1}]$. This technical difficulty can be overcome by defining the upper and lower Riemann sums in terms of suprema and infima rather than maxima and minima. The concept of a supremum and infimum will be introduced in MATH1241.

Remark 8.1.10. One should think about why Definition 8.1.5 is a good definition for the area under the graph of a function. The following questions need to be answered:

- If the area under the graph of f is well-defined by Definition 8.1.5, would it agree with any other reasonable definition for the area under the graph of f?
- Is Definition 8.1.5 consistent with the axioms of area listed in Subsection 8.1.1?

The answer to each of these questions is 'yes,' but we leave it to the reader to ponder why this might be so. Some of the results given over the next few sections do confirm that our definition of area under a graph is consistent with the axioms of area.

8.2 Integration using Riemann sums

(Ref: SH10 §5.2)

In this section we calculate the area under the graph of functions by computing Riemann sums. Our notation for the upper and lower Riemann sums follows that of equations (8.4) and (8.5).

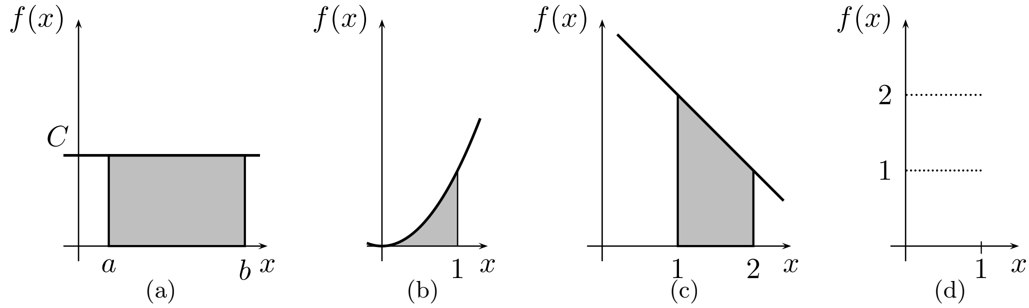


Figure 8.8: Diagrams for examples in Section 8.2.

Example 8.2.1. Suppose that $f(x) = C$ where C is a positive constant. Show that f is Riemann integrable on $[a, b]$ and calculate the area under the graph of f from a to b (see Figure 8.8 (a)).

Solution. Suppose that $\mathcal{P}$ is a partition of $[a, b]$ and that $\mathcal{P} = \{a_0, a_1, \ldots, a_n\}$. Since f is constant, $\overline{f}_k = \underline{f}_k = C$ for every k between 1 and n. Therefore

$$\begin{aligned}
\overline{S}_{\mathcal{P}}(f) &= \sum_{k=1}^{n} \overline{f}_k (a_k - a_{k-1}) \\
&= \sum_{k=1}^{n} C(a_k - a_{k-1}) \\
&= C(a_1 - a_0 + a_2 - a_1 + a_3 - a_2 + \cdots + a_{n-1} - a_{n-2} + a_n - a_{n-1}) \\
&= C(a_k - a_0) \\
&= C(b - a).
\end{aligned}$$

The same calculation gives $\underline{S}_{\mathcal{P}}(f) = C(b - a)$. Since $\mathcal{P}$ is an arbitrary partition of $[a, b]$, we have

$$\underline{S}_{\mathcal{P}}(f) = C(b - a) = \overline{S}_{\mathcal{P}} \qquad \text{for every partition } \mathcal{P} \text{ of } [a, b].$$

Hence there is a unique number A satisfying (8.7), namely $A = C(b - a)$. Therefore f is Riemann integrable,

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

and the area under the graph of f is $C(b - a)$. $\square$

Remark 8.2.2. The area under the graph of f is consistent with axiom (A5).

There are many other kinds of functions which are Riemann integrable. One of the largest classes consists of those functions that are bounded and piecewise continuous.

Definition 8.2.3. A function $f : [a, b] \to \mathbb{R}$ is said to be *piecewise continuous* if it is continuous on $[a, b]$ at all except perhaps a finite number of points.

Figure 8.9 illustrates the graphs of two functions f and g that are piecewise continuous on $[a, b]$. The function f is bounded on $[a, b]$ while the function g is not. According to the next theorem, the area of the shaded region under the graph of f is well-defined.

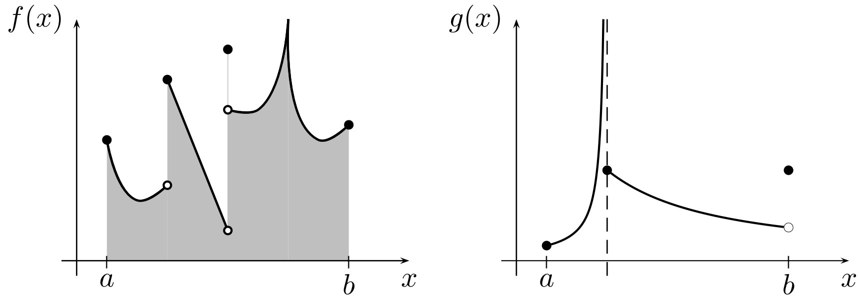


Figure 8.9: Piecewise continuous functions

Theorem 8.2.4. *If f is bounded and piecewise continuous on $[a, b]$ then f is integrable on $[a, b]$.*

Since the proof of this theorem is difficult, we omit it.

In practice, it often hard to prove from first principles that a function f is Riemann integrable. However, if we know in advance that f is integrable on $[a, b]$ then it is much easier to calculate its definite integral from a to b. All one needs to do is show that

$$\lim_{n\to\infty} \overline{S}_{\mathcal{P}_n}(f) = \lim_{n\to\infty} \underline{S}_{\mathcal{P}_n}(f)$$

for some sequence of partitions $\mathcal{P}_n$ of $[a, b]$. If I denotes this common limit, then

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

Example 8.2.5. If $f(x) = x^2$, find the area under the graph of f from 0 to 1 (see Figure 8.8 (b)).

Solution. We first note that f is integrable on $[0, 1]$ since it is bounded and continuous on $[0, 1]$. Suppose that

$$\mathcal{P}_n = \{0, \tfrac{1}{n}, \tfrac{2}{n}, \tfrac{3}{n}, \ldots, \tfrac{n-1}{n}, \tfrac{n}{n}.\}$$

Then

$$\overline{S}_{\mathcal{P}_n}(f) = \frac{1}{3} + \frac{1}{2n} + \frac{1}{6n^2} \qquad \text{and} \qquad \underline{S}_{\mathcal{P}_n}(f) = \frac{1}{3} - \frac{1}{2n} + \frac{1}{6n^2}$$

(see Example 8.1.2 for these calculations). Since

$$\lim_{n\to\infty} \overline{S}_{\mathcal{P}_n}(f) = \frac{1}{3} = \lim_{n\to\infty} \underline{S}_{\mathcal{P}_n}(f),$$

we conclude that

$$\int_0^1 f(x)\,dx = \frac{1}{3}.$$

So the area underneath the graph of f from 0 to 1 is $\frac{1}{3}$. □

Example 8.2.6. Suppose that $f(x) = 3 - x$. Calculate $\int_1^2 f(x)\,dx$ (see Figure 8.8 (c)).

Our final example shows that some functions are not Riemann integrable. For such a function f, the area under the graph of f is *not* defined according to Riemann integration.

Example 8.2.7. Suppose that $f : [0, 1] \to \mathbb{R}$ is defined by

$$f(x) = \begin{cases} 2 & \text{if } x \in \mathbb{Q} \\ 1 & \text{if } x \notin \mathbb{Q}. \end{cases}$$

(see Figure 8.8 (d)). Show that f is not Riemann integrable.

Solution. Suppose that $\mathcal{P}_n$ is an arbitrary partition of $[0, 1]$. Then $\overline{f}_k = 2$ while $\underline{f}_k = 1$ for every k between 1 and n. Hence

$$\underline{S}_{\mathcal{P}_n}(f) = 1 \qquad \text{and} \qquad \overline{S}_{\mathcal{P}_n}(f) = 2.$$

(Exercise: show that this is the case by following a method similar to Example 8.2.1). Therefore there is no *unique* real number I satisfying (8.7). We conclude that f is not Riemann integrable. □

Remark 8.2.8. A more sophisticated definition of the integral, known as the *Lebesgue integral*, was developed at the beginning of the twentieth century by the French mathematician Henri Lebesgue. Lebesgue's definition of the integral allows us to integrate some functions which are not Riemann integrable. For example, if f is the function of Example 8.2.7 then

- f is not Riemann integrable and the Riemann integral $\displaystyle\int_0^1 f(x)\,dx$ is not defined;
- f is Lebesgue integrable and the Lebesgue integral $\displaystyle\int_0^1 f(x)\,dx$ is equal to 1.

The Lebesgue integral is studied in some third year mathematics courses.

8.3 The Riemann integral and signed area

(Ref: SH10 §5.5)

The definite integral

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

can only be interpreted as the area under the graph of f from a to b if $f(x) \geq 0$ for all x in $[a, b]$. How should we interpret the definite integral if the integrand f has negative values on $[a, b]$?

Consider the function f graphed in Figure 8.10 (a). The upper sum from a to b, illusrated in Figure 8.10 (b), approximates area(Ω_1). However, since $f(x) < 0$ when $b < x < c$, each term in the upper sum from b to c will be negative. Hence the upper sum from b to c approximates $-$area(Ω_2). By increasing the number of subdivisions of $[a, c]$ and decreasing the size of each subdivison, these approximations get better and better. It follows that

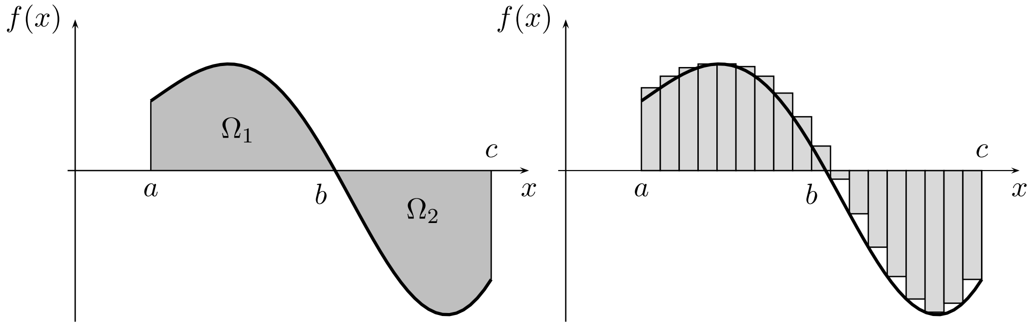
$$\int_a^c f(x)\,dx = \text{area}(\Omega_1) - \text{area}(\Omega_2).$$

We call this quantity the *signed area* under the graph of f from a to c.

Thus there are two ways to calculate the (unsigned or absolute) area of the entire shaded region Ω of Figure 8.10 (a):

- $\text{area}(\Omega) = \displaystyle\int_a^c |f(x)|\,dx$; or
- $\text{area}(\Omega) = \displaystyle\int_a^b f(x)\,dx - \int_b^c f(x)\,dx.$

This kind of calculation will be illustrated in Section 8.6.



(a) (b)

Figure 8.10: Signed area and the definite integral

8.4 Basic properties of the Riemann integral

(Ref: SH10 §5.8)

The following proposition gives a list of some of the basic properties of the Riemann integral.

Proposition 8.4.1 (Basic properties of the Riemann integral). *Suppose that f and g are integrable functions over $[a, b]$.*

(i) If α and β are real numbers then $\alpha f + \beta g$ is integrable and

$$\int_a^b (\alpha f + \beta g)(x)\, dx = \alpha \int_a^b f(x)\, dx + \beta \int_a^b g(x)\, dx.$$

(ii) If $a < c < b$ then

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

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

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

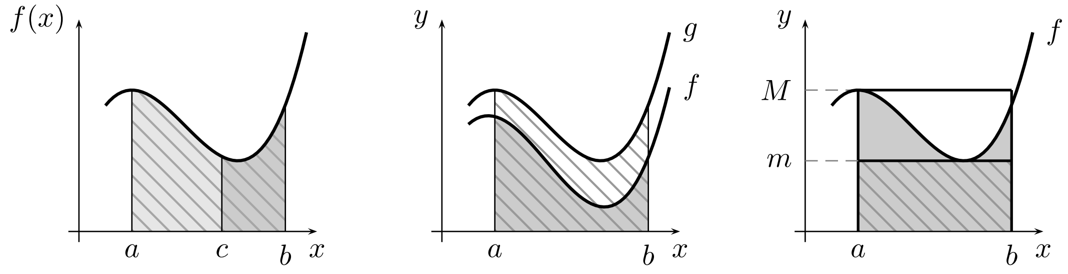
(v) If $m \leq f(x) \leq M$ for all x in $[a, b]$ then

$$m(b - a) \leq \int_a^b f(x)\, dx \leq M(b - a).$$

(vi) If $|f|$ is integrable on $[a, b]$ then

$$\left| \int_a^b f(x)\, dx \right| \leq \int_a^b |f(x)|\, dx.$$

Each of these properties can be interpreted geometrically (see Figure 8.11). For example, under the assumption that $f(x) \geq 0$, property (iv) says that if the graph of f lies beneath the graph of g then the area under the graph of f will be less than the area under the graph of g. Except for (i), each of these properties is easy to prove.



(a) Prop 8.4.1(ii) (b) Prop 8.4.1 (iv) (c) Prop 8.4.1 (v)

Figure 8.11: Geometric interpretation of Proposition 8.4.1.

Proof. We omit the proof of (i) and (ii).

(iii) If $f(x) \geq 0$ for all x in $[a, b]$, then each term in lower Riemann sum (8.5) is non-negative for any partition $\mathcal{P}$ of $[a, b]$. Hence

$$0 \leq \underline{S}_{\mathcal{P}}(f) \leq \int_a^b f(x)\,dx$$

by Definition 8.1.6.

(iv) Suppose that $h = g - f$. Then $h(x) \geq 0$ for all x in $[a, b]$ and by applying (iii) and (i) we have

$$0 \leq \int_a^b h(x)\,dx = \int_a^b \big(g(x) - f(x)\big)\,dx = \int_a^b g(x)\,dx - \int_a^b f(x)\,dx.$$

Rearranging the inequality gives the result.

(v) By (iv) we have

$$\int_a^b m\,dx \leq \int_a^b f(x)\,dx \leq \int_a^b M\,dx.$$

By Example 8.2.1, the integral on the left is $m(b-a)$ and the integral on the right is $M(b-a)$, completing the proof.

(vi) Now

$$-|f(x)| \leq f(x) \leq |f(x)| \qquad \forall x \in [a, b],$$

so

$$\int_a^b -|f(x)|\,dx \leq \int_a^b f(x)\,dx \leq \int_a^b |f(x)|\,dx$$

by (iv). By (i), $\int_a^b -|f(x)|\,dx = -\int_a^b |f(x)|\,dx$ and so

$$-\int_a^b |f(x)|\,dx \leq \int_a^b f(x)\,dx \leq \int_a^b |f(x)|\,dx$$

which is equivalent to

$$\left|\int_a^b f(x)\,dx\right| \leq \int_a^b |f(x)|\,dx$$

by a property of the absolute value function. □

So far the definite integral

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

has only been defined when $a < b$. We now extend this definition.

Definition 8.4.2. Suppose that $b < a$ and that f in integrable on $[b, a]$. Then we define

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

to be the real number

$$-\int_b^a f\,dx$$

and we define

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

to be 0.

This definition gives a more general version of Proposition 8.4.1 (ii). We omit the proof.

Proposition 8.4.3. *Suppose that a, b and c are real numbers and that f is integrable over some interval containing a, b and c. Then*

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

8.5 The first fundamental theorem of calculus

In this section we present a surprising theoretical result that indicates a strong connection between integration (calculating areas) and differentiation (calculating gradients of tangents). This result has enormous ramifications for the problem of calculating area.

We begin by studying the area under the graph of a function. Suppose that f is a continuous function on the interval $[a, b]$. Since f is integrable on $[a, b]$ (see Theorem 8.2.4) we can define the function $F : [a, b] \to \mathbb{R}$ by the formula

$$F(x) = \int_a^x f(t)\,dt \qquad \forall x \in [a, b]. \tag{8.8}$$

For x in $[a, b]$, the number $F(x)$ represents the (signed) area captured underneath the graph of x on the interval $[a, x]$ (see Figure 8.12 (a)). Thus F is an 'area function'.

(Note that the limit x in (8.8) is a different symbol from the 'dummy variable' t. It would be incorrect and confusing to write

$$F(x) = \int_a^x f(x)\,dx$$

for formula (8.8).)

We investigate some of the properties of F. First, it is easy to see that if $f(t) \geq 0$ for all t in $[a, b]$ then the shaded area under f increases as x increases. That is, if f is a non-negative function then F is an increasing function.

Second, we ask whether F continuous on $[a, b]$ and differentiable on (a, b). Since differentiability on (a, b) implies continuity on (a, b), we address the question of differentiability first. Consider the difference quotient

$$\frac{F(x + h) - F(x)}{h}$$

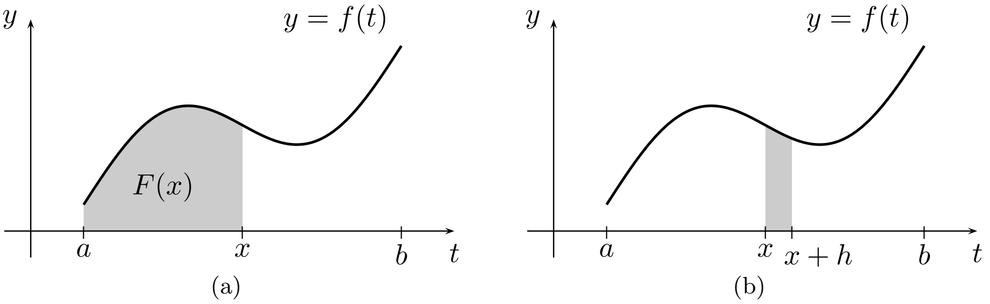


Figure 8.12: The area function F

when $x \in (a, b)$ and h is small. Now, if h is a small positive number then Figure 8.12 (b) shows that

$$\begin{aligned} F(x+h) - F(x) &= (\text{area under } f \text{ on the interval } [x, x+h]) \\ &\approx f(x) \times h, \end{aligned}$$

since $f(t) \approx f(x)$ for all t in $[x, x+h]$ by the continuity of f. Dividing both sides by h gives

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

As h gets smaller, this approximation gets better and so we expect that

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

A similar argument may be repeated when h is a small negative number to obtain the corresponding left-hand limit. Therefore it seems true that

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

This intuitive argument can be made rigorous (see the proof at the end of this section). Hence we conclude that F is continuous and differentable on (a, b) and that $F'(x) = f(x)$ for all x in (a, b).

The fact that a continuous function f and its 'area function' F are related by the differential equation

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

has far-reaching consequences. This result is so important that it has come to be known as the *fundamental theorem of calculus.*

Theorem 8.5.1 (The first fundamental theorem of calculus)**.** *If f is continuous function defined on $[a, b]$ then the function $F : [a, b] \to \mathbb{R}$, defined by*

$$F(x) = \int_a^x f(t)\, dt, \tag{8.9}$$

is continuous on $[a, b]$, differentiable on (a, b) and has derivative F' given by

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

for all x in (a, b).

The first fundamental theorem of calculus is a deep mathematical result which relates differentiation to integration. From the statement of the theorem we note the following points.

- The fact that F satisfies the equation $F' = f$ on (a, b) says that F is an antiderivative of f on (a, b). Hence every continuous function f has an antiderivative F, given by formula (8.9).

- Since any two antiderivatives of f differ by a constant (see Theorem 5.9.4), every antiderivative of f is of the form
$$F + \text{constant},$$
where F is given by the integral formula (8.9). Hence the process of integration and of antidifferentiation are essentially the same. (This is surprising from a geometric point of view, where integration is used to calculate areas while differentiation is used to calculate gradients of tangents.)

- The above point suggests that integration and differentiation are inverse processes. If one takes a function f, integrates it and differentiates the result, then one obtains f again. That is, differentiation undoes what integration does to f. This is precisely expressed by the formula
$$f(x) = \frac{d}{dx}\left(\int_a^x f(t)\,dt\right).$$
Whether or not the converse statement is true (that is, that integration undoes what differentiation does to a function f) will be discussed in the next section.

Thus the first fundamental theorem of calculus suggests that there may be a different method for calculating the area under the graph of f; rather than integrating via limits of Riemann sums, try integrating via antidifferentiation. This approach is investigated in the next section.

We end this section by presenting a proof of the first fundamental theorem of calculus. The proof of the differentiability of F is based on the intuitive argument given immediately prior to the statement of Theorem 8.5.1.

Proof of Theorem 8.5.1. We begin by showing that F is differentiable on (a, b). Suppose that $x \in (a, b)$ and consider a small positive number h. By Proposition 8.4.1 (ii) one can show that

$$F(x + h) - F(x) = \int_x^{x+h} f(t)\,dt. \tag{8.10}$$

Since f is continuous on $[a, b]$, it attains a minimum value m_h and a maximum value M_h on $[x, x+h]$; that is,

$$m_h \le f(t) \le M_h \qquad \forall t \in [x, x + h].$$

By Proposition 8.4.1 (v) (see also Figure 8.13),

$$m_h h \le \int_x^{x+h} f(t)\,dt \le M_h h$$

and hence

$$m_h h \le F(x + h) - F(x) \le M_h h.$$

by (8.10). Now $h > 0$ so

$$m_h \le \frac{F(x + h) - F(x)}{h} \le M_h.$$

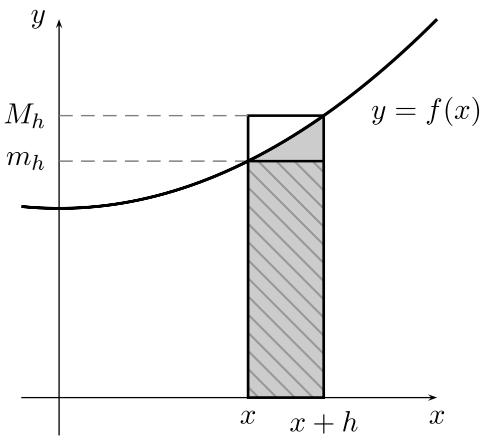


Figure 8.13: Diagram for the proof of Theorem 8.5.1.

Since f is continuous on $[x, x+h]$,

$$\lim_{h\to 0^+} m_h = f(x) = \lim_{h\to 0^+} M_h$$

and hence

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

by the pinching theorem (Theorem 2.5.9).

In a similar manner, one can show that

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

Hence F is differentiable on (a,b) and $F'(x) = f(x)$ for all x in (a,b).

To complete the proof, we need to show that F is continuous on $[a,b]$. Since F is differentiable on (a,b), it is also continuous on (a,b) by Theorem 4.5.1. So it suffices to show that F is continuous at the endpoints of $[a,b]$.

We will show that F is continuous at a. Since f is continuous on $[a,b]$ there is a positive constant M such that $|f(x)| \le M$ for all x in $[a,b]$ (see Corollary 3.4.9). Suppose that $x \in (a,b)$. By using the properties of Proposition 8.4.1,

$$\begin{aligned}
|F(x)-F(a)| &= \left|\int_a^x f(t)\,dt - \int_a^a f(t)\,dt\right| \\
&= \left|\int_a^x f(t)\,dt\right| && \text{(Definition 8.4.2)} \\
&\le \int_a^x |f(t)|\,dt && \text{(Proposition 8.4.1 (vi))} \\
&\le \int_a^x M\,dt && (f \text{ is bounded}) \\
&= M|x-a| && \text{(Example 8.2.1).}
\end{aligned}$$

Hence

$$|F(x)-F(a)| \le M|x-a| \to 0$$

as $x \to a^+$, or in other words, $\lim_{x\to a^+} F(x) = F(a)$. So F is continuous at a.

The proof of continuity at the endpoint b is similar. $\square$

8.6 The second fundamental theorem of calculus

(Ref: SH10 §5.4)

The first fundamental theorem of calculus says that the processes of integration and antidifferentiation are essentially the same. The second fundamental theorem of calculus gives a fast way of calculating integrals by exploiting this similarity.

Theorem 8.6.1 (The second fundamental theorem of calculus)**.** *Suppose that f is a continuous function on $[a,b]$. If F is an antiderivative of f on $[a,b]$ then*

$$\int_a^b f(t)\,dt = F(b) - F(a). \tag{8.11}$$

The second fundamental theorem of calculus is an astonishing result. To calculate the signed area under the graph of a continuous function f, all one needs to know is the value of any antiderivative of f at the endpoints a and b. To illustrate the power of this result, compare the lengthy calculation of Riemann sums in Example 8.1.2 with the example below.

Example 8.6.2. Suppose that $f(x) = x^2$ for all real numbers x. Calculate the area of the region bounded by the x-axis, the graph of f and lines $x = 0$ and $x = 1$.

Solution. Essentially we are being asked to evaluate

$$\int_0^1 f(t)\,dt$$

where $f(t) = t^2$ (see Figure 8.4 (a)). An antiderivative F of f is given by $F(t) = \frac{t^3}{3}$. Hence, by the second fundamental theorem of calculus,

$$\int_0^1 f(t)\,dt = F(1) - F(0) = \frac{1^3}{3} - \frac{0^3}{3} = \frac{1}{3}.$$

So the area of the region is 1/3. $\square$

Remark 8.6.3. The expression $F(b) - F(a)$ appearing in (8.11) is used so frequently that it is often abbreviated to

$$F(x)\Big|_a^b \qquad \text{or} \qquad \Big[F(x)\Big]_a^b.$$

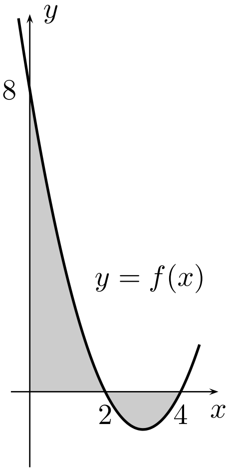
Thus the calculation of Example 8.6.2 may be written as

$$\int_0^1 t^2\,dt = \left[\frac{t^3}{3}\right]_0^1 = \frac{1}{3} - 0 = \frac{1}{3}.$$

When calculating areas, it is important to bear in mind that the integral in (8.11) represents the *signed* area.

Example 8.6.4. Calculate the area of the region bounded by the x-axis, the lines $x = 0$ and $x = 4$ and the function f given by

$$f(x) = (x-2)(x-4).$$



Solution. The above sketch shows that part of the region lies beneath the x-axis. So the area we seek corresponds to the integral

$$\begin{aligned}
\int_0^4 |f(x)|\,dx &= \int_0^2 f(x)\,dx + \int_2^4 (-f(x))\,dx \\
&= \int_0^2 x^2 - 6x + 8\,dx - \int_2^4 x^2 - 6x + 8\,dx \\
&= \left[\frac{x^3}{3} - 3x^2 + 8x\right]_0^2 - \left[\frac{x^3}{3} - 3x^2 + 8x\right]_2^4 \\
&= \left[\left(\frac{8}{3} - 12 + 16\right) - 0\right] \\
&\qquad - \left[\left(\frac{64}{3} - 48 + 32\right) - \left(\frac{8}{3} - 12 + 16\right)\right] \\
&= 8.
\end{aligned}$$

So the area is 8 square units. □

Note that this is a different problem to: Evaluate $\displaystyle\int_0^4 f(x)\,dx$, whose solution is:

$$\begin{aligned}
\int_0^4 f(x)\,dx &= \int_0^4 x^2 - 6x + 8\,dx \\
&= \left[\frac{x^3}{3} - 3x^2 + 8x\right]_0^4 \\
&= \frac{16}{3}.
\end{aligned}$$

We now present a proof of the second fundamental theorem of calculus. The tools used in the proof are the first fundamental theorem of calculus and a corollary of the mean value theorem.

Proof of Theorem 8.6.1. Suppose that f is continuous on $[a,b]$ and that F is an antiderivative of f on $[a,b]$. Define a function $G : [a,b] \to \mathbb{R}$ by the formula

$$G(x) = \int_a^x f(t)\,dt \qquad \forall x \in [a,b].$$

By the first fundamental theorem of calculus (see Theorem 8.5.1), G is continuous on $[a, b]$ and $G'(x) = f(x)$ for all x in (a, b). Hence G is also an antiderivative of f on $[a, b]$. By Theorem 5.9.4, there is a constant C such that

$$G(x) = F(x) + C \tag{8.12}$$

for all x in $[a, b]$. Note that

$$G(a) = \int_a^a f(t)\,dt = 0.$$

So if $x = a$ then equation (8.12) becomes

$$0 = F(a) + C.$$

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

$$\begin{aligned}\int_a^b f(t)\,dt &= G(b)\\ &= F(b) + C \qquad\qquad \text{(by (8.12))}\\ &= F(b) - F(a)\end{aligned}$$

as required. ☐

We return to the question of whether integration and differentiation are inverse processes. Suppose that f is continuous on $[a, b]$. The first fundamental theorem of calculus states that if one integrates a function f and differentiates the result, then one obtains f again. This is expressed by the formula

$$f(x) = \frac{d}{dx}\left(\int_a^x f(t)\,dt\right) \qquad \forall x \in (a, b).$$

To be able to say that integration and differentiation are inverses operations of each other, we must ask whether the converse is true. That is, if one differentiates a function f and integrates the result, then does one obtain f again? In precise terms, we would like the formula

$$f(x) = \int_a^x f'(t)\,dt \qquad \forall x \in (a, b) \tag{8.13}$$

to hold. Unfortunately, this converse is false. The following corollary to the second fundamental theorem of calculus says that one obtains f again only when $f(a) = 0$. Otherwise formula (8.13) is 'out by a constant.'

Corollary 8.6.5. *Suppose that f is continuous on $[a, b]$ and has a continuous derivative on (a, b). Then*

$$\int_a^x f'(t)\,dt = f(x) - f(a)$$

for all x in $[a, b]$

Proof. Apply the second fundamental theorem of calculus to the function f' on the interval $[a, x]$. ☐

The second fundamental theorem of calculus allows a fruitful interaction between Riemann sums and antidifferentiation.

Example 8.6.6. Suppose that $f(x) = \dfrac{1}{1+x^2}$. By considering the lower Riemann sum of f with respect to the partition

$$\{0, \tfrac{1}{n}, \tfrac{2}{n}, \dots, \tfrac{n-1}{n}, \tfrac{n}{n}\}$$

of $[0,1]$, show that

$$\frac{\pi}{4} = \lim_{n\to\infty} \sum_{k=1}^{n} \frac{n}{n^2+k^2} = \lim_{n\to\infty} \left(\frac{n}{n^2+1^2} + \frac{n}{n^2+2^2} + \frac{n}{n^2+3^2} + \cdots + \frac{n}{n^2+n^2} \right).$$

Solution. Let $\mathcal{P}_n$ denote the above partition of $[0,1]$. Since f is a decreasing function on $[0,1]$, the lower Riemann sum is given by

$$\begin{aligned} \underline{S}_{\mathcal{P}_n}(f) &= \sum_{k=1}^{n} \frac{1}{n} \frac{1}{1+(k/n)^2} \\ &= \sum_{k=1}^{n} \frac{n}{n^2+k^2}. \end{aligned}$$

Since f is bounded and continuous, it is integrable on $[0,1]$ and hence

$$\lim_{n\to\infty} \underline{S}_{\mathcal{P}_n}(f) = \int_0^1 \frac{1}{1+x^2}\,dx$$

(see Remark 8.6.7). But by the second fundamental theorem of calculus,

$$\int_0^1 \frac{1}{1+x^2}\,dx = \Big[\tan^{-1}(x)\Big]_0^1 = \frac{\pi}{4}.$$

Hence

$$\frac{\pi}{4} = \lim_{n\to\infty} \sum_{k=1}^{n} \frac{n}{n^2+k^2}$$

as desired. □

Remark 8.6.7. The solution to Example 8.6.6 uses the following result. Suppose that f is Riemann integrable on $[a,b]$ and that $\{\mathcal{P}_n\}_{n=1}^{\infty}$ is a sequence of partitions of $[a,b]$. Let s_n denote the maximum size of the subintervals generated by $\mathcal{P}_n$. If $\lim\limits_{n\to 0} s_n = 0$ then

$$\lim_{n\to\infty} \underline{S}_{\mathcal{P}_n}(f) = \int_a^b f(x)\,dx = \lim_{n\to\infty} \overline{S}_{\mathcal{P}_n}(f).$$

8.7 Indefinite integrals

(Ref: SH10 §5.6)

The fundamental theorems of calculus say that integration and antidifferentiation are, in a certain sense, the same process. In particular, if a F is an antiderivative of f then

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

for some suitable constant C. If we have no particular interest in the interval $[a, x]$ and merely want to indicate that F is an antiderivative of f, then we write

$$\int f(t)\,dt = F(t) + C. \tag{8.14}$$

An integral expressed in this way, without limits, is called an *indefinite integral*. The constant C in (8.14) is called the *constant of integration*.

The table of antiderivatives given in Section 5.9 is equivalent to the following table of indefinite integrals.

Indefinite integrals	
$\int x^r\,dx = \frac{1}{r+1}x^{r+1} + C,$	where r is a rational number and $r \neq -1$
$\int \sin x\,dx = -\cos x + C$	
$\int \cos x\,dx = \sin x + C$	
$\int e^{ax}\,dx = \frac{1}{a}e^{ax} + C$	
$\int \frac{f'(x)}{f(x)}\,dx = \ln\lvert f(x)\rvert + C$	

Since differentiation and indefinite integration (or antidifferentiation) are inverse processes, they have many analogous properties. For example, differentiation behaves *linearly*, which means that

$$(f + g)'(x) = f'(x) + g'(x)$$

and

$$(\alpha f)'(x) = \alpha \times f'(x)$$

for every real number α and every differentiable function f and g. The next proposition says that indefinite integration is also behaves linearly. (We have already shown in Proposition 8.4.1 that definite integration is linear.)

Proposition 8.7.1. *Suppose that f and g are integrable functions and that α is a real constant. Then*

$$\int (f + g)(x)\,dx = \int f(x)\,dx + \int g(x)\,dx \tag{8.15}$$

and

$$\int (\alpha f)(x)\,dx = \alpha \int f(x)\,dx. \tag{8.16}$$

Proof. We prove (8.15) only and leave the proof of 8.16 as an exercise. Suppose that F, G and H are antiderivatives of f, g and $f + g$ respectively. Then

$$H'(x) = (f + g)(x) = f(x) + g(x) = F'(x) + G'(x)$$

by the linearity of differentiation. By expressing the equation $H' = F' + G'$ in terms of indefinite integrals we obtain (8.15) as desired. $\square$

Example 8.7.2. Find

$$\int (10x^4 + 8\sin x - 9e^x)\,dx.$$

Proof. By Proposition 8.7.1,

$$\begin{aligned}\int (10x^4 + 8\sin x - 9e^x)\,dx &= 10\int x^4\,dx + 8\int \sin x\,dx - 9\int e^x\,dx \\ &= 10\left(\frac{1}{5}x^5 + C_1\right) + 8\,(-\cos x + C_2) - 9\,(e^x + C_3) \\ &= 2x^5 - 8\cos x - 9e^x + C\end{aligned}$$

where C_1, C_2 and C_3 are constants of integration and $C = 10C_1 + 8C_2 - 9C_3$. □

Remark 8.7.3. It is standard practice to combine of all constants of integration and simply write '$+C$'. Thus the calculation in Example 8.7.2 may be written as

$$\begin{aligned}\int (10x^4 + 8\sin x - 9e^x)\,dx &= 10\int x^4\,dx + 8\int \sin x\,dx - 9\int e^x\,dx \\ &= 10 \times \frac{1}{5}x^5 + 8(-\cos x) - 9e^x + C \\ &= 2x^5 - 8\cos x - 9e^x + C.\end{aligned}$$

8.8 Integration by substitution

(Ref: SH10 §5.7)

When differentiating the composition of two functions, we apply the chain rule. By reversing this process with integration, we obtain a technique known as *integration by substitution.*

The basic idea is the following. Suppose that F is an antiderivative for f. Then

$$\int f(g(x)).g'(x)\,dx = F(g(x)) + C, \tag{8.17}$$

as can be easily verified by differentiating the right-hand side via the chain rule. On the other hand,

$$\int f(u)\,du = F(u) + C. \tag{8.18}$$

If we make the substitution $u = g(x)$ then the right-hand sides of (8.17) and (8.18) are the same. This suggests the substitution $du = g'(x)\,dx$, so that the left-hand side of (8.17) is transformed into the left-hand side of (8.18). These mechanical substitutions can be used to transform the complicated integral into a simpler one.

Example 8.8.1. Calculate the indefinite integral

$$\int (2x - 3)\cos(x^2 - 3x + 4)\,dx$$

by the method of substitution.

Solution. We begin with the substitution

$$u = x^2 - 3x + 4.$$

Since

$$\frac{du}{dx} = 2x - 3,$$

we also make the substitution

$$du = (2x - 3)\,dx.$$

Then

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

We can verify our answer by differentiating $\sin(x^2 - 3x + 4)$ via the chain rule. □

A similar technique can be used to evaluate definite integrals.

Example 8.8.2. Evaluate the definite integral

$$\int_0^2 3x^2\sqrt{1+x^3}\,dx.$$

Solution. If $u = 1 + x^3$ then $\frac{du}{dx} = 3x^2$. Hence we make the following substitutions:

$$\begin{aligned} u &= 1 + x^3 \\ du &= 3x^2\,dx \\ u &= 1 \text{ (when } x = 0) \\ u &= 9 \text{ (when } x = 2). \end{aligned}$$

Thus

$$\begin{aligned}\int_0^2 3x^2\sqrt{1+x^3}\,dx &= \int_1^9 \sqrt{u}\,du \\ &= \left[\frac{2}{3}u^{3/2}\right]_1^9 \\ &= \frac{2}{3}(9^{3/2} - 1^{3/2}) \\ &= \frac{52}{3}.\end{aligned}$$

□

In general, the reversal of the chain rule via integration gives rise to the *change of variables formula*

$$\boxed{\int_a^b f(g(x))g'(x)\,dx = \int_{g(a)}^{g(b)} f(u)\,du,} \tag{8.19}$$

The substitutions

$$u = g(x), \qquad du = g'(x)\,dx,$$

are only to be understood within this context. Precise conditions under which formula (8.19) holds is given by the following theorem.

Theorem 8.8.3 (Change of variables formula)**.** *Suppose that g is a differentiable function such that g' is continuous on $[a, b]$. If f is continuous on any interval I containing $g(a)$ and $g(b)$ then the change of variables formula (8.19) holds.*

Proof. Since f is continuous on I it has an antiderivative $F : I \to \mathbb{R}$ by the first fundamental theorem of calculus. By the chain rule (Theorem 4.2.2),

$$\begin{aligned}(F \circ g)'(x) &= F'(g(x))g'(x)\\ &= f(g(x))g'(x).\end{aligned}$$

So by two applications of the second fundamental theorem of calculus,

$$\begin{aligned}\int_a^b f(g(x))g'(x)\,dx &= (F \circ g)(b) - (F \circ g)(a)\\ &= F(g(b)) - F(g(a))\\ &= \int_{g(a)}^{g(b)} f(u)\,du,\end{aligned}$$

completing the proof. □

Finding a fruitful substitution is not always easy (and sometimes not even possible). In principle, one looks for a function g in the integrand whose derivative g' is also in the integrand. Once such a function g is identified, try the substitution $u = g(x)$. As illustrated below, it may be necessary to manipulate the integrand to implement this strategy.

Example 8.8.4. Evaluate

$$\int \frac{e^{\sqrt{x}}}{\sqrt{x}}\,dx.$$

Solution. Note that

$$\frac{d}{dx}\big(\sqrt{x}\big) = \frac{1}{2\sqrt{x}}$$

and so if $g(x) = \sqrt{x}$ then $g'(x)$ appears in the integrand (up to the constant factor $1/2$). So the substitution we use is

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

Hence

$$\begin{aligned}\int \frac{e^{\sqrt{x}}}{\sqrt{x}}\,dx &= 2\int \frac{e^{\sqrt{x}}}{2\sqrt{x}}\,dx\\ &= 2\int e^u\,du\\ &= 2e^u + C\\ &= 2e^{\sqrt{x}} + C.\end{aligned}$$

□

There are occasions when we make a substitution of the form $x = g(u)$ rather than $u = g(x)$.

Example 8.8.5. Evaluate the integral I, where

$$I = \int_4^9 \frac{dx}{2+\sqrt{x}}.$$

Solution. If the integral were

$$\int_4^9 \frac{dx}{2+x},$$

then it would be easy to evaluate. The difficulty with I clearly lies with the square root appearing in the integrand. To remove the square root, we use the substitution $x = u^2$. This gives

$$\begin{aligned} x &= u^2 \\ dx &= 2u\,du \\ u &= 2 \text{ when } x = 4 \\ u &= 3 \text{ when } x = 9. \end{aligned}$$

(Note that we are implicitly assuming that $u > 0$. It would be incorrect to take the limits in the variable u to be -2 and 3, for example.) Hence

$$\begin{aligned} I &= \int_2^3 \frac{2u\,du}{2+u} \\ &= 2\int_2^3 \frac{u}{2+u}\,du \\ &= 2\int_2^3 \frac{2+u-2}{2+u}\,du \\ &= 2\int_2^3 1 - \frac{2}{2+u}\,du \\ &= 2\Big[u - 2\ln|2+u|\Big]_2^3 \\ &= 2 + 4\ln\left(\frac{4}{5}\right). \end{aligned}$$

$\square$

If an integrand possesses symmetry, then the corresponding definite integral may be easy to evaluate. For example, if f is an odd function and we integrate f over a balanced interval $[-a, a]$, then the 'negative area' cancels out the 'positive area' and consequently the integral is zero (see Figure 8.14). This result is proved using integration by substitution.

Proposition 8.8.6. *Suppose that f is a continuous function and a is a real number.*

(i) If f is even then

$$\int_{-a}^a f(x)\,dx = 2\int_0^a f(x)\,dx.$$

(ii) If f is odd then

$$\int_{-a}^a f(x)\,dx = 0.$$

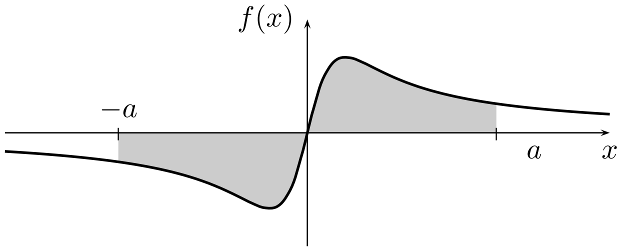


Figure 8.14: Integration with an odd function f over $[-a, a]$

(iii) If f is periodic with period T then

$$\int_a^{a+T} f(x)\,dx = \int_0^T f(x)\,dx.$$

Proof. We prove (ii) only and leave the proofs of (i) and (iii) as an exercise. Suppose that f is an odd function. By breaking up the integral and using a substitution we have

$$\begin{aligned}
\int_{-a}^{a} f(x)\,dx &= \int_{-a}^{0} f(x)\,dx + \int_0^a f(x)\,dx \\
&= -\int_a^0 f(-u)\,du + \int_0^a f(x)\,dx && \text{(substituting } u = -x) \\
&= \int_0^a f(-u)\,du + \int_0^a f(x)\,dx \\
&= -\int_0^a f(u)\,du + \int_0^a f(x)\,dx && \text{(since } f(-u) = f(u)) \\
&= 0,
\end{aligned}$$

completing the proof of (ii). □

Example 8.8.7. Find

$$\int_{\frac{\pi}{17}}^{\frac{\pi}{17}+2\pi} \sin^5 x\,dx.$$

Solution. If $f(x) = \sin^5 x$ then f is both odd and periodic with period 2π. Hence

$$\begin{aligned}
\int_{\frac{\pi}{17}}^{\frac{\pi}{17}+2\pi} \sin^5 x\,dx &= \int_0^{2\pi} \sin^5 x\,dx && \text{(by Proposition 8.8.6 (iii) when } a = \tfrac{\pi}{17}) \\
&= \int_{-\pi}^{\pi} \sin^5 x\,dx && \text{(by Proposition 8.8.6 (iii) when } a = -\pi) \\
&= 0 && \text{(by Proposition 8.8.6 (ii)).}
\end{aligned}$$

□

8.9 Integration by parts

(Ref: SH10 §8.2)

Suppose that u and v are two functions of a variable x. The product rule for differentiation gives

$$\frac{d}{dx}(uv) = v\frac{du}{dx} + u\frac{dv}{dx}.$$

By integrating both sides with respect to x we obtain

$$uv = \int v\frac{du}{dx}\,dx + \int u\frac{dv}{dx}\,dx.$$

Rearranging this equation gives the *integration by parts* formula

$$\boxed{\int u\frac{dv}{dx}\,dx = uv - \int v\frac{du}{dx}\,dx.} \tag{8.20}$$

For definite integrals the formula becomes

$$\boxed{\int_a^b u\frac{dv}{dx}\,dx = \Big[uv\Big]_a^b - \int_a^b v\frac{du}{dx}\,dx.} \tag{8.21}$$

by the second fundamental theorem of calculus.

Integration by parts allows one to evaluate the integral of a product of two functions. Its success depends on choosing u and v such that

$$\int_a^b v\frac{du}{dx}\,dx \quad \text{is easier to calculate than} \quad \int_a^b u\frac{dv}{dx}\,dx.$$

Once u and v are chosen, simply apply formula (8.20) or (8.21).

Remark 8.9.1. The shorthand version

$$\boxed{\int uv' = uv - \int vu'}$$

of formulae (8.20) and (8.21) is easier to remember.

Example 8.9.2. Evaluate the integral

$$\int xe^{2x}\,dx.$$

Solution. We set

$$\begin{aligned} u &= x & v &= \tfrac{1}{2}e^{2x} \\ u' &= 1 & v' &= e^{2x}. \end{aligned}$$

Then integration by parts (see formula (8.20)) gives

$$\begin{aligned} \int xe^{2x}\,dx &= \tfrac{1}{2}xe^{2x} - \int \tfrac{1}{2}e^{2x} \times 1\,dx \\ &= \tfrac{1}{2}xe^{2x} - \tfrac{1}{4}e^{2x} + C, \end{aligned}$$

completing the solution.

(Note that if we had instead set

$$\begin{aligned} u &= e^{2x} & v &= \tfrac{1}{2}x^2 \\ u' &= 2e^{2x} & v' &= x. \end{aligned}$$

then integration by parts gives

$$\int xe^{2x}\,dx = \tfrac{1}{2}x^2e^{2x} - \int x^2e^{2x}\,dx$$

and the integral on the right-hand side is harder to evaluate than the one we started with.) ☐

Sometimes integration by parts must be applied several times to evaluate the integral.

Example 8.9.3. Find the integral I, where

$$I = \int_0^\pi x^2 \sin x\,dx.$$

Example 8.9.4. Evaluate the indefinite integral I, where

$$I = \int e^x \sin x\,dx$$

Integration by parts can also be used to evaluate some integrals of the form $\int f(x)\,dx$. The trick is to rewrite the integral as

$$\int f(x) \times 1\,dx.$$

Example 8.9.5. Evaluate the integral

$$\int \cos^{-1} x\,dx.$$

8.10 Improper integrals

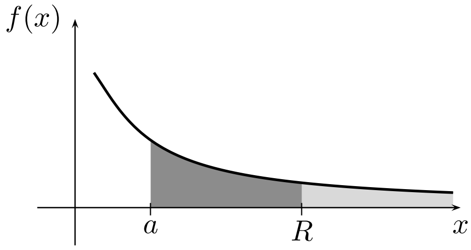
(Ref: SH10 §11.7)

So far we have defined definite integrals for bounded functions over finite intervals. In this section we look at integrals over infinite intervals. Such integrals have many applications in probability, statistics, physics and engineering.

To give meaning to the *improper integral*,

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

we examine the behaviour of $\int_a^R f(x)\,dx$ as $R \to \infty$, as illustrated below.



Definition 8.10.1. (a) Suppose that there is a real number L such that
$$\int_a^R f(x)\,dx \to L$$
as $R \to \infty$. Then f is said to be *integrable* over $[a, \infty)$ and the integral $\int_a^\infty f(x)\,dx$ is said to be *convergent*. Morever, we write
$$\int_a^\infty f(x)\,dx = L.$$

(b) Suppose that
$$\int_a^R f(x)\,dx$$
does not have a limit as $R \to \infty$. Then f we say that f *not integrable* over (a, ∞) and we say that the integral $\int_a^\infty f(x)\,dx$ is *divergent*.

One can write down a similar definition for improper integrals of the form $\int_{-\infty}^b f(x)\,dx$.

Example 8.10.2. Evaluate the following improper integrals or show that they diverge:

(a) $\displaystyle\int_0^\infty \frac{1}{x^2+1}\,dx$;

(b) $\displaystyle\int_1^\infty \frac{1}{\sqrt{x}}\,dx$;

(c) $\displaystyle\int_{-\infty}^0 e^{2x}\,dx$.

Solution. (a) If R is a real number then
$$\begin{aligned}\int_0^R \frac{1}{x^2+1}\,dx &= \Big[\tan^{-1}\Big]_0^R \\ &= \tan^{-1} R - 0 \\ &\to \frac{\pi}{2}\end{aligned}$$
as $R \to \infty$. Hence the improper integral converges and
$$\int_0^\infty \frac{1}{x^2+1}\,dx = \frac{\pi}{2}.$$

(b) If R is a real number then
$$\begin{aligned}\int_1^R x^{-1/2}\,dx &= \Big[2x^{1/2}\Big]_1^R \\ &= 2\sqrt{R} - 2 \\ &\to \infty\end{aligned}$$
as $R \to \infty$. Hence the improper integral diverges.

□

We now consider improper integrals whose interval of integration is the entire real line.

Definition 8.10.3. We say f is integrable over $(-\infty, \infty)$ if f is integrable over both $(-\infty, 0)$ and $(0, \infty)$. In this case we write
$$\int_{-\infty}^{\infty} f(x)\,dx = \int_{-\infty}^{0} f(x)\,dx + \int_{0}^{\infty} f(x)\,dx.$$
If f is not integrable on either of the intervals $(-\infty, 0)$ or $(0, \infty)$, then we say that the improper integral
$$\int_{-\infty}^{\infty} f(x)\,dx$$
diverges.

The following example shows that the improper integral $\displaystyle\int_{-\infty}^{\infty} f(x)\,dx$ may diverge even if $\displaystyle\lim_{R\to\infty}\int_{-R}^{R} f(x)\,dx$ exists.

Example 8.10.4. Consider the improper integral
$$\int_{-\infty}^{\infty} x\,dx.$$
If R is a positive number then
$$\int_{-R}^{R} x\,dx = 0$$
since the function f, given by $f(x) = x$, is odd (see Proposition 8.8.6). Hence
$$\lim_{R\to\infty}\int_{-R}^{R} x\,dx = 0.$$
However
$$\int_{0}^{R} x\,dx = \left[\tfrac{1}{2}x^2\right]_0^R = \frac{R^2}{2} \to \infty$$
as $R \to \infty$. Hence
$$\int_{0}^{\infty} x\,dx$$
diverges, and by Definition 8.10.3
$$\int_{-\infty}^{\infty} x\,dx$$
diverges also.

The following proposition determines the convergence or divergence of improper integrals whose integrand is a power of x. Such integrals will be used frequently in Section 8.11

Proposition 8.10.5 (Convergence and divergence of p-integrals)**.** *The improper integral*
$$\int_{1}^{\infty} \frac{1}{x^p}\,dx$$
is convergent if $p > 1$ and divergent if $p \leq 1$.

Proof. If $p \neq 1$ then

$$\begin{aligned}\int_1^R x^{-p}\,dx &= \left[\frac{x^{1-p}}{1-p}\right]_1^R \\ &= \frac{R^{1-p}-1}{1-p} \\ &\to \begin{cases} \frac{1}{p-1} & \text{when } 1-p<0 \\ \infty & \text{when } 1-p>0 \end{cases}\end{aligned}$$

as $R \to \infty$. Hence the integral converges when $1-p<0$ (that is, when $p>1$) and diverges when $1-p>0$ (that is, when $p<1$).

In the case when $p=1$, we have

$$\int_1^R \frac{1}{x}\,dx = \Big[\ln x\Big]_1^R = \ln R - \ln 1 \to \infty \quad \text{as } R \to \infty.$$

Hence the integral diverges when $p=1$. ☐

Remark 8.10.6. While you may be familiar with the fact that the ln function is unbounded, the proof that $\ln R \to \infty$ as $R \to \infty$ requires the use of Riemann sums and is given later in §9.2.

8.11 Comparison tests for improper integrals

(Ref: SH10 §11.7)

Some of the most important improper integrals, such as

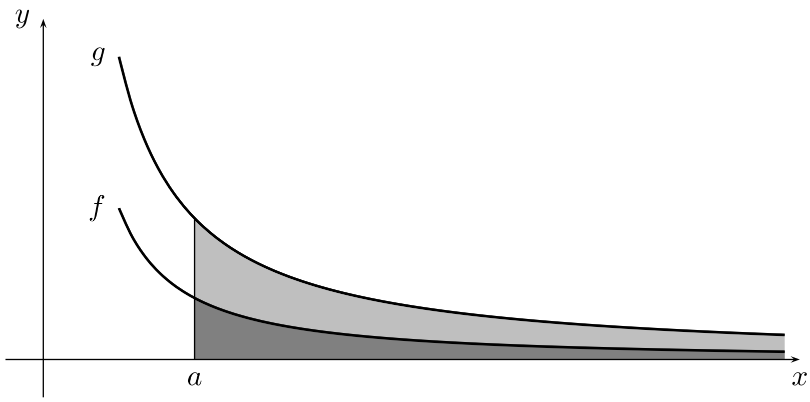
$$\int_0^\infty e^{-x^2}\,dx$$

have integrands without any antiderivative in the elementary functions. In these cases, determining convergence via a computation like

$$\int_a^\infty f(x)\,dx = \lim_{R\to\infty}\int_a^R f(x)\,dx = \lim_{R\to\infty} F(R) - F(a),$$

(where F is an antiderivative of f) is not a viable strategy. Instead, we compare the integral of f with another integral $\int_a^\infty g(x)\,dx$ whose behaviour is known. The basic idea is as follows.

Suppose that $0 \le f(x) \le g(x)$ whenever $x > a$, as illustrated below.



Obviously the area under the graph of g is greater than that under the graph of f. Hence if the area under g is finite then the area under f is also finite. If the area under f is infinite then so too is the area under g. The next theorem expresses this conclusion in terms of improper integrals.

Theorem 8.11.1 (The comparison test). *Suppose that f and g are integrable functions and that $0 \le f(x) \le g(x)$ whenever $x > a$.*

(i) If $\displaystyle\int_a^\infty g(x)\,dx$ converges then $\displaystyle\int_a^\infty f(x)\,dx$ converges.

(ii) If $\displaystyle\int_a^\infty f(x)\,dx$ diverges then $\displaystyle\int_a^\infty g(x)\,dx$ diverges.

Proof. We begin by proving (i). Suppose that f and g are integrable functions, that $0 \le f(x) \le g(x)$ for all x in $[a, \infty)$ and that $\int_0^\infty g(x)\,dx$ converges. Then

$$0 \le \int_a^R f(x)\,dx \le \int_a^R g(x)\,dx$$

whenever $R > 0$ by Proposition 8.4.1. Since the integral

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

increases as R increases and is bounded above by $\int_a^\infty g(x)\,dx$, it has a limit as $R \to \infty$. Hence $\int_a^\infty f(x)\,dx$ is also convergent.

The proof of (ii) is left as an exercise. □

When applying the comparison test, we often compare an improper integral I with

$$\int_1^\infty \frac{1}{x^p}\,dx,$$

since the behaviour of this second integral is known (see Proposition 8.10.5). The value of p is chosen by analysing the dominant terms appearing in the integrand of I. The next example illustrates this procedure.

Example 8.11.2. Determine whether or not the following improper integrals converge.

(a) $\displaystyle\int_1^\infty \frac{1}{x^3+1}\,dx$

(b) $\displaystyle\int_2^\infty \frac{x+2}{x^{3/2}-1}\,dx.$

Solution. (a) When x is large, the dominant term in the denominator of

$$\frac{1}{x^3+1}$$

is x^3. So intuitively,

$$\frac{1}{x^3+1} \approx \frac{1}{x^3}$$

for large x. Consequently we aim to compare

$$\int_1^\infty \frac{1}{x^3+1}\,dx \qquad \text{with} \qquad \int_1^\infty \frac{1}{x^3}\,dx.$$

To make this comparison rigorous, we seek an appropriate inequality. Note that $x^3+1 > x^3$ for all $x > 1$ and so

$$\frac{1}{x^3+1} < \frac{1}{x^3}$$

for all $x > 1$. Since

$$\int_1^\infty \frac{1}{x^3}\,dx$$

converges (see Proposition 8.10.5),

$$\int_1^\infty \frac{1}{x^3+1}\,dx$$

also converges by the comparison test.

□

There are many examples where the 'dominant term analysis' is straightforward but it is difficult to obtain an appropriate inequality for successful use of the comparison test. This difficulty can sometimes be overcome by multiplying one of the integrands by a 'fudge factor'. The name 'fudge factor' might sounds like we're about to 'cheat' and do something that is not mathematically rigourous, but despite the name we really will be finding an inequality that will rigourously prove convergence or divergence of our improper integral.

Example 8.11.3. Determine whether or not the the improper integral

$$\int_2^\infty \frac{1}{x^{3/2}-1}\,dx.$$

converges.

Solution. Dominant term analysis suggests that

$$\frac{1}{x^{3/2}-1} \approx \frac{1}{x^{3/2}}$$

when x is large. So we compare

$$\int_2^\infty \frac{1}{x^{3/2}-1}\,dx \qquad \text{with} \qquad \int_2^\infty \frac{1}{x^{3/2}}\,dx.$$

Now the integral to the right converges (by Proposition 8.10.5) but it is clear that

$$\frac{1}{x^{3/2}-1} > \frac{1}{x^{3/2}}$$

when $x > 2$, so we cannot immediately apply the comparison test.

However, by multiplying the integrand on the right with the 'fudge factor' 2, we see that

$$\frac{1}{x^{3/2}-1} < \frac{2}{x^{3/2}} \tag{8.22}$$

when $x > 2$. (To see why, observe that when $x > 2$,

$$x^{3/2} - 1 > x^{3/2} - \tfrac{1}{2}x^{3/2} = \tfrac{1}{2}x^{3/2}$$

whence (8.22) follows.) Since

$$\int_2^\infty \frac{2}{x^{3/2}}\,dx$$

converges,

$$\int_2^\infty \frac{1}{x^{3/2}-1}\,dx.$$

must also. □

Finding a fudge factor can be a lot of work in itself. Later we give a method for determining the convergence of this integral that does not rely on the use of fudge factors.

Dominant term analysis is not always straightforward, and at times may not even be useful. One must develop an intuition of how functions decay at infinity and of what comparisons to use.

Example 8.11.4. Determine whether or not the following improper integrals converge.

(a) $\displaystyle\int_1^\infty e^{-\sqrt{x}}\,dx$

(b) $\displaystyle\int_1^\infty \frac{\sin x + 2}{\sqrt{x}+1}\,dx.$

Solution. (a) Dominant term analysis does not apply since there is only one term appearing in the integrand. Instead, we recognise that the decay of the exponential $e^{-\sqrt{x}}$ is very rapid; it is certainly much faster than the decay of $\frac{1}{x^2}$. This can be verified by the following limit calculation:

$$\begin{aligned}
\lim_{x\to\infty} \frac{e^{-\sqrt{x}}}{\frac{1}{x^2}} &= \lim_{x\to\infty} \frac{x^2}{e^{x^{1/2}}} \\
&= \lim_{x\to\infty} \frac{2x}{\frac{1}{2}x^{-1/2}e^{x^{1/2}}} && \text{(by l'Hôpital's rule)} \\
&= \lim_{x\to\infty} \frac{4x^{3/2}}{e^{x^{1/2}}} \\
&= \lim_{x\to\infty} \frac{4.\frac{3}{2}x^{1/2}}{\frac{1}{2}x^{-1/2}e^{x^{1/2}}} && \text{(by l'Hôpital's rule)} \\
&= \lim_{x\to\infty} \frac{12x}{e^{x^{1/2}}} \\
&\vdots && \text{(two more applications of l'Hôpital's rule)} \\
&= \lim_{x\to\infty} \frac{24}{e^{x^{1/2}}} \\
&= 0.
\end{aligned}$$

Hence we conclude that

$$e^{-\sqrt{x}} < \frac{1}{x^2}$$

whenever x is sufficiently large. Since

$$\int_1^\infty \frac{1}{x^2}\,dx$$

converges,

$$\int_1^\infty e^{-\sqrt{x}}\,dx$$

also converges by the comparison test.

(b) While the numerator of

$$\frac{\sin x + 2}{\sqrt{x}+1}$$

has no dominant term, intuitively we see that, since the numerator is bounded and the dominant term of the denominator is $\sqrt{x}$, we should compare the integrand with

$$\frac{1}{\sqrt{x}}.$$

Since

$$\int_1^\infty \frac{1}{\sqrt{x}}\,dx$$

diverges, we seek an inequality of the form

$$f(x) \le \frac{\sin x + 2}{\sqrt{x}+1}$$

where $f(x)$ is $1/\sqrt{x}$ with an appropriate fudge factor.

Now, $-1 \le \sin x$, and so

$$\begin{aligned}\frac{\sin x + 2}{\sqrt{x}+1} &\ge \frac{1}{\sqrt{x}+1}\\ &\ge \frac{1}{2\sqrt{x}}\end{aligned}$$

for all $x \ge 1$. Since

$$\int_1^\infty \frac{1}{2\sqrt{x}}\,dx$$

diverges (see Proposition 8.10.5),

$$\int_1^\infty \frac{\sin x + 2}{\sqrt{x}+1}\,dx$$

also diverges by the comparison test. □

The next theorem provides an alternative approach for constructing a comparison of integrals. It is particularly useful if dominant term analysis is straightforward but one wants to avoid the use of fudge factors.

Theorem 8.11.5 (The limit form of the comparison test)**.** *Suppose that f and g are non-negative and bounded on $[a,\infty)$. If*

$$\lim_{x\to\infty} \frac{f(x)}{g(x)} = L$$

and $0 < L < \infty$ *then either*

$$\int_a^\infty f(x)\,dx \qquad \textit{and} \qquad \int_a^\infty g(x)\,dx \qquad \textit{both converge}$$

or

$$\int_a^\infty f(x)\,dx \qquad \textit{and} \qquad \int_a^\infty g(x)\,dx \qquad \textit{both diverge.}$$

The hypothesis

$$\lim_{x\to\infty} \frac{f(x)}{g(x)} = L$$

indicates that f and g exhibit similar behaviour at infinity. The proof of the theorem will be given at the end of this section.

Example 8.11.6. Determine whether or not the following improper integrals converge.

(a) $\displaystyle\int_1^\infty \frac{1}{\sqrt{1+x+x^2}}\,dx$

(b) $\displaystyle\int_2^\infty \frac{x^2+3x}{5x^4-2}\,dx.$

Solution. (a) Suppose that $f(x) = \dfrac{1}{\sqrt{1+x+x^2}}$. We need to find a suitable function g to compare f with. As x gets large, the dominant term under the square root is x^2. Therefore it seems that

$$\frac{1}{\sqrt{1+x+x^2}} \approx \frac{1}{\sqrt{x^2}} = \frac{1}{x}$$

when x is large. So define g by $g(x) = \frac{1}{x}$. The following calculation verifies that our choice of g is suitable:

$$\lim_{x\to\infty} \frac{f(x)}{g(x)} = \lim_{x\to\infty} \frac{x}{\sqrt{1+x+x^2}} = \lim_{x\to\infty} \sqrt{\frac{x^2}{1+x+x^2}} = \lim_{x\to\infty} \sqrt{\frac{1}{\frac{1}{x^2}+\frac{1}{x}+1}} = 1.$$

Since this limit is a positive real number, we can apply the limit form of the comparison test. By Proposition 8.10.5, $\int_1^\infty g(x)\,dx$ diverges. Hence $\int_1^\infty f(x)\,dx$ also diverges.

□

Remark 8.11.7. As a general rule, students will find the limit form of the comparison test easier to use than the inequality form of the comparison test when the integrand is a rational function, or is a ratio of roots of polynomials.

Proof of Theorem 8.11.1 (MATH1141 only). Suppose that f and g are non-negative and bounded on $[a,\infty)$. (It follows, for later use, that if $a < M < R$ and R is increasing then the integrals

$$\int_M^R f(x)\,dx \qquad \text{and} \qquad \int_M^R g(x)\,dx,$$

are also increasing.) Suppose also that

$$\lim_{x\to\infty} \frac{f(x)}{g(x)} = L$$

where $0 < L < \infty$. Choose a positive number ϵ such that $0 < \epsilon < L$. By Definition 2.2.1, there is a positive real number M such that

$$\left|\frac{f(x)}{g(x)} - L\right| < \epsilon$$

whenever $x > M$. Hence

$$-\epsilon < \frac{f(x)}{g(x)} - L < \epsilon \qquad \forall x > M$$

and rearranging gives

$$0 < (L-\epsilon)g(x) < f(x) < (L+\epsilon)g(x) \qquad \forall x > M.$$

By integrating with respect to x we obtain

$$0 < (L-\epsilon)\int_M^R g(x)\,dx < \int_M^R f(x)\,dx < (L+\epsilon)\int_M^R g(x)\,dx$$

whenever $R > M$. By letting R approach infinity it is seen that

- if $\int_M^\infty g(x)\,dx$ is convergent then $\int_M^R f(x)$ is also convergent (since $\int_M^R f(x)$ increases as R increases and it is bounded above by $(L+\epsilon)\int_M^\infty g(x)\,dx$);

- if $\int_M^\infty f(x)\,dx$ is convergent then $\int_M^R g(x)$ is also convergent (since $\int_M^R g(x)$ is increases as R increases and is bounded above by $\frac{1}{L-\epsilon}\int_M^\infty f(x)\,dx$);

- if $\int_M^\infty f(x)\,dx$ is divergent then $\int_M^R g(x)\,dx \to \infty$ as $R \to \infty$; and

- if $\displaystyle\int_M^\infty g(x)\,dx$ is divergent then $\int_M^R f(x)\,dx \to \infty$ as $R \to \infty$.

Hence either

$$\int_M^\infty f(x)\,dx \qquad \text{and} \qquad \int_M^\infty g(x)\,dx \qquad \text{both converge}$$

or

$$\int_M^\infty f(x)\,dx \qquad \text{and} \qquad \int_M^\infty g(x)\,dx \qquad \text{both diverge.}$$

The proof of the theorem now easily follows by first observing that

$$\int_a^R f(x)\,dx = \int_a^M f(x)\,dx + \int_M^R f(x)\,dx$$

and

$$\int_a^R g(x)\,dx = \int_a^M g(x)\,dx + \int_M^R g(x)\,dx$$

whenever $a < M < R$, and then letting R approach infinity. ☐

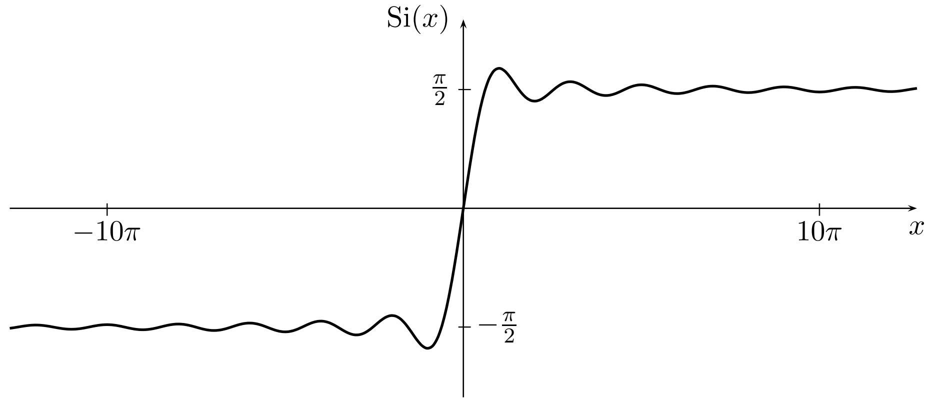


Figure 8.15: The Si function.

8.12 Functions defined by an integral

Many important functions are defined by an integral. If f is a continuous function then we may define a function F by

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

We note the following points.

- The derivative of F may be calculated by the first fundamental theorem of calculus.
- The value of $F(x)$ may be approximated by computing Riemann sums.
- Symmetries of f can help in understanding the behaviour of F (see, for example, Proposition 8.8.6).
- The behaviour of $F(x)$ as $x \to \infty$ or as $x \to -\infty$ may be ascertained by studying an improper integral of f.

We give one example. Others may be found in the problem set for Chapter 8 and at the beginning of Chapter 9.

Consider the function $f : \mathbb{R} \to \mathbb{R}$ given by

$$f(t) = \begin{cases} \frac{\sin t}{t} & \text{if } t \neq 0 \\ 1 & \text{if } t = 0. \end{cases}$$

Since $\lim\limits_{t\to 0} \dfrac{\sin t}{t} = 1$, the function f is continuous everywhere. We can now define the function $\mathrm{Si} : \mathbb{R} \to \mathbb{R}$ by

$$\mathrm{Si}(x) = \int_0^x f(t)\,dt \qquad \forall x \in \mathbb{R}.$$

The name Si comes from 'sine integral.' This function is used in electrical engineering for signal processing and in surveying for the Global Positioning System (GPS). Its graph is plotted in Figure 8.15.

Example 8.12.1. Consider the functions f and Si defined above.

(a) Find the value of $\mathrm{Si}(0)$.

(b) Show that Si is an odd function.

(c) Find and classify the critical points of Si on $(0, 3\pi)$.

(d) A function $G : \mathbb{R} \to \mathbb{R}$ is defined by the formula

$$G(x) = \int_0^{x^2} f(t)\,dt.$$

Calculate $G'(x)$.

Solution. (a) $\mathrm{Si}(0) = \displaystyle\int_0^0 f(t)\,dt = 0.$

(b) We need to show that $\mathrm{Si}(-x) = -\mathrm{Si}(x)$. Now

$$\begin{aligned}\mathrm{Si}(-x) &= \int_0^{-x} \frac{\sin t}{t}\,dt \\ &= \int_0^{x} \frac{\sin(-u)}{-u}\,(-du) \qquad \text{(using the substitution } t=-u) \\ &= -\int_0^{x} \frac{-\sin u}{-u}\,du \\ &= -\int_0^{x} \frac{\sin u}{u}\,du \\ &= -\mathrm{Si}(x).\end{aligned}$$

Hence Si is an odd function.

(c) By the first fundamental theorem of calculus,

$$\mathrm{Si}'(x) = \frac{\sin x}{x}$$

for all x in $(0, 3\pi)$. Critical points occur when $\mathrm{Si}'(x) = 0$, that is, when $x = \pi$ or when $x = 2\pi$. To classify the critical points, we calculate the second derivative:

$$\mathrm{Si}''(x) = \frac{x\cos x - \sin x}{x^2}.$$

Now $\mathrm{Si}''(\pi) = -\frac{1}{\pi}$ and $\mathrm{Si}''(2\pi) = \frac{1}{2\pi}$. Therefore Si has a local maximum point when $x = \pi$ and a local minimum point when $x = 2\pi$.

(d) Note that $G(x) = \mathrm{Si}(x^2)$, so to calculate $G'(x)$ we use the chain rule:

$$G'(x) = \frac{d}{dx}\big(\mathrm{Si}(x^2)\big) = 2x\ \mathrm{Si}'(x^2) = 2x \times f(x^2) = \frac{2x\sin(x^2)}{x^2} = \frac{2\sin(x^2)}{x}.$$

(If this calculation is too dense, then consider

$$y = G(x), \qquad u = x^2$$

so that

$$y = \int_0^u f(t)\,dt \qquad \text{and} \qquad \frac{dy}{du} = f(u)$$

by the first fundamental theorem of calculus. Then by the chain rule,

$$\frac{dy}{dx} = \frac{dy}{du}\,\frac{du}{dx} = f(u) \times 2x = 2xf(x^2) = \frac{2x\sin(x^2)}{x^2} = \frac{2\sin(x^2)}{x}.$$

Hence $G'(x) = \frac{2\sin(x^2)}{x}$.) $\square$

8.13 Maple notes

The command `int(f(x),x);` will cause Maple to attempt to find the indefinite integral $\int f(x)\,dx$, and the command `int(f(x),x=a..b);` is used to find definite integrals $\int_a^b f(x)\,dx$. If Maple is unable to find a primitive, then numerical integration can be carried out by combining the `int` and `evalf` commands.

For example,

```
> int(1/(x*log(x)^2),x);
```

$$-(\ln(x))^{-1}$$

```
> int(1/(x*log(x)^2),x=2..infinity);
```

$$(\ln(2))^{-1}$$

```
> int( exp(-x^2)*ln(x), x=0..1 );
```

$$\int_0^1 e^{-x^2}\ln(x)\,dx$$

```
> evalf(%);
```

$$-0.9059404763$$

Problems for Chapter 8

Problems 8.1 : Area and the Riemann Integral and 8.2 : Integration using Riemann sums

1. [**R**] [**V**]

   a) By taking the partition $\mathcal{P}_n = \left\{0, \dfrac{1}{n}, \dfrac{2}{n}, \ldots, 1\right\}$ of the interval [0,1], calculate the lower sum $\underline{S}_{\mathcal{P}_n}(f)$ and the upper sum $\overline{S}_{\mathcal{P}_n}(f)$ for each function f.

      i) $f(x) = 1$

      ii) $f(x) = x$

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

      [You may need $\displaystyle\sum_{k=1}^{n} k^2 = \tfrac{1}{6}n(n+1)(2n+1)$.]

      iv) $f(x) = x^3$

      [You may need $\displaystyle\sum_{k=1}^{n} k^3 = \tfrac{1}{4}n^2(n+1)^2$.]

      v) $f(x) = \begin{cases} 1 & \text{if } x \in \mathbb{Q} \\ 0 & \text{if } x \notin \mathbb{Q} \end{cases}$

   b) By taking the limit as $n \to \infty$ for each sum $\underline{S}_{\mathcal{P}_n}(f)$ and $\overline{S}_{\mathcal{P}_n}(f)$ calculated in (a), either calculate $\displaystyle\int_0^1 f(x)\,dx$, or show that f is not Riemann integrable.

2. [**R**] An electrical signal $S(t)$ has its amplitude $|S(t)|$ tested (sampled) every $\frac{1}{10}$ of a second. It is desired to estimate the energy over a period of half a second, given exactly by

$$\left(\int_0^{\frac{1}{2}} |S(t)|^2\,dt\right)^{\frac{1}{2}}.$$

   The results of the measurement are shown in the following table:

t	.1	.2	.3	.4	.5
$\lvert S(t)\rvert$	60	50	50	45	55
$e(t)$	5	3	7	4	10

   a) Using the above data for $S(t)$, set up an appropriate Riemann sum and compute an approximate value for the energy.

   b) It is known that the signal varies by an amount of at most $\pm e(t)$, as shown above, in each $\frac{1}{10}$ second period. Calculate upper and lower bounds for the energy.

3. **[X]** Consider the partition $\mathcal{P}_n$ of $[1,2]$, given by $\mathcal{P}_n = \{q^0, q^1, q^2, \ldots, q^n\}$ where $q^n = 2$. (Notice that (i) the divisions are not of equal width and (ii) $1 < q < 2$ and $q \to 1$ as $n \to \infty$.) If $f(x) = x^j$ for some positive integer j, then evaluate the integral

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

by calculating the limit $\lim_{n\to\infty} \underline{S}_{\mathcal{P}_n}(f)$ of the corresponding lower Riemann sums.

Problems 8.3 : The Riemann integral and signed area and 8.4 : Basic properties of the Riemann integral

4. **[R]** Find the area of the region bounded by the line $y = x$ and the parabola $y = x^2 - 2$.

5. **[R]** Find

   a) $\displaystyle\int_4^9 \frac{x^3 - x}{x^{3/2}}\,dx$ b) $\displaystyle\int_{-4}^2 |x|\,dx.$

6. **[H]** Find a function f which satisfies the integral equation

$$\int_0^x tf(t)\,dt = \int_x^0 (t^2+1)f(t)\,dt + x.$$

7. **[R]** Explain why $\displaystyle\int_{-1}^1 \frac{1}{x^2}\,dx = \left[-\frac{1}{x}\right]_{-1}^1 = -1 - 1 = -2$ is not valid.

8. **[H]**

   a) Suppose that f is a continuous increasing (and hence invertible) function on $[a,b]$. If $c = f(a)$, $d = f(b)$ and $a, b, c, d \geq 0$, then explain why

$$\int_c^d f^{-1}(t)\,dt = bd - ac - \int_a^b f(x)\,dx.$$

   b) Use this to find $\displaystyle\int_{1/2}^1 \sin^{-1} x\,dx.$

9. **[H]** Suppose that $U'(x) = u(x)$.

   a) Find $V'(x)$ if $V(x) = (a-x)U(x) + \displaystyle\int_0^x U(t)\,dt$ where a is a constant.

   b) Hence show that $\displaystyle\int_0^a U(x)\,dx = aU(0) + \int_0^a (a-x)u(x)\,dx.$

Problems 8.5 : The first fundamental theorem of calculus

10. [**H**] Suppose that $f(t) = \lfloor t \rfloor$ and $F(x) = \displaystyle\int_0^x f(t)\,dt$, where $\lfloor t \rfloor$ is the greatest integer less than or equal to t. Use a graph of f to sketch F on the interval $[-1, 3]$. Is F continuous? Where is F differentiable?

11. [**H**] Suppose that $f(t) = \sin(t^2)$. Sketch the graph of f on the interval $[0, 3]$. Use this to sketch the graph of F on the interval $[0, 3]$, where $F(x) = \displaystyle\int_0^x f(t)\,dt$. Indicate where F has local maxima and minima.

12. [**R**] [**V**] Find $F'(x)$ for each function $F : \mathbb{R} \to \mathbb{R}$ given below.

    a) $F(x) = \displaystyle\int_0^x \sin(t^2)\,dt$ b) $F(x) = \displaystyle\int_0^{x^3} \sin(t^2)\,dt$

    c) $F(x) = \displaystyle\int_{x^3}^1 \sin(t^2)\,dt$ d) $F(x) = \displaystyle\int_x^{x^3} \sin(t^2)\,dt$

13. [**R**] Find $\dfrac{d}{dx}\displaystyle\int_x^4 (5-4t)^5\,dt$.

Problems 8.6 : The second fundamental theorem of calculus

14. [**R**]

    a) Suppose that $f(x) = \dfrac{1}{x}$. By considering the lower Riemann sum for f with respect to the partition
    $$\left\{\frac{n}{n}, \frac{n+1}{n}, \frac{n+2}{n}, \ldots, \frac{2n}{n}\right\}$$
    of $[1, 2]$, show that
    $$\ln 2 = \lim_{n\to\infty}\left(\frac{1}{n+1} + \frac{1}{n+2} + \cdots + \frac{1}{2n}\right).$$

    b) Suppose that $f(x) = \dfrac{1}{\sqrt{1-x^2}}$.

       i) Show that f is increasing on the interval $[0, \frac{1}{2}]$.

       ii) Find the upper Riemann sum for f with respect to the partition
       $$\left\{\frac{0}{2n}, \frac{1}{2n}, \frac{2}{2n}, \frac{3}{2n}, \ldots, \frac{n}{2n}\right\}$$
       of $[0, \frac{1}{2}]$.

       iii) Hence evaluate
       $$\lim_{n\to\infty}\left(\frac{1}{\sqrt{4n^2-1^2}} + \frac{1}{\sqrt{4n^2-2^2}} + \frac{1}{\sqrt{4n^2-3^2}} + \cdots + \frac{1}{\sqrt{4n^2-n^2}}\right).$$

Problems 8.7 : Indefinite integrals

15. **[R]** Evaluate the following integrals by *inspection.*

a) $\displaystyle\int x\, e^{x^2}\, dx$ b) $\displaystyle\int \frac{\sin\sqrt{x}}{\sqrt{x}}\, dx$

c) $\displaystyle\int_0^1 2x(1+x^2)^3\, dx$ d) $\displaystyle\int_{-a}^{a} x^2\sqrt{a^3-x^3}\, dx \quad (a>0)$

e) $\displaystyle\int_0^{\pi/2} \cos^3 x\, \sin x\, dx$ f) **[H]** $\displaystyle\int_{-1}^{0} \sqrt{t^2+t^4}\, dt$

Problems 8.8 : Integration by substitution

16. **[R]** Use a substitution to evaluate the following integrals.

a) $\displaystyle\int \frac{dx}{1+\sqrt{x}}$ b) $\displaystyle\int x(5x-1)^{19}\, dx$

c) **[V]** $\displaystyle\int \frac{1-x}{(1+x)^3}\, dx$ d) **[V]** $\displaystyle\int_0^4 \frac{dx}{5+\sqrt{x}}$

17. **[X]** Use the substitution $u = t - t^{-1}$ to find $\displaystyle\int \frac{1+t^2}{1+t^4}\, dt.$

Problems 8.9 : Integration by parts

18. **[R]** Use integration by parts to evaluate the following integrals.

a) $\displaystyle\int_0^1 x\, e^{5x}\, dx$ b) $\displaystyle\int x^2 \cos x\, dx$ c) $\displaystyle\int \ln x\, dx$

d) $\displaystyle\int_0^{0.5} \sin^{-1} x\, dx$ e) $\displaystyle\int_1^{e} x^7 \ln x\, dx$ f) $\displaystyle\int_0^{\pi} x^2 \cos 2x\, dx$

g) **[V]** $\displaystyle\int e^x \cos x\, dx$ h) $\displaystyle\int \tan^{-1} x\, dx$ i) **[H]** $\displaystyle\int_0^{\pi/4} \sec^3\theta\, d\theta$

Problems 8.10 : Improper integrals

19. **[R]** **[V]** Evaluate the following improper integrals or show that they diverge.

a) $\displaystyle\int_0^{\infty} e^{-5x} dx$ b) $\displaystyle\int_{-\infty}^{1} e^{-0.01x} dx$ c) $\displaystyle\int_0^{\infty} \frac{dx}{4+x^2}$

d) $\displaystyle\int_{-\infty}^{\infty} x^3\, e^{-x^4} dx$ e) $\displaystyle\int_2^{\infty} \frac{dx}{(x-1)^{3/2}}$ f) $\displaystyle\int_e^{\infty} \frac{dx}{x\,\ln\, x}$

20. **[H]** Prove that $\displaystyle\int_0^{\infty} x^n e^{-x} dx = n!$ whenever $n = 0, 1, 2, \ldots.$

21. **[H]**

a) Find $\displaystyle\lim_{R\to\infty}\int_{-R}^{R}\frac{x}{1+x^2}\,dx$.

b) Find $\displaystyle\lim_{R\to\infty}\int_{-R}^{2R}\frac{x}{1+x^2}\,dx$.

c) Does $\displaystyle\int_{-\infty}^{\infty}\frac{x}{1+x^2}\,dx$ converge? Explain.

Problems 8.11 : Comparison tests for improper integrals

22. [**R**] Use the inequality form of the comparison test to determine whether or not the following improper integrals converge.

a) [**V**] $\displaystyle\int_1^{\infty}\frac{1}{\sqrt{1+x^4}}\,dx$ b) $\displaystyle\int_2^{\infty}\frac{1}{\sqrt[3]{x^2-x}}\,dx$ c) $\displaystyle\int_2^{\infty}\frac{1}{\ln x}\,dx$

23. [**R**] Use the limit form of the comparison test to determine whether or not the following improper integrals converge.

a) $\displaystyle\int_2^{\infty}\frac{x}{2x^3-1}\,dx$ b) $\displaystyle\int_1^{\infty}\frac{2x-1}{x^2+2}\,dx$ c) $\displaystyle\int_2^{\infty}\frac{1}{\sqrt{x^6-1}}\,dx$

24. [**R**] Use a comparison test to determine whether or not the following improper integrals converge.

a) $\displaystyle\int_1^{\infty}\frac{3x+\sin x+2}{2x^3-x+8}\,dx$ b) $\displaystyle\int_4^{\infty}\frac{4x^3-x+5}{x^4-x^2+1}\,dx$ c) [**H**] $\displaystyle\int_2^{\infty}\frac{\ln t}{t^{3/2}}\,dt$

25. [**H**] Find all real numbers s such that the improper integral

$$\int_1^{\infty}\frac{x^s}{1+x}\,dx$$

is convergent.

26. [**H**] Find all real numbers p such that $\displaystyle\int_2^{\infty}\frac{1}{x(\ln x)^p}\,dx$ converges.

27. [**H**] For which pairs of numbers (a,b) does the improper integral $\displaystyle\int_1^{\infty}\frac{x^b}{(1+x^2)^a}\,dx$ converge?

Problems 8.12 : Functions defined by an integral

28. [**R**] [**V**] Given a positive real number x, let $\pi(x)$ denote the number of primes less than or equal to x. The function Li with domain $(1,\infty)$ is given by

$$\mathrm{Li}(x)=\int_2^x\frac{1}{\ln t}\,dt$$

and is known as the 'logarithmic integral function'. It has the property that

$$\frac{\mathrm{Li}(x)}{\pi(x)} \approx 1$$

when x is sufficiently large.

a) Evaluate $\pi(10)$, $\pi(20)$ and $\pi(3.14159)$.

b) Suppose that $x > 0$. What does $\dfrac{\pi(x)}{x}$ represent?

c) Find $\frac{d}{dx}\,\mathrm{Li}(x)$ and $\mathrm{Li}(2)$.

d) By applying the mean value theorem to Li on the interval $[2, 10^6]$, find a lower bound for $\mathrm{Li}(10^6)$.

e) If x is large then

$$\frac{\pi(x)}{x} \approx \frac{\pi(x)}{x}\frac{\mathrm{Li}(x)}{\pi(x)} = \frac{\mathrm{Li}(x)}{x}.$$

Using this approximation and your answer to part (d), find an approximate lower bound for $\dfrac{\pi(10^6)}{10^6}$.

Note: There are $78,498$ primes less than one million so the actual value of $\dfrac{\pi(10^6)}{10^6}$ is 0.078498.

29. [**R**] The function erf : $\mathbb{R} \to \mathbb{R}$ is defined by the formula

$$\mathrm{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2}\,dt.$$

The function erf is an error function and can be used to calculate the probability that a measurement has an error in a given range of values.

a) Calculate $\mathrm{erf}'(x)$.

b) Explain why erf is an increasing function on $\mathbb{R}$.

c) [**H**] Show that erf is an odd function.

d) i) By calculating Riemann sums with respect to the partition $\{0, \frac{1}{4}, \frac{1}{2}, \frac{3}{4}, 1\}$, find upper and lower bounds for $\mathrm{erf}(1)$.

ii) Explain why $e^{-t^2} < e^{-t}$ whenever $t > 1$.

iii) Hence show that $\displaystyle\int_1^\infty e^{-t^2}\,dt$ converges and find an upper bound for this improper integral.

iv) Using your answers to (i) and (iii), find an upper bound for $\displaystyle\lim_{x\to\infty} \mathrm{erf}(x)$. (In fact, $\displaystyle\lim_{x\to\infty} \mathrm{erf}(x) = 1$ but this is not so easy to prove.)

e) Sketch the graph of erf.

f) Explain why erf has an inverse function erf^{-1} and sketch its graph.

Chapter 9

The logarithmic and exponential functions

The volume of a colony of bacteria on an agar plate is 1 cubic millimetre and doubles every day. Let $V(t)$ denote the volume of bacteria after t days. It is clear that

$$V(0) = 1, \qquad V(1) = 2, \qquad V(2) = 4, \qquad V(3) = 8, \qquad V(4) = 16,$$

and so on. In general

$$V(t) = 2^t$$

if t is a nonnegative integer. The corresponding points are plotted in Figure 9.1 (a).

If one wanted to determine the volume of the colony part-way through a day, it would be natural to use the formula

$$V(t) = 2^t,$$

where t is a nonnegative real number. This formula makes sense when t is a rational number. If $t = p/q$ where p and q are integers and $q > 0$, then 2^t is defined to be the unique positive qth root of 2^p. However, if t is an irrational number, this definition does not apply. So what do we mean by $2^{\sqrt{3}}$ or 2^π? We have not yet seen a definition for such numbers.

That this shortcoming in our current definition of 2^t is a serious problem can be illustrated graphically. The graph corresponding to $V(t) = 2^t$, where t is confined to the nonnegative rational numbers, has an infinite number of 'gaps' (see Figure 9.1 (b)). Clearly this is not a satisfactory state of affairs. One of the aims of this chapter is to rectify this deficiency and show that, when t is an irrational number, 2^t can be defined in such a way that the graph of $V : [0, \infty) \to \mathbb{R}$ is continuous (see in Figure 9.1 (c)).

Our plan of attack is the following.

1. In Section 9.2, we define the function ln as an integral (by the fundamental theorem of calculus).

2. In Section 9.3, we define the function exp as the inverse of ln (by the inverse function theorem).

3. In Section 9.4, we define the number b^t, where $b > 0$ and t is an irrational number, by a formula involving both ln and exp.

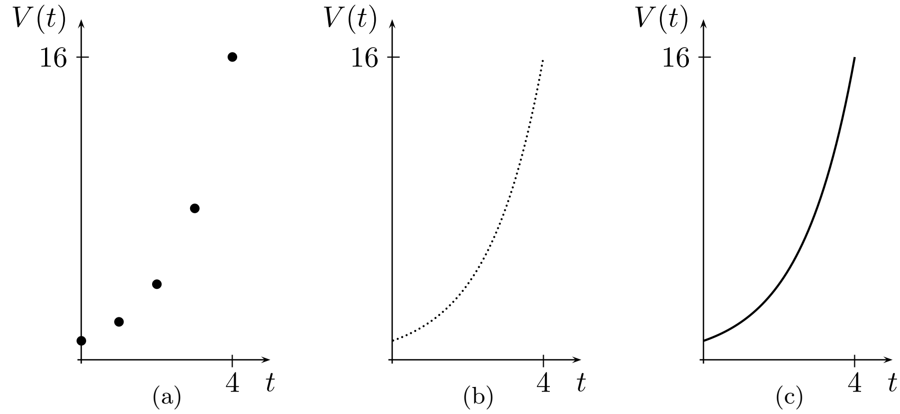


Figure 9.1: Growth in volume of bacteria on an agar plate.

We shall see that when the definition of b^t (for rational t) is combined with the definition of b^t (for irrational t), the resulting function $f : \mathbb{R} \to \mathbb{R}$, given by

$$f(t) = b^t \qquad \forall t \in \mathbb{R},$$

is continuous.

9.1 Powers and logarithms

The content of this section is revision, and is placed here primarily as reference. It summarises a few known definitions and facts about powers and logarithms.

Definition 9.1.1. Suppose that b is a positive real number. If p and q are integers and $q > 0$ then $b^{p/q}$ is defined to be the unique positive qth root of b^p.

The following is an elementary definition of the logarithm to the base b.

Definition 9.1.2. Suppose that b is a positive real number not equal to one, c is a rational number and $a = b^c$. Then $\log_b a$ is defined by the formula

$$c = \log_b a.$$

Note that, according to Definition 9.1.2, one cannot write down $\log_b a$ for *any* positive number a (even if a is rational); $\log_b a$ is only defined if $a = b^c$ for some rational number c. (This shortcoming in the domain of the function $\log_b$ will be rectified later in the chapter when Definition 9.1.2 is superseded by a superior definition.)

The elementary properties of logarithms are summarised below: if $x > 0$, $y > 0$ and $b > 0$ with $b \neq 1$ then

$$\begin{aligned} \log_b 1 &= 0, & \log_b b &= 1, \\ \log_b(xy) &= \log_b x + \log_b y, & \log_b\left(\tfrac{x}{y}\right) &= \log_b x - \log_b y, \\ \log_b(b^r) &= r, & \log_b(x^r) &= r\log_b x. \end{aligned} \tag{9.1}$$

Remark 9.1.3. In the properties listed above, it is implicitly assumed that r is a rational number, and that x and y are of the form b^c for some rational number c. One of the aims of this chapter is to remove these awkward assumptions.

9.2 The natural logarithm function

(Ref: SH10 §7.2)

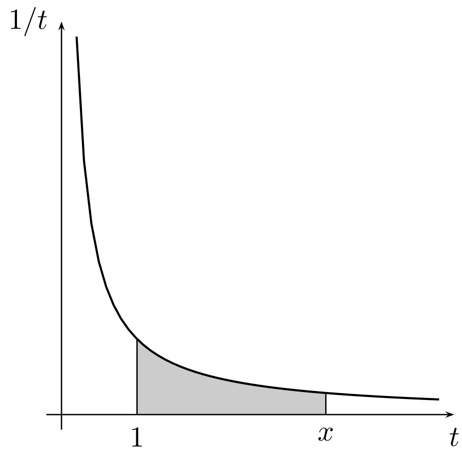
The ln function will already be familiar to you from high school. However, it was probably never defined rigorously. In order to address the problems identified in the introduction to Chapter 9, it will be necessary to start from scratch. Fortunately, by the fundamental theorem of calculus, this is not difficult.

Definition 9.2.1. The function $\ln : (0, \infty) \to \mathbb{R}$ is defined by the formula

$$\ln(x) = \int_1^x \frac{1}{t}\,dt.$$

We read ln as either 'ell en' or as 'log'.

Since the function ln is defined by an integral, $\ln x$ is simply the area of the shaded region shown below in the case when $x > 1$.



One can also immediately see, by the first fundamental theorem of calculus, that

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

This and other properties of ln are stated in the proposition below.

Proposition 9.2.2. *The function* ln *has the following properties:*

(i) ln *is differentiable on* $(0,\infty)$ *and* $\dfrac{d}{dx}\big(\ln x\big) = \dfrac{1}{x}$*;*

(ii) $\ln x > 0$ *if* $x > 1$, $\ln 1 = 0$ *and* $\ln x < 0$ *if* $0 < x < 1$*;*

(iii) $\ln x \to -\infty$ *as* $x \to 0^+$ *and* $\ln x \to \infty$ *as* $x \to \infty$*;*

(iv) $\ln(xy) = \ln(x) + \ln(y)$ *for all positive real numbers* x *and* y*;*

(v) $\ln\left(\frac{x}{y}\right) = \ln(x) - \ln(y)$ *for all positive real numbers* x *and* y*; and*

(vi) $\ln(x^r) = r\ln(x)$ *whenever* r *is a rational number and* x *is a positive real number.*

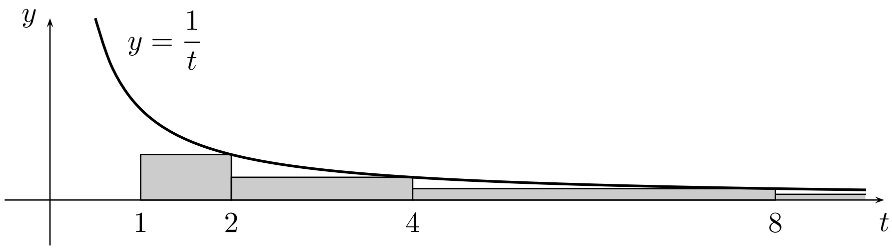
Proof. (i) Apply the first fundamental theorem of calculus (Theorem 8.5.1) to the definition of ln.

(ii) This follows from Definition 8.4.2 and the fact that $\frac{1}{t} > 0$ when $t > 0$.

(iii) We need to show that the improper integral $\int_1^\infty \frac{1}{t}\,dt$ diverges to infinity. The diagram below shows that

$$\int_1^2 \frac{dt}{t} \geq 1 \times \frac{1}{2}, \qquad \int_2^4 \frac{dt}{t} \geq 2 \times \frac{1}{4}, \qquad \int_4^8 \frac{dt}{t} \geq 4 \times \frac{1}{8}$$

and so on.



Hence

$$\begin{aligned}\int_1^{2^n} \frac{dt}{t} &\geq \underbrace{\frac{1}{2} + \frac{1}{2} + \cdots + \frac{1}{2}}_{n \text{ terms}} \\ &= \frac{n}{2} \\ &\to \infty\end{aligned}$$

as $n \to \infty$. Hence the integral diverges and therefore $\ln x \to \infty$ as $x \to \infty$.

This argument can be adapted to show that

$$\int_{2^{-n}}^1 \frac{dt}{t} \to -\infty$$

as $n \to \infty$. Hence $\ln x \to -\infty$ as $x \to 0^+$.

(iv) Suppose that y is some fixed positive real number and that $x \in (0,\infty)$. Then, by the chain rule and part (i),

$$\frac{d}{dx}\big(\ln(xy)\big) = y\ln'(xy) = y \times \frac{1}{xy} = \frac{1}{x} = \frac{d}{dx}\big(\ln(x)\big).$$

Hence

$$\ln(xy) = \ln(x) + C$$

for some constant C. When $x = 1$ we obtain

$$\ln(y) = \ln(1) + C = 0 + C = C$$

by part (ii). Hence

$$\ln(xy) = \ln(x) + C = \ln(x) + \ln(y)$$

as required.

(v) The proof uses the same technique as the proof of (vi) and is left to the reader.

(vi) This proof also uses the same technique as the proof of (iv). □

The fact that $\ln(1) = 0$ and that ln satisfies properties (iv), (v) and (vi) suggests that ln coincides with one of the logarithm functions $\log_b$ for some base b (with the advantage that $\mathrm{Dom}(\ln) = (0, \infty)$). We now aim to identify the base.

The function ln is increasing and continuous with $\mathrm{Range}(\ln) = \mathbb{R}$ (see Proposition 9.2.2 (i) and (ii)). Hence, by the intermediate value theorem, there is exactly one real number x that satisfies the equation

$$\ln(x) = 1.$$

We call this solution e (after Leonard Euler, who was probably the greatest mathematician of the eighteenth century).

Definition 9.2.3. The real number e is defined to be the unique number x that satisfies the equation

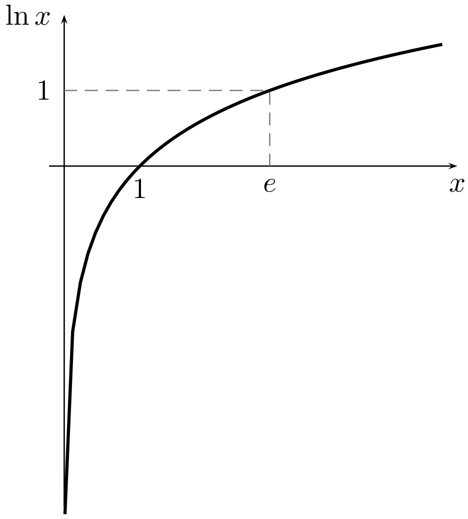
$$\int_1^x \frac{1}{t}\, dt = 1.$$

Thus $\ln(e) = 1$. One can show that e is irrational and that $e \approx 2.71828$. From Proposition 9.2.2, we see now that

$$\ln(e^r) = r$$

for every rational number r. This equation shows that ln coincides with the logarithm to the base e. Since the number e arises naturally in many contexts, ln is usually called the *natural logarithm.*

By using the properties of ln given by Propoistion 9.2.2, one can draw the graph of ln.



We note that, although $\ln(x) \to \infty$ as $x \to \infty$, we have $\frac{d}{dx}\ln(x) = \frac{1}{x} \to 0$ as $x \to \infty$. This means the graph of ln becomes flatter towards infinity and hence approaches infinity very slowly.

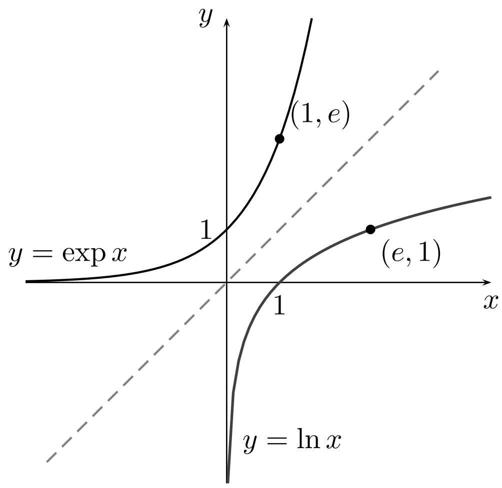
9.3 The exponential function

(Ref: SH10 §7.4)

In Section 9.2 it was shown that $\frac{d}{dx}(\ln x) > 0$ for all $x > 0$. Therefore, by the inverse function theorem (Theorem 6.4.1), ln has an inverse function.

Definition 9.3.1. The function $\exp : \mathbb{R} \to (0, \infty)$ is defined to be the inverse function of $\ln : (0, \infty) \to \mathbb{R}$.

By reflecting the graph of ln in the line $y = x$, one obtains the graph of exp.



Proposition 9.3.2. *The function* $\exp : \mathbb{R} \to (0, \infty)$ *has the following properties:*

(i) $\exp(\ln x) = x$ *for all* x *in* $(0, \infty)$ *and* $\ln(\exp x) = x$ *for all* x *in* $\mathbb{R}$*;*

(ii) $\exp(1) = e$ *and* $\exp(0) = 1$*;*

(iii) $\exp x \to \infty$ *as* $x \to \infty$ *and* $\exp x \to 0$ *as* $x \to -\infty$*.*

(iv) $\exp$ *is differentiable on* $\mathbb{R}$ *and* $\frac{d}{dx}(\exp x) = \exp x$ *for all* x *in* $\mathbb{R}$*;*

(v) $\exp(x+y) = \exp(x)\exp(y)$ *for all* x *and* y *in* $\mathbb{R}$*; and*

(vi) $\exp(rx) = (\exp x)^r$ *for every real number* x *and every rational number* r*.*

Proof. The proof of (i) follows immediately from Definition 9.3.1 and the proof of (ii) and (iii) then follow from (i) and the continuity of ln and exp.

That exp is differentiable on $\mathbb{R}$ follows from the inverse function theorem. To complete the proof of (iii), we differentiate both sides of the equation

$$\ln(\exp x) = x$$

with respect to x to obtain

$$\frac{1}{\exp x} \times \frac{d}{dx}(\exp x) = 1$$

by the chain rule. Rearranging gives $\frac{d}{dx}(\exp x) = \exp x$ as required.

If x and y are real numbers then

$$\begin{aligned}
\exp(x+y) &= \exp\Big(\ln(\exp x) + \ln(\exp y)\Big) && \text{(by part (i))} \\
&= \exp\Big(\ln(\exp x \times \exp y)\Big) && \text{(by Proposition 9.2.2 (iv))} \\
&= \exp x \times \exp y && \text{(by part (i)).}
\end{aligned}$$

This completes the proof of (iv)

To prove (v), suppose that r is a rational number and x is a real number. Then

$$\begin{aligned}
\exp(rx) &= \exp\Big(r\ln(\exp x)\Big) && \text{(by part (i))} \\
&= \exp\Big(\ln\big((\exp x)^r\big)\Big) && \text{(by Proposition 9.2.2 (v))} \\
&(\exp x)^r && \text{(by part (i)).}
\end{aligned}$$

This completes the proof of the proposition. □

Note that by Proposition 9.3.2 (ii) and (v) we have

$$\exp(r) = (\exp 1)^r = e^r \tag{9.2}$$

for every rational number r. This inspires the following definition.

Definition 9.3.3. Suppose that x is an irrational number. Then we define the real number e^x by the formula

$$e^x = \exp x.$$

Thus Definition 9.3.3 extends Definition 9.1.1, in the case when $b = e$, to include irrational powers. Moreover, since $\exp x = e^x$ and exp is a continuous function, the function $f : \mathbb{R} \to \mathbb{R}$, given by $f(x) = e^x$, is also continuous. We have thus made significant progress towards the problem outlined at the beginning of the chapter. Finally, since $\exp x = e^x$, we call exp the *exponential function.*

9.4 Exponentials and logarithms with other bases

(Ref: SH10 §7.5)

At the end of the last section, we gave a definition of e^x, where x is an irrational number. Moreover, if $f : \mathbb{R} \to \mathbb{R}$ is given by $f(x) = e^x$, then f is a continuous and differentiable function. In this section, we do the same for b^x, where b is any positive real number.

To begin, consider a rational number r and any positive real number b. Then

$$b^r = \exp\big(\ln(b^r)\big) = \exp\big(r \ln b\big) = e^{r \ln b} \tag{9.3}$$

by the definitions and properties established in the last two sections. This inspires the following definition.

Definition 9.4.1. Suppose that b is a positive real number and x is an irrational number. We define the number b^x by the formula

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

By combining equation (9.3) and Definition 9.4.1 we see that

$$b^x = e^{x \ln b} \qquad \forall x \in \mathbb{R}. \tag{9.4}$$

Since exp and ln are continuous functions, the function $f : \mathbb{R} \to \mathbb{R}$, given by

$$f(x) = b^x \qquad \forall x \in \mathbb{R},$$

is also continuous (see Proposition 3.1.2). This solves the problem articulated at the beginning of the chapter. Thus

$$2^{\sqrt{3}} = e^{\sqrt{3} \ln 2} \simeq 1.492106\ldots.$$

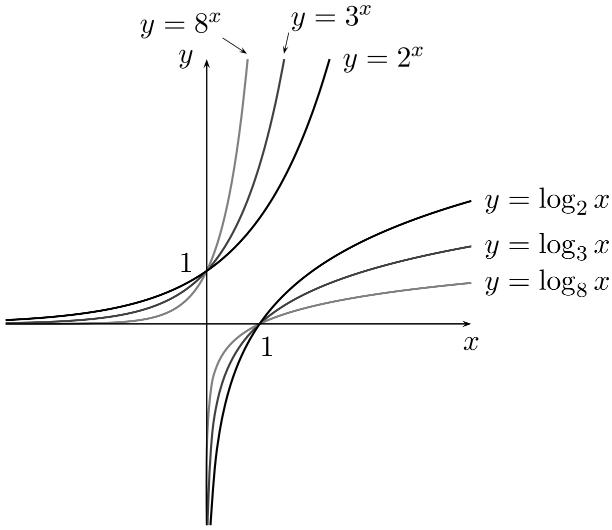
Moreover, we can discard our old definition of $\log_b$ (see Definition 9.1.2), together with its technical difficulties (see Remark 9.1.3), and replace it with a simpler and more powerful definition.

Definition 9.4.2. Suppose that b is a positive real number, $b \neq 1$. Then the function $\log_b : (0, \infty) \to \mathbb{R}$ is defined to be the inverse of the function $f : \mathbb{R} \to (0, \infty)$, where

$$f(x) = b^x$$

for all x in $\mathbb{R}$.

We leave it as an exercise to show that the function f of Definition 9.4.2 is one-to-one and hence invertible. The graphs of various exponential functions and their corresponding logarithmic functions are shown below.



The following proposition gives a formula for $\log_b$ in terms of ln. Its proof is left as an exercise.

Proposition 9.4.3. *Suppose that b is a positive real number, $b \neq 1$. Then*

$$\log_b x = \frac{\ln x}{\ln b} \qquad \forall x \in (0, \infty). \tag{9.5}$$

Using equations (9.4) and (9.5), one can easily establish familiar properties for the exponentials and logarithms to the base b, with the advantage that we no longer need the technical restrictions of Remark 9.1.3.

9.5 Integration and the ln function

We have already seen that

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

provided that $x > 0$. It follows from the chain rule that

$$\frac{d}{dx}\big(\ln(f(x))\big) = \frac{1}{f(x)} \times f'(x) = \frac{f'(x)}{f(x)},$$

provided that f is differentiable and $f(x) > 0$. Hence

$$\int \frac{f'(x)}{f(x)}\,dx = \ln(f(x)) + C.$$

It is not difficult to remove the restriction that $f(x) > 0$. To take a simple example, suppose that $x \neq 0$. Then

$$\begin{aligned}\frac{d}{dx}\big(\ln|x|\big) &= \begin{cases} \frac{d}{dx}\big(\ln x\big) & \text{if } x > 0 \\ \frac{d}{dx}\big(\ln(-x)\big) & \text{if } x < 0 \end{cases} \\ &= \begin{cases} \frac{1}{x} & \text{if } x > 0 \\ \frac{1}{-x} \times (-1) & \text{if } x < 0 \end{cases} \\ &= \frac{1}{x}.\end{aligned}$$

Thus

$$\int \frac{1}{x}\,dx = \ln|x| + C$$

provided x is restricted to an interval not containing zero. By generalising this calculation we obtain the formula

$$\boxed{\int \frac{f'(x)}{f(x)}\,dx = \ln|f(x)| + C,}$$

provided that f is never zero over the interval of integration.

Example 9.5.1. Find

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

(b) $\displaystyle\int \sec x\,dx$.

Solution. (a) By using the fact that $\tan x = \frac{\sin x}{\cos x}$ we have

$$\begin{aligned}\int \tan x\,dx &= -\int \frac{-\sin x}{\cos x}\,dx \\ &= -\ln|\cos x| + C.\end{aligned}$$

(The indefinite integral of cot can be found similarly.)

(b) This example involves a special trick:

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

(A similar trick can be used to find the indefinite integral of cosec.) □

9.6 Logarithmic differentiation

(Ref: SH10 §7.3)

Logarithms are powerful because they transform powers into products, products into sums and quotients into differences. In this section, we illustrate the use of the logarithm for differentiating functions that are defined by (combinations of) powers, products or quotients.

Example 9.6.1. Find $\frac{dy}{dx}$ in each case:

(a) $y = 10^x$

(b) $y = \left(\dfrac{(3x^2+4)(x+2)}{x^3+5x}\right)^{3/5}$

(c) $y = (\sin x)^{\cos x}$.

Solution. (a) If we take ln of both sides of the equation $y = 10^x$ then

$$\begin{aligned}\ln y &= \ln 10^x\\ &= x\ln 10.\end{aligned}$$

By differentiating with respect to x, we obtain

$$\frac{1}{y}\frac{dy}{dx} = \ln 10$$

and hence

$$\begin{aligned}\frac{dy}{dx} &= (\ln 10)\times y\\ &= (\ln 10)10^x.\end{aligned}$$

Alternatively: $\frac{d}{dx}(10^x) = \frac{d}{dx}(e^{x\ln 10}) = \ln 10 e^{x\ln 10} = (\ln 10)10^x$. $\square$

Remark 9.6.2. Recall that one can only take the logarithm of a positive number. Hence our solution to Example 9.6.1 (b) and (c) is only valid when $y > 0$.

9.7 Indeterminate forms with powers

(Ref: SH10 §11.6)

Consider the limits

$$\lim_{x\to 0^+} x^x \qquad \text{and} \qquad \lim_{x\to\infty} x^{1/x}. \tag{9.6}$$

The first limit is of the form 0^0 while the second is of the form ∞^0. Both of these forms are examples of indeterminate forms and hence we cannot say what the corresponding limits are in each case without further calculation. Since each limit involves a power, it is natural to first take the logarithm of the limit and then bring l'Hôpital's rule into play.

Example 9.7.1. Evaluate the limit

$$\lim_{x\to 0^+} x^{2x}.$$

Solution. The limit is an indeterminate form of the type 0^0. By taking the natural logarithm, we can transform the limit into an indeterminate form of the type $\frac{\infty}{\infty}$:

$$\begin{aligned}\lim_{x\to 0^+} x^{2x} &= \lim_{x\to 0^+} \exp\Big(\ln x^{2x}\Big) && \text{(since ln and exp are inverses)}\\ &= \lim_{x\to 0^+} \exp\Big(2x\ln x\Big)\\ &= \exp\Big(\lim_{x\to 0^+} 2x\ln x\Big) && \text{(since exp is continuous)}\\ &= \exp\Big(\lim_{x\to 0^+}\frac{\ln x}{1/(2x)}\Big).\end{aligned}$$

We can now apply l'Hôpital's rule to the problem. By differentiating the numerator and denominator and then simplifying we obtain

$$\begin{aligned}
\lim_{x\to 0^+} x^{2x} &= \exp\left(\lim_{x\to 0^+} \frac{1/x}{-1/(2x^2)}\right) \\
&= \exp\left(\lim_{x\to 0^+} -2x\right) \\
&= \exp(0) \\
&= 1.
\end{aligned}$$

□

Each of the limits in (9.6) can be evaluated by using the ideas in this example.

Remark 9.7.2. The limit of Example 9.7.1 is of the form 0^0 and has the value 1. The limit

$$\lim_{x\to 0^+} \left(e^{-\frac{1}{x}}\right)^x$$

is also of the form 0^0. However, its value is not 1:

$$\lim_{x\to 0^+} \left(e^{-\frac{1}{x}}\right)^x = \lim_{x\to 0^+} e^{-1} = \frac{1}{e}.$$

The fact that limits of the type 0^0 can have different values explains why 0^0 is an indeterminate form.

Problems for Chapter 9

Problems 9.1 : Powers and logarithms and 9.2 : The natural logarithm function

1. **[R]**

   a) Write down the definition of $\ln x$, where $x > 0$.

   b) Explain why $\dfrac{d}{dx}\ln x = \dfrac{1}{x}$ whenever $x > 0$.

   c) Suppose that r is a rational number and that x and y are positive real numbers.

      i) By first differentiating $\ln(xy)$ with respect to x, show that $\ln(xy) = \ln x + \ln y$.

      ii) Use the same technique to show that

$$\ln\left(\frac{x}{y}\right) = \ln x - \ln y \qquad \text{and} \qquad \ln(x^r) = r\ln x.$$

2. **[R]**

   a) Prove, using upper and lower Riemann sums and the definition of $\ln x$, that $\ln 2 < 1 < \ln 4$, and hence that $2 < e < 4$.

   b) **[H]** Use Maple and the method of part (a) to prove that $\frac{5}{2} < e < 3$. How many partition points do you need?

3. **[R]** Find the derivatives of

   a) $f(x) = \ln\sqrt{x^3+1}$ b) $g(x) = e^{|x|}$

   c) $h(x) = \ln(\ln(\ln x))$ d) $q(x) = e^{\ln(x^5+6)}$

Problems 9.3 : The exponential function and 9.5 : Integration and the ln function

4. **[R]** Find

   a) $\displaystyle\int \frac{e^{2x}}{1+e^{2x}}\,dx$ b) $\displaystyle\int \frac{e^{1/x}}{x^2}dx$ c) $\displaystyle\int 3^x\,dx$

   d) $\displaystyle\int \frac{e^{\sqrt{x}}}{8\sqrt{x}}\,dx$ e) $\displaystyle\int \frac{\ln x}{x}\,dx$ f) $\displaystyle\int \cot x\,dx.$

   (Hint for part (f): express cot *in terms of* sin *and* cos.*)*

5. **[R]** Sketch the curves

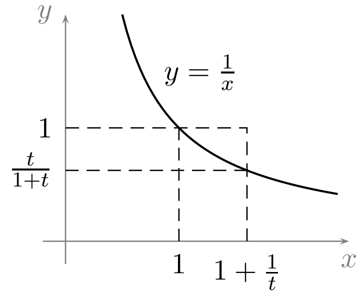
   a) $y = \ln(1+e^x)$ b) $y = \dfrac{(e^x+x)}{(e^x-x)}.$

6. [**R**]

   a) Sketch the curve $y = \dfrac{\ln x}{x}$, noting any turning points and asymptotes.

   b) By using (a) or otherwise, prove that $\pi^e < e^\pi$.

7. [**R**] [**V**]



   a) From the graph, explain why $\dfrac{1}{1+t} \le \ln\left(1 + \dfrac{1}{t}\right) \le \dfrac{1}{t}$ whenever $t \ge 0$.

   b) Deduce that $\lim\limits_{t\to\infty} \ln\left(1 + \dfrac{1}{t}\right)^t = 1$ and hence find the value of $\lim\limits_{t\to\infty}\left(1 + \dfrac{1}{t}\right)^t$.

Problems 9.6 : Logarithmic differentiation

8. [**R**] Use logarithmic differentiation to find $\dfrac{dy}{dx}$ if

   a) $y = 3^x$ b) $y = \left(\dfrac{x^3 - 3}{1 + x^2}\right)^{1/5}$

   c) $y = (\sin x)^{\sin x}$ d) $y = \sin(x^{\sin x})$.

Problems 9.7 : Indeterminate forms with powers

9. [**R**] Calculate the following limits:

   a) [**V**] $\lim\limits_{x\to\infty} \dfrac{\ln x}{x^a}$, $a > 0$ b) $\lim\limits_{x\to 0+} x^a \ln x$, $a > 0$

   c) $\lim\limits_{x\to 0+} x^x$ d) $\lim\limits_{x\to 0+} x^{2/\ln x}$

   e) $\lim\limits_{x\to\infty} x^{1/x}$ f) $\lim\limits_{x\to\infty} a^{1/x}$, $a > 0$

   g) $\lim\limits_{x\to\infty} \left(1 + \dfrac{a}{x}\right)^x$ h) $\lim\limits_{x\to\infty} x^{100} e^{-x}$

   i) $\lim\limits_{x\to\infty} p(x)\, e^{-x}$, where p is any polynomial.

10. [**H**] Prove that the functions $f : (-1, \infty) \to \mathbb{R}$ and $g : (-1, \infty) \to \mathbb{R}$, given by

$$f(x) = \ln(1 + x) - \left(x - \frac{x^2}{2}\right) \qquad \text{and} \qquad g(x) = \left(x - \frac{x^2}{2} + \frac{x^3}{3}\right) - \ln(1 + x),$$

are increasing on $(0, \infty)$. Deduce that

$$x - \frac{x^2}{2} < \ln(1 + x) < x - \frac{x^2}{2} + \frac{x^3}{3}$$

whenever $x > 0$.

Chapter 10

The hyperbolic functions

A telecommunications company wishes to suspend a cable from a sequence of telegraph poles in a cost effective manner. There are at least two competing considerations. First, it is desirable to minimise the length of cable used. Second, it is desirable to minimise the amount of tension in the cable (since an increase in tension requires an increase in cable strength). The configuration in Figure 10.1 (a) requires a short length of strong cable while that in Figure 10.1 (b) reduces the tension in the cable at the cost of increasing its length. Is there an optimal configuration for this problem?

In order to solve this problem, one needs an equation which describes the curve of a suspended cable. Such a curve is called a *catenary*, after the Latin word for 'chain.' Galileo claimed that every catenary was a parabola, but this was later disproved in 1669. Two decades later, the Swiss mathematician Jakob Bernoulli issued a public challenge to find the general equation that describes a catenary. Three mathematicians Gottfried Leibniz, Christiaan Huygens and Johann Bernoulli (Jakob's brother) independently derived the correct equation:

$$y = \frac{a}{2}\left(e^{x/a} + e^{-x/a}\right) \tag{10.1}$$

(see Figure 10.1 (c)). In this chapter we study the function described by equation (10.1) and other functions related to it. Such functions are important to mathematics, engineering, physics (especially the theory of relativity) and architecture.

We note from the outset that a large number of new formulae will be introduced in this chapter. However, students do not need to memorise all of them. Those which should be memorised are listed in Section 10.7.

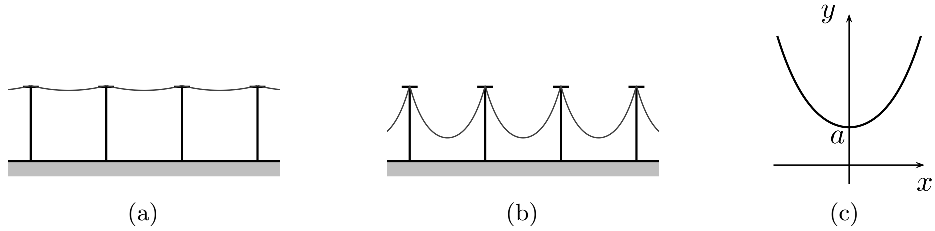


Figure 10.1: Catenaries.

10.1 Hyperbolic sine and cosine functions

(Ref: SH10 §7.8)

We mentioned in the introduction that the equation describing a freely suspended chain is given by $y = f(x)$, where

$$f(x) = \frac{a}{2}\big(e^{x/a} + e^{-x/a}\big).$$

To begin we study the simplest case: when $a = 1$.

Definition 10.1.1. The *hyperbolic cosine function* $\cosh : \mathbb{R} \to \mathbb{R}$ is defined by the formula

$$\cosh x = \tfrac{1}{2}(e^x + e^{-x}) \qquad \forall x \in \mathbb{R}.$$

The reasoning behind the name 'hyperbolic cosine' will gradually become apparent. Since cosh is a linear combination of exponential functions, it is differentiable and

$$\frac{d}{dx}\big(\cosh x\big) = \frac{1}{2}\frac{d}{dx}\big(e^x + e^{-x}\big) = \frac{1}{2}(e^x - e^{-x}).$$

The derivative of cosh is important in its own right and is given a special name.

Definition 10.1.2. The *hyperbolic sine function* $\sinh : \mathbb{R} \to \mathbb{R}$ is defined by the formula

$$\sinh x = \tfrac{1}{2}(e^x - e^{-x}) \qquad \forall x \in \mathbb{R}.$$

Some people pronounce sinh as 'shine.' One can easily show that the derivative of sinh is cosh. Thus we have

$$\boxed{\frac{d}{dx}(\sinh x) = \cosh x \qquad \text{and} \qquad \frac{d}{dx}(\cosh x) = \sinh x.} \tag{10.2}$$

We now consider the graphs of these functions.

Example 10.1.3. Sketch the graph of sinh.

Solution. We begin by identifying a few key features of the sinh function.

- Since
$$\sinh(-x) = \tfrac{1}{2}(e^{-x} - e^x) = -\tfrac{1}{2}(e^x - e^{-x}) = -\sinh x,$$
we see that sinh is an odd function. As a corollary, $\sinh(0) = 0$.

- Since
$$\frac{d}{dx}(\sinh x) = \cosh x = \tfrac{1}{2}(e^x + e^{-x}) > 0,$$
sinh is an increasing function.

- Since $\lim\limits_{x\to\infty} e^{-x} = 0$, $\sinh x$ gets arbitrarily close to $\frac{1}{2}e^x$ as $x \to \infty$.

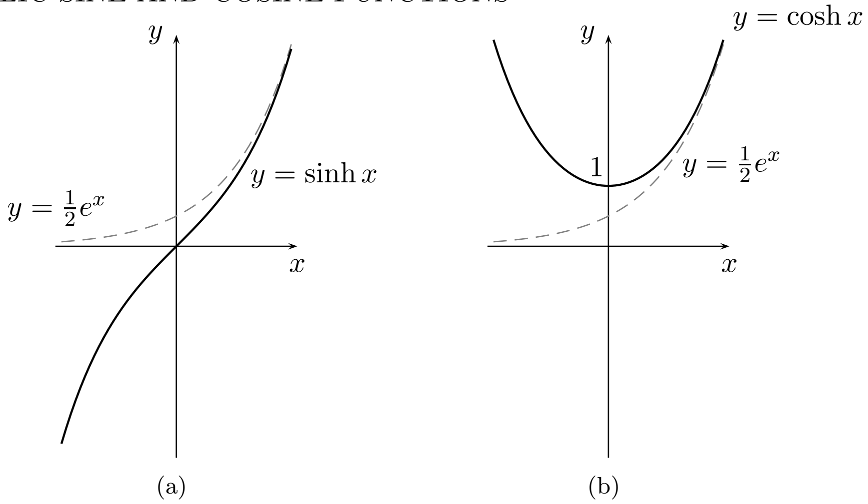


Figure 10.2: Graphs of sinh and cosh.

From these observations we obtain a sketch of the graph of sinh (see Figure 10.2). □

Example 10.1.4. Sketch the graph of cosh.

Solution. Some properties of the cosh function are listed below:

- cosh is an even function;
- $\cosh(0) = 1$;
- cosh is decreasing on $(-\infty, 0)$, stationary at 0 and increasing on $(0, \infty)$;
- $\cosh x \geq 1$ for all x in $\mathbb{R}$;
- $\cosh x$ gets arbitrarily close to $\frac{1}{2}e^x$ as $x \to \infty$.

The reader should verify each of these properties by appealing to Definition 10.1.1. From these observations, we obtain a sketch of the graph of cosh (see Figure 10.2). □

The functions cosh and sinh are related by the following important identity.

Proposition 10.1.5. *If x is a real number then*

$$\cosh^2 x - \sinh^2 x = 1.$$

Proof. If $x \in \mathbb{R}$ then

$$\begin{aligned}
\cosh^2 x - \sinh^2 x &= \left(\tfrac{1}{2}(e^x + e^{-x})\right)^2 - \left(\tfrac{1}{2}(e^x - e^{-x})\right)^2 \\
&= \frac{1}{4} \times 2e^x \times 2e^{-x} \qquad \text{difference of two squares} \\
&= 1,
\end{aligned}$$

completing the proof. □

We are now in a position to appreciate why these functions are called 'hyperbolic cosine' and 'hyperbolic sine.' First, we examine their similarity to the trigonometric cosine and sine functions. The sine and cosine functions have the following properties:

$$\frac{d}{dx}(\cos x) = -\sin x \qquad \frac{d}{dx}(\sin x) = \cos x$$

cos is an even function   sin is an odd function

$$\cos^2 x + \sin^2 x = 1.$$

The functions sinh and cosh have analogous properties (with the occasional adjustment of a negative sign):

$$\frac{d}{dx}(\cosh x) = \sinh x \qquad \frac{d}{dx}(\sinh x) = \cosh x$$

cosh is an even function   sinh is an odd function

$$\cosh^2 x - \sinh^2 x = 1.$$

The next example explains the origin of the name *hyperbolic* in 'hyperbolic cosine.'

Example 10.1.6. Sketch the curve defined by the parametric equation

$$\begin{cases} x(t) = \cosh t \\ y(t) = \sinh t, \end{cases} \tag{10.3}$$

where $t \in \mathbb{R}$.

Solution. The parameter t in equation (10.3) can be eliminated by Proposition 10.1.5 to obtain

$$\begin{aligned} 1 &= \cosh^2 t - \sinh^2 t \\ &= \big(x(t)\big)^2 - \big(y(t)\big)^2, \end{aligned}$$

or simply

$$x^2 - y^2 = 1.$$

This equation describes an hyperbola. Since $\cosh t > 0$ for every real number t, it follows that $x > 0$ and hence the curve is the branch of the hyperbola that lies in the right-half plane (see the black curve in Figure 10.3 (b)). Its asymptotes are the lines $y = x$ and $y = -x$.

(The other branch of the hyperbola is parameterised by

$$\begin{cases} x(t) = -\cosh t \\ y(t) = \sinh t \end{cases}$$

where $t \in \mathbb{R}$, and is shown in Figure 10.3 (b) in gray.) $\square$

Hence we use trigonometric sine and cosine to paramaterise a circle or ellipse, and hyperbolic sine and cosine to parameterise an hyperbola.

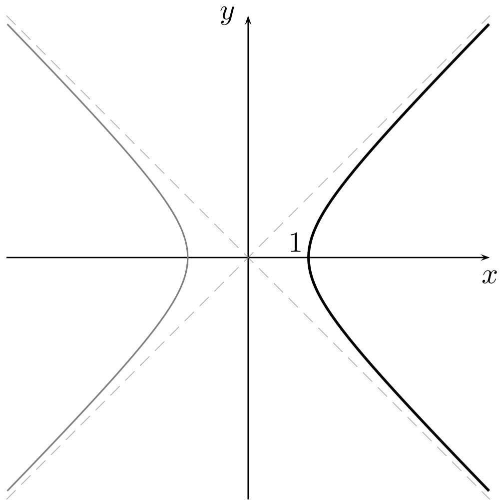


Figure 10.3: Sketch for Example 10.1.6.

10.2 Other hyperbolic functions

(Ref: SH10 §7.9)

Other hyperbolic functions are defined in analogy to the trigonometric functions. Thus we have

$$\tanh x = \frac{\sinh x}{\cosh x}, \qquad \coth x = \frac{\cosh x}{\sinh x},$$
$$\operatorname{sech} x = \frac{1}{\cosh x}, \qquad \operatorname{cosech} x = \frac{1}{\sinh x}.$$

Some people pronounce tanh as 'than.'

The derivative of each of these functions is easy to compute.

Example 10.2.1. Compute the derivative of tanh.

Solution. By the definition of tanh,

$$\begin{aligned}
\frac{d}{dx}(\tanh x) &= \frac{d}{dx}\left(\frac{\sinh x}{\cosh x}\right) \\
&= \frac{\cosh x \, \frac{d}{dx}(\sinh x) - \sinh x \, \frac{d}{dx}(\cosh x)}{\cosh^2 x} && \text{(by the quotient rule)} \\
&= \frac{\cosh^2 x - \sinh^2 x}{\cosh^2 x} && \text{(by (10.2))} \\
&= \frac{1}{\cosh^2 x} && \text{(by Proposition 10.1.5)} \\
&= \operatorname{sech}^2 x.
\end{aligned}$$

□

A complete list of derivatives will be given in Section 10.4.

Example 10.2.2. Sketch the graph of the function tanh.

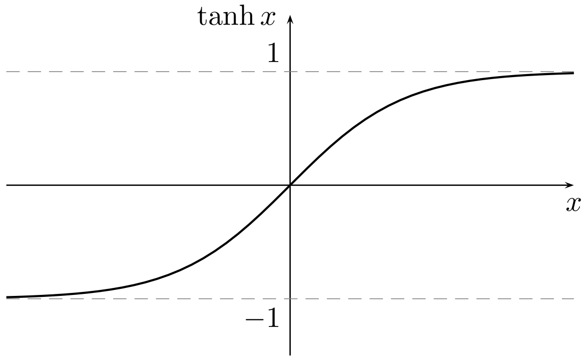


Figure 10.4: The graph of tanh.

Solution. Since $\tanh x = \frac{\sinh x}{\cosh x}$ and $\cosh x > 0$ for all x in $\mathbb{R}$, we see that $\mathrm{Dom}(\tanh) = \mathbb{R}$. Some properties of tanh are listed below.

- By using the fact that cosh is even and sinh is odd, we see that

$$\tanh(-x) = \frac{\sinh(-x)}{\cosh(-x)} = \frac{-\sinh x}{\cosh x} = -\tanh(x).$$

  Hence tanh is an odd function. It follows that $\tanh(0) = 0$.

- Since $\frac{d}{dx}(\tanh x) = \operatorname{sech}^2 x > 0$ for all x in $\mathbb{R}$, the function tanh is increasing everywhere.

- By writting $\tanh x$ in terms of exponentials we find that

$$\begin{aligned} \tanh x &= \frac{\sinh x}{\cosh x} \\ &= \frac{e^x - e^{-x}}{e^x + e^{-x}} \\ &= \frac{1 - e^{-2x}}{1 + e^{-x}} \\ &\to 1 \end{aligned}$$

  as $x \to \infty$.

- Finally, the slope of the graph of tanh at the origin is

$$\frac{d}{dx}(\tanh x)\Big|_{x=0} = \operatorname{sech}^2 0 = 1.$$

These features are shown in the following sketch in Figure 10.4. □

10.3 Hyperbolic identities

(Ref: SH10 §7.8)

It has already been observed that some identities involving cosh and sinh are *almost* analogous to those for cos and sin. This analogy can also be seen in the following larger list of identities. There are the 'difference of squares' identities

$$\begin{aligned}\cosh^2 x - \sinh^2 x &= 1,\\ 1 - \tanh^2 x &= \operatorname{sech}^2 x,\\ \coth^2 x - 1 &= \operatorname{cosech}^2 x,\end{aligned}$$

the 'sum and difference' formulae

$$\begin{aligned}\sinh(x \pm y) &= \sinh x \cosh y \pm \cosh x \sinh y,\\ \cosh(x \pm y) &= \cosh x \cosh y \pm \sinh x \sinh y,\\ \tanh(x \pm y) &= \frac{\tanh x \pm \tanh y}{1 \pm \tanh x \tanh y},\end{aligned}$$

and the 'double-angle' formulae

$$\begin{aligned}\sinh(2x) &= 2 \sinh x \cosh x,\\ \cosh(2x) &= \cosh^2 x + \sinh^2 x,\\ \tanh(2x) &= \frac{2 \tanh x}{1 + \tanh^2 x}.\end{aligned}$$

The following remark gives an easy method for remembering these identities.

Remark 10.3.1. There is a convenient mnemonic for these hyperbolic identities, provided that you know the corresponding trigonometric identities. Simply replace cos with cosh and sin with $i \sinh$. For example,

$$\begin{aligned}\cos^2 x + \sin^2 x = 1 \quad &\longrightarrow \quad \cosh^2 x + i^2 \sinh^2 x = 1\\ &\longrightarrow \quad \cosh^2 x - \sinh^2 x = 1,\end{aligned}$$

and

$$\begin{aligned}&\sin(x + y) = \sin x \cos y + \cos x \sin y\\ &\qquad \longrightarrow \quad i \sinh(x + y) = i \sinh x \cosh y + i \cosh x \sinh y\\ &\qquad \longrightarrow \quad \sinh(x + y) = \sinh x \cosh y + \cosh x \sinh y.\end{aligned}$$

Since $\tan = \frac{\sin}{\cos}$ we replace tan with $\frac{i \sinh}{\cosh} = i \tanh$. Similarly sec is replaced with sech. Thus

$$\begin{aligned}1 + \tan^2 x = \sec^2 x \quad &\longrightarrow \quad 1 + i^2 \tanh^2 x = \operatorname{sech}^2 x\\ &\longrightarrow \quad 1 - \tanh^2 x = \operatorname{sech}^2 x.\end{aligned}$$

We emphasise that Remark 10.3.1 provides a way to *remember* each formula, not to *prove* it. Proofs are constructed using techniques that are illustrated in the following example.

Example 10.3.2. (a) Prove the identity

$$\cosh(x + y) = \cosh x \cosh y + \sinh x \sinh y.$$

(b) Hence show that

$$\cosh(x - y) = \cosh x \cosh y - \sinh x \sinh y.$$

(c) By using the result of (a), prove the double-angle formula for cosh.

(d) By assuming that

$$\sinh(x+y) = \sinh x \cosh y + \cosh x \sinh y, \tag{10.4}$$

use the result of (a) to show that

$$\tanh(x+y) = \frac{\tanh x + \tanh y}{1 + \tanh x \tanh y}.$$

Solution. (a) By the definition of cosh,

$$\text{LHS} = \tfrac{1}{2}\big(e^{x+y} + e^{-(x+y)}\big)$$

while

$$\begin{aligned}
\text{RHS} &= \tfrac{1}{2}\big(e^{x} + e^{-x}\big)\tfrac{1}{2}\big(e^{y} + e^{-y}\big) + \tfrac{1}{2}\big(e^{x} - e^{-x}\big)\tfrac{1}{2}\big(e^{y} - e^{-y}\big) \\
&= \tfrac{1}{4}\big(e^{x+y} + e^{x-y} + e^{-x+y} + e^{-x-y}\big) + \tfrac{1}{4}\big(e^{x+y} - e^{x-y} - e^{-x+y} + e^{-x-y}\big) \\
&= \tfrac{1}{2}e^{x+y} + \tfrac{1}{2}e^{-x-y} \\
&= \text{LHS}.
\end{aligned}$$

This proves the identity.

□

Remark 10.3.3. Note that in Example 10.3.2, each proof started by simplifying *either* the left-hand side of the equation, *or* the right-hand side of the equation. It would be wrong to begin the proof of (a) by

$$\begin{aligned}
\cosh(x+y) &= \cosh x \cosh y + \sinh x \sinh y \\
\tfrac{1}{2}\big(e^{x+y} + e^{-(x+y)}\big) &= \tfrac{1}{2}\big(e^{x} + e^{-x}\big)\tfrac{1}{2}\big(e^{y} + e^{-y}\big) + \tfrac{1}{2}\big(e^{x} - e^{-x}\big)\tfrac{1}{2}\big(e^{y} - e^{-y}\big) \\
\cdots &= \cdots
\end{aligned}$$

Such a method is flawed and can be used to 'prove' erroneous statements such as '$1 = 0$'.

10.4 Hyperbolic derivatives and integrals

(Ref: SH10 §§7.8, 7.9)

Below is a list of derivatives for the hyperbolic functions:

$$\begin{aligned}
\frac{d}{dx}\big(\sinh x\big) &= \cosh x & \frac{d}{dx}\big(\cosh x\big) &= \sinh x \\
\frac{d}{dx}\big(\tanh x\big) &= \operatorname{sech}^2 x & \frac{d}{dx}\big(\coth x\big) &= -\operatorname{cosech}^2 x \\
\frac{d}{dx}\big(\operatorname{sech} x\big) &= -\operatorname{sech} x \tanh x & \frac{d}{dx}\big(\operatorname{cosech} x\big) &= -\operatorname{cosech} x \coth x.
\end{aligned}$$

These derivatives can be established by using the definition of each function, as in Example 10.2.1. The derivatives of sinh, cosh and tanh are the most important to remember.

Corresponding to these derivatives are the indefinite integrals

$$\int \sinh x \, dx = \cosh x + C, \qquad \int \operatorname{sech}^2 x \, dx = \tanh x + C$$

and so on.

Example 10.4.1. Calculate the derivative of the function f given by

$$f(x) = \ln(\text{sech}(2x)).$$

Solution. First we have

$$\frac{d}{dx}\big(\text{sech } 2x\big) = -2 \text{ sech}(2x)\tanh(2x).$$

Hence the chain rules gives

$$f'(x) = \frac{-2 \text{ sech}(2x)\tanh(2x)}{\text{sech}(2x)} = -2\tanh(2x).$$

$\square$

Example 10.4.2. Find the following integrals:

(a) $\displaystyle\int_0^{\frac{1}{5}\ln 3} \sinh 5x \, dx$

(b) $\displaystyle\int x \cosh x \, dx$

(c) $\displaystyle\int \frac{\text{sech}^2(\sqrt{x})}{\sqrt{x}} \, dx$

(d) $\displaystyle\int e^x \sinh x \, dx.$

Solution. (a) We have

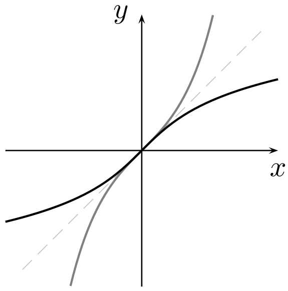
$$\begin{aligned}\int_0^{\frac{1}{5}\ln 3} \sinh 5x \, dx &= \Big[\tfrac{1}{5}\cosh 5x\Big]_0^{\frac{1}{5}\ln 3} \\ &= \tfrac{1}{10}\Big[e^{5x} + e^{-5x}\Big]_0^{\frac{1}{5}\ln 3} \\ &= \tfrac{1}{10}\Big(e^{\ln 3} + e^{-\ln 3} - e^0 - e^0\Big) \\ &= \tfrac{1}{10}\big(3 + \tfrac{1}{3} - 1 - 1\big) \\ &= \tfrac{2}{15}.\end{aligned}$$

$\square$

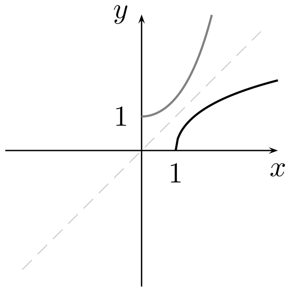
10.5 The inverse hyperbolic functions

(Ref: SH10 §7.9)

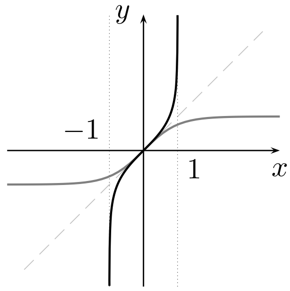
The functions $\sinh : \mathbb{R} \to \mathbb{R}$ and $\tanh : \mathbb{R} \to (-1, 1)$ are both increasing on $\mathbb{R}$ and hence have well-defined inverses. However, $\cosh : \mathbb{R} \to [1, \infty)$ is not one-to-one on $\mathbb{R}$. To obtain an inverse, one must restrict the domain of cosh. The following definition gives the conventional restriction.



(a) sinh and $\sinh^{-1}$



(b) cosh and $\cosh^{-1}$



(c) tanh and $\tanh^{-1}$

Figure 10.5: Graphs of the hyperbolic functions (in gray) and their inverses (in black).

Definition 10.5.1. The function $\cosh^{-1} : [1, \infty) \to [0, \infty)$ is defined to be the inverse function of the restricted hyperbolic cosine function

$$\cosh : [0, \infty) \to [1, \infty).$$

Graphs of the functions $\sinh^{-1}$, $\cosh^{-1}$ and $\tanh^{-1}$ are illustrated in Figure 10.5.

Expressions involving the hyperbolic functions and their inverses can be simplified using the identities of Section 10.3. It is important to note that

$$\cosh^{-1}(\cosh x) = x$$

only when $x \geq 0$.

Example 10.5.2. Simplify

(a) $\sinh\left(\cosh^{-1} \frac{4}{3}\right)$,

(b) $\cosh^{-1}(\cosh(-7))$, and

(c) $\cosh(2 \sinh^{-1} 3)$.

Solution. (a) The identity

$$\cosh^2 t - \sinh^2 t = 1$$

can be rearranged to give

$$\sinh^2 t = \cosh^2 t - 1.$$

By substituting $\cosh^{-1} \frac{4}{3}$ for t, we obtain

$$\sinh^2\left(\cosh^{-1} \tfrac{4}{3}\right) = \left(\tfrac{4}{3}\right)^2 - 1 = \tfrac{7}{9}.$$

We need to decide whether to take the positive or negative square root. Since $\cosh^{-1}$ is a nonnegative function, we have $\cosh^{-1} \frac{4}{3} > 0$ and hence

$$\sinh(\cosh^{-1} \tfrac{4}{3}) > 0.$$

Therefore

$$\sinh\left(\cosh^{-1} \tfrac{4}{3}\right) = +\tfrac{\sqrt{7}}{3}.$$

(b) We cannot write $\cosh^{-1}(\cosh(-7)) = -7$ because $\cosh^{-1}$ is only the inverse function of cosh with domain restricted to $[0, \infty)$. However, since cosh is an even function,

$$\cosh^{-1}(\cosh(-7)) = \cosh^{-1}(\cosh(7)) = 7.$$

(c) We use the double-angle formula

$$\cosh(2t) = 1 + 2\sinh^2 t$$

for cosh:

$$\begin{aligned}\cosh(2\sinh^{-1} 3) &= 1 + 2\big(\sinh(\sinh^{-1} 3)\big)^2\\ &= 1 + 2 \times 3^2\\ &= 19.\end{aligned}$$

□

Each of the functions sinh, cosh and tanh can be expressed in terms of the exponential function. It is not surprising, then, that their inverses can be expressed in terms of the natural logarithm.

Proposition 10.5.3. *The following identities hold:*

$$\begin{aligned}\sinh^{-1} x &= \ln(x + \sqrt{x^2+1}) && \forall x \in \mathbb{R},\\ \cosh^{-1} x &= \ln(x + \sqrt{x^2-1}) && \forall x \in [1, \infty),\\ \tanh^{-1} x &= \frac{1}{2}\ln\left(\frac{1+x}{1-x}\right) && \forall x \in (-1, 1).\end{aligned}$$

Proof. We prove the formula for $\sinh^{-1}$ only; the others can be proved similarly.

Suppose that $y = \sinh^{-1} x$. Then

$$\sinh y = x,$$

which means that

$$\tfrac{1}{2}(e^y - e^{-y}) = x.$$

If we multiply this equation through by $2e^y$ then we obtain

$$e^{2y} - 1 = 2xe^y,$$

which is a quadratic equation in e^y. Further rearrangement gives

$$(e^y)^2 - 2x(e^y) = 1,$$

to which we complete the square to obtain

$$(e^y - x)^2 = 1 + x^2.$$

Since $e^y > 0$, we take the positive square root. Thus

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

(Note that this may also be obtained by using the quadratic formula rather than completing the square). Taking the natural logarithm of both sides of this equation gives

$$y = \ln(x + \sqrt{x^2+1})$$

as required. □

By the inverse function theorem, the functions

$$\sinh^{-1} : \mathbb{R} \to \mathbb{R}, \qquad \cosh^{-1} : (1, \infty) \to (0, \infty) \qquad \text{and} \qquad \tanh^{-1} : (-1, 1) \to \mathbb{R}$$

are differentiable on their respective domains. A formula for each derivative is given in the next proposition.

Proposition 10.5.4. *The derivatives of* $\sinh^{-1}$, $\cosh^{-1}$ *and* $\tanh^{-1}$ *are given by*

$$\begin{aligned}
\frac{d}{dx}\big(\sinh^{-1} x\big) &= \frac{1}{\sqrt{x^2+1}}, \\
\frac{d}{dx}\big(\cosh^{-1} x\big) &= \frac{1}{\sqrt{x^2-1}}, \\
\frac{d}{dx}\big(\tanh^{-1} x\big) &= \frac{1}{1-x^2}.
\end{aligned}$$

Proof. We prove the formula for $\frac{d}{dx}(\sinh^{-1} x)$ only; the others can be proved similarly.

Suppose that $y = \sinh^{-1} x$, Then

$$\sinh y = x$$

and differentiating with respect to x gives

$$\cosh y \, \frac{dy}{dx} = 1.$$

Hence

$$\frac{dy}{dx} = \frac{1}{\cosh y} = \frac{1}{\cosh(\sinh^{-1} x)}. \tag{10.5}$$

The identity $\cosh^2 t - \sinh^2 t = 1$ gives

$$\cosh t = \sqrt{1 + \sinh^2 t}$$

(where we have taken the positive square root since $\cosh t > 0$ for all t). Hence

$$\cosh(\sinh^{-1} x) = \sqrt{1 + x^2}.$$

If we substitute this into (10.5) then we obtain

$$\frac{d}{dx}\big(\sinh^{-1} x\big) = \frac{1}{\sqrt{1+x^2}}$$

as required. $\square$

Remark 10.5.5. Students should be familiar with the proofs of Propositions 10.5.3 and 10.5.4.

10.6 Integration leading to the inverse hyperbolic functions

When propositions 10.5.3 and 10.5.4 for $\sinh^{-1}$ are restated in terms of indefinite integration we obtain

$$\int \frac{dx}{\sqrt{x^2+1}} = \sinh^{-1} x + C = \ln(x + \sqrt{x^2+1}) + C.$$

A more general version may be obtained by observing that, when $a > 0$,

$$\frac{d}{dx}\left(\sinh^{-1}\frac{x}{a}\right) = \frac{1}{a}\,\frac{1}{\sqrt{\left(\frac{x}{a}\right)^2 + 1}} = \frac{1}{\sqrt{x^2 + a^2}}.$$

In this manner, one is able to prove the formulae

$$\begin{aligned}
\int \frac{dx}{\sqrt{x^2+a^2}} &= \sinh^{-1}\frac{x}{a} + C \\
&= \ln(x + \sqrt{x^2+a^2}) + (C - \ln a), && a > 0,
\end{aligned}$$

$$\begin{aligned}
\int \frac{dx}{\sqrt{x^2-a^2}} &= \cosh^{-1}\frac{x}{a} + C \\
&= \ln(x + \sqrt{x^2-a^2}) + (C - \ln a), && x \geq a > 0,
\end{aligned}$$

$$\begin{aligned}
\int \frac{dx}{a^2-x^2} &= \begin{cases} \frac{1}{a}\tanh^{-1}\frac{x}{a} + C, & |x| < a \\ \frac{1}{a}\coth^{-1}\frac{x}{a} + C, & |x| > a > 0 \end{cases} \\
&= \frac{1}{2a}\ln\left|\frac{a+x}{a-x}\right| + C, && x^2 \neq a^2.
\end{aligned}$$

These formulae are included in the table of standard integrals that is issued at the final examination.

Example 10.6.1. Find

(a) $\displaystyle\int \frac{dx}{\sqrt{1+9x^2}}$

(b) $\displaystyle\int \frac{dx}{\sqrt{x^2-2x+10}}$.

Solution. (a) A little algebraic manipulation shows that

$$\frac{1}{\sqrt{1+9x^2}} = \frac{1}{3\sqrt{(1/3)^2 + x^2}}.$$

Hence

$$\begin{aligned}
\int \frac{dx}{\sqrt{1+9x^2}} &= \frac{1}{3}\int \frac{dx}{\sqrt{(1/3)^2+x^2}} \\
&= \frac{1}{3}\sinh^{-1}\frac{x}{1/3} + C \\
&= \frac{1}{3}\sinh^{-1}(3x) + C.
\end{aligned}$$

Alternatively, one could approach this problem by using the substitution $u = 3x$.

(b) The technique we apply to this example is known as 'completing the square':

$$x^2 - 2x + 10 = x^2 - 2x + 1 + 9 = (x-1)^2 + 3^2 = u^2 + 3^2,$$

where $u = x - 1$. Hence

$$\begin{aligned}\int \frac{dx}{\sqrt{x^2 - 2x + 10}} &= \int \frac{du}{\sqrt{u^2 + 3^2}}\\ &= \sinh^{-1} \frac{u}{3} + C\\ &= \sinh^{-1} \frac{x-1}{3} + C.\end{aligned}$$

□

10.7 A summary of important hyperbolic formulae

A large number of formulae was introduced in this chapter. Below is a list of formulae, with the more important appearing earlier in the list, that students are expected to know.

- The definitions of sinh and cosh:

$$\sinh x = \tfrac{1}{2}(e^x - e^{-x}) \qquad \cosh x = \tfrac{1}{2}(e^x + e^{-x})$$

- The derivatives of sin and cosh:

$$\frac{d}{dx}(\sinh x) = \cosh x \qquad \frac{d}{dx}(\cosh x) = \sinh x$$

- The definition of tanh:

$$\tanh x = \frac{\sinh x}{\cosh x}$$

- The hyperbolic identity

$$\cosh^2 x - \sinh^2 x = 1$$

- The derivative of $\sinh^{-1}$:

$$\frac{d}{dx}(\sinh^{-1} x) = \frac{1}{\sqrt{1+x^2}}$$

- The derivative of tanh:

$$\frac{d}{dx}(\tanh x) = \operatorname{sech}^2 x.$$

Other formulae need not be memorised, though students will do well to note the following points.

- Derivatives for the inverse hyperbolic functions can be read from the table of standard integrals.
- Identities which express the inverse hyperbolic functions in terms of the natural logarithm (see Proposition 10.5.3) can be read from the table of standard integrals.
- The definition of coth, sech and cosech are in exact analogy with the definition of cot, sec and cosec.
- The hyperbolic identities listed in Section 10.3 are easily remembered from their trigonometric counterparts (see Remark 10.3.1).

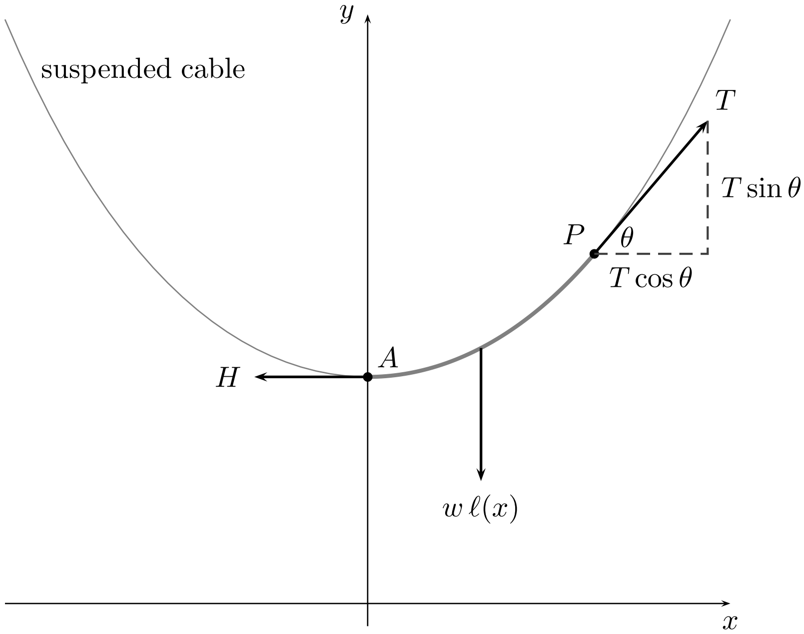


Figure 10.6: Forces acting on a suspended cable.

- It is not expected that students will memorise the derivatives of coth, sech and cosech. However, they can be easily derived from first principles provided that the derivatives of sinh and cosh are known.

A table of standard integrals will be issued at the final examination.

10.8 (Appendix):The catenary

We end this chapter showing that the cosh function describes the curve of a suspended cable.

Suppose that a cable of uniform mass is suspended from two fixed points. To describe the equation of the resulting curve in terms of x and y, we fix a coordinate system such that vertex of the curve passes through the point $A(0, y_0)$ as in Figure 10.6. (The ordinate y_0 can be any real number; its choice only fixes the coordinate system and does not alter the shape of the catenary.) Consider a point $P(x, y)$ that lies on the curve and suppose that

- T is the tension in the cable at $P(x, y)$,
- H is the tension in the cable at $A(0, y_0)$,
- $\ell(x)$ is the length of the cable from $A(0, y_0)$ to $P(x, y)$, and
- w is the weight per unit length of the cable.

(see Figure 10.6). Thus the gravitational force acting on the cable along the segment AP is given by $w\,\ell(x)$. We note that the quantities T, θ and ℓ depend on x but that w and H do not. We also make the natural assumption that the catenary is described by a differentiable function.

Theorem 10.8.1. *There is a choice of coordinate system such that the catenary in Figure 10.6 is described by the equation*

$$y = \frac{H}{w}\cosh\left(\frac{wx}{H}\right).$$

To prove the theorem, it suffices to show that if the catenary is described by $y = f(x)$, where f is a differentiable function, then

$$f(x) = \frac{H}{w}\cosh\left(\frac{wx}{H}\right)$$

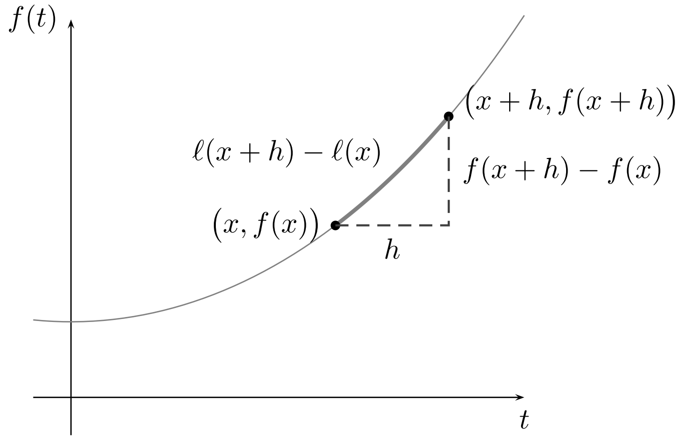
for a particular choice of y_0.

We begin with a general formula that relates a function f to its arc length function ℓ by a differential equation.

Lemma 10.8.2. *Suppose that f is a differentiable function. Then the arc length function ℓ is also differentiable and*

$$\ell'(x) = \sqrt{1 + \big(f'(x)\big)^2} \tag{10.6}$$

Sketch proof. Suppose that h is a small nonzero real number and consider the diagram below.



By Pythagoras' theorem,

$$\ell(x+h) - \ell(x) \approx \sqrt{h^2 + \big(f(x+h) - f(x)\big)^2}.$$

If we divide both sides by h then

$$\frac{\ell(x+h) - \ell(x)}{h} \approx \sqrt{1 + \left(\frac{f(x+h) - f(x)}{h}\right)^2}.$$

As $h \to 0$ the approximation gets better and we obtain

$$\ell'(x) = \sqrt{1 + \big(f'(x)\big)^2}$$

as required. ☐

The next lemma uses facts specific to the catenary.

Lemma 10.8.3. *If the equation of the catenary in Figure 10.6 is given by $y = f(x)$, then f satisfies the following conditions:*

$$f(0) = y_0, \tag{10.7}$$
$$f'(0) = 0, \tag{10.8}$$
$$f''(x) = \frac{w}{H}\sqrt{1 + \big(f'(x)\big)^2}. \tag{10.9}$$

Proof. The first two conditions are a result of our choice of coordinate system and the assumption that f is differentiable. It remains to establish the third condition.

Since the cable is stationary, the forces $w\,\ell(x)$ and $T\sin\theta$ that act vertically cancel each other out. That is,

$$T\sin\theta = w\,\ell(x).$$

Similarly, the horizontal forces H and $T\cos\theta$ also cancel each other out and so

$$T\cos\theta = H.$$

Using these equations, and the fact that $f'(x) = \tan\theta$ at the point $P(x,y)$, gives

$$f'(x) = \tan\theta = \frac{T\sin\theta}{T\cos\theta} = \frac{w\,\ell(x)}{H}.$$

So far we have

$$f'(x) = \frac{w\,\ell(x)}{H}.$$

Since w and H are constants, differentiating this equation with respect to x gives

$$f''(x) = \frac{w}{H}\,\ell'(x).$$

Therefore

$$f''(x) = \frac{w}{H}\sqrt{1 + \big(f'(x)\big)^2}.$$

by equation (10.6). □

Proof of Theorem 10.8.1. It remains to find all functions f that satisfy the conditions of Lemma 10.8.3. If we let u denote $f'(x)$ and b denote w/H, then equation (10.9) becomes

$$\frac{du}{dx} = b\sqrt{1+u^2}.$$

Hence

$$\frac{dx}{du} = \frac{1}{b\sqrt{1+u^2}}.$$

Every solution to this differential equation is of the form

$$x = \frac{1}{b}\sinh^{-1} u + C_1.$$

Hence we have

$$f'(x) = u = \sinh(bx - bC_1).$$

Enforcing condition (10.8) shows that either $b = 0$ or $C_1 = 0$. Now if $b = 0$ then $f'(x) = 0$ and f is simply a constant function (that is, the cable is pulled tight and exhibits no sag). On the other hand, if $b \neq 0$ then $C_1 = 0$ and hence

$$f'(x) = \sinh(bx).$$

Therefore

$$f(x) = \frac{\cosh(bx)}{b} + C_2$$

and enforcing the condition (10.7) yields $C_2 = \frac{1}{b} - y_0$. If we choose y_0 to be $1/b$ then $C_2 = 0$ and the function f that describes the catenary is given by

$$f(x) = \frac{\cosh(bx)}{b} = \frac{H}{w}\cosh\left(\frac{wx}{H}\right),$$

proving the theorem. □

10.9 Maple notes

Maple knows about the hyperbolic functions `sinh`, `cosh`, `tanh`, `csch`, `sech`, and `coth`, and inverse hyperbolic functions `arcsinh`, `arccosh`, `arctanh`, `arccsch`, `arcsech`, and `arccoth`. For example,

```
> int(arctanh(x),x);
```

$$x \operatorname{arctanh}(x) + \frac{1}{2} \ln\left(1 - x^2\right)$$

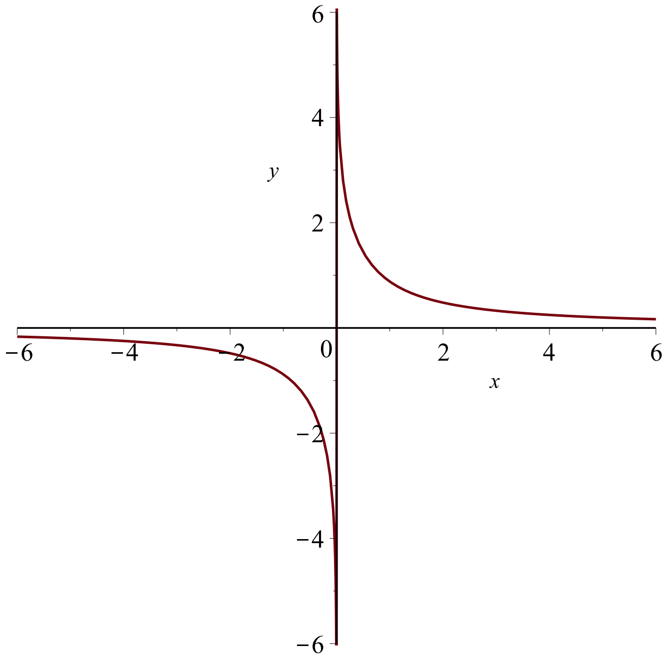
```
> # which is closely analogous to
> int(arctan(x),x);
```

$$x \arctan(x) - \frac{1}{2} \ln\left(1 + x^2\right)$$

```
> # BUT the following result does not follow the analogy between trig and hyperbolic
functions
> int(sech(x),x);
```

$$\arctan(\sinh(x))$$

```
> plot(arccsch(x),x=-6..6, y=-6..6, discont=true, numpoints=100);
```



Problems for Chapter 10

Problems 10.1 : Hyperbolic sine and cosine functions and 10.2 : Other hyperbolic functions

1. [**R**] [**V**] Define $\sinh x$ and $\cosh x$. Hence show that

   a) $\dfrac{d}{dx}(\cosh 6x) = 6\sinh 6x$;

   b) $\ln(\sinh x) < x - \ln 2$ whenever $x > 0$.

2. [**R**] By expressing the following hyperbolic functions in terms of $\sinh x$ and $\cosh x$, find the derivative of each function f given below.

   a) $f(x) = \tanh x$ b) $f(x) = \text{sech } x$ c) $f(x) = \coth x$

3. [**R**] In each case, find $f'(x)$.

   a) $f(x) = \sinh(3x^2)$ b) $f(x) = \cosh(\frac{1}{x})$ c) $f(x) = \sinh(\ln x)$

Problems 10.3 : Hyperbolic identities

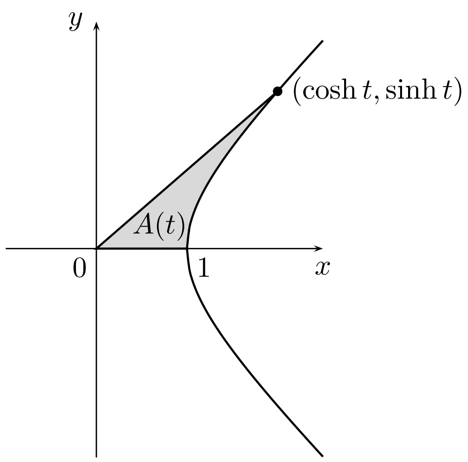
4. [**R**]

   a) Given the formula $\sinh(A+B) = \sinh A\cosh B + \cosh A\sinh B$, find a formula for $\sinh 2x$. By differentiation or otherwise, find a formula for $\cosh 2x$.

   b) [**H**] Using the results of part (a), express $\sinh 3x$ as a cubic polynomial in $\sinh x$. Hence, or otherwise, find $\displaystyle\int \sinh^3 x\, dx$.

5. [**R**] Show that $\cosh x + \sinh x = e^x$. Deduce that $(\cosh x + \sinh x)^n = \cosh nx + \sinh nx$.

6. [**H**] [**V**] Consider the hyperbola $x^2 - y^2 = 1$, where $x \geq 1$.



   a) Using the definitions of cosh and sinh, prove that, for every real number t, the point $(\cosh t, \sinh t)$ lies on the hyperbola.

b) When $t > 0$, let $A(t)$ denote the shaded region in the diagram. Explain why

$$A(t) = \frac{1}{2}\cosh t \sinh t - \int_1^{\cosh t} \sqrt{x^2 - 1}\, dx.$$

c) By first calculating $A'(t)$, prove that $A(t) = \dfrac{t}{2}$.

Problems 10.4 : Hyperbolic derivatives and integrals

7. [**R**] Evaluate the following integrals.

a) $\displaystyle\int \cosh(4x)\, dx$ b) $\displaystyle\int_0^{\frac{1}{3}\ln 2} \sinh 3x\, dx$

c) $\displaystyle\int \cosh^2 x\, dx$ d) $\displaystyle\int \frac{\sinh(\sqrt{x})}{\sqrt{x}}\, dx$

Problems 10.5 : The inverse hyperbolic functions

8. [**R**] Simplify $\cosh(\sinh^{-1}(3/4))$, $\cosh^{-1}(\cosh(-3))$ and $\sinh(\tanh^{-1}(5/13))$.

9. [**R**] Show that

a) $\dfrac{d}{dx}\big(\cosh^{-1} x\big) = \dfrac{1}{\sqrt{x^2 - 1}}$, for $x > 1$ b) $\dfrac{d}{dx}\big(\tanh^{-1} x\big) = \dfrac{1}{1 - x^2}$.

10. [**R**] Show that

a) $\cosh^{-1} x = \ln(x + \sqrt{x^2 - 1}) \qquad \forall x \in [1, \infty)$

b) [**V**] $\tanh^{-1} x = \dfrac{1}{2}\ln\left(\dfrac{1+x}{1-x}\right) \qquad \forall x \in (-1, 1)$.

11. [**R**] Find $\dfrac{dy}{dx}$ if

a) $y = \sinh^{-1}(2x)$

b) $y = \tanh^{-1}(1/x)$

c) $y = \cosh^{-1}(\sec x)$ whenever $0 < x < \pi/2$.

Problems 10.6 : Integration leading to the inverse hyperbolic functions

12. [**R**] Find

a) [**V**] $\displaystyle\int \frac{dx}{\sqrt{1 + 4x^2}}$ b) $\displaystyle\int_0^{1/2} \frac{dx}{1 - x^2}$ c) $\displaystyle\int \frac{dx}{\sqrt{x^2 + 4x + 13}}$.

13. [**X**] Sketch the function sech^{-1}. What is its maximal domain? For $y = \operatorname{sech}^{-1} x$, show that

a) $\dfrac{dy}{dx} = \dfrac{-1}{x\sqrt{1 - x^2}}$ b) $y = \ln\left(\dfrac{1 + \sqrt{1 - x^2}}{x}\right)$.

Answers to selected problems

Chapter 1

1. a) The set of integers between $-\pi$ and π.

   c) The empty set.

3. Answer for both: the interior and boundary of the triangle with vertices at $(0,0)$, $(2,0)$ and $(2,4)$.

4. a) $x < 0$ or $x > 1$ b) $1 < x < 2$ c) $x < -2$ or $x > 0$

   d) $-1 < x < 1$ e) $-2 \le x < 1$ or $x \ge 3$

5. a) $-4 < x < 2$ b) $x < -5$ or $x > 1$

   c) $-1 < x < -1/3$ d) $x > 0$

6. c) From (a) we have $x^2 + \dfrac{1}{x^2} \ge 2$ with equality if and only if $x = \pm 1$.

7. a) F b) F c) T d) T e) F

8. Hint: $(x^2 + y^2)^2 \ge 4x^2y^2$.

10. a) $-\sqrt{5} \le x \le \sqrt{5}$; $0 \le y \le \sqrt{5}$

    b) $x \le -\sqrt{5}$ or $x \ge \sqrt{5}$; $y \ge 0$

    c) $x \ne 8$; $y \ne 0$

    d) $[1, \infty)$; $[0, \infty)$

    e) $(1, \infty)$; $(0, \infty)$

    f) $\{x \in \mathbb{R} : 2n\pi \le x \le (2n+1)\pi;\ n \in \mathbb{Z}\}$; $[0, 1]$

    g) The union of the intervals $[-\frac{7\pi}{6} + 2k\pi, \frac{\pi}{6} + 2k\pi]$ where $k \in \mathbb{Z}$; $0 \le y \le \sqrt{3}$

    h) $\{x \in \mathbb{R} : x \ne (2n+1)\pi/2,\ n$ an integer$\}$; $[1, \infty)$

    i) $\mathbb{R}$; $[-1, \infty)$

11. a) 22 b) $x^2 + 10x + 22$ c) 6 d) $x^2 + 2$

12. a) $x - 1 + 1/\sqrt{x-1}$ b) $\sqrt{x-1}$ c) $(x-1)^{3/2}$ d) $(1/\sqrt{x-1}) - 1$

16. $[4, 13]$

17. $x = 1, 7$

18. a) If $p(x) = a_0 + a_1x + \cdots + a_nx^n$ then $p(q(x)) = a_0 + a_1q(x) + a_2(q(x))^2 + \cdots + a_n(q(x))^n$. Products and sums of polynomials are again polynomials.

    b) Yes.

Chapter 2

1. a) 1 b) 2 c) 0
   d) Doesn't exist ($\to \infty$). e) 5 f) Doesn't exist.

2. a) 0 b) 0

4. b) 0

5. a) 1 b) $M = 10$ (best possible) c) $M = 1/\sqrt{\epsilon}$ will do.

6. a) 4 b) 0 c) 0 d) 0 e) 0

7. a) Not necessarily, as the information given indicates only that the inequality holds for a subset of (ϵ^{-1}, ∞).

   b) Yes. In fact one can prove that $\lim\limits_{x\to\infty} g(x) = 5$ from the definition of the limit by taking M to be $\frac{1}{\epsilon}$.

8. a) 50 metres per second b) $5 \ln 50 \approx 19.56$ seconds after leaving the plane.

9. a) Yes. If limit of $f(x)$ as $x \to \infty$ does not exist and $f(x) \neq 0$, then $\lim\limits_{x\to\infty} (f(x) - f(x)) = 0$ and $\lim\limits_{x\to\infty} (f(x)/f(x)) = 1$.

   b) Yes, since $g(x) = (f(x) + g(x)) - f(x)$.

   c) No, as in (b).

   d) No. For example if $f(x) = 0$ for all x and $\lim\limits_{x\to\infty} g(x)$ does not exist, we have $\lim\limits_{x\to\infty} (f(x)g(x)) = 0$.

10. a) 10 b) 4 c) 3 d) $-1/9$

11. a) -1 b) 1 c) No

12. a) Doesn't exist. b) Doesn't exist. c) Doesn't exist. d) Doesn't exist.

13. a) 0 b) 0

14. a) $|CB| = \theta$, $|CA| = \sin\theta$, $|DB| = \tan\theta$.

15. Neither the left-hand nor right-hand limits exist due to wild oscillatory behaviour.

Chapter 3

1. b) Yes

2. a) Continuous everywhere. b) Continuous everywhere except at $\pi/2$.

3. $k = 8$

5. Use the intermediate value theorem.

9. a) Yes b) Yes c) No d) Yes

Chapter 4

2. a) $5(4x^3 + 21x^6)$ b) $(4x^3-2)(4x^2+2x+4)+(x^4-2x)(8x+2)$
   c) $(16y - y^4)/(y^3 + 8)^2$ d) $(2x^2 - 4)/(x^2 - 4)^{1/2}$
   e) $-4/(t^2 - 4)^{3/2}$ f) $3\cos 3y + 12\cos 2y \sin 2y$
   g) $(4x^3 - x^4)e^{-x}$ h) $x\ln(x^3 + 1) + 3x^2(x^2 + 1)/2(x^3 + 1)$
   i) $\sec^2 x$ j) $-\tan x$

3. a) 0 b) 0 c) $f'(0) = 0$

4. a) i) $x \neq 0$ ii) all x b) i) all x ii) all x
   c) i) $x \neq -2$ ii) $x \neq -2$

7. $2pf'(a)$

8. a) $x + 17\pi + \cos 2x$ b) $1 - 2\sin 2x$ c) $2-x^2+\cos 2(2-x^2)$
   d) $1 - 2\sin 2(2 - x^2)$ e) $-2x(1 - 2\sin 2(2 - x^2))$

9. a) $\dfrac{dy}{dx} = \dfrac{3x^2 - y}{x - 3y^2}$ b) $\dfrac{dy}{dx} = (y - 4x\sqrt{xy})/(4y\sqrt{xy} - x)$

11. $y = 2$

12. a) (i) $b = 0$ (ii) $a = 1$, $b = 0$
    b) (i) $b = 1$ (ii) $a = 2$, $b = 1$.

13. $a = 1$, $b = 0$

14. a) $f(8.01) \approx f(8) = 2$
    b) i) $y = (x - 8)/12 + 2$
       ii) $f(8.01) \approx (8.01 - 8)/12 + 2 = 2 + \frac{1}{1200}$
    c) The approximation in (b) is much better.

15. $\sqrt{3}\,ar/2$

16. $7/8\ \mathrm{m}^2/\mathrm{s}$

17. a) $\frac{1}{8\pi}$ cm/s
    b) $\frac{32000\pi}{81}$cm^3

18. a) $\frac{dh}{dt} = \frac{2}{125\pi}$ mm/s when $h = 50\,$mm.

Chapter 5

1. $\sqrt{\frac{7}{3}}$ b) $\frac{1}{2}$

5. c) 0

7. a) By the Mean Value Theorem, for some c with $16 < c < 17$, $\sqrt{17} - \sqrt{16} = \frac{1}{2\sqrt{c}} < \frac{1}{2\sqrt{16}} = 0.125$.
   b) 0.008
   c) 2×10^{-6}.

8. -1, 1 and 4 are stationary points; 4 is a local minimum point;-1 is a local maximum point.

9. No

10. a) $11, -61$ b) $3, -253$ c) $27/256, -750$
    d) $250, -54$ e) $2, 0$

11. $(12/13, 18/13)$

12. $p_n'(x) = p_{n-1}(x)$, and if $p_{n-1}(x) = 0$ then $p_n(x) = x^n/n!$. These hints are all you need!

13. a) $(400)/(4 + \pi)$, $100\pi/(4 + \pi)$ b) $0, 100$

14. The greatest distance is $a + 2$; the least distance is $\begin{cases} \sqrt{1 - a^2/3} & \text{if } 0 \le a \le 3/2 \\ |a - 2| & \text{if } a > 3/2. \end{cases}$

15. $a = \frac{\pi}{2}, x = \frac{3\pi}{4}, \frac{7\pi}{4}$; $a = \frac{3\pi}{2}, x = \frac{\pi}{4}, \frac{5\pi}{4}$. The Maple commands
```
with(plots):
animate(plot,[cos(a) + 2*cos(2*x) + cos(4*x-a), x=0..2*Pi],a=0..2*Pi);
```
should confirm your answers.

17. Three real zeros

18. a) $f(t) = -\cos t + t^2/2 + 3$ b) No

19. a) 0 b) 8/3

20. a) $\frac{1}{3}$ b) $\frac{m}{n}$ c) -1 d) $-\frac{1}{2}$ e) $\frac{1}{4}$ f) $\frac{1}{3}$

21. a) $\to 0$ b) $\to \infty$ c) $\to 0$
    d) $\to 1$ e) $\to 1$ f) $\to \frac{3}{2}$

22. Combine the two fractions and apply l'Hôpital twice only. You will need to simplify the quotient obtained after the first application of l'Hôpital. Maple can confirm your answer.

23. $(a,b) = (-\sqrt{2}, \sqrt{2})$ or $(\sqrt{2}, -\sqrt{2})$

26. a) $-1/2$
    b) $a = -1/2$, $b = 1$

27. c) $a = b = 0$

Chapter 6

2. a) $f^{-1}(x) = \dfrac{1}{3}(x-1)$
   b) $g^{-1}(x) = -\sqrt{x-1}$, $\quad \mathrm{Dom}(g^{-1}) = [1, \infty)$,
   $\mathrm{Range}(g^{-1}) = (-\infty, 0]$, $\quad (g^{-1})'(x) = \frac{-1}{2\sqrt{x-1}}$

4. b) $1/3$

5. b) The restriction of f to $(-\infty, -1]$ has an inverse with domain $(-\infty, 3]$, the restriction of f to $[-1, 1]$ has an inverse with domain $[-1, 3]$, and the restriction of f to $[1, \infty)$ has an inverse with domain $[-1, \infty)$.

6. a) No b) Yes

7. a) The graph is symmetric about $x = -\frac{1}{2}$, which surely gives a local maximum of $f(x)$. There will be four (maximal) intervals where f will have an inverse. Try this exercise on `Maple`. The commands `plot, diff` and `solve` should suffice.
   b) f is one-to-one; $f^{-1}(x) = x^{1/17} - 1$ is not differentiable when $x = 0$.
   c) I can be one of four intervals.

8. a) $\pi/3$ b) $2/5$ c) $-\pi/3$ d) $\pi/3$
   e) $4/5$ f) $3/\sqrt{34}$ g) $\pi/3$ h) $\pi - x$

11. a) $-2/\sqrt{1-4x^2}$ b) $1/(2\sqrt{x-x^2}\,)$ c) $2/(4x^2 - 12x + 10)$

12. Differentiate; $-1 \le x \le 1$; $\pi/2$.

13. b) $f(x) = \pi/2$ when $x > 0$ and $f(x) = -\pi/2$ when $x < 0$.

14. b) The derivative of the inverse is $-1/x\sqrt{x^2-1}$ when $x > 1$.

16. $a = \pi/2$, $b = 0$

17. a) 24π km/min b) $104\pi/3$ km/min

18. $\sqrt{48}$ metres

Chapter 7

1. $[-1,5]$, $[0,3]$, upper half of circle.

2. a) period $2\pi/3$, odd b) period 3π, neither
   c) not periodic, even d) period $\pi/3$, odd
   e) period π, even f) 2π

3. odd, even, neither, odd, odd, even.

4. The asymptotes are

   a) $x = 3$, $y = x + 2$ b) $x = -1$, $y = x - 1$ c) $x = -3$, $x = 2$, $y = x - 1$.

6. a) $x \geq 3$, $-\frac{1}{3} < x \leq \frac{1}{3}$ b) $x = -\frac{1}{3}, x = -3, y = 1$ c) $(1, -\frac{1}{4})$, $(-1, -4)$ d) Domain: $x \neq 3, -\frac{1}{3}$, Range: $(-\infty, -4], [-\frac{1}{4}, \infty)$.

7. a) $\dfrac{x^2}{16} + \dfrac{y^2}{25} = 1$, ellipse b) $\dfrac{x^2}{9} - \dfrac{y^2}{4} = 1$, hyperbola
   c) $y = x^{2/3}$ d) spiral

8. a) ii) $(2,0)$ iii) -1
   b) ii) $(5,0)$, $(-1,0)$ iii) $4t^3/3$
   c) ii) $(1,0)$, $(-1,0)$ iii) $-\cot t$

9. a) $3x - 27y + 52 = 0$ b) $\frac{1}{9}$

10. a) $y = 3x^{\frac{2}{3}}$.

11. b) Hint: the length of one particular arc of the larger circle equals the length of one arc on the smaller circle.
    d) $x^{2/3} + y^{2/3} = 1$

12. a) $\mathbf{p}(t) = \mathbf{a} + t(\mathbf{b} - \mathbf{a}), \mathbf{p}(0) = \mathbf{a}, \mathbf{p}(1) = \mathbf{b}$ b) $y = 4 - x, \mathbf{q}(1/2)$ is the coodinate vector for the midpoint of B and C c) $p_0(t) = (1-t)^2, p_1(t) = 2t(1-t), p_2(t) = t^2$

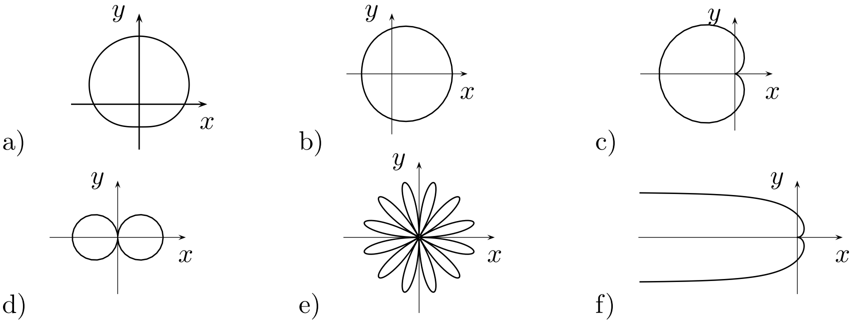
13. a) $(3,0)$ b) $(-3\sqrt{3}, -3)$ c) $(\sqrt{2}, -\sqrt{2})$

14. a) $(3, \pi)$ b) $(\sqrt{2}, -3\pi/4)$ c) $(4, 2\pi/3)$
    d) $(1, \pi/2)$ e) $(4, 5\pi/6)$ f) $(4, -5\pi/6)$

15. a) Circle, centre (0,0), radius 4
    b) A ray in the second quadrant
    c) A spiral of Archimedes

16. a) Circle, centre (0,3), radius 3
    b) Circle, centre (1,0), radius 1

17. The following sketches are a guide to shape only.



19. $\dfrac{(x-2)^2}{9} + \dfrac{y^2}{5} = 1$

Chapter 8

1. a) i) $\overline{S}_{\mathcal{P}_n}(f) = \underline{S}_{\mathcal{P}_n}(f) = 1$
   ii) $\underline{S}_{\mathcal{P}_n}(f) = \frac{1}{2}\left(1 - \frac{1}{n}\right)$, $\quad \overline{S}_{\mathcal{P}_n}(f) = \frac{1}{2}\left(1 + \frac{1}{n}\right)$
   iii) $\underline{S}_{\mathcal{P}_n}(f) = \frac{1}{6}\left(1 - \frac{1}{n}\right)\left(2 - \frac{1}{n}\right)$, $\quad \overline{S}_{\mathcal{P}_n}(f) = \frac{1}{6}\left(1 + \frac{1}{n}\right)\left(2 + \frac{1}{n}\right)$
   v) $\overline{S}_{\mathcal{P}_n}(f) = 1$, $\quad \underline{S}_{\mathcal{P}_n}(f) = 0$

   b) i) 1 ii) $\frac{1}{2}$ iii) $\frac{1}{3}$ iv) $\frac{1}{4}$ (v) Not Riemann integrable

2. a) $\sqrt{1365} = 36.95$
   b) $\sqrt{1690.9} = 41.12$ and the lower bound is $\sqrt{1078.9} = 32.85$

4. 4.5

5. a) 82.4 b) 10

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

7. $\frac{1}{x}$ is not differentiable on all of $[-1, 1]$ so the FTC doesn't apply.

8. a) Draw a picture! b) $5\pi/12 - \sqrt{3}/2$

10. F is continuous everywhere, but not differentiable at the integers.

12. a) $\sin x^2$ b) $3x^2 \sin x^6$ c) $-3x^2 \sin x^6$ d) $3x^2 \sin x^6 - \sin x^2$

13. $-(5 - 4x)^5$

14. biii) $\frac{\pi}{6}$.

15. a) $\frac{1}{2}\, e^{x^2} + C$ b) $-2\cos\sqrt{x} + C$ c) $15/4$
    d) $4\sqrt{2}\, a^{9/2}/9$ e) $1/4$ f) $(2\sqrt{2} - 1)/3$

16. a) $2\sqrt{x} - 2\ln(1 + \sqrt{x}) + C$ b) $\frac{1}{25}\left(\frac{1}{21}(5x-1)^{21} + \frac{1}{20}(5x-1)^{20}\right) + C$
    c) $x/(x+1)^2 + C$ d) $4 - 10\,\ln(7/5)$

17. $\dfrac{1}{\sqrt{2}}\tan^{-1}\left(\dfrac{t^2-1}{\sqrt{2}\,t}\right)$ for $t \neq 0$

18. a) $\frac{4e^5+1}{25}$ b) $x^2\sin x + 2x\,\cos x - 2\,\sin x + C$
c) $x(\ln(x)-1)+C$ d) $\frac{\pi}{12}+\frac{\sqrt{3}}{2}-1$
e) $\frac{7e^8+1}{64}$ f) $\frac{\pi}{2}$
g) $\frac{e^x}{2}(\cos x + \sin x)$ h) $x\tan^{-1}x - \ln\sqrt{1+x^2}+C$
i) $\frac{\sqrt{2}}{2}+\frac{1}{2}\ln(1+\sqrt{2})$

19. a) $1/5$ b) diverges c) $\pi/4$
d) 0 e) 2 f) diverges

21. a) 0 b) $\ln 2$ c) No

22. a) convergent b) divergent c) divergent

23. a) convergent b) divergent c) convergent

24. a) convergent b) divergent c) convergent

25. $s < 0$

26. $p > 1$

27. The integral converges whenever $2a - b > 1$.

28. a) 4, 8, 2
c) $\mathrm{Li}'(x) = \frac{1}{\ln x} > 0$ so Li is an increasing function; $\mathrm{Li}(2) = 0$.
d) $\mathrm{Li}(10^6) \geq \frac{10^6-2}{6\ln 10}$.
e) $\frac{\pi(10^6)}{x} \gtrapprox 0.07238$.

29. a) $\frac{2}{\sqrt{\pi}}e^{-x^2}$
d) (i) $0.749 < \mathrm{erf}(1) < 0.928$ (iii) $1/e$ (iv) 1.344

Chapter 9

2. a) A partition into 7 equal parts will suffice

3. a) $3x^2/2(x^3+1)$ b) e^x for $x > 0$, $-e^{-x}$ for $x < 0$
c) $\frac{1}{(\ln(\ln\,x))(\ln\,x)x}$ d) $5x^4$ (where $x > -6^{1/5}$)

4. a) $\frac{1}{2}\ln(1+e^{2x})$ b) $-e^{1/x}$
c) $3^x/\ln 3$ d) $\frac{e^{\sqrt{x}}}{4}$
e) $\frac{(\ln x)^2}{2}$ f) $\ln|\sin x|$

7. b) e

8. a) $3^x \ln 3$ b) $\left(\dfrac{x^3-3}{x^2+1}\right)^{1/5}\left(\dfrac{3x^2}{5(x^3-3)}-\dfrac{2x}{5(1+x^2)}\right)$

   c) $(\sin x)^{\sin x}\cos x\,(1+\ln(\sin x))$ d) $\cos(x^{\sin x})\,x^{\sin x}\left(\cos x\ln x+\dfrac{\sin x}{x}\right)$

9. a) 0 b) 0 c) 1 d) e^2 e) 1

   f) 1 g) e^a h) 0 i) 0

Chapter 10

2. a) $\operatorname{sech}^2 x$ b) $-\operatorname{sech}x\tanh x$ c) $-\operatorname{cosech}^2 x$

3. a) $6x\cosh(3x^2)$ b) $\frac{-\sinh(1/x)}{x^2}$ c) $\frac{1}{2}+\frac{1}{2x^2}$

4. a) $\sinh 2x = 2\cosh x\sinh x\,;$ $\cosh 2x = \cosh^2 x+\sinh^2 x$

   b) $\frac{1}{4}(\frac{1}{3}\cosh 3x - 3\cosh x)$ or $\frac{1}{3}\cosh^3 x - \cosh x$

7. a) $\frac{\sinh 4x}{4}$ b) $\frac{1}{12}$ c) $(2x+\sinh 2x)/4$ d) $2\cosh\sqrt{x}$

8. 5/4, 3, 5/12

11. a) $2/\sqrt{1+4x^2}$ b) $\frac{1}{1-x^2}$ for $|x|>1$ c) $\sec x$

12. a) $\frac{1}{2}\sinh^{-1}2x$ b) $\tanh^{-1}\frac{1}{2}=\frac{1}{2}\ln 3$ c) $\sinh^{-1}\left(\frac{x+2}{3}\right)$

Index