MATH1231 Mathematics 1B
MATH1241 Higher Mathematics 1B

CALCULUS NOTES

Preface

Please read carefully.

These Notes form the basis for the calculus strand of MATH1231 and MATH1241. However, not all of the material in these Notes is included in the MATH1231 or MATH1241 calculus syllabuses. A detailed syllabus is given, commencing on page (ix) of these Notes.

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 semester and continue to work steadily throughout the semester. 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 MATH1241.

6. 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. The problems set in tests and exams will be similar to the problems given in these notes. Further information on the problems and class tests is on pages (xi).

7. You will be expected to use the computer algebra package Maple in tests and understand Maple syntax and output for the end of semester 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 FOR MATH1231 MATHEMATICS 1B

In this syllabus the references to the textbook are *not* intended as a definition of what you will be expected to know. They are just a guide to finding relevant material. Some parts of the subject are not covered in the textbook and some parts of the textbook (even in the sections mentioned in the references below) are not included in the subject. The scope of the course is defined by the content of the lectures and problem sheets. The approximate lecture time for each section is given below. References to the 8th and 10th editions of Salas & Hille are shown as SH8 and SH10.

	SH8	SH10
1. **Functions of several variables.** (3 hours)		
Contours and level curves, partial derivatives.	14.1-14.4	15.1-15.4
Mixed derivative theorem, increment estimation.	14.6	15.6
Chain rules, tangent planes.		
2. **Integration techniques.** (4 hours)		
Trigonometric integrals and reduction formulae.	8.3	8.3
Trigonometric and hyperbolic substitutions.	8.4	8.4
Rational functions and partial fractions.	8.5	8.5
Further substitutions.	8.6	8.6
3. **Ordinary differential equations.** (6 hours)		
Particular, general, explicit and implicit solutions.	18.1	
1st order equations: separable, linear, exact.	8.9, 18.2, 15.9	9.1, 9.2, 19.1, 19.2
Modelling with odes		9.1, 9.2
2nd order linear equations with constant coeffts:		
homogeneous, non-homogeneous (undetermined coeffts).	18.3, 18.4	9.3, 19.4
4. **Taylor series.** (7 hours)		
Taylor polynomials, Taylor's theorem.	11.5	12.6, 12.7
Application to stationary points.		
<u>*Sequences*</u>: convergence and divergence;	10.2, 10.3	11.2-11.4
combination of sequences.		
<u>*Series*</u>: partial sums; convergence;	11.1, 11.2	12.1, 12.2
kth term test for divergence;		
integral,comparison and ratio tests;	11.1-11.3	12.3,12.4
alternating series (Leibniz' test);		
absolute and conditional convergence;	11.4	12.5
rearrangement of series.		
Taylor and Maclaurin series.	11.6	12.7
<u>*Power series*</u>: radius and interval	11.7, 11.8	12.8, 12.9
of convergence; operations on power series.		
5. **Applications of integration.** (3 hours)		
Average value of a function.	5.8	5.9
Arc length.	9.8	10.7
Arc length in polar coordinates.	9.5, 9.8	10.7
Area of surfaces of revolution.	9.9	10.8

CALCULUS SYLLABUS FOR MATH1241 HIGHER MATHEMATICS 1B

This is the syllabus for *Higher* Mathematics 1B.
In this syllabus the references to the textbook are *not* intended as a definition of what you will be expected to know. They are just a guide to finding relevant material. Some parts of the subject are not covered in the textbook and some parts of the textbook (even in the sections mentioned in the references below) are not included in the subject. The scope of the subject is defined by the content of the lectures and problem sheets. The approximate lecture time for each section is given below. References to the 8th and 10th editions of Salas & Hille are shown under SH8 and SH10 and references to *Calculus* by M. Spivak under Sp.

	SH8	SH10	Sp
1. **Functions of several variables.** (3 hours)			
Contours and level curves, partial derivatives.	14.1-14.4	15.1-15.4	
Mixed derivative theorem, increment estimation.	14.6	15.6	
Chain rules, tangent planes.			
2. **Integration techniques.** (4 hours)			
Trigonometric integrals and reduction formulae.	8.3	8.3	18
Trigonometric and hyperbolic substitutions.	8.4	8.4	18
Rational functions and partial fractions.	8.5	8.5	18
Further substitutions.	8.6	8.6	18
3. **Ordinary differential equations.** (6 hours)			
Particular, general, explicit and implicit solutions.	18.1		
1st order equations: separable, linear, exact.	8.9, 18.2, 15.9	9.1, 9.2, 19.1, 19.2	
Modelling with odes		9.1, 9.2	
2nd order linear equations with constant coeffts:			
homogeneous, non-homogeneous (undetermined coeffts).	18.3, 18.4	9.3, 19.4	
4. **Taylor series.** (7 hours)			
Taylor polynomials, Taylor's theorem.	11.5	12.6, 12.7	
Application to stationary points.			
<u>*Sequences*</u>: convergence and divergence;	10.2, 10.3	11.2-11.4	21
combination of sequences.			
Upper, lower bounds, sup and inf,	10.1-10.3	11.1	8, 21
bounded monotonic sequences.	10.2	11.1	
Recursively defined sequences.			
<u>*Series*</u>: partial sums; convergence;	11.1, 11.2	12.1, 12.2	22
kth term test for divergence;			
comparison, integral, ratio and root tests;	11.1-11.3	12.3,12.4	22
alternating series (Leibniz' test);			
absolute and conditional convergence;	11.4	12.5	22
rearrangement of series.			
Taylor and Maclaurin series.	11.6	12.7	19
<u>*Power series*</u>: radius and interval	11.7, 11.8	12.8, 12.9	23
of convergence; operations on power series.			
5. **Applications of integration.** (3 hours)			
Average value of a function.	5.8	5.9	
Arc length in Cartesian and polar coordinates.	9.5, 9.8	10.7	
Area of surfaces of revolution.	9.9	10.8	

PROBLEM SETS

The Calculus problems are located at the end of each chapter of the Calculus Notes booklet.

To help you decide which problems to try first, each problem is marked with an [**R**], an [**H**] or an [**HH**]. A few problems are marked with an [**X**] for MATH1241 students.

All students should make sure that they attempt the questions marked [**R**]. The problems marked [**H**] or [**HH**] are intended as a challenge for students in MATH1231 as well as MATH1241. Some harder parts of [**R**] problems are marked with a star. Any problems which depend on work covered only in MATH1241 are marked [**X**]. Problems marked with [**V**] have a video solution available on Moodle.

WEEKLY CALCULUS PROBLEM SCHEDULE

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. The following table gives the range of questions suitable for each week. In addition it suggests specific recommended problems to do before your classroom tutorials.

The Online Tutorials will develop your problem solving skills, and give you examples of mathematical writing. Online Tutorials help build your understanding from lectures towards solving problems on your own. Because this overlaps with the skills developed through homework, there are fewer recommended homework in Online Tutorial weeks.

WEEKLY CALCULUS HOMEWORK SCHEDULE

Week	**Calculus problems**		**Recommended Homework**
	Chapter	**Problems up to**	**Problems**
1	No tutorial.		
2	1	4	1(c), 3(c), 4(d)
3	1	18	7,12
	2	5	1(e), 1(f), 1(k), 2(c), 3(b), 3(e)
4	2	17	7, 15(a), 15(d), 17(b), 17(e)
5	2	22	18(c), 22(b), 22(c), 22(i)
6	3	17 (Test 1)	1(h), 4(d), 6, 8(a), 9(c)
7	3	32	20, 23, 30(a), 31(b), 32(c)
8	3	44	33(a), 37, 38, 40
9	4	18	4, 6, 12(c), 12(e)
10	4	32	19, 20, 23, 25(b), 26(a), 27(c)
11	4	42 (Test 2)	34(c), 35, 41(d), 42(a)
12	4	49	44(a), 45, 46, 48
13	5	13	2, 3(c), 5, 7, 10(a), 13

WEEKLY MATH1231 CALCULUS TUTORIAL SCHEDULE

The main reason for having tutorials is to give you a chance to tackle and discuss problems which you find difficult or don't fully understand.

There are two kinds of tutorials: Online and Classroom. Calculus Online Tutorials are delivered using MapleTA. These can be completed from home, are available for a two week period, and are due on Sunday night in weeks 3,5,7,9,11 and 13. Calculus Classroom tutorials are delivered in a classroom by a calculus tutor. The topics covered in a classroom tutorial are flexible, and you can (and should) ask your tutor to cover any homework topics you find difficult. You may also be asked to present solutions to homework questions to the rest of the class.

The following table lists the topics covered in each tutorial.

Week	**Location**	**Topics Covered**
2	Classroom	1.1 : Sketching simple surfaces in $\mathbb{R}^3$ 1.2 : Partial differentiation 1.3 : Tangent planes to surfaces
3	Online	1.4 : The total differential approximation 1.5 : Chain rules
4	Classroom	2.1 : Trigonometric integrals 2.2 : Reduction formulae
5	Online	2.3 : Trigonometric and hyperbolic substitutions 2.4 : Integrating rational functions 2.5 : Other substitutions
6	Classroom	3.3 : Separable ODEs 3.4 : First order linear ODEs 3.5 : Exact ODEs
7	Online	3.7 : Modelling with first order ODEs
8	Classroom	3.8 : Second order linear ODEs with constant coefficients
9	Online	4.1 : Taylor polynomials 4.2 : Taylor's theorem
10	Classroom	4.3 : Sequences 4.4 : Infinite series
11	Online	4.5 : Tests for series convergence 4.6 : Taylor series
12	Classroom	4.7 : Power series 4.8 : Manipulation of power series
13	Online	5.1 : The average value of a function 5.2 : The arc length of a curve 5.3 : The speed of a moving particle 5.4 : Surface area

WEEKLY MATH1241 CALCULUS TUTORIAL SCHEDULE

MATH1241 Tutorials cover the same material as MATH1231, only in greater detail. The tutorial structure is more flexible, which is designed to allow for classroom discussion. Only a subset of the recommended discussion questions will be discussed in your classroom tutorial, which are held every even week starting in week 2. Online Tutorial questions for calculus are due at Sunday 23:59 every odd week starting in week 3.

weeks	**Chapter**	**Online Tutorial**	**Recommended Classroom Discussion Questions**
2	1	None	1c, 3c, 4d
3 and 4	1	7, 18	12
	2	3b,10	1,2c, 3e, 7, 17b, 17e
5 and 6	2	18, 22i	22b, 22c
	3	1h, 6	4d, 8a, 9c
7 and 8	3	20, 30a	23, 31b, 32c
	3	33a, 38	37, 40
9 and 10	4	4, 12e	6, 12c
		19, 25b	20, 23, 26a, 27c
11 and 12	4	34c, 41d	35, 40, 42a
		45, 48	44a, 46
13	5	2, 5, 7, 10a	None

Chapter 1

Functions of several variables

Most functions which arise in real world applications depend on more than one variable. In this chapter, we give a brief introduction to functions of two (or more) variables. Examples include functions defined by the following well known formulae:

- $A(b,h) = \frac{1}{2}bh$ (the area of a triangle);
- $D(x,y) = \sqrt{x^2+y^2}$ (the distance of a point (x,y) from the origin);
- $\ell(x,y) = 2(x+y)$ (the perimeter of a rectangle of dimensions x and y units); and
- $F(m_1, m_2, \mathbf{x}_1, \mathbf{x}_2) = \dfrac{Gm_1m_2}{|\mathbf{x}_1 - \mathbf{x}_2|^2}$ (the gravitational force between two bodies of mass m_1 and m_2 positioned at points $P(\mathbf{x_1})$ and $Q(\mathbf{x}_2)$).

Geometrically, a function of two variables represents a surface in $\mathbb{R}^3$, just as a function of one variable represents a curve in $\mathbb{R}^2$. We will introduce the *partial derivative* of a function of two variables and use this to calculate the equation of the tangent plane to a surface at a point. Other applications of the partial derivative include function and error estimation.

1.1 Sketching simple surfaces in $\mathbb{R}^3$

(Ref: SH10 §15.2, 15.3)

The graph of a function f of one variable, given by $y = f(x)$, gives rise to a curve in $\mathbb{R}^2$. The graph of a function F of two variables, given by $z = F(x,y)$, gives rise to a surface in $\mathbb{R}^3$. In this section we introduce some simple techniques for sketching surfaces given by an equation $z = F(x,y)$.

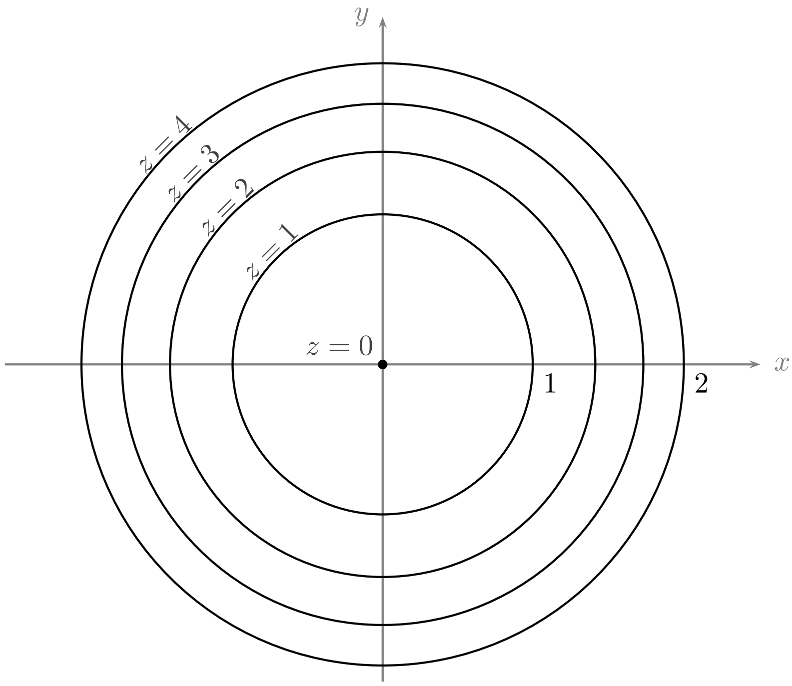
Sketching a surface in $\mathbb{R}^3$ can be challenging because the sketch must be represented in $\mathbb{R}^2$. Topographic maps solve this problem by using *contour lines* to represent the height (above sea level) of the surface of the earth at various points. We adapt this idea to sketching a surface $z = F(x,y)$ described by a function F. Here, z represents the height of the surface above the xy-plane, and the contours of the surface are defined as follows.

Definition 1.1.1. A *contour* or *level curve* of a function $F : \mathbb{R}^2 \to \mathbb{R}$ is a curve in $\mathbb{R}^2$ corresponding to an equation of the form $F(x,y) = C$, where C is a constant.

For each level curve, the corresponding value of C gives the height of the curve above the xy-plane.

Example 1.1.2. Sketch level curves for the function $F : \mathbb{R}^2 \to \mathbb{R}$, where $F(x, y) = x^2 + y^2$.

Solution. The level curves of F are of the form $x^2 + y^2 = C$, where C is a *nonnegative* constant (since if $C < 0$ then there is no solution to the equation $x^2 + y^2 = C$). The level curves given by $x^2 + y^2 = 0$, $x^2 + y^2 = 1$, $x^2 + y^2 = 2$, $x^2 + y^2 = 3$ and $x^2 + y^2 = 4$ are shown below.

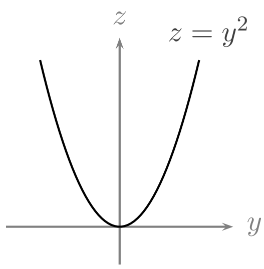


Level curves can be interpreted in the following way. If one walked around the circle $x^2 + y^2 = 4$, then one would remain a constant height of 4 units above the xy-plane. If one started at the origin and walked 2 units 'east,' then one would rise from 'sea level' to 4 units above sea level. ☐

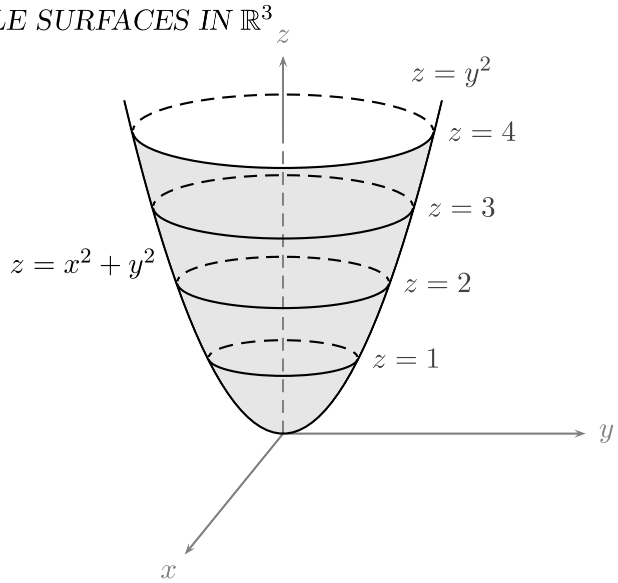
Some simple surfaces may be sketched in $\mathbb{R}^3$ by using the level curves of the surface and considering the intersection of the surface with the yz-plane. The final sketch is drawn using perspective.

Example 1.1.3. Sketch the surface in $\mathbb{R}^3$ described by the equation $z = x^2 + y^2$.

Solution. The level curves for the surface are circles (see the sketch in the previous example). When $x = 0$, we have $z = y^2$. This gives the intersection of the surface with the yz-plane. The profile of this intersection is shown below.



The yz-profile and level curves help us produce the sketch of the surface.



□

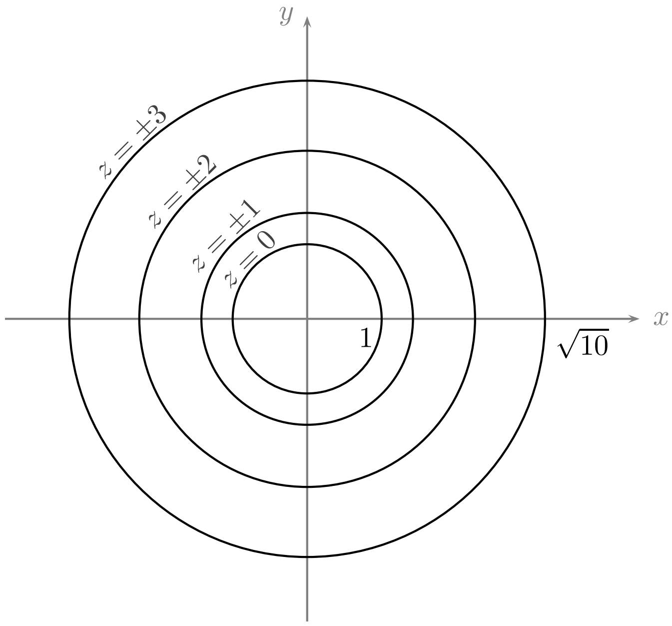
Remark 1.1.4. The previous diagram illustrates the conventional orientation for the x-, y- and z-axes.

Example 1.1.5. A surface in $\mathbb{R}^3$ is described by the equation $x^2 + y^2 - z^2 = 1$. Sketch some level curves and hence sketch the surface in $\mathbb{R}^3$.

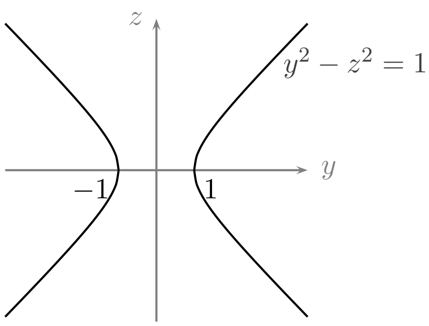
Solution. Each level curve is obtained by setting z equal to C, for some constant C.

z	level curve
0	$x^2 + y^2 = 1$
± 1	$x^2 + y^2 = 2$
± 2	$x^2 + y^2 = 5$
± 3	$x^2 + y^2 = 10$

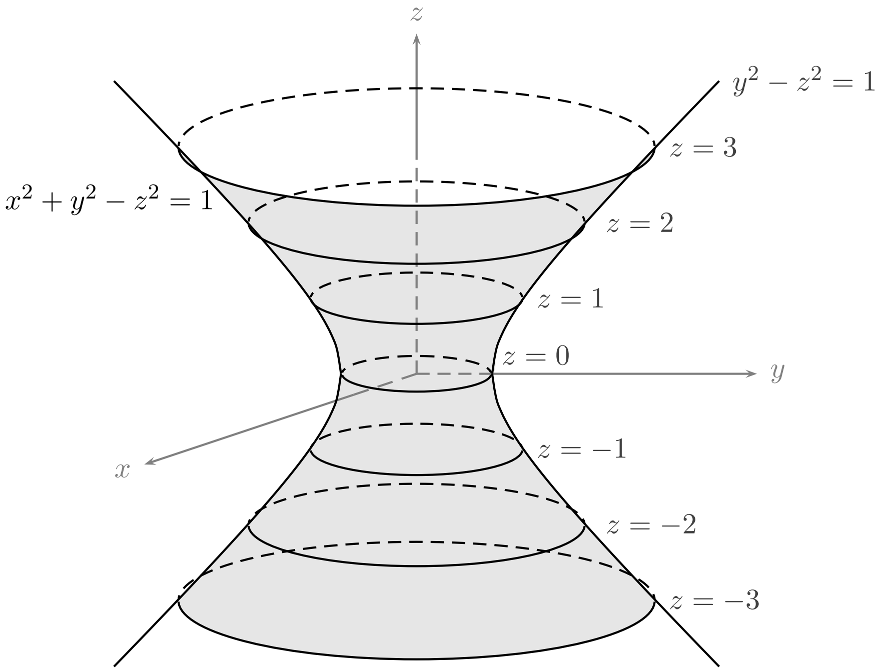
Each level curve is a circle. Those given in the table are sketched below.



If $x = 0$ then $y^2 - z^2 = 1$, which is a hyperbola.



Putting these two sketches together gives the following surface.



This surface is called a *hyperboloid of one sheet*. □

1.2 Partial differentiation

(Ref: SH10 §15.4, 15.6)

Suppose that f is a differentiable function of one variable. In MATH1131 we

- investigated techniques for calculating the derivative of f,
- interpreted the derivative in terms of the rate of change of f,
- used the derivative to calculate the equation of the tangent line to the graph of f at a point a, and
- used the tangent line to give a *linear approximation* to f near a.

Over the next few sections, we generalise some of these ideas and techniques to functions F of two variables. In particular, we introduce the notion of a *partial derivative* and then

- show how to calculate the partial derivatives of F,
- interpret the partial derivatives in terms of the rate of change of F,
- use the partial derivatives to calculate the equation of the tangent plane to the graph of F at a point (a, b), and
- use the tangent plane to give a *linear approximation* for F near (a, b).

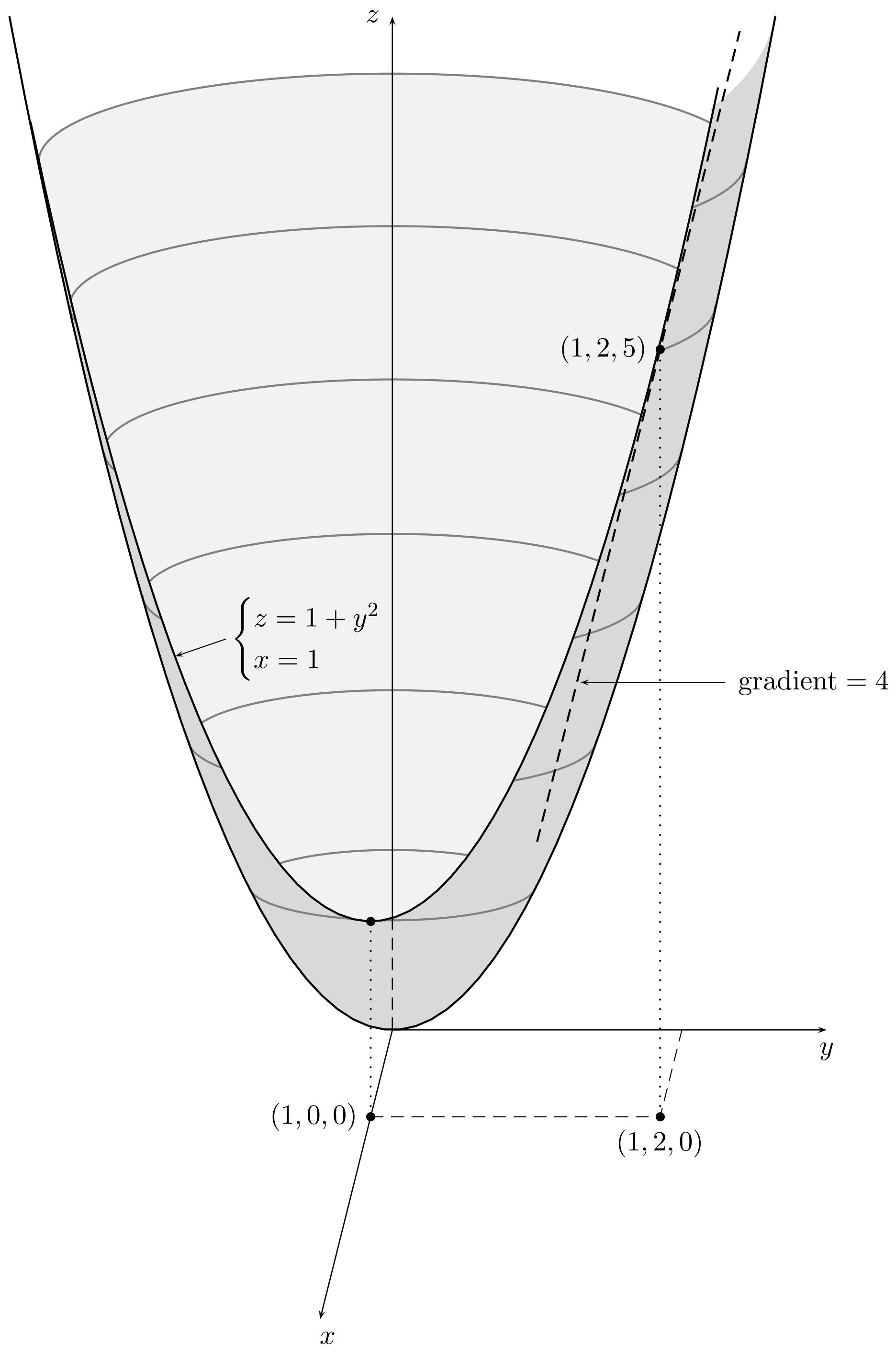


Figure 1.1: The intersection of the surface $z = x^2 + y^2$ with the plane $x = 1$. The dashed line is the tangent line to the cross-section at the point $(1, 2, 5)$.

Our discussion will assume that we have an *intuitive* idea of what is meant by a 'tangent plane to a surface.' A rigorous discussion of the existence of tangent planes is connected with the formal definition of the *derivative* (as opposed to *partial derivative*) of a function of two variables. This rigour is pursued in some second year courses.

To introduce the notion of a partial derivative, consider the function F given by

$$F(x, y) = x^2 + y^2.$$

The surface $z = x^2 + y^2$, corresponding to the graph of F, was sketched in Example 1.1.3. Our immediate goal is to quantify the rate of change of $F(x, y)$ at the point $(1, 2)$.

First, we find the rate of change of $F(x, y)$ at $(1, 2)$ in the y-direction. This means that we hold x constant and only allow y to vary. Since we are considering the rate of change at the point $(1, 2)$, we have $x = 1$ and

$$z = F(x, y) = F(1, y) = 1 + y^2$$

(since x is constant but y is not). Geometrically, we may interpret $z = 1 + y^2$ as the intersection of the surface $z = x^2 + y^2$ with the plane $x = 1$. This intersection is shown in Figure 1.1. The rate of change of $F(x, y)$ at $(1, 2)$ in the y-direction is equal to the gradient of the dashed tangent line shown in the figure. This gradient may be calculated in the usual way:

$$\text{gradient} = \frac{d}{dy}(1 + y^2)\Big|_{y=2} = 2y\Big|_{y=2} = 4.$$

Therefore the rate of change of $F(x, y)$ at $(1, 2)$ in the y-direction is 4.

A fast way of finding this rate of change is the following. First, differentiate F with respect to y, treating x as a constant, to obtain

$$F_y(x, y) = 2y.$$

Then the rate of change of $F(x, y)$ at $(1, 2)$ in the y-direction is

$$F_y(1, 2) = 2 \times 2 = 4.$$

The function F_y is called the *partial derivative of F with respect to y.*

In a similar manner, one may find the rate of change of $F(x, y)$ at $(1, 2)$ in the x-direction. First, differentiate F with respect to x, treating y as a constant, to obtain

$$F_x(x, y) = 2x.$$

Then the rate of change of $F(x, y)$ at $(1, 2)$ in the x-direction is

$$F_x(1, 2) = 2 \times 1 = 2.$$

The function F_x is called the *partial derivative of F with respect to x.*

The partial derivatives of a function F may be defined formally by using limits.

Definition 1.2.1. Suppose that F is a function of two variables x and y. The partial derivatives of F with respect to x and y are defined by

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

and

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

wherever these limits exist.

Remark 1.2.2. If F is a function of x and y, then $F_x(x,y)$ is calculated by treating y as a constant and differentiating F with respect to x. On the other hand, $F_y(x,y)$ is calculated by treating x as a constant and differentiating F with respect to y.

Remark 1.2.3. As illustrated at the beginning of this section, $F_y(a,b)$ gives the rate of change of F at the point (a,b) in the y-direction. Interpreted geometrically, the number $F_y(a,b)$ is the gradient of the tangent to the cross-section at (a,b) when the surface $z = F(x,y)$ is intersected with the plane $x = a$. Similarly, $F_x(a,b)$ gives the rate of change of F at the point (a,b) in the x-direction.

Example 1.2.4. Suppose that $F(x,y) = x^2y + 2y + 4$. Find F_x and F_y.

Solution. To calculate F_x, we treat y as a constant and differentiate F with respect to x:

$$F_x(x,y) = 2xy.$$

To calculate F_y, we treat x as a constant and differentiate F with respect to y:

$$F_y(x,y) = x^2 + 2.$$

□

Partial derivatives may be denoted in a variety of ways.

Remark 1.2.5 (Notation)**.** Suppose that a function F has partial derivatives F_x and F_y. Then F_x may be denoted by

$$\frac{\partial F}{\partial x} \qquad \text{or} \qquad D_1F,$$

while F_y may be denoted by

$$\frac{\partial F}{\partial y} \qquad \text{or} \qquad D_2F.$$

The notation involving the 'curly d' is used most frequently. However, the notation D_1F and D_2F is less ambiguous. (For example, to calculate $D_1F(y,x)$, we differentiate F with respect to its first variable and then evaluate this partial derivative at the point (y,x). In 'curly d' notation, one would write $\frac{\partial F}{\partial x}(y,x)$. Here, the first x represents the first variable of the function while the x in parentheses represents the second ordinate of the point (y,x).)

Example 1.2.6. Suppose that

$$F(x,y) = 3e^{xy^3}\sin y.$$

Find $\frac{\partial F}{\partial x}$ and $\frac{\partial F}{\partial y}$.

Solution. To calculate $\frac{\partial F}{\partial x}$, we treat y as a constant and differentiate F with respect to x:

$$\frac{\partial F}{\partial x} = 3y^3e^{xy^3}\sin y.$$

To calculate $\frac{\partial F}{\partial y}$, we treat x as a constant and differentiate F with respect to y:

$$\begin{aligned}
\frac{\partial F}{\partial y} &= 3\sin(y)\frac{\partial}{\partial y}\left(e^{xy^3}\right) + 3e^{xy^3}\frac{\partial}{\partial y}(\sin y) && \text{(by using the product rule)}\\
&= 3\sin(y)\left(3xy^2e^{xy^3}\right) + 3e^{xy^3}(\cos y)\\
&= 3e^{xy^3}\left(3xy^2\sin y + \cos y\right).
\end{aligned}$$

□

So far we have seen examples of first order partial derivatives. Second order partial derivatives can be computed by differentiating first order partial derivatives. Some notation is given below:

$$\begin{array}{ccll} \dfrac{\partial^2 F}{\partial x^2} & \text{means} & \dfrac{\partial}{\partial x}\left(\dfrac{\partial F}{\partial x}\right); & \\ \dfrac{\partial^2 F}{\partial y^2} & \text{means} & \dfrac{\partial}{\partial y}\left(\dfrac{\partial F}{\partial y}\right); & \\ \dfrac{\partial^2 F}{\partial x \partial y} & \text{means} & \dfrac{\partial}{\partial x}\left(\dfrac{\partial F}{\partial y}\right); & \text{and} \\ \dfrac{\partial^2 F}{\partial y \partial x} & \text{means} & \dfrac{\partial}{\partial y}\left(\dfrac{\partial F}{\partial x}\right). & \end{array}$$

Second order partial derivatives may be denoted using other notation. For example,

$$\frac{\partial^2 F}{\partial x^2} = F_{xx} = D_1^2 F$$

while

$$\frac{\partial^2 F}{\partial y \partial x} = F_{xy} = D_2 D_1 F.$$

Example 1.2.7. Compute all second order partial derivatives of F, where $F(x, y) = 3xy^4 + e^x \sin y$.

Solution. The first order partial derivatives are given by

$$F_x(x, y) = 3y^4 + e^x \sin y \qquad \text{and} \qquad F_y(x, y) = 12xy^3 + e^x \cos y.$$

The second order partial derivatives are obtained by furher differentiation. Thus

$$\begin{aligned} F_{xx}(x, y) &= \frac{\partial}{\partial x}\big(3y^4 + e^x \sin y\big) = e^x \sin y \\ F_{yy}(x, y) &= \frac{\partial}{\partial y}\big(12xy^3 + e^x \cos y\big) = 36xy^2 - e^x \sin y \\ F_{yx}(x, y) &= \frac{\partial}{\partial x}\big(12xy^3 + e^x \cos y\big) = 12y^3 + e^x \cos y \\ F_{xy}(x, y) &= \frac{\partial}{\partial y}\big(3y^4 + e^x \sin y\big) = 12y^3 + e^x \cos y. \end{aligned}$$

□

Note in the previous example that $F_{xy} = F_{yx}$. This is no accident.

Theorem 1.2.8 (The mixed derivative theorem)**.** *Suppose that F is a function of two variables. If F and all its first and second order partial derivatives are continuous then*

$$\frac{\partial^2 F}{\partial x \partial y} = \frac{\partial^2 F}{\partial y \partial x}.$$

Two questions arise. First, what does it mean for a function F of two variables to be continuous? Second, why is the theorem true? We give a partial answer to each of these questions in the following remarks.

Remark 1.2.9. Suppose that F is a function of two variables. While a formal definition for continuity is not given in this course, the following rule may often be used to verify that F is continuous.

> If F can be constructed by combining (via function addition, multiplication, division and composition) a finite number of continuous functions of a single variable, then F is continuous on its domain.

Thus the function F, given by

$$F(x, y) = 3xy^4 + e^x \sin y,$$

is continuous since

$$F(x, y) = f(x)g(y) + h(x)k(y),$$

where the functions given by $f(t) = 3t$, $g(t) = t^4$, $h(t) = e^t$ and $k(t) = \sin t$ are continuous. Similarly, the function G given by

$$G(x, y) = e^{3x^4 y}$$

is continuous since $G(x, y) = h(g(x)f(y))$. Suffice to say, most functions of two variables given in this course are continuous on their domains.

Remark 1.2.10. To begin proving the mixed derivative theorem, one begins with the definition of the partial derivative. Now

$$\begin{aligned}
F_{xy}(a, b) &= \lim_{h_2 \to 0} \frac{F_x(a, b + h_2) - F_x(a, b)}{h_2} \\
&= \lim_{h_2 \to 0} \lim_{h_1 \to 0} \frac{\frac{F(a+h_1, b+h_2) - F(a, b+h_2)}{h_1} - \frac{F(a+h_1, b) - F(a, b)}{h_1}}{h_2} \\
&= \lim_{h_2 \to 0} \lim_{h_1 \to 0} \frac{F(a + h_1, b + h_2) - F(a, b + h_2) - F(a + h_1, b) + F(a, b)}{h_1 h_2}.
\end{aligned}$$

Similarly,

$$F_{yx}(a, b) = \lim_{h_1 \to 0} \lim_{h_2 \to 0} \frac{F(a + h_1, b + h_2) - F(a + h_1, b) - F(a, b + h_2) + F(a, b)}{h_1 h_2}.$$

To show that $F_{yx}(a, b) = F_{xy}(a, b)$, one need only swap the order in which the limits are taken. This step can be justified by the continuity hypothesis of the mixed derivative theorem. The details are not easy and will be omitted here.

1.3 Tangent planes to surfaces

In this section, we give an heuristic derivation of the Cartesian equation for the tangent plane to a surface at a given point. Our treatment relies on intuition rather than rigour; a full presentation requires a rigorous account of what is meant by a 'tangent plane' and conditions under which a tangent plane exists. Such questions are tackled in some second year courses.

As a side benefit of our derivation for equation of the tangent plane, we also obtain a formula for a *normal vector* to the surface at a given point. A normal vector is defined as follows. Suppose that a surface has a tangent plane at a point P. We say that a vector $\mathbf{n}$ is *normal* to the surface at P if $\mathbf{n}$ is normal to the tangent plane to the surface at P.

The main ideas for deriving a formula for the tangent planes and normal vectors are contained in the following example.

Example 1.3.1. Suppose that $F(x, y) = x^2 + y^2$. By using vector geometry, find the Cartesian equation of the tangent plane to the surface $z = F(x, y)$ at the point where $(x, y, z) = (1, 2, 5)$. Find also a vector $\mathbf{n}$ that is normal to the surface at this point.

Solution. We will complete the solution in four steps.

Step 1. First, intersect the surface $z = x^2 + y^2$ with the plane $x = 1$, as in Figure 1.1. This gives the cross-sectional profile

$$\begin{cases} z = 1 + y^2 \\ x = 1 \end{cases}$$

as illustrated in Figure 1.1. The dashed line passing through the point $(1, 2, 5)$ lies in the plane $x = 1$ and is tangent to the parabola $z = 1 + y^2$. Its gradient is

$$F_y(1, 2) = 4.$$

By using the point-gradient formula for a straight line, the Cartesian equation for this tangent is given by

$$z - 5 = 4(y - 2), \qquad x = 1.$$

If $\lambda = y - 2$ then the equation of the tangent line in parametric vector form is

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 5 \end{pmatrix} + \lambda \begin{pmatrix} 0 \\ 1 \\ 4 \end{pmatrix} \tag{1.1}$$

whenever $\mu \in \mathbb{R}$. Note that this line lies in the tangent plane to the surface at $(1, 2, 5)$.

Step 2. Now we intersect the surface with the plane $y = 2$ and repeat the method of Step 1. The cross-sectional profile of this intersection is given by

$$\begin{cases} z = x^2 + 4 \\ y = 2. \end{cases}$$

Consider the line that lies in the plane $y = 2$ and is tangent to $z = x^2 + 4$ at $(1, 2, 5)$. Its gradient is given by

$$F_x(1, 2) = 2$$

and hence its Cartesian equation by

$$z - 5 = 2(x - 1), \qquad y = 2,$$

If $\mu = x - 1$ then we find that the parametric vector form of the tangent line is

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 5 \end{pmatrix} + \mu \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix} \tag{1.2}$$

whenever $\lambda \in \mathbb{R}$. Note that this line also lies in the tangent plane to the surface at $(1, 2, 5)$.

Step 3. Since the lines given by (1.1) and (1.2) both lie in the tangent plane, and since the vectors

$$\begin{pmatrix} 0 \\ 1 \\ 4 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix}$$

are nonparallel, the tangent plane to the surface at $(1,2,5)$ is given by

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 5 \end{pmatrix} + \lambda \begin{pmatrix} 0 \\ 1 \\ 4 \end{pmatrix} + \mu \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix}$$

where λ and μ are arbitrary real numbers.

Step 4. We now convert the parametric vector form of the plane to point-normal form. A vector $\mathbf{n}$ that is normal to the plane is given by

$$\mathbf{n} = \begin{pmatrix} 0 \\ 1 \\ 4 \end{pmatrix} \times \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix} = \begin{pmatrix} 2 \\ 4 \\ -1 \end{pmatrix}.$$

Hence the point-normal form for the tangent plane is

$$\begin{pmatrix} 2 \\ 4 \\ -1 \end{pmatrix} \cdot \begin{pmatrix} x-1 \\ y-2 \\ z-5 \end{pmatrix} = 0.$$

By expanding the dot product, one obtains $z = 5 + 2(x-1) + 4(y-2)$, which simplifies to

$$2x + 4y - z = 5.$$

This is the cartesian form of the plane, and the vector $(2,4,-1)^T$ is normal to the surface at the point $(1,2,5)$. $\square$

Note in the above example that $F_x(1,2) = 2$ and $F_y(1,2) = 4$. Each of these numbers appear as coefficients in the equation of the tangent plane and as components of the normal vector. By generalising the above example, one obtains the formulae in the following proposition.

Proposition 1.3.2. *Suppose that F is a function of two variables and (x_0, y_0, z_0) is a point that lies on the surface $z = F(x,y)$. If the surface has a tangent plane at the point (x_0, y_0, z_0), then the tangent plane is given by the equation*

$$z = z_0 + F_x(x_0, y_0)(x - x_0) + F_y(x_0, y_0)(y - y_0)$$

and a normal vector to the surface at (x_0, y_0, z_0) is given by

$$\begin{pmatrix} F_x(x_0, y_0) \\ F_y(x_0, y_0) \\ -1 \end{pmatrix}.$$

Remark 1.3.3. The above proposition is similar to what is already known for functions of a single variable. In fact, we have the following. Suppose that f is a function of one variable and (x_0, y_0) is a point that lies on the curve $y = f(x)$. If f is differentiable at x_0, then the tangent line to the curve at (x_0, y_0) is given by the equation

$$y = y_0 + f'(x_0)(x - x_0)$$

and a normal vector to the curve at (x_0, y_0) is given by

$$\begin{pmatrix} f'(x_0) \\ -1 \end{pmatrix}.$$

Remark 1.3.4. The proposition implicitly assumes that the first order partial derivatives for F exist at (x_0, y_0). It can be shown that if a tangent plane exists then these first order partial derivatives also exist. Hence this implicit assumption causes no problems. (Be warned however, that the existence of first order partial derivatives at (x_0, y_0) does not guarantee the existence of a tangent plane at (x_0, y_0, z_0).)

Example 1.3.5. Suppose that $F(x, y) = \sin(\pi x y^2)$. Write down the Cartesian equation for the tangent plane to the surface $z = F(x, y)$ at the point $(2, -1, 0)$ and find a vector that is normal to the surface at this point.

Solution. The partial derivatives of F are given by

$$F_x(x, y) = \pi y^2 \cos(\pi x y^2) \qquad \text{and} \qquad F_y(x, y) = 2\pi x y \cos(\pi x y^2).$$

Hence

$$F_x(2, -1) = \pi \qquad \text{and} \qquad F_y(2, -1) = -4\pi.$$

So the equation of the tangent plane is given by

$$z = 0 + \pi(x - 2) - 4\pi(y + 1) \text{ or } \pi x - 4\pi y - z = 6\pi.$$

and a normal vector is

$$\begin{pmatrix} \pi \\ -4\pi \\ -1 \end{pmatrix}.$$

□

Notes:

- One could also use the point normal form directly to find the equation of the tangent plane. Thus the vector normal to the plane is $\mathbf{n} = \begin{pmatrix} f_x \\ f_y \\ -1 \end{pmatrix}$ (evaluated at the point P), and so the equation of the tangent plane is $(\mathbf{x} - P).\mathbf{n} = 0$.

  Hence, in the previous example, we could simply expand $\left(\mathbf{x} - \begin{pmatrix} 2 \\ -1 \\ 0 \end{pmatrix}\right) . \begin{pmatrix} \pi \\ -4\pi \\ -1 \end{pmatrix} = 0$ to find the equation of the tangent plane.

- More generally, if a surface is given in the form $g(x, y, z) = 0$, then the vector $\mathbf{n} = \begin{pmatrix} g_x \\ g_y \\ g_z \end{pmatrix}$ evaluated at some point P on the surface, is a vector normal to the surface at P.

Example 1.3.6. Find the equation of the tangent plane to the ellipsoid $\frac{x^2}{4} + \frac{y^2}{2} + \frac{z^2}{8} = 1$ at the point $P(0, 1, 2)$

Solution. A vector normal to the ellipsoid at P is given by

$$\mathbf{n} = \begin{pmatrix} \frac{x}{2} \\ y \\ \frac{z}{4} \end{pmatrix}_P = \begin{pmatrix} 0 \\ 1 \\ 1/2 \end{pmatrix}.$$

Hence the equation of the tangent has the form

$$0x + 1y + \frac{1}{2}z = d$$

for some constant d. Substituting in P, we find that $d = 2$ and so the desired equation is $2y + z = 4$. $\square$

1.4 The total differential approximation

Suppose that f is a differentiable function of one variable. The equation of the tangent to the graph of f at a point x_0 is given by

$$y = y_0 + f'(x_0)(x - x_0),$$

where $y_0 = f(x_0)$. When x is close to x_0, the tangent line is close to the graph of f. In other words,

$$f(x) \approx f(x_0) + f'(x_0)(x - x_0).$$

Now writing $\Delta x = x - x_0$ and $\Delta f = f(x) - f(x_0)$ we have

$$\Delta f \approx f'(x_0)\Delta x.$$

This is called the *differential approximation* to Δf. Our goal is to generalise this idea to functions of two variables.

Suppose that a surface given by $z = F(x, y)$ has a tangent plane at the point (x_0, y_0, z_0). The equation of the tangent is given by

$$z = z_0 + F_x(x_0, y_0)(x - x_0) + F_y(x_0, y_0)(y - y_0).$$

If (x, y) is near (x_0, y_0) then

$$F(x, y) \approx z_0 + F_x(x_0, y_0)(x - x_0) + F_y(x_0, y_0)(y - y_0).$$

Since (x_0, y_0, z_0) lies on the surface, $z_0 = F(x_0, y_0)$. Thus

$$F(x, y) - F(x_0, y_0) \approx F_x(x_0, y_0)(x - x_0) + F_y(x_0, y_0)(y - y_0). \tag{1.3}$$

Let Δx denote the difference $x - x_0$ between x and x_0 and let Δy denote the difference $y - y_0$ between y and y_0. Then (1.3) becomes

$$F(x_0 + \Delta x, y_0 + \Delta y) - F(x_0, y_0) \approx F_x(x_0, y_0)\Delta x + F_y(x_0, y_0)\Delta y.$$

The left-hand side of this approximation represents the change in F when x_0 and y_0 are changed by Δx and Δy respectively. We call this the *increment* in F and denote it by ΔF; that is,

$$\Delta F = F(x_0 + \Delta x, y_0 + \Delta y) - F(x_0, y_0).$$

Hence

$$\boxed{\Delta F \approx F_x(x_0, y_0)(x - x_0) + F_y(x_0, y_0)(y - y_0).}$$

This formula is called the *total differential approximation to* ΔF. By supressing the point of evaluation, the total differential approximation may be written as

$$\boxed{\Delta F \approx \frac{\partial F}{\partial x}\Delta x + \frac{\partial F}{\partial y}\Delta y.}$$

The approximation improves when Δx and Δy are smaller.

The total differential approximation can be used to estimate the change in the output of a function given changes to each of the inputs.

Example 1.4.1. The ideal gas law asserts that the pressure P, volume V and temperature T of an ideal gas are related by the formula

$$PV = kT,$$

where k is a constant. If the temperature is increased by 4% and the volume is decreased by 5%, estimate the percentage increase in pressure.

Solution. We have

$$P = \frac{kT}{V}, \qquad \frac{\partial P}{\partial T} = \frac{k}{V} \qquad \text{and} \qquad \frac{\partial P}{\partial V} = -\frac{kT}{V^2}.$$

Now the temperature is increased by 4%, so $\Delta T = 0.04T$. Similarly, the volume is decreased by 5% and so $\Delta V = -0.05V$. By the total differential approximation,

$$\begin{aligned}
\Delta P &\approx \frac{\partial P}{\partial T}\Delta T + \frac{\partial P}{\partial V}\Delta V \\
&= \frac{k}{V} \times 0.04T - \frac{kT}{V^2} \times (-0.05V) \\
&= 0.04P + 0.05P && (\text{since } P = \tfrac{kT}{V}) \\
&= 0.09P.
\end{aligned}$$

Hence the pressure increases by approximately 9%. □

In science, engineering, psychology, economics and so on, measurements are often made that are not exact. Any quantities calculated from these measurements will also contain errors. The total differential approximation can give us an idea how bad such errors can get. Given a function F of two variables x and y, one can interpret ΔF as the error in the output given errors Δx and Δy in the inputs. Typically, one does not know the precise value of Δx and Δy, but sometimes one can find an upper bound for the absolute errors $|\Delta x|$ and $|\Delta y|$. The total differential approximation then gives an approximate upper bound for the absolute error $|\Delta F|$ in F:

$$\begin{aligned}
|\Delta F| &\approx \left|\frac{\partial F}{\partial x}\Delta x + \frac{\partial F}{\partial y}\Delta y.\right| \\
&\le \left|\frac{\partial F}{\partial x}\right| |\Delta x| + \left|\frac{\partial F}{\partial y}\right| |\Delta y|,
\end{aligned}$$

(where we have used the triangle inequality in the last step). The next example illustrates these ideas.

Example 1.4.2. The dimensions of a cylinder are measured to the nearest millimeter using a measuring tape. The circumference is measured to be 22.0 cm and height is measured to be 15.0 cm. Use these measurements to (a) estimate the volume of the cylinder, and (b) estimate an upper bound for the percentage error in your answer to part (a).

Solution. (a) Let r, C and h denote the radius, circumference and height respectively. Then $C = 2\pi r$ and so

$$V = \pi r^2 h = \pi \left(\frac{C}{2\pi}\right)^2 h = \frac{C^2 h}{4\pi}.$$

By using the measurements $C = 22$ and $h = 15$, one finds that

$$V = \frac{1815}{\pi}.$$

And so the volume is estimated to be $\frac{1815}{\pi}\,\mathrm{cm}^3$, which is approximately $577.73\,\mathrm{cm}^3$.

(b) The absolute error in each measurement is at most 0.5 mm, which is 0.05 cm. Let ΔC and Δh denote the error in each measurement. Then

$$|\Delta C| \le 0.05 \qquad \text{and} \qquad |\Delta h| \le 0.05.$$

The increment ΔV is the error in our calculation for the volume. Now

$$V = \frac{C^2 h}{4\pi}, \qquad \frac{\partial V}{\partial C} = \frac{Ch}{2\pi} \qquad \text{and} \qquad \frac{\partial V}{\partial h} = \frac{C^2}{4\pi}.$$

So the total differential approximation (when $C = 22$ and $h = 15$) is

$$\begin{aligned}
\Delta V &\approx \frac{\partial V}{\partial C}\Delta C + \frac{\partial V}{\partial h}\Delta h \\
&= \frac{Ch}{2\pi}\Delta C + \frac{C^2}{4\pi}\Delta h \\
&= \frac{165}{\pi}\Delta C + \frac{121}{\pi}\Delta h
\end{aligned}$$

If we take absolute values of both sides then

$$\begin{aligned}
|\Delta V| &\approx \left|\frac{165}{\pi}\Delta C + \frac{121}{\pi}\Delta h\right| \\
&\le \frac{165}{\pi}|\Delta C| + \frac{121}{\pi}|\Delta h| \qquad \text{(by the triangle inequality)} \\
&\le \frac{165}{\pi} \times 0.05 + \frac{121}{\pi} \times 0.05 \\
&= \frac{286}{20\pi}.
\end{aligned}$$

So an upper bound for the absolute error in V is approximately $\dfrac{286}{20\pi}$ (that is, approximately $4.55\,\mathrm{cm}^3$). An upper bound for the percentage error is given by

$$\begin{aligned}
\frac{\max|\Delta V|}{V} \times 100\% &\approx \frac{286}{20\pi} \cdot \frac{\pi}{1815} \times 100\% \\
&= \frac{26}{33}\,\%.
\end{aligned}$$

Hence the percentage error is no more than about 0.79%. $\square$

1.5 Chain rules

If f and g are functions of one variable, then the derivative of $f \circ g$ may be calculated using the chain rule for functions of one variable. In this section, we study compositions of functions of more than one variable. To calculate their partial derivatives, we use a chain rule for functions of more than one variable.

Suppose that F is a function of x and y and that x and y are each functions of t. A small change Δt in t produces a corresponding change Δx and Δy in x and y. These changes in turn produce a corresponding change ΔF in F. By the total differential approximation,

$$\Delta F \approx \frac{\partial F}{\partial x}\Delta x + \frac{\partial F}{\partial y}\Delta y,$$

and this approximation gets better as Δx and Δy approach zero. If we divide through by Δt then

$$\frac{\Delta F}{\Delta t} \approx \frac{\partial F}{\partial x}\frac{\Delta x}{\Delta t} + \frac{\partial F}{\partial y}\frac{\Delta y}{\Delta t}.$$

As $\Delta t \to 0$,

$$\frac{\Delta x}{\Delta t} = \frac{x(t+\Delta t) - x(t)}{\Delta t} \to \frac{dx}{dt}$$

and

$$\frac{\Delta y}{\Delta t} = \frac{y(t+\Delta t) - y(t)}{\Delta t} \to \frac{dy}{dt}.$$

Finally, if we view F as a function of t then, by a similar argument, $\dfrac{\Delta F}{\Delta t} \to \dfrac{dF}{dt}$. So in the limit, the total differential approximation becomes

$$\boxed{\frac{dF}{dt} = \frac{\partial F}{\partial x}\frac{dx}{dt} + \frac{\partial F}{\partial y}\frac{dy}{dt}.} \tag{1.4}$$

This is an example of a chain rule for a function of two variables.

The above chain rule must be interpreted properly. First, each derivative must be evaluated at a correct point. Second, the F appearing on the left-hand side is a function of one variable t, while the F that appears on the right-hand side is a function of two variables x and y. Technically, these are two different functions. The next theorem expresses chain rule (1.4) without these ambiguities.

Theorem 1.5.1. *Suppose that F is a function of two variables and that x and y are both functions of one variable. Define the function ϕ by $\phi(t) = F(x(t), y(t))$ and the point (x_0, y_0) by $(x_0, y_0) = (x(t_0), y(t_0))$. If x and y are both differentiable at t_0 and the partial derivatives of F exist and are continuous at (x_0, y_0), then ϕ is differentiable at t_0 and*

$$\phi'(t_0) = D_1F(x_0, y_0)x'(t_0) + D_2F(x_0, y_0)y'(t_0). \tag{1.5}$$

Remark 1.5.2. Formulae (1.4) and (1.5) are equivalent. The former is easier to remember while the latter is more precise. To remember the rule, consider the following *chain diagram*.

$$\begin{tikzcd} & x \arrow[dr, "\frac{dx}{dt}"] & \\ F \arrow[ur, "\frac{\partial F}{\partial x}"] \arrow[dr, "\frac{\partial F}{\partial y}"'] & & t \\ & y \arrow[ur, "\frac{dy}{dt}"'] & \end{tikzcd}$$

To construct the diagram, draw an arrow from each function to each of its variables. Then $\frac{dF}{dt}$ is the sum of all paths (left to right) from F to t, where the derivatives are multiplied across each path.

Example 1.5.3. The potential energy E of a particle at point (x, y) is given by $E(x, y) = \sin(\pi x^2 y)$. If the x-ordinate of the particle is increasing at a rate of 3 units per second, and the y-ordinate of the particle is decreasing at a rate of 2 units per second, find the rate of change of potential energy when the particle has coordinate $(-1, 2)$.

Solution. Since the ordinates x and y of the particle change with time t, we may view x and y as functions of t. We are told that

$$\frac{dx}{dt} = 3 \qquad \text{and} \qquad \frac{dy}{dt} = -2.$$

By the chain rule,

$$\begin{aligned} \frac{dE}{dt} &= \frac{\partial E}{\partial x}\frac{dx}{dt} + \frac{\partial E}{\partial y}\frac{dy}{dt} \\ &= 2\pi xy\cos(\pi x^2 y) \times 3 + \pi x^2 \cos(\pi x^2 y) \times (-2) \\ &= 2\pi x(3y - x)\cos(\pi x^2 y). \end{aligned}$$

When $(x, y) = (-1, 2)$,

$$\frac{dE}{dt} = -14\pi.$$

So the rate of change of E at $(-1, 2)$ is -14π. □

We now examine the case when F is a function of x and y, where each of x and y is a function of both s and t. This situation is sometimes written as

$$F = F(x, y), \qquad x = x(s, t) \qquad \text{and} \qquad y = y(s, t).$$

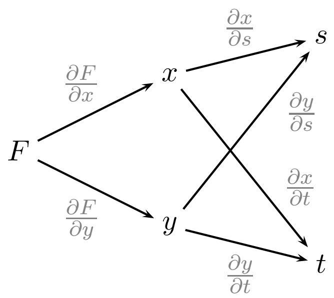
If we treat s as a constant and differentiate F with respect to t, then chain rule (1.4) gives

$$\boxed{\frac{\partial F}{\partial t} = \frac{\partial F}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial F}{\partial y}\frac{\partial y}{\partial t}.}$$

Similarly, if we treat t as a constant and differentiate F with respect to s, then chain rule (1.4) gives

$$\boxed{\frac{\partial F}{\partial s} = \frac{\partial F}{\partial x}\frac{\partial x}{\partial s} + \frac{\partial F}{\partial y}\frac{\partial y}{\partial s}.}$$

Each of these new chain rules may be remembered using the following chain diagram.



For example, to remember the rule for $\frac{\partial F}{\partial s}$, simply sum all paths (left to right) from F to s, where the derivatives are multiplied across each path.

Example 1.5.4. Suppose that $z = F(x, y)$. Express the point (x, y) in terms of polar coordinates (r, θ). Hence express $\frac{\partial z}{\partial r}$ and $\frac{\partial z}{\partial \theta}$ in terms of x, y, F_x and F_y. Finally, show that the partial derivatives satisfy the equation

$$r\frac{\partial z}{\partial r} + \frac{\partial z}{\partial \theta} = (x - y)\frac{\partial F}{\partial x} + (x + y)\frac{\partial F}{\partial y}.$$

Solution. We have

$$z = F(x, y), \qquad x = r\cos\theta, \qquad y = r\sin\theta \qquad \text{and} \qquad r^2 = x^2 + y^2.$$

So the chain rule gives

$$\begin{aligned}
\frac{\partial z}{\partial r} &= \frac{\partial z}{\partial x}\frac{\partial x}{\partial r} + \frac{\partial z}{\partial y}\frac{\partial y}{\partial r} \\
&= F_x(x, y)\cos\theta + F_y(x, y)\sin\theta \\
&= \frac{x}{r}\,F_x(x, y) + \frac{y}{r}\,F_y(x, y) \\
&= \frac{x}{\sqrt{x^2 + y^2}}\,F_x(x, y) + \frac{y}{\sqrt{x^2 + y^2}}\,F_y(x, y)
\end{aligned}$$

and

$$\begin{aligned}
\frac{\partial z}{\partial \theta} &= \frac{\partial z}{\partial x}\frac{\partial x}{\partial \theta} + \frac{\partial z}{\partial y}\frac{\partial y}{\partial \theta} \\
&= -F_x(x, y)r\sin\theta + F_y(x, y)r\cos\theta \\
&= -y\,F_x(x, y) + x\,F_y(x, y).
\end{aligned}$$

Finally,

$$\begin{aligned}
r\frac{\partial z}{\partial r} + \frac{\partial z}{\partial \theta} &= r\left(\frac{x}{\sqrt{x^2 + y^2}}\frac{\partial F}{\partial x} + \frac{y}{\sqrt{x^2 + y^2}}\frac{\partial F}{\partial y}\right) + \left(-y\frac{\partial F}{\partial x} + x\frac{\partial F}{\partial y}\right) \\
&= x\frac{\partial F}{\partial x} + y\frac{\partial F}{\partial y} - y\frac{\partial F}{\partial x} + x\frac{\partial F}{\partial y} \\
&= (x - y)\frac{\partial F}{\partial x} + (x + y)\frac{\partial F}{\partial y},
\end{aligned}$$

as required. □

We present one more useful version of the chain rule. Suppose that F is a function of u and that u is a function of both x and y. This is sometimes written as

$$F = F(u) \qquad \text{and} \qquad u = u(x, y).$$

The corresponding chain rules are

$$\boxed{\frac{\partial F}{\partial x} = \frac{dF}{du}\frac{\partial u}{\partial x}}$$

and

$$\boxed{\frac{\partial F}{\partial y} = \frac{dF}{du}\frac{\partial u}{\partial y}}.$$

The chain diagram coresponding to this situation is illustrated below.

$$\begin{tikzcd}
 & & x \\
F \arrow[r, "\frac{dF}{du}"] & u \arrow[ur, "\frac{\partial u}{\partial x}"] \arrow[dr, "\frac{\partial u}{\partial y}"'] & \\
 & & y
\end{tikzcd}$$

These chain rules may be easily written down after sketching the chain diagram.

1.6 Functions of more than two variables

Until now we have only discussed functions of two variables. In this section, the ideas met in this chapter are generalised to functions of three variables. We present a summary only.

Suppose that F is a function of three variables x, y and z. The partial derivatives of F are defined by

$$\begin{aligned}
F_x(x,y,z) &= \lim_{h\to 0} \frac{F(x+h,y,z)-F(x,y,z)}{h} \\
F_y(x,y,z) &= \lim_{h\to 0} \frac{F(x,y+h,z)-F(x,y,z)}{h} \\
F_z(x,y,z) &= \lim_{h\to 0} \frac{F(x,y,z+h)-F(x,y,z)}{h}
\end{aligned}$$

wherever these limits exist. Equivalent notation for each of these partial derivatives is given below:

$$F_x = \frac{\partial F}{\partial x} = D_1 F, \qquad F_y = \frac{\partial F}{\partial y} = D_2 F, \qquad \text{and} \qquad F_z = \frac{\partial F}{\partial z} = D_3 F.$$

If (a,b,c) is a point in $\mathbb{R}^3$ then $F_x(a,b,c)$ is the rate of change of F in the x-direction at (a,b,c). Similarly, $F_z(a,b,c)$ is the rate of change of F in the z-direction at (a,b,c).

The partial derivatives of F are calculated by differentiating F with respect to one variable and treating the other variables as constants. For example, if

$$F(x,y,z) = e^{2x} z \cos y$$

then

$$F_x(x,y,z) = 2e^{2x} z \cos y, \qquad F_y(x,y,z) = -e^{2x} z \sin y \qquad \text{and} \qquad F_z(x,y,z) = e^{2x} \cos y.$$

A 'hyper-surface' in $\mathbb{R}^4$ is the natural generalisation of a surface in $\mathbb{R}^3$. A function F of three variables can be used to define a hyper-surface $w = F(x,y,z)$ in $\mathbb{R}^4$. Given a point (x_0, y_0, z_0, w_0) on the surface, the equation of the 'tangent plane' to the surface at this point is given by

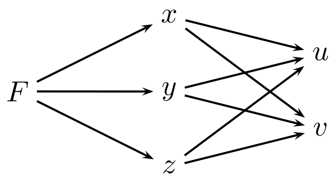
$$w = w_0 + F_x(x_0,y_0,z_0)(x-x_0) + F_y(x_0,y_0,z_0)(y-y_0) + F_z(x_0,y_0,z_0)(z-z_0)$$

(assuming, of course, that that a 'tangent plane' to the surface exists at this point).

The *total differential approximation* ΔF is given by

$$\boxed{\Delta F \approx \frac{\partial F}{\partial x}\Delta x + \frac{\partial F}{\partial y}\Delta y + \frac{\partial F}{\partial z}\Delta z.}$$

The chain rules are easily written down using chain diagrams. For example, suppose that F is a function of x, y and z and that x, y and z are each functions of both u and v. The corresponding chain diagram is shown below.



So the chain rule for $\dfrac{\partial F}{\partial u}$ is given by

$$\frac{\partial F}{\partial u} = \frac{\partial F}{\partial x}\frac{\partial x}{\partial u} + \frac{\partial F}{\partial y}\frac{\partial y}{\partial u} + \frac{\partial F}{\partial z}\frac{\partial z}{\partial u},$$

while the chain rule for $\dfrac{\partial F}{\partial v}$ is given by

$$\frac{\partial F}{\partial v} = \frac{\partial F}{\partial x}\frac{\partial x}{\partial v} + \frac{\partial F}{\partial y}\frac{\partial y}{\partial v} + \frac{\partial F}{\partial z}\frac{\partial z}{\partial v}.$$

The generalisation of each of these formulae to a function of four (or more) variables should be obvious.

1.7 Maple notes

The `plot3d` command is useful for visualizing the graphs of functions of several variables.
The MAPLE `diff` command carries out partial differentiation: `diff(f(x,y), x);` computes $\dfrac{\partial f}{\partial x}$, and `diff(f(x,y), x,y);` calculates $\dfrac{\partial^2 f}{\partial y \partial x}$. For example,

```
> diff(x^3*y-sin(y^2),x);
```

$$3x^2y$$

```
> diff(x^3*y-sin(y^2),y$2);
```

$$4\sin(y^2)y^2 - 2\cos(y^2)$$

Chapter 2

Integration techniques

Many real world problems, such as

- calculating the area of a region
- locating the centre of a region,
- calculating the volume, surface area and centre of mass of a solid,
- calculating the length of a curve,
- determining the probability that a certain event occurs,
- analysing the harmonics of a musical instrument,
- determining the solution to models of various physical phenomena, and
- calculating the work done by a force,

boil down to evaluating an appropriate integral. Some of these applications were explored in last semester's course while others will be discussed later in this course or in second year. But what these applications demand is mastery of integration. In this chapter we work towards that goal by examining techniques for integrating various types of integrals that arise when solving real world problems. The hard work done here will pay off when applications of such integrals are studied.

2.1 Trigonometric integrals

(Ref: SH10 §8.3)

In this section we focus specifically on integrals involving the trigonometric functions.

2.1.1 Integrating powers of sine and cosine

The first class of trigonometric integrals considered consist of integrals of the form

$$\int \cos^m x \sin^n x \, dx, \tag{2.1}$$

where m and n are non-negative integers. There are essentially two cases: (i) either m or n (or both) are odd; or (ii) both m and n are even. We'll begin with the first case.

Case (i). Suppose that m is odd in (2.1). Then we use the substitution $u = \sin x$ along with the identity

$$\sin^2 x + \cos^2 x = 1$$

to evaluate the integral.

Example 2.1.1. Evaluate the integral $\displaystyle\int \cos^3 x \sin^4 x \, dx$.

Solution. The substitution

$$u = \sin x, \qquad du = \cos x \, dx$$

yields

$$\begin{aligned}
\int \cos^3 x \sin^4 x \, dx &= \int \cos^2 x \sin^4 x \cos x \, dx \\
&= \int (1 - \sin^2 x) \sin^4 x \cos x \, dx \\
&= \int (1 - u^2) u^4 \, du \\
&= \int u^4 - u^6 \, du \\
&= \frac{u^5}{5} - \frac{u^7}{7} + C \\
&= \frac{\sin^5 x}{5} - \frac{\sin^7 x}{7} + C.
\end{aligned}$$

□

If n is odd in (2.1) then we use the substitution $u = \cos x$ and follow the same strategy. (If both m and n are odd, then either of the substitutions $u = \sin x$ or $u = \cos x$ will work.)

Example 2.1.2. Evaluate the integral $\displaystyle\int \cos^6 x \sin^5 x \, dx$.

Solution. This time we use the substitution

$$u = \cos x, \qquad du = -\sin x \, dx$$

to evaluate the integral:

$$\begin{aligned}
\int \cos^6 x \sin^5 x \, dx &= -\int \cos^6 x (\sin^2 x)^2 (-\sin x) \, dx \\
&= -\int \cos^6 x (1 - \cos^2 x)^2 (-\sin x) \, dx \\
&= -\int u^6 (1 - u^2)^2 \, du \\
&= -\int u^6 - 2u^8 + u^{10} \, du \\
&= -\left(\frac{u^7}{7} - \frac{2u^9}{9} + \frac{u^{11}}{11} \right) + C \\
&= -\frac{\cos^7 x}{7} + \frac{2\cos^9 x}{9} - \frac{\cos^{11} x}{11} + C.
\end{aligned}$$

□

Case (ii). The case where *both* m and n are even in (2.1) requires an entirely different approach. This time we use the identities

$$\cos^2 x = \frac{1+\cos 2x}{2} \qquad \text{and} \qquad \sin^2 x = \frac{1-\cos 2x}{2} \tag{2.2}$$

to change integral (2.1) into a sum of integrals of the form

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

We then repeat the methods of Case (i) or Case (ii) until each integral in the sum is easy to compute.

Example 2.1.3. Evaluate $\displaystyle\int \sin^2 x\,dx$.

Solution. The second identity in (2.2) gives

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

□

The next example is much harder.

Example 2.1.4. Evaluate $\displaystyle\int \sin^2 x\cos^4 x\,dx$.

Solution. The identities (2.2) give

$$\begin{aligned}\int \sin^2 x\cos^4 x\,dx &= \int \left(\frac{1-\cos 2x}{2}\right)\left(\frac{1+\cos 2x}{2}\right)^2 dx\\ &= \frac{1}{8}\int (1-\cos 2x)(1+\cos 2x)^2\,dx\\ &= \frac{1}{8}\int 1+\cos 2x - \cos^2 2x - \cos^3 2x\,dx\\ &= \frac{x}{8} + \frac{\sin 2x}{16} - \frac{1}{8}\int \cos^2 2x\,dx - \frac{1}{8}\int \cos^3 2x\,dx.\end{aligned} \tag{2.3}$$

The first integrand of (2.3) is an even power of $\cos 2x$ and is evaluated using the first identity in (2.2):

$$\begin{aligned}\int \cos^2 2x\,dx &= \frac{1}{2}\int 1+\cos 4x\,dx\\ &= \frac{x}{2} + \frac{\sin 4x}{8} + C_1.\end{aligned}$$

The second integrand of (2.3) is an odd power of $\cos 2x$. The substitution

$$u = \sin 2x, \qquad du = 2\cos 2x\,dx$$

gives

$$\begin{aligned}\int \cos^3 2x\,dx &= \frac{1}{2}\int (1-\sin^2 2x)2\cos 2x\,dx\\ &= \frac{1}{2}\int 1-u^2\,du\\ &= \frac{u}{2}-\frac{u^3}{6}+C_2\\ &= \frac{\sin 2x}{2}-\frac{\sin^3 2x}{6}+C_2.\end{aligned}$$

Following from (2.3) we obtain

$$\begin{aligned}\int \sin^2 x\cos^4 x\,dx &= \frac{x}{8}+\frac{\sin 2x}{16}-\frac{x}{16}-\frac{\sin 4x}{64}-\frac{\sin 2x}{16}+\frac{\sin^3 2x}{48}+C\\ &= \frac{x}{16}-\frac{\sin 4x}{64}+\frac{\sin^3 2x}{48}+C.\end{aligned}$$

□

2.1.2 Integrating multiple angles of sine and cosine

The next class of trigonometric integrals consists of integrals of the form

$$\int \cos mx\sin nx\,dx, \qquad \int \cos mx\cos nx\,dx \qquad \text{or} \qquad \int \sin mx\sin nx\,dx, \tag{2.4}$$

where m and n are real numbers. These have many applications, including analysis of waves, musical harmonics and distribution of heat in solids. Such applications are discussed in some second year courses.

To evaluate the integrals in (2.4) we need the following trigonometric identities.

Lemma 2.1.5. *Suppose that A and B are real numbers. Then*

$$\sin A\cos B = \tfrac{1}{2}\big(\sin(A+B)+\sin(A-B)\big) \tag{2.5}$$
$$\cos A\cos B = \tfrac{1}{2}\big(\cos(A-B)+\cos(A+B)\big) \tag{2.6}$$
$$\sin A\sin B = \tfrac{1}{2}\big(\cos(A-B)-\cos(A+B)\big) \tag{2.7}$$

Proof. We only prove the first identity; the other proofs are similar. We begin with the sum and difference formulae

$$\begin{aligned}\sin(A+B) &= \sin A\cos B+\cos A\sin B\\ \sin(A-B) &= \sin A\cos B-\cos A\sin B.\end{aligned}$$

Adding the two identities gives

$$\sin(A+B)+\sin(A-B) = 2\sin A\cos B,$$

whereupon dividing by 2 establishes (2.5). □

Example 2.1.6. Evaluate $\int \cos 5x\cos 3x\,dx$.

Solution. Identity (2.6) implies that

$$\begin{aligned}\int \cos 5x \cos 3x \, dx &= \frac{1}{2}\int \cos(5x-3x)+\cos(5x+3x)\, dx \\ &= \frac{1}{2}\int \cos(2x)+\cos(8x)\, dx \\ &= \frac{\sin 2x}{4}+\frac{\sin 8x}{16}+C.\end{aligned}$$

□

2.1.3 Integrating powers of tan and sec

Students will not be expected to evaluate difficult integrals involving powers of tan and sec. However, it is expected that they will be able to use the facts that

$$\tan^2 x + 1 = \sec^2 x, \qquad \frac{d}{dx}\tan x = \sec^2 x \qquad \text{and} \qquad \frac{d}{dx}\sec x = \tan x \sec x$$

to find suitable substitutions or strategies.

Example 2.1.7. Evaluate $\displaystyle\int \tan^2 x\, dx$.

Proof. The Pythagorean identity gives

$$\begin{aligned}\int \tan^2 x\, dx &= \int \sec^2 x - 1\, dx \\ &= \tan x - x + C.\end{aligned}$$

□

Example 2.1.8. Evaluate $\displaystyle\int \sec^4 x \tan x\, dx$.

Proof. The substitution

$$u = \sec x, \qquad du = \sec x \tan x\, dx$$

yields

$$\begin{aligned}\int \sec^4 x \tan x\, dx &= \int u^3\, du \\ &= \frac{\sec^4 x}{4} + C.\end{aligned}$$

(Of course, one can always by-pass the substitution and integrate by inspection.) □

2.2 Reduction formulae

We begin with an example.

Example 2.2.1. Suppose that I_n is defined by

$$I_n = \int_0^{\pi/4} \tan^n x \, dx$$

whenever $n \geq 0$. Show that

$$I_n = \frac{1}{n-1} - I_{n-2} \qquad \forall n \geq 2. \tag{2.8}$$

Hence evaluate

$$\int_0^{\pi/4} \tan^6 x \, dx.$$

Proof. By the identity $\tan^2 x = \sec^2 x - 1$,

$$\begin{aligned}
I_n &= \int_0^{\pi/4} \tan^{n-2} x \tan^2 x \, dx \\
&= \int_0^{\pi/4} \tan^{n-2} x (\sec^2 x - 1) \, dx \\
&= \int_0^{\pi/4} \tan^{n-2} x \sec^2 x \, dx - \int_0^{\pi/4} \tan^{n-2} x \, dx \\
&= \left[\frac{\tan^{n-1} x}{n-1} \right]_0^{\pi/4} - I_{n-2} \\
&= \frac{1}{n-1} - I_{n-2}
\end{aligned}$$

as required.

Using (2.8), we see that

$$\begin{aligned}
\int_0^{\pi/2} \tan^6 x \, dx &= I_6 \\
&= \frac{1}{5} - I_4 \\
&= \frac{1}{5} - \left(\frac{1}{3} - I_2 \right) \\
&= \frac{1}{5} - \frac{1}{3} + \left(\frac{1}{1} - I_0 \right) \\
&= \frac{1}{5} - \frac{1}{3} + \frac{1}{1} - \int_0^{\pi/4} dx \\
&= \frac{1}{5} - \frac{1}{3} + \frac{1}{1} - \frac{\pi}{4} \\
&= \frac{13}{15} - \frac{\pi}{4}.
\end{aligned}$$

(Note that I_0 must be evaluated directly, since formula (2.8) is only valid when $n \geq 2$.) □

Formula (2.8) is an example of a *reduction formula*, since it expresses an integral in terms of a 'smaller' integral of the same type. As illustrated above, once a reduction formula is known, integrals of that type may be evaluated rapidly. Although not the case with the previous example, most reduction formulae are proved using integration by parts.

Example 2.2.2. Suppose that

$$I_n = \int \sin^n x\, dx$$

whenever $n \geq 0$. Show that

$$I_n = -\frac{\sin^{n-1} x \cos x}{n} + \frac{n-1}{n} I_{n-2} \qquad \forall n \geq 2.$$

Solution. If we apply integration by parts with

$$\begin{aligned} u &= \sin^{n-1} x & v &= -\cos x \\ u' &= (n-1)\sin^{n-2} x \cos x & v' &= \sin x \end{aligned}$$

then

$$\begin{aligned} I_n &= \int \sin^{n-1} x \sin x\, dx \\ &= -\sin^{n-1} x \cos x + (n-1) \int \sin^{n-2} x \cos x \cos x\, dx \\ &= -\sin^{n-1} x \cos x + (n-1) \int \sin^{n-2} x (1 - \sin^2 x)\, dx \\ &= -\sin^{n-1} x \cos x + (n-1) \int \sin^{n-2} x\, dx - (n-1) \int \sin^n x\, dx \\ &= -\sin^{n-1} x \cos x + (n-1) I_{n-2} - (n-1) I_n. \end{aligned}$$

If we gather the I_n terms to the left-hand side then

$$nI_n = -\sin^{n-1} x \cos x + (n-1) I_{n-2}.$$

Dividing both sides by n gives the result. □

In the final example, the reduction formula has two parameters (m and n) instead of one.

Example 2.2.3. Suppose that

$$I_{m,n} = \int_0^{\pi/2} \cos^m x \sin^n x\, dx \tag{2.9}$$

whenever m and n are nonnegative integers.

(a) [**X**] Show that

$$I_{m,n} = \begin{cases} \left(\frac{m-1}{m+n}\right) I_{m-2,n} & \text{provided that } m \geq 2 \\ \left(\frac{n-1}{m+n}\right) I_{m,n-2} & \text{provided that } n \geq 2. \end{cases} \tag{2.10}$$

(b) [**R**] Using the result of (a), evaluate $\displaystyle\int_0^{\pi/2} \cos^4 x \sin^6 x\, dx$.

Solution. (a) Integration by parts with

$$u = \cos^{m-1} x \qquad v = \frac{\sin^{n+1} x}{n+1}$$
$$u' = -(m-1)\cos^{m-2} x \sin x \qquad v' = \sin^n x \cos x$$

gives

$$\begin{aligned} I_{m,n} &= \left[\frac{\cos^{m-1} x \sin^{n+1} x}{n+1}\right]_0^{\pi/2} + \frac{m-1}{n+1}\int_0^{\pi/2} \cos^{m-2} x \sin^{n+2} x\, dx \\ &= \frac{m-1}{n+1}\int_0^{\pi/2} \sin^n x \cos^{m-2} x(1-\cos^2 x)\, dx \\ &= \frac{m-1}{n+1} I_{m-2,n} - \frac{m-1}{n+1} I_{m,n}. \end{aligned}$$

By bringing the $I_{m,n}$ terms to the left-hand side and rearranging, the first formula is obtained. The second formula is proved similarly.

(b) The first formula in (2.10) allows us to reduce the first parameter:

$$\begin{aligned} \int_0^{\pi/2} \cos^4 x \sin^6 x\, dx &= I_{4,6} \\ &= \frac{3}{10} I_{2,6} \\ &= \frac{3}{10}\cdot\frac{1}{8} I_{0,6}. \end{aligned}$$

The second formula in (2.10) allows us to reduce the second parameter:

$$\begin{aligned} \int_0^{\pi/2} \cos^4 x \sin^6 x\, dx &= \frac{3}{10}\cdot\frac{1}{8} I_{0,6} \\ &= \frac{3}{10}\cdot\frac{1}{8}\cdot\frac{5}{6} I_{0,4} \\ &= \frac{3}{10}\cdot\frac{1}{8}\cdot\frac{5}{6}\cdot\frac{3}{4} I_{0,2} \\ &= \frac{3}{10}\cdot\frac{1}{8}\cdot\frac{5}{6}\cdot\frac{3}{4}\cdot\frac{1}{2} I_{0,0}. \end{aligned}$$

Finally, we note from (2.9) that $I_{0,0} = \frac{\pi}{2}$. Hence

$$\int_0^{\pi/2} \cos^4 x \sin^6 x\, dx = \frac{3}{10}\cdot\frac{1}{8}\cdot\frac{5}{6}\cdot\frac{3}{4}\cdot\frac{1}{2}\cdot\frac{\pi}{2} = \frac{3\pi}{512},$$

completing the problem. □

2.2.1 [X] Application: the irrationality of π

In this subsection, we use a reduction formula to prove that π is an irrational number. Students studying MATH1231 may want to skip to the next section if this does not interest them.

Although the number π has been studied for over 2000 years, it was only in 1770 that it was shown (by Johann Heinrich Lambert) that π is an irrational number. The proof we give is simpler than Lambert's proof and is similar to a proof discovered in the twentieth century. The main idea (as with many other irrationality proofs) is to assume that π is a rational number and find a contradition. In our proof, a contradiction arises by showing that a certain definite integral, which is known to lie in the interval $(0,1)$, is an integer if one assumes that π is a rational number.

Suppose that q is a positive integer. For each natural number n, define the integral I_n by

$$I_n = \frac{q^{2n}}{n!}\int_{-\pi/2}^{\pi/2}\left(\frac{\pi^2}{4}-x^2\right)^n \cos x\,dx. \tag{2.11}$$

Lemma 2.2.4. *Suppose that q is a positive integer and I_n is defined as above. If $n \geq 2$ then*

$$I_n = (4n-2)q^2 I_{n-1} - q^4\pi^2 I_{n-2}. \tag{2.12}$$

Moreover,

$$I_1 = 4q^2 \qquad \textit{and} \qquad I_0 = 2.$$

Proof. We give an outline proof only; students should be able to fill in the details. Suppose that $n \geq 2$. Integration by parts with

$$u = \left(\frac{\pi^2}{4}-x^2\right)^n \qquad v = \sin x$$

$$u' = -2nx\left(\frac{\pi^2}{4}-x^2\right)^{n-1} \qquad v' = \cos x$$

yields

$$I_n = \frac{2nq^{2n}}{n!}\int_{-\pi/2}^{\pi/2} x\left(\frac{\pi^2}{4}-x^2\right)^{n-1}\sin x\,dx.$$

A second application of integration by parts with

$$u = x\left(\frac{\pi^2}{4}-x^2\right)^{n-1} \qquad v' = \sin x$$

gives

$$\begin{aligned}
I_n &= \frac{2q^{2n}}{(n-1)!}\int_{-\pi/2}^{\pi/2}\left(\frac{\pi^2}{4}-x^2\right)^{n-1}\cos x\,dx - \frac{4q^{2n}}{(n-2)!}\int_{-\pi/2}^{\pi/2} x^2\left(\frac{\pi^2}{4}-x^2\right)^{n-2}\cos x\,dx \\
&= 2q^2 I_{n-1} - \frac{4q^{2n}}{(n-2)!}\int_{-\pi/2}^{\pi/2} x^2\left(\frac{\pi^2}{4}-x^2\right)^{n-2}\cos x\,dx.
\end{aligned}$$

In the right-most integrand, write x^2 as $\dfrac{\pi^2}{4}-\left(\dfrac{\pi^2}{4}-x^2\right)$. Hence

$$\begin{aligned}
I_n &= 2q^2 I_{n-1} - \frac{4q^{2n}}{(n-2)!}\left(\frac{\pi^2}{4}\int_{-\pi/2}^{\pi/2}\left(\frac{\pi^2}{4}-x^2\right)^{n-2}\cos x\,dx - \int_{-\pi/2}^{\pi/2}\left(\frac{\pi^2}{4}-x^2\right)^{n-1}\cos x\,dx\right) \\
&= 2q^2 I_{n-1} - q^4\pi^2 I_{n-2} + 4(n-1)q^2 I_{n-1}.
\end{aligned}$$

If we gather both I_{n-1} terms then we obtain (2.12) as required. The proof that $I_1 = 4q^2$ and $I_0 = 2$ is straightforward and is left as an exercise. $\square$

The next lemma will be used to show that $0 < I_n < 1$ for sufficiently large n.

Lemma 2.2.5. *If $a > 0$ then* $\lim_{n\to\infty} \dfrac{a^n}{n!} = 0$.

Proof. Take N to be any integer greater than $2a$. If $n > N$ then

$$\begin{aligned}\frac{a^n}{n!} &= \frac{a^N}{N!} \cdot \frac{a^{n-N}}{(N+1)(N+2)\dots n} \\ &= \frac{a^N}{N!} \cdot \frac{a}{N+1} \cdot \frac{a}{N+2} \cdots \frac{a}{n}.\end{aligned}$$

Now $\dfrac{a^N}{N!}$ is some fixed number and

$$\frac{a}{N+1} < \frac{1}{2}, \quad \frac{a}{N+2} < \frac{1}{2}, \quad \dots, \quad \frac{a}{n} < \frac{1}{2},$$

so

$$0 < \frac{a^n}{n!} < \frac{a^N}{N!}\left(\frac{1}{2}\right)^{n-N}.$$

As $n \to \infty$, the right hand side approaches 0 and hence $\lim_{n\to\infty} \dfrac{a^n}{n!} = 0$ by a sequence version of the pinching theorem.

□

Theorem 2.2.6. *The number π is irrational.*

Proof. Suppose that $\pi = \frac{p}{q}$ where p and q are positive integers and consider the integral I_n defined by (2.11) whenever $n \geq 0$.

First we argue by mathematical induction that I_n is an integer for every value of n. By Lemma 2.2.4, I_0 and I_1 are integers. Suppose inductively that I_{k-2} and I_{k-1} are integers whenever $k \geq 2$. By (2.12) and the assumption that $\pi = p/q$,

$$I_k = (4n-2)q^2 I_{k-1} - p^2 q^2 I_{k-2}$$

and so I_k is also an integer. Hence I_n is an integer whenever $n \geq 0$.

On the other hand, it is not hard to see that

$$0 < \left(\frac{\pi^2}{4} - x^2\right)^n \cos x \leq \left(\frac{\pi^2}{4}\right)^n$$

whenever $n \geq 0$ and $-\frac{\pi}{2} < x < \frac{\pi}{2}$. Hence

$$\begin{aligned}0 < I_n &< \frac{q^{2n}}{n!} \int_{-\pi/2}^{\pi/2} \left(\frac{\pi^2}{4}\right)^n dx \\ &= \frac{q^{2n}}{n!} \left(\frac{\pi^2}{4}\right)^n \pi \\ &= \frac{p}{q} \cdot \frac{(p^2/4)^n}{n!}. \qquad (2.13)\end{aligned}$$

As $n \to \infty$, the expression in (2.13) approaches 0 by Lemma 2.2.5. Hence $0 < I_n < 1$ whenever n is sufficiently large. In particular, there is a large value of n for which I_n is not an integer, giving a contradiction.

Hence we conclude that π is an irrational number. □

2.3 Trigonometric and hyperbolic substitutions

(Ref: SH10 §8.4)

Many integrals can be evaluated by finding the right substitution, but unfortunately there is no general systematic way to do this. Integrals involving square roots of quadratics often yield to trigonometric or hyperbolic substitutions.

The following table indicates which substitution can be tried for integrals containing an expression of the form $\sqrt{\pm x^2 \pm a^2}$.

Expression in integrand	Trigonometric substitution	Hyperbolic substitution
$\sqrt{a^2 - x^2}$	$x = a \sin\theta$	$x = a \tanh\theta$
$\sqrt{a^2 + x^2}$	$x = a \tan\theta$	$x = a \sinh\theta$
$\sqrt{x^2 - a^2}$	$x = a \sec\theta$	$x = a \cosh\theta$

Whether or not a trigonometric substitution is more efficient than a hyperbolic substitution depends on the particular integral. In general, trigonometric substitutions are favoured because once integration is completed in the variable θ, it is easier to restate the result in terms of x.

Example 2.3.1. Evaluate $\displaystyle\int \sqrt{1 - x^2}\, dx$.

Solution. The substitution

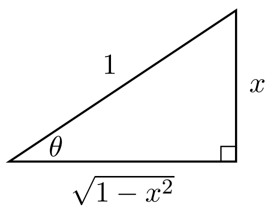
$$x = \sin\theta \qquad\qquad dx = \cos\theta\, d\theta$$

yields

$$\begin{aligned}
\int \sqrt{1-x^2}\, dx &= \int \sqrt{1-\sin^2\theta}\cos\theta\, d\theta \\
&= \int \sqrt{\cos^2\theta}\cos\theta\, d\theta && \text{(since } \sin^2\theta + \cos^2\theta = 1) \\
&= \int \cos^2\theta\, d\theta \\
&= \frac{1}{2}\int 1 + \cos 2\theta\, d\theta && \text{(by the double-angle formula for cosine)} \\
&= \frac{1}{2}\left(\theta + \frac{\sin 2\theta}{2}\right) + C \\
&= \frac{1}{2}(\theta + \sin\theta\cos\theta) + C && \text{(by the double-angle formula for sin).}
\end{aligned}$$

To state our answer in terms of x, it is easiest to draw a triangle.

$$\sin\theta = \frac{x}{1}$$



We see that $\theta = \sin^{-1} x$ and $\cos\theta = \dfrac{\sqrt{1-x^2}}{1}$. Hence

$$\int \sqrt{1-x^2}\,dx = \frac{1}{2}\left(\sin^{-1} x + x\sqrt{1-x^2}\right) + C.$$

□

Example 2.3.2. Evaluate $\displaystyle\int \frac{dx}{(4+x^2)^{3/2}}$.

Solution. By using the substitution

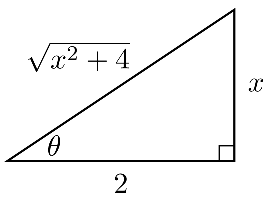
$$x = 2\tan\theta \qquad\qquad dx = 2\sec^2\theta\,d\theta$$

and the identity $\tan^2\theta + 1 = \sec^2\theta$, we have

$$\begin{aligned}
\int \frac{dx}{(4+x^2)^{3/2}} &= \int \frac{2\sec^2\theta\,d\theta}{\left(\sqrt{4\tan^2\theta+4}\,\right)^3} \\
&= \int \frac{2\sec^2\theta\,d\theta}{\left(2\sqrt{\tan^2\theta+1}\,\right)^3} \\
&= \int \frac{2\sec^2\theta\,d\theta}{(2\sec\theta)^3} \\
&= \frac{1}{4}\int \frac{d\theta}{\sec\theta} \\
&= \frac{1}{4}\int \cos\theta\,d\theta \\
&= \frac{\sin\theta}{4} + C.
\end{aligned}$$

To write the answer in terms of x, consider the following triangle.

$$\tan\theta = \frac{x}{2}$$



Thus $\sin\theta = \dfrac{x}{\sqrt{x^2+4}}$ and hence

$$\int \frac{dx}{(4+x^2)^{3/2}} = \frac{x}{4\sqrt{x^2+4}} + C.$$

□

Example 2.3.3. Use the substitution $x = 3\cosh\theta$ to evaluate $\displaystyle\int \frac{x^3\,dx}{\sqrt{x^2-9}}$.

Solution. The substitution

$$x = 3\cosh\theta \qquad\qquad dx = 3\sinh\theta$$

and the identity $\cosh^2\theta - \sinh^2\theta = 1$ gives $\sqrt{x^2-9} = 3\sinh\theta$. Hence

$$\begin{aligned}
\int \frac{x^3\,dx}{\sqrt{x^2-9}} &= \int \frac{3^4\cosh^3\theta\sinh\theta\,d\theta}{\sqrt{9\cosh^2\theta-9}} \\
&= \int \frac{3^4\cosh^3\theta\sinh\theta\,d\theta}{3\sinh\theta} && \text{(since } \cosh^2\theta-\sinh^2\theta=1) \\
&= 27\int \cosh^3\theta\,d\theta \\
&= 27\int \cosh\theta\cosh^2\theta\,d\theta \\
&= 27\int \cosh\theta(1+\sinh^2\theta)\,d\theta && \text{(since } \cosh^2\theta-\sinh^2\theta=1) \\
&= 27\int 1+u^2\,dx && \text{(using the substitution } u=\sinh\theta) \\
&= 27\big(u+\tfrac{1}{3}u^3\big)+C.
\end{aligned}$$

As was observed above, $\sqrt{x^2-9} = 3\sinh\theta$ and hence

$$u = \sinh\theta = \tfrac{1}{3}\sqrt{x^2-9}.$$

Therefore

$$\begin{aligned}
\int \frac{x^3\,dx}{\sqrt{x^2-9}} &= 27\big(u+\tfrac{1}{3}u^3\big)+C \\
&= 9\sqrt{x^2-9}+\tfrac{1}{3}(\sqrt{x^2-9})^3+C.
\end{aligned}$$

□

Exercise: Evaluate the integral in Example 2.3.3 by using an appropriate trigonometric substitution.

2.4 Integrating rational functions

(Ref: SH10 §8.5)

The main result of this section is that every rational function has an antiderivative among the elementary functions. Moreover, there is a systematic way of finding this antiderivative.

Before we begin, we remind the reader that a rational function f is of the form

$$f(x) = \frac{p(x)}{q(x)},$$

where p and q are polynomials. We say that f is *proper* if the **degree** of the denominator q is greater than the **degree** of the numerator p. We say that f is *improper* if the degree of the denominator q is less than or equal to the degree of the numerator p. We say that a quadratic polynomial is *irreducible* if it has no real linear factors. (Equivalently, a quadratic ax^2+bx+c is irreducible if its discriminant b^2-4ac is negative.)

Before articulating the general strategy for integrating a rational function, we revise some known tactics for integrating simpler examples.

Example 2.4.1. Evaluate $\displaystyle\int \frac{x}{x^2+2x+10}\,dx$.

Solution. The first tactic is to rewrite integrand so that the derivative of the denominator is sitting on the numerator:

$$\begin{aligned}\int \frac{x}{x^2+2x+10}\,dx &= \frac{1}{2}\int \frac{2x}{x^2+2x+10}\,dx \\ &= \frac{1}{2}\int \frac{(2x+2)-2}{x^2+2x+10}\,dx \\ &= \frac{1}{2}\int \frac{2x+2}{x^2+2x+10} - \frac{2}{x^2+2x+10}\,dx.\end{aligned}$$

The first term can now be integrated using the ln function. To integrate the second term, we complete the square in the denominator:

$$\begin{aligned}x^2+2x+10 &= x^2+2x+1+9 \\ &= (x+1)^2+3^2.\end{aligned}$$

Hence

$$\begin{aligned}\int \frac{x}{x^2+2x+10}\,dx &= \frac{1}{2}\int \frac{2x+2}{x^2+2x+10}\,dx - \int \frac{1}{(x+1)^2+3^2}\,dx \\ &= \frac{1}{2}\ln|x^2+2x+10| - \frac{1}{3}\tan^{-1}\left(\frac{x+1}{3}\right) + C.\end{aligned}$$

□

The integrand of Example 2.4.1 is a proper rational function whose denominator is an irreducible quadratic. Any such function can be integrated using the techniques illustrated in that example. We turn now to study a general strategy for integrating *any* rational function.

2.4.1 The overall strategy

In this subsection we give an overview of the approach to integrating rational functions. The basic procedure is summarised below, afterwards illustrated with an example.

1. If the rational function is improper, then use polynomial division to write f as the sum of a polynomial and a proper rational function. Since the polynomial is easy to integrate, we need only focus on integrating a proper rational function.

2. It can be shown using algebra that every proper rational function f can be written as a unique sum of functions of the form

$$\frac{A}{(x-a)^k} \qquad \text{and} \qquad \frac{Bx+C}{(x^2+bx+c)^k}, \tag{2.14}$$

   where the quadratic x^2+bx+c is *irreducible*. This sum is called the *partial fractions decomposition* of f. We discuss how to find the partial fractions decomposition in the next subsection.

3. Now we only need to integrate functions of the form given by (2.14). By completing the square, using a substitution or performing simple algebraic manipulation, these can be integrated by the standard formulae

$$\begin{aligned}\int x^k\,dx &= \frac{x^{k+1}}{k+1} + C, \qquad k \neq -1\\ \int \frac{g'(x)}{g(x)}\,dx &= \ln|g(x)| + C\\ \int \frac{dx}{a^2+x^2}\,dx &= \frac{1}{a}\tan^{-1}\frac{x}{a} + C.\end{aligned}$$

Example 2.4.2. Find $\displaystyle\int \frac{x^4-5x^3+12x^2-21x+35}{x^3-3x^2+4x-12}\,dx.$

Solution. Denote the integrand by $f(x)$.

Step 1. Note that f is improper. So polynomial division gives

$$\begin{array}{r l}
 & x + 2 \\ \cline{2-2}
x^3-3x^2+4x-12 \,\big) & x^4-5x^3+12x^2-21x+35 \\
 & \underline{x^4-3x^3+4x^2-12x\phantom{{}+35}} \\
 & -2x^3+8x^2-9x+35 \\
 & \underline{-2x^3+6x^2-8x+24} \\
 & 2x^2-x+11
\end{array}$$

and hence

$$f(x) = x + 2 + \frac{2x^2-x+11}{x^3-3x^2+4x-12}.$$

Note that rational expression on the far right-hand side is proper.

Step 2. The partial fractions decomposition of $\dfrac{2x^2-x+11}{x^3-3x^2+4x-12}$ is given by

$$\frac{2x^2-x+11}{x^3-3x^2+4x-12} = \frac{2}{x-3} - \frac{1}{x^2+4}. \tag{2.15}$$

(Note that the quadratic x^2+4 is irreducible.) It is not hard to verify that (2.15) is true; the question is, How does one find such a decomposition? We answer this question in Subsection 2.4.2.

Step 3. The results of Steps 1 and 2 give

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

To integrate f, we need only integrate each term in the sum. Hence

$$\begin{aligned}\int \frac{x^4-5x^3+12x^2-21x+35}{x^3-3x^2+4x-12}\,dx &= \int\left(x-2+\frac{2}{x-3}-\frac{1}{x^2+4}\right)dx\\ &= \tfrac{1}{2}x^2 - 2x + 2\ln|x-3| - \tfrac{1}{2}\tan^{-1}\tfrac{x}{2} + C,\end{aligned}$$

completing the problem. □

In the next subsection, we focus on finding the partial fractions decomposition of a proper rational function.

2.4.2 Partial fractions decompositions

To find the partial fractions decomposition of a proper rational function $\frac{p}{q}$, we factorise the denominator q as much as possible; that is, we express q as a product of real linear factors and real irreducible quadratic factors. The form of the partial fractions decomposition is determined by this factorisation. There are several cases, depending on the type of factorisation.

Case 1: The denominator splits into distinct linear factors. Examples of two such rational functions and the form of their partial fractions decompositions are given below:

$$\begin{aligned}\frac{x-3}{(x-1)(x-2)} &= \frac{A}{x-1}+\frac{B}{x-2}\\ \frac{x^2-x+7}{x(2x+1)(x-3)} &= \frac{A}{x}+\frac{B}{2x+1}+\frac{C}{x-3}.\end{aligned}$$

The constants A, B and C in each case can be determined using the following method.

Example 2.4.3. Find the partial fractions decomposition of $\dfrac{7x-1}{x^2-2x-3}$.

Solution. By factorising we find that $x^2-2x-3=(x-3)(x+1)$. So the partial fractions decomposition takes the form

$$\frac{7x-1}{(x-3)(x+1)} = \frac{A}{x-3}+\frac{B}{x+1},$$

where A and B are constants to be determined. To find A and B, multiply through by $(x-3)(x+1)$ to obtain the polynomial equation

$$7x-1 = A(x+1)+B(x-3) \qquad \forall x\in\mathbb{R}.$$

Since this identity is true for all values x, the values of A and B are easily determined by choosing suitable values of x:

$$\begin{array}{ccccc} x=3 & \Rightarrow & 7\times 3-1 = A(3+1) & \Rightarrow & A=5\\ x=-1 & \Rightarrow & 7\times(-1)-1 = B(-1-3) & \Rightarrow & B=2.\end{array}$$

Hence the partial fractions decomposition is given by

$$\frac{7x-1}{x^2-2x-3} = \frac{7x-1}{(x-3)(x+1)} = \frac{5}{x-3}+\frac{2}{x+1}.$$

This may be easily verified by rewritting the right-hand side over a common denominator and simplifying. □

Case 2: The denominator has a repeated linear factor. Examples of two such rational functions and the form of their partial fractions decompositions are given below:

$$\begin{aligned}\frac{x^2+1}{(x+4)^3} &= \frac{A}{x+4}+\frac{B}{(x+4)^2}+\frac{C}{(x+4)^3}\\ \frac{x^2-2}{(x-1)(x-2)^2} &= \frac{A}{x-1}+\frac{B}{x-2}+\frac{C}{(x-2)^2}.\end{aligned}$$

Note carefully how the repeated factors appear on the right-hand side. The constants A, B and C in each case can be determined using the following method.

Example 2.4.4. Find the partial fractions decomposition of $\dfrac{x^2-3x+8}{x(x-2)^2}$.

Solution. The partial fractions decomposition takes the form

$$\frac{x^2-3x+8}{x(x-2)^2} = \frac{A}{x} + \frac{B}{x-2} + \frac{C}{(x-2)^2},$$

where A, B and C are constants. To find these constants, we multiply through by $x(x-2)^2$ to obtain

$$x^2-3x+8 = A(x-2)^2 + Bx(x-2) + Cx \qquad \forall x \in \mathbb{R}.$$

Now substitute the obvious values for x to determine the values of A and C:

$$\begin{array}{lllll} x=2 & \Rightarrow & 6=2C & \Rightarrow & C=3 \\ x=0 & \Rightarrow & 8=4A & \Rightarrow & A=2. \end{array}$$

To determine B, we can substitute any other value for x. However, it is best to choose a small integer to keep the arithmetic simple:

$$\begin{array}{lllll} x=1 & \Rightarrow & 6=A-B+C & \Rightarrow & B=A+C-6=-1. \end{array}$$

(Alternately, one can find B by noting that

$$x^2-3x+8 = 2(x-2)^2 + Bx(x-2) + 3x \qquad \forall x \in \mathbb{R}$$

and comparing coefficients for x^2.) Hence we obtain the partial fractions decomposition

$$\frac{x^2-3x+8}{x(x-2)^2} = \frac{2}{x} - \frac{1}{x-2} + \frac{3}{(x-2)^2}.$$

□

Case 3: The denominator has an irreducible quadratic factor. Examples of two such rational functions and the form of their partial fractions decompositions are given below:

$$\frac{x^2+x}{(x-1)(x^2+9)} = \frac{A}{x-1} + \frac{Bx+C}{x^2+9}$$
$$\frac{x^3-2x+4}{(x^2+5)(x^2+x+1)} = \frac{Ax+B}{x^2+5} + \frac{Cx+D}{x^2+x+1}$$

Note carefully how the irreducible quadratic appears on the right-hand side. As before, the constants A, B, C and D in each case can be determined by algebra.

Example 2.4.5. Find the partial fractions decomposition of $\dfrac{4x^2+2x+1}{(x+1)(x^2+x+1)}$.

Solution. The partial fractions decomposition takes the form

$$\frac{4x^2+2x+1}{(x+1)(x^2+x+1)} = \frac{A}{x+1} + \frac{Bx+C}{x^2+x+1},$$

where A, B and C are constants. Multiplying through by $(x+1)(x^2+x+1)$ gives

$$4x^2+2x+1 = A(x^2+x+1) + (Bx+C)(x+1) \qquad \forall x \in \mathbb{R}.$$

Now substitute suitable values for x:

$$\begin{array}{lllll} x=-1 & \Rightarrow & 3=A & \Rightarrow & A=3 \\ x=0 & \Rightarrow & 1=A+C & \Rightarrow & C=1-A=-2 \\ x=1 & \Rightarrow & 7=3A+2(B+C) & \Rightarrow & B=1. \end{array}$$

(Alternatively, after finding A, we could compare the coefficients of the x^2 terms on both sides to deduce that $B=1$, and compare the constant terms to deduce that $C=-2$.) Hence

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

is the partial fractions decomposition. □

Case 4: The denominator has repeated irreducible quadratic factor. This case rarely appears in first year mathematics courses because it is more computationally intensive. Nevertheless, for completeness the basic form of decomposition is illustrated below:

$$\begin{aligned} \frac{x^2+x}{(x^2+9)^3} &= \frac{Ax+B}{x^2+9}+\frac{Cx+D}{(x^2+9)^2}+\frac{Ex+F}{(x^2+9)^3} \\ \frac{x^3-2x+4}{(x-2)(x^2+x+1)^2} &= \frac{A}{x-2}+\frac{Bx+C}{x^2+x+1}+\frac{Dx+E}{(x^2+x+1)^2}. \end{aligned}$$

As before, the constants appearing in each example can be determined by algebra.

The final example tests our ability to generalise each of these cases to rational functions whose denominators have many factors of different types.

Example 2.4.6. Write down the *form* of partial fractions decomposition for the rational function given by

$$\frac{4x^4-3x^2+x-9}{x^3(x-7)(x^2+3)^2(x^2+x+2)}.$$

(You are not required to evaluate the constant coefficients.)

Solution. The partial fractions decomposition is given by

$$\frac{4x^4-3x^2+x-9}{x^3(x-7)(x^2+3)^2(x^2+x+2)}=\frac{A}{x}+\frac{B}{x^2}+\frac{C}{x^3}+\frac{D}{x-7}+\frac{Ex+F}{x^3+3}+\frac{Gx+H}{(x^2+3)^2}+\frac{Ix+J}{x^2+x+2}$$

where $A,B,\ldots,J$ are real constants. □

Remark 2.4.7. It is important to check that the denominator of the rational function has been completely factorised before writing down the form of partial fractions decomposition. In particular, one should check that every quadratic factor appearing in the factorisation is irreducible.

2.4.3 Integrating rational functions: two examples

In this subsection we illustrate how techniques discussed in the previous subsections are applied.

Example 2.4.8. Find $\displaystyle\int \frac{8x^3-12x^2-13x-5}{2x^2-3x-2}\,dx$.

Solution. We follow the steps outlined in Subsection 2.4.1. Denote the integrand by $f(x)$.

Step 1. Since f is improper, we begin with polynomial division. This gives

$$\begin{array}{r@{\,}l} & 4x \\ 2x^2 - 3x - 2\,\big) & \overline{8x^3 - 12x^2 - 13x - 5} \\ & \underline{8x^3 - 12x^2 - 8x\phantom{{}-5}} \\ & - 5x - 5 \end{array}$$

whence

$$\begin{aligned} f(x) &= 4x + \frac{-5x-5}{2x^2-3x-2} \\ &= 4x - \frac{5x+5}{2x^2-3x-2}. \end{aligned}$$

Step 2. To find the partial fractions decomposition of

$$\frac{5x+5}{2x^2-3x-2}$$

we factorise the denominator:

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

Hence the decomposition is given by

$$\frac{5x+5}{(2x+1)(x-2)} = \frac{A}{2x+1} + \frac{B}{x-2},$$

where A and B are constants. Multiplying through by $(2x+1)(x-2)$ gives

$$5x + 5 = A(x-2) + B(2x+1)$$

and by using the substitution $x = 2$ we deduce that $B = 3$. By comparing coefficients of x on both sides it is easy to see that $A = -1$. Hence

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

Step 3. By the previous two steps,

$$f(x) = 4x + \frac{1}{2x+1} - \frac{3}{x-2}.$$

Integrating gives

$$\int f(x)\,dx = 2x^2 + \tfrac{1}{2}\ln|2x+1| - 3\ln|x-2| + C,$$

completing our answer. $\square$

Example 2.4.9. Find $\displaystyle\int \frac{4x^2 - 15x + 29}{(x-5)(x^2-4x+13)}\,dx.$

Solution. The integrand is a proper rational function and its denominator completely factorised in the real numbers. So we immediately look for its partial fractions decomposition, which is of the form

$$\frac{4x^2-15x+29}{(x-5)(x^2-4x+13)} = \frac{A}{x-5} + \frac{Bx+C}{x^2-4x+13}$$

for some real constants A, B and C. Hence

$$4x^2-15x+29 = A(x^2-4x+13) + (Bx+C)(x-5),$$

from which appropriate substitutions allow the evaluation of the unknown constants:

$$\begin{array}{lllll} x=5 & \Rightarrow & 54 = 18A & \Rightarrow & A=3 \\ x=0 & \Rightarrow & 29 = 13A - 5C & \Rightarrow & C=2 \\ x=1 & \Rightarrow & 18 = 10A - 4(B+C) & \Rightarrow & B=1. \end{array}$$

Hence

$$\frac{4x^2-15x+29}{(x-5)(x^2-4x+13)} = \frac{3}{x-5} + \frac{x+2}{x^2-4x+13}.$$

The first term of the decomposition is easy to integrate. We therefore focus on the second term:

$$\begin{aligned} \int \frac{x+2}{x^2-4x+13}\,dx &= \frac{1}{2}\int \frac{2x+4}{x^2-4x+13}\,dx \\ &= \frac{1}{2}\left[\int \frac{2x-4}{x^2-4x+13}\,dx + \int \frac{8}{x^2-4x+13}\,dx\right] \\ &= \frac{1}{2}\int \frac{2x-4}{x^2-4x+13}\,dx - 4\int \frac{1}{x^2-4x+13}\,dx \\ &= \frac{1}{2}\ln|x^2-4x+13| - 4\int \frac{1}{(x-2)^2+9}\,dx \\ &= \frac{1}{2}\ln|x^2-4x+13| - \frac{4}{3}\tan^{-1}\left(\frac{x-2}{3}\right) + C. \end{aligned}$$

Putting everything together gives

$$\int \frac{4x^2-15x+29}{(x-5)(x^2-4x+13)}\,dx = 3\ln|x-5| + \frac{1}{2}\ln|x^2-4x+13| - \frac{4}{3}\tan^{-1}\left(\frac{x-2}{3}\right) + C.$$

□

2.5 Other substitutions

(Ref: SH10 §8.6)

The method of partial fractions allows us, in principle, to find an antiderivative, among the elementary functions, for any given rational function. So given a 'non-standard' integral, a sound technique for integration is to look for a substitution that will convert the given integral into the integral of rational function. Choosing a good substitution is often a matter of experience and a little inspiration.

Example 2.5.1. Evaluate the following integrals.

(a) $\displaystyle\int \frac{dx}{1+x^{1/4}}$

(b) $\displaystyle\int \frac{x^{1/2}}{x^{1/3}+x^{1/4}}\,dx$

(c) $\displaystyle\int \frac{dx}{\sqrt{e^{2x}-1}}$

Proof. (a) The aim is to replace the fractional power $x^{1/4}$ with something more convenient. The obvious substitution to use is $x = u^4$, which leads to the substitution $dx = 4u^3\,du$. Hence

$$\int \frac{dx}{1+x^{1/4}} = \int \frac{4u^3\,du}{1+u}$$

and we now have the integral of a rational function. Polynomial division gives

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

(This result may also be obtained by writing

$$\begin{aligned} u^3 &= (u^3+1) - 1 \\ &= (u+1)(u^2-u+1) - 1, \end{aligned}$$

thus avoiding the use of polynomial long division.) Consequently,

$$\begin{aligned} \int \frac{dx}{1+x^{1/4}} &= 4\left(\frac{u^3}{3} - \frac{u^2}{2} + u - \ln|1+u|\right) + C \\ &= \frac{4x^{3/4}}{3} - 2x^{1/2} + 4x^{1/4} - 4\ln\left|1+x^{1/4}\right| + C. \end{aligned}$$

(b) We aim to remove the fractional powers $x^{1/2}$, $x^{1/3}$ and $x^{1/4}$. The lowest common multiple of 2, 3 and 4 is 12, so we choose the substitution $x = u^{12}$. Hence

$$\begin{aligned} \int \frac{x^{1/2}}{x^{1/3}+x^{1/4}}\,dx &= \int \frac{u^6}{u^4+u^3}\,12u^{11}\,du \\ &= 12\int \frac{u^{14}}{u+1}\,du. \end{aligned}$$

From here we either use polynomial division, or the standard factorisation

$$u^n - 1 = (u+1)\big(u^{n-1} - u^{n-2} + u^{n-3} - \cdots + (-1)^{n+1}\big)$$

to obtain

$$\int \frac{x^{1/2}}{x^{1/3}+x^{1/4}}\,dx = 12\int u^{13} - u^{12} + u^{11} - \ldots + u - 1 + \frac{1}{u+1}\,du.$$

It is easy to evaluate the integral from here.

(c) One option is to use the substitution

$$u = e^x \qquad du = e^x\,dx$$

so that

$$\begin{aligned}\int \frac{dx}{\sqrt{e^{2x}-1}} &= \int \frac{e^x\,dx}{e^x\sqrt{e^{2x}-1}}\\ &= \int \frac{du}{u\sqrt{u^2-1}}.\end{aligned}$$

From here the integral can be evaluated using the substitution $u = \sec\theta$ or $u = \cosh\theta$. This is left as an exercise.

A better approach is to remove the square root from the very first substitution. The substitution $u^2 = e^{2x} - 1$ implies that

$$2u\frac{du}{dx} = 2e^{2x} = 2(u^2+1),$$

which leads to the substitution

$$dx = \frac{u\,du}{u^2+1}.$$

Hence

$$\begin{aligned}\int \frac{dx}{\sqrt{e^{2x}-1}} &= \int \frac{u\,du}{u(u^2+1)}\\ &= \int \frac{du}{u^2+1}\\ &= \tan^{-1} u + C\\ &= \tan^{-1}\sqrt{e^{2x}-1} + C.\end{aligned}$$

Thus the second method is more efficient than the first. □

As seen in the last example, there may more than one method to evaluate a given integral. Choosing the most efficient substitution to use is not always easy, but intuition can be developed with time, experience and practice. The next example is therefore left to the student as an exercise.

Example 2.5.2. Evaluate $\displaystyle\int_0^1 \frac{x^3}{(4+x^2)^{5/2}}\,dx$ by

(i) using the substitution $x = 2\tan\theta$ (since the integrand involves $\sqrt{4+x^2}$);

(ii) using the substitution $x = 2\sinh\theta$ (since the integrand involves $\sqrt{4+x^2}$);

(iii) using the substitution $u^2 = 4 + x^2$ (aiming for a rational function);

(iv) using the substitution $u = 4 + x^2$.

Which method works best?

2.6 Maple notes

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

`convert(f, parfrac, x);` performs a partial fraction decomposition of the rational function `f` in the variable `x`. For example,

```
> convert(x^2/(x+2), parfrac, x);
```

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

```
> convert(x/(x-b)^2, parfrac, x);
```

$$\frac{b}{(x-b)^2} + \frac{1}{x-b}$$

Chapter 3

Ordinary differential equations

In many practical applications (in physics, economics, social sciences, engineering, applied science, mathematics and so on), information is known about the relationship between a quantity and its rates of change, but one may not have an exact formula for the quantity itself. For example, a simple population model states that the rate of change of a population, at any given time, is proportional to size of the population itself. If we write $P(t)$ for the population at time t, then we arrive at the equation

$$\frac{dP}{dt} = kP,$$

where k is the constant of proportionality. An equation, such as the one given above, which involves one (or more) of the derivatives of a function, is called a *differential equation*. Some other simple examples include

- $\dfrac{d^2x}{dx^2} = -k^2x$, which is used to describe the displacement x from the origin of a particle undergoing simple harmonic motion;
- $\dfrac{dT}{dt} = k(T - 20)$, which is used to describe how the temperature T of an object changes in room temperature; and
- $\dfrac{dx}{dy} - 0.08y = 0.05(60000 + 1000t)$, which is used to describe how the amount y (in dollars) of a particular investment changes in time (see Example 3.4.3).

If possible, the aim from here is to find an explicit formula (or formulae) describing the unknown function (respectively P, x, T and y in the examples above) appearing in each differential equation. The primary goal of this chapter is to examine some techniques for obtaining a formula for a function given a differential equation for that function.

3.1 An introduction

We begin with a definition.

Definition 3.1.1. An *ordinary differential equation* is an equation expressed in terms of exactly one independent variable and one (or more) of the derivatives of a function of this variable. The *order* of an ordinary differential equation is the order of the highest derivative present.

For example, the equation

$$\frac{d^3y}{dx^3} + \sin x\,\frac{dy}{dx} = 3x^2y$$

is an ordinary differential equation of order 3; the independent variable is x and y is assumed to be a function of x. The equation

$$\left(\frac{d^2x}{dt^2}\right)^{3/2} + \frac{dx}{dt} - tx = 0$$

is an ordinary differential equation of order 2; the independent variable is t and x is assumed to be a function of t.

In these notes, the term 'ordinary differential equation' will often be abbreviated as ODE. Such equations are called 'ordinary' because they involve ordinary derivatives. This is to distinguish them from differential equations that involve partial derivatives. (The study of 'partial differential equations' will be introduced in some second year courses.)

Ordinary differential equations can be written in several ways using a variety of notations. For example, each of the equations

$$\begin{aligned}\frac{d^2y}{dx^2} + 4x\,\frac{dy}{dx} &= e^x\\ f''(x) + 4xf'(x) &= e^x\\ y'' + 4xy' &= e^x\end{aligned}$$

represent the same ODE.

Definition 3.1.2. A *solution* to an nth order ordinary differential equation is a function which is n-times differentiable and satisfies the given equation.

The next example illustrates this definition as well as introducing the terms 'particular solution' and 'general solution.'

Example 3.1.3. Consider the ODE

$$\frac{dy}{dx} = x^2 + 5.$$

Then the function y, given by

$$y(x) = \frac{x^3}{3} + 5x,$$

is a solution to the ODE. Note that if $y(x) = \frac{x^3}{3} + 5x + 6$ or if $y(x) = \frac{x^3}{3} + 5x - 45$, then y is also a solution to the ODE. Each of these solutions is called a *particular solution* to the ODE. Using the mean value theorem (see Section 5.9 of the MATH1131 calculus notes), it is easily shown that *every* particular solution y to the ODE can be written in the form

$$y(x) = \frac{x^3}{3} + 5x + C, \tag{3.1}$$

where $C \in \mathbb{R}$. The family of solutions given by (3.1), where $C \in \mathbb{R}$, is called the *general solution* to the ODE.

In the above example, solution y could be expressed explicitly as a function of the independent variable x. Hence we obtained an *explicit solution* to the ODE. However, this cannot always be done, as the following example illustrates. Sometimes we must settle for an *implicit solution* to the ODE.

Example 3.1.4. Show that y, given implicitly by the equation

$$y^2 = \cos(x^2 + y^2), \tag{3.2}$$

is a particular solution to the ODE

$$2x\sin(x^2 + y^2) + \big(2y\sin(x^2 + y^2) + 2y\big)\frac{dy}{dx} = 0.$$

Proof. To verify that y solves the ODE, we first need to calculate $\frac{dy}{dx}$. Implicit differentiation of (3.2) with respect to x gives

$$2y\,\frac{dy}{dx} = -\sin(x^2 + y^2) \times \left(2x + 2y\,\frac{dy}{dx}\right)$$

(where we have used the chain rule to obtain the right-hand side). Hence

$$\big(2y + 2y\sin(x^2 + y^2)\big)\frac{dy}{dx} = -2x\sin(x^2 + y^2).$$

By simple rearrangement it is easily seen that

$$2x\sin(x^2 + y^2) + \big(2y\sin(x^2 + y^2) + 2y\big)\frac{dy}{dx} = 0,$$

and hence the ODE is satisfied. (This ODE shall be revisited again in Section 3.5, where we shall *find* the general solution, instead of merely verifying that a given function is a solution.) $\square$

For some differential equations, it may not even be possible to find an implicit solution. *If* it is possible to prove that an solution exists, then mathematicians and scientists must often settle for working with an approximate solution to the ODE. However, such issues will not concern us in this course.

3.2 Initial value problems

In most practical applications where ODEs are used, information is also known about the value of the unknown function and its derivatives at a particular point. This information, together with an ODE, forms an *initial value problem.*

> **Definition 3.2.1.** An *initial value problem* is an nth order ODE together with a set of values of the solution and its first $(n-1)$ derivatives at some fixed point x_0. These values are called the *initial conditions* of the initial value problem.

For example,

- $\dfrac{dy^2}{dx^2} + 5x\,\dfrac{dy}{dx} + y = \sin x, \qquad y(0) = 2, \quad \left.\dfrac{dy}{dx}\right|_{x=0} = 7;$
- $f'(t) - e^{2t} f(t) = 3t^2, \qquad f(1) = 5$; and
- $y''' + 3y'' + 4y = \cosh x, \qquad y''(\pi) = 2, \quad y'(\pi) = 0, \quad y(\pi) = -1$

are all initial value problems. The term 'initial value problem' is often abbreviated as IVP.

To solve an initial value problem, we usually try to find a general solution to the ODE (which is expressed using unspecified constants) and then determine the values of these constants by imposing the initial conditions.

Example 3.2.2. Solve the IVP

$$\frac{d^2y}{dx^2} = 6x, \qquad y'(0) = 2, \quad y(0) = -1.$$

Solution. Integrating the ODE once gives

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

where $C \in \mathbb{R}$. By imposing the initial condition $y'(0) = 2$, we deduce that $C = 2$. Hence

$$\frac{dy}{dx} = 3x^2 + 2.$$

Integrating again gives

$$y = x^3 + 2x + D,$$

where $D \in \mathbb{R}$. The initial condition $y(0) = -1$ implies that $D = -1$. Hence the solution y to the IVP is given by

$$y = x^3 + 2x - 1.$$

Note that this solution is valid for all x in $\mathbb{R}$; that is, the solution y is defined on $\mathbb{R}$. $\square$

Not every initial value problem is as straightforward to solve as the example above. Solving an IVP is, in general, very difficult, and the following questions arise.

(a) Does the IVP have a solution?

(b) Does it have a unique solution?

(c) If initial values are given at the point a, then how far on either side of a does the solution extend?

The following two examples show that care must be taken in answering such questions, even for IVPs that appear to be 'simple.'

Example 3.2.3. Solve the initial values problem

$$\frac{dy}{dx} = \sqrt{y}, \qquad y(0) = 0.$$

Solution. If we assume that $y(x) \neq 0$ then the ODE can be written as

$$\frac{dx}{dy} = \frac{1}{\sqrt{y}}, \tag{3.3}$$

from which we find that $x = 2\sqrt{y} + C$, where C is a real number. When $x = 0$ we have that $y = 0$ and hence $C = 0$. Rearranging gives

$$y(x) = \frac{x^2}{4}.$$

However, note that $y(x) = 0$ is also a solution to the IVP. Hence the IVP does not have a unique solution. □

Example 3.2.4. Solve the initial values problem

$$\frac{dy}{dx} = \frac{1}{x}, \qquad y(1) = 2.$$

How far does the solution extend on either side of the point 1?

Solution. First, the ODE implies that $\frac{dy}{dx}$ does not exist at 0. However, on the interval $(0, \infty)$, we obtain the general solution

$$y(x) = \ln x + C,$$

where $C \in \mathbb{R}$. The initial condition implies that $C = 2$ and so

$$y(x) = \ln x + 2$$

whenever $x > 0$. Hence we have found solution that extends to the interval $(0, \infty)$.

(Note that we could give a *family* of solutions defined on the set $\{x \in \mathbb{R} : x \neq 0\}$ by

$$y(x) = \begin{cases} \ln|x| + D & \text{if } x < 0 \\ \ln|x| + 2 & \text{if } x > 0, \end{cases}$$

where $D \in \mathbb{R}$. However, for most practical applications such a solution would not be used on $(-\infty, 0)$ because of the break in the domain of y at 0.) □

3.3 Separable ODEs

(Ref: SH10 §9.2)

A separable ODE is a differential equation where the two variables involved (say x and y) can be separated so that all the y's are on one side of the equation and all x's are on the other. We give an example and then state the general form.

Example 3.3.1. Solve the initial value problem

$$\frac{dy}{dx} = y^2(1+x^2), \qquad y(0) = 1.$$

Solution. First we separate the variables x and y to obtain

$$\frac{1}{y^2}\,dy = (1+x^2)\,dx. \tag{3.4}$$

Integrating gives

$$\int \frac{1}{y^2}\,dy = \int (1+x^2)\,dx, \tag{3.5}$$

whence

$$-\frac{1}{y} = x + \frac{x^3}{3} + C,$$

where $C \in \mathbb{R}$. This gives an *implicit* solution to the ODE.

We now impose the initial condition to evaluate the constant C. When $x = 0$ and $y = 1$ we obtain

$$-\frac{1}{1} = 0 + \frac{0^3}{3} + C,$$

whence $C = -1$ and

$$-\frac{1}{y} = x + \frac{x^3}{3} - 1.$$

Finally, in this particular example, one can easily make y the subject to obtain the *explicit* solution

$$\begin{aligned} y &= \frac{-1}{x + \frac{x^3}{3} - 1} \\ &= \frac{-3}{3x + x^3 - 3}, \end{aligned}$$

thus completing the problem. □

Remark 3.3.2. Equation (3.4) makes sense within the context of integration (see equation (3.5)). The fact that *this* kind of symbolic manipulation with dy and dx 'works' can be traced back to the chain rule, which is the basis for implicit differentiation. (To make this clear, note in the example above that

$$\frac{1}{y^2}\frac{dy}{dx} = 1 + x^2$$

and so integrating both sides with respect to x gives

$$\int \frac{1}{y^2}\frac{dy}{dx}\,dx = \int 1 + x^2\,dx.$$

Hence

$$-\frac{1}{y} = x + \frac{x^3}{3} + C, \tag{3.6}$$

which may be easily verified by (implicitly) differentiating both sides of (3.6) respect to x.) Students should not think that they can manipulate the symbols dy and dx in other ways and still obtain valid results.

In general, a separable ODE is one that can be written in the form

$$\frac{dy}{dx} = \frac{g(x)}{h(y)}. \tag{3.7}$$

To solve this equation, write

$$h(y)\,dy = g(x)\,dx,$$

and then integrate both sides to obtain an implicit solution

$$H(y) = G(x) + C,$$

where C is the constant of integration. Whenever possible, isolate y on the left-hand side to find the explicit solution. If initial conditions are given, then the constant C can be determined.

Example 3.3.3. Solve the equation

$$\sinh y \cos^2 x \frac{dy}{dx} = \tan x + 4.$$

Proof. We separate the variables to obtain

$$\sinh y\,dy = \frac{\tan x + 4}{\cos^2 x}\,dx.$$

Simplification followed by integration gives

$$\begin{aligned} \sinh y\,dy &= (\tan x \sec^2 x + 4\sec^2 x)\,dx \\ \int \sinh y\,dy &= \int (\tan x \sec^2 x + 4\sec^2 x)\,dx \\ \cosh y &= \tfrac{1}{2}\tan^2 x + 4\tan x + C, \end{aligned} \tag{3.8}$$

where $C \in \mathbb{R}$. In this case it is best to leave the solution in implicit form (3.8), since cosh is not a one-to-one function. □

We end with an application of this method to the real world.

Example 3.3.4 (Newton's law of cooling)**.** (a) Newton's law of cooling states that the rate of heat loss of a body is proportional to the difference in temperatures between the body and its surroundings. Set up an ODE to model this law and solve it.

(b) A hot object is placed into a room of temperature $20°C$. Unfortunately the object is too hot for the thermometer to measure its initial temperature. However, after 6 minutes, the temperature of the object was measured as $80°C$, and after eight minutes as $50°C$. What was the original temperature of the object?

Proof. (a) Suppose that

- T is the temperature of the object at time t,
- A is the ambient temperature (that is, the temperature of the surroundings), and
- k is the constant of proportionality.

Newton's law of cooling implies that

$$\frac{dT}{dt} = k(T - A).$$

To solve this equation, we separate variables:

$$\begin{aligned}\frac{1}{T-A}\,dT &= k\,dt \\ \int \frac{1}{T-A}\,dT &= \int k\,dt \\ \ln(T-A) &= kt + C,\end{aligned}$$

where C is the constant of integration. By taking exponentials of both sides we obtain

$$T - A = e^{kt+C}.$$

Hence

$$T = A + Ke^{kt}, \tag{3.9}$$

where $K = e^C > 0$.

(b) We have

$$A = 20, \qquad T(6) = 80, \qquad \text{and} \qquad T(8) = 50.$$

Hence (3.9) implies that

$$\begin{cases} 80 = 20 + Ke^{6k} \\ 50 = 20 + Ke^{8k}, \end{cases}$$

which can be solved to give

$$k = -\frac{\ln 2}{2} \qquad \text{and} \qquad K = 480.$$

Hence

$$T(0) = 20 + 480e^0 = 500.$$

So the initial temperature of the object was $500°C$. □

3.4 First order linear ODEs

(Ref: SH10 §9.1)

A first order linear ODE can be written in the form

$$\frac{dy}{dx} + f(x)y = g(x), \tag{3.10}$$

where f and g are given functions of a single variable x. The ODE is called *linear* since there are no non-linear terms (such as y^2, $\sin y$ or $\sqrt{y'}$) involving y or its derivative y'.

A very slick method for solving first order linear ODEs is summarised in the steps below.

1. Write the ODE in the form (3.10).

2. Calculate $e^{\int f(x)\,dx}$ (ignoring the constant of integration). We denote this by $h(x)$ and call it the *integrating factor*.

3. Multiply (3.10) by the integrating factor $h(x)$ to obtain

$$h(x)\frac{dy}{dx} + h(x)f(x)y = g(x)h(x).$$

   By using the product rule for differentiation, the left-hand side can now be rewritten so that

$$\frac{d}{dx}\big(h(x)\,y\big) = g(x)h(x)$$

   (this is easily seen in the examples that follow).

4. Integrate both sides and then rearrange for y to solve the ODE. Don't forget the constant of integration!

Example 3.4.1. Solve $\dfrac{dy}{dx} + 3y = e^{-x}$.

Solution. The ODE is already in the form (3.10) with $f(x)$ equal to 3. The integrating factor h is therefore given by

$$h(x) = e^{\int 3\,dx} = e^{3x}.$$

Multiplying the ODE by the integrating factor e^{3x} gives

$$e^{3x}\frac{dy}{dx} + 3e^{3x}y = e^{2x}.$$

By the product rule, we can contract the left hand side to obtain

$$\frac{d}{dx}\big(e^{3x}\,y\big) = e^{2x}$$

(this step is easy to check by working backwards). Integrating gives

$$e^{3x}y = \tfrac{1}{2}e^{2x} + C,$$

where $C \in \mathbb{R}$. Now divide by e^{3x} to obtain the explicit solution

$$y = \tfrac{1}{2}e^{-x} + Ce^{-3x},$$

where $C \in \mathbb{R}$. (Note that if the constant of integration is accidently omitted then we lose half the solution!) □

Example 3.4.2. Solve the IVP

$$(x-1)^3\frac{dy}{dx} + 4(x-1)^2y = x+1, \qquad y(0) = 2. \tag{3.11}$$

Solution. First rewrite the ODE into the standard form (3.10) to obtain

$$\frac{dy}{dx} + 4(x-1)^{-1}y = \frac{x+1}{(x-1)^3}. \tag{3.12}$$

The integrating factor h is given by

$$h(x) = e^{\int 4(x-1)^{-1}\,dx} = e^{4\ln(x-1)} = e^{\ln(x-1)^4} = (x-1)^4.$$

(Note that it is important to simplify $h(x)$ before proceeding with the method.)

Multiplying (3.12) by the integrating factor gives

$$(x-1)^4\frac{dy}{dx} + 4(x-1)^3y = x^2-1,$$

from which we obtain

$$\frac{d}{dx}\big((x-1)^4\,y\big) = x^2-1$$

by the product rule. Integrating gives

$$(x-1)^4y = \tfrac{1}{3}x^3 - x + C,$$

where $C \in \mathbb{R}$. To evaluate C, we impose the initial condition $y(0) = 2$ and find that

$$(0-1)^4 2 = \tfrac{1}{3}0^3 - 0 + C.$$

Hence $C = 2$. Therefore the solution of the IVP is given by

$$y = \frac{\frac{1}{3}x^3 - x + 2}{(x-1)^4} = \frac{x^3-3x+6}{3(x-1)^4}.$$

(Note that the solution is valid when $x \in (-\infty, 1)$ or $x \in (1, \infty)$ but not when $x = 1$. In fact, it is not hard to see from (3.11) that there is no real-valued function y satisfying the ODE with 1 in its domain.) □

The following example shows an application of a first order linear ODE to a real world problem.

Example 3.4.3. An investor has a salary of \$60,000 per year which is expected to increase at a rate of \$1000 per annum. Suppose that an initial deposit of \$1000 is invested in a program that pays 8% per annum, and that the investor deposits 5% of their salary each year. Find the amount invested after t years.

Solution. We will approximate the situation by assuming that interest is calculated continuously and that deposits are made continuously.

Let $y(t)$ denote the dollars invested after t tears. Then

$$\begin{array}{ccccc} \dfrac{dy}{dt} & = & 0.08y & + & 0.05(60000+1000t) \\ \\ (\text{ rate of increase of investment} & = & 8\% \text{ of investment} & + & 5\% \text{ of salary }). \end{array}$$

This is a first order linear ODE. If we rewrite the ODE as

$$\frac{dy}{dt} - 0.08y = 0.05(60000+1000t),$$

then we see that the integrating factor h is given by

$$h(t) = e^{\int -0.08\,dt} = e^{-0.08t}.$$

After multiplying the ODE by the integrating factor and contracting the left-hand side by the product rule, one obtains

$$\frac{d}{dt}(e^{-0.08t}y) = 0.05(60000 + 1000t)e^{-0.08t}.$$

Integration (where we use integration by parts for the right-hand side) and rearrangement gives

$$y(t) = -625t - 46312.5 + Ce^{0.08t},$$

where C is the constant of integration. Imposing the initial condition $y(0) = 1000$ yields the final solution

$$y(t) = 45312.5\,e^{0.08t} - 625t - 46312.5,$$

where $t \geq 0$.

To illustrate, note that after 10 years the investment totals about $y(10) \approx 51507.86$ dollars. $\square$

3.5 Exact ODEs

(Ref: SH10 §19.2)

In this section we examine another approach to solving (some) first order ODEs.

To begin, suppose that H is a function of two variables x and y satisfying the equation

$$H(x, y) = C,$$

where C is a real constant. If we consider y as a function of x and differentiate both sides with respect to x, then the chain rule gives

$$\frac{\partial H}{\partial x} + \frac{\partial H}{\partial y}\frac{dy}{dx} = 0.$$

If $\dfrac{\partial H}{\partial x}$ and $\dfrac{\partial H}{\partial y}$ are denoted by F and G respectively, then we obtain the differential equation

$$F(x, y) + G(x, y)\frac{dy}{dx} = 0. \tag{3.13}$$

Of course, if F and G are defined as above, then $H(x, y) = C$ is a solution to (3.13).

Conversely, suppose that we want to solve a differential equation of the form (3.13). The above discussion shows that *if* there exists a function H of two variables such that $F = \frac{\partial H}{\partial x}$ and $G = \frac{\partial H}{\partial y}$ *then* the solution is given by $H(x, y) = C$, where $C \in \mathbb{R}$. The difficulty is, this condition on F and G is not so easy to verify. Fortunately, there is an easier condition. Recall, by the theorem on mixed partial derivatives, that if H is a 'nice' function then

$$\frac{\partial^2 H}{\partial y \partial x} = \frac{\partial^2 H}{\partial x \partial y},$$

or in other words,

$$\frac{\partial F}{\partial y} = \frac{\partial G}{\partial x}.$$

This second condition on F and G is known as the condition for *exactness* and is much easier to verify.

We summarise these observations in the next definition and theorem.

Definition 3.5.1. An ordinary differential equation of the form

$$F(x,y) + G(x,y)\frac{dy}{dx} = 0$$

is called *exact* if

$$\frac{\partial F}{\partial y} = \frac{\partial G}{\partial x}.$$

Theorem 3.5.2. *Suppose that an ordinary differential equation of the form (3.13) is exact. Then the solution to (3.13) is given by $H(x,y) = C$, where C is a constant and where H is a function satisfying the equations*

$$\frac{\partial H}{\partial x} = F \qquad \text{and} \qquad \frac{\partial H}{\partial y} = G.$$

Remark 3.5.3. The differential equation (3.13) is equivalent to

$$\frac{dy}{dx} = -\frac{F(x,y)}{G(x,y)}$$

or

$$F(x,y)dx + G(x,y)dy = 0.$$

The left-hand side of the second expression is an example of a *differential form.*

Example 3.5.4. Show that the differential equation

$$\frac{dy}{dx} = -\frac{2x+y+1}{2y+x+1}$$

is exact, and hence find its solution.

Solution. First we rewrite the differential equation to obtain

$$(2x+y+1) + (2y+x+1)\frac{dy}{dx} = 0.$$

Write $F = 2x+y+1$ and $G = 2y+x+1$. Then

$$\frac{\partial F}{\partial y} = 1 = \frac{\partial G}{\partial x},$$

so the differential equation is exact. Hence there exists a function H satisfying

$$\frac{\partial H}{\partial x} = F(x,y) = 2x+y+1, \tag{3.14}$$

$$\frac{\partial H}{\partial y} = G(x,y) = 2y+x+1. \tag{3.15}$$

To find H, we begin by integrating (3.14) with respect to x (and treating y as a constant), so that

$$H(x,y) = x^2 + xy + x + C_1(y), \tag{3.16}$$

where the 'constant of integration' $C_1(y)$ is a function of y. Similarly, integrating (3.15) with respect to y (and treating x as a constant) gives

$$H(x,y) = y^2 + xy + y + C_2(x), \tag{3.17}$$

where the 'constant of integration' $C_2(x)$ is a function of x. A comparison of (3.16) and (3.17) shows that

$$H(x,y) = x^2 + xy + y^2 + x + y. \tag{3.18}$$

Hence the solution to the differential equation is given by

$$x^2 + xy + y^2 + x + y = C, \tag{3.19}$$

where C is a real constant.

Note that, since (3.19) is a quadratic in y, we could rewrite the solution explicitly for y to give

$$y = \frac{-(x+1) \pm \sqrt{(x+1)^2 - 4(x^2 + x - C)}}{2}, \qquad C \in \mathbb{R}$$

(this gives *two* functions for every value of C). However, in this case the solution is probably better left in implicit form, as in (3.19). □

Remark 3.5.5. Technically, (3.18) should read

$$H(x,y) = x^2 + xy + y^2 + x + y + K,$$

where K is an arbitrary constant. However, then the solution to the ODE is given by

$$x^2 + xy + y^2 + x + y + K = C_0,$$

where C_0 is yet another constant. By combining the constants K and C_0 on the right-hand side, this solution is equivalent to (3.19). Hence is is customary to ignore the constant K.

The next example follows the same overall strategy, but illustrates a slightly different approach to finding the function H.

Example 3.5.6. Solve the differential equation

$$2x\sin(x^2 + y^2) + \big(2y\sin(x^2 + y^2) + 2y\big)\frac{dy}{dx} = 0.$$

Solution. Write $F(x,y) = 2x\sin(x^2 + y^2)$ and $G(x,y) = 2y\sin(x^2 + y^2) + 2y$. It is easily seen that

$$\frac{\partial F}{\partial y} = 4xy\cos(x^2 + y^2) = \frac{\partial G}{\partial x},$$

and so the equation is exact. Hence we look for a function H satisfying

$$\frac{\partial H}{\partial x} = F(x,y) = 2x\sin(x^2 + y^2), \tag{3.20}$$

$$\frac{\partial H}{\partial y} = G(x,y) = 2y\sin(x^2 + y^2) + 2y. \tag{3.21}$$

As with the previous example, we integrate (3.20) with respect to x (and treating y as a constant) to obtain

$$H(x, y) = -\cos(x^2 + y^2) + C_1(y), \tag{3.22}$$

where the 'constant of integration' $C_1(y)$ is a function of y. So now we only need to determine $C_1(y)$. To do so, differentiating (3.22) with respect to y (and treating x as a constant) gives

$$\frac{\partial H}{\partial y} = 2y\sin(x^2 + y^2) + C_1'(y).$$

Comparing this with (3.21) shows that $C_1'(y) = 2y$, whence $C_1(y) = y^2$. (Here we omit the constant of integration for reasons given in Remark 3.5.5.) Hence

$$H(x, y) = -\cos(x^2 + y^2) + y^2$$

and the solution to the differential equation is given by

$$-\cos(x^2 + y^2) + y^2 = C,$$

where $C \in \mathbb{R}$. (Note that in this example it is not possible to give an explicit expression for y in terms of x.) $\square$

The final example gives an interesting variation on this theme.

Example 3.5.7. Solve the differential equation

$$(e^x - \sin y)\, dx + \cos y\, dy = 0. \tag{3.23}$$

Solution. Suppose that $F(x, y) = e^x - \sin y$ and $G(x, y) = \cos y$. Since

$$\frac{\partial F}{\partial y} = -\cos y \qquad \text{and} \qquad \frac{\partial G}{\partial x} = 0,$$

the differential equation is *not* exact. What happens if we use the method of the last two examples regardless?

Suppose that there is a function H such that

$$\frac{\partial H}{\partial x} = F(x, y) = e^x - \sin y, \tag{3.24}$$
$$\frac{\partial H}{\partial y} = G(x, y) = \cos y. \tag{3.25}$$

Integrating (3.24) with respect to x gives

$$H(x, y) = e^x - x\sin y + C_1(y),$$

where $C_1(y)$ is *independent of* x. Partial differentiation with respect to y yields

$$\frac{\partial H}{\partial y} = -x\cos y + C_1'(y).$$

If we compare this with (3.25), then we conclude that $C_1'(y) = (1 + x)\cos y$, which contradicts the fact that $C_1(y)$ is independent of x. Hence no such function H exists.

Fortunately, not all is lost. If we multiply (3.23) through by the function e^{-x}, then we obtain

$$(1 - e^{-x} \sin y)\, dx + e^{-x} \cos y\, dy = 0. \tag{3.26}$$

Since the integrating factor e^{-x} is never zero, solutions to (3.26) will also be solutions to (3.23). Moreover, it is easily verified that (3.26) is exact. Hence we can now (successfully) use the method for solving exact differential equations to obtain the solution

$$x + e^{-x} \sin y = C,$$

where $C \in \mathbb{R}$. Details are left to the reader. □

Remark 3.5.8. As illustrated in the previous example, if an ODE of the form

$$F(x, y)\, dx + G(x, y)\, dy = 0$$

is not exact, then it may be possible to transform it into an exact ODE by multiplying through by a suitable function. In general, finding such a function is difficult and lies beyond the scope of this course.

3.6 Solving ODEs by using a change of variable [X]

(Ref: SH10 §19.1)

This section is for MATH1241 students only. We have seen in previous sections how to solve separable, linear and exact first order ODEs. While not all first order ODEs are among these types, some can be transformed into one of these types by a suitable change of variables. We illustrate the principle with two examples.

Example 3.6.1. Use the substitution $y(x) = x \cdot v(x)$ to solve the differential equation

$$\frac{dy}{dx} = \frac{xy - y^2}{x^2}. \tag{3.27}$$

Solution. The idea is to transform (3.27) into a separable ODE involving v and x. Once the general solution for v is found, then it is easy to write down the solution for y.

Using the substitution $y(x) = x \cdot v(x)$ and the product rule for differentiation, we see that

$$\frac{dy}{dx} = \frac{d}{dx}(xv) = v\frac{dx}{dx} + x\frac{dv}{dx} = v + x\frac{dv}{dx}.$$

Hence (3.27) becomes

$$v + x\frac{dv}{dx} = \frac{x(xv) - (xv)^2}{x^2}.$$

If we simplify the right-hand side then

$$v + x\frac{dv}{dx} = v - v^2,$$

and hence we obtain the separable ODE

$$-\frac{dv}{v^2} = \frac{dx}{x}.$$

Integrating both sides gives

$$\frac{1}{v} = \ln|x| + C$$

which implies that

$$v = \frac{1}{\ln|x| + C}$$

where $C \in \mathbb{R}$. Now $v = \dfrac{y}{x}$ and so

$$y = \frac{x}{\ln|x| + C}$$

gives the general solution to (3.27). □

Example 3.6.2. Solve the equation

$$\frac{dy}{dt} + 2y + y^2t^2e^{2t} = 0 \tag{3.28}$$

by using the substitution $z = 1/y$.

Solution. Note that (3.28) is not a first order linear ODE because of the term involving y^2. In this example we will see that the nonlinear equation in y and t becomes a linear ODE in z and t under the transformation $z = 1/y$.

To make use of the substitution $z = 1/y$, we need to express $\frac{dy}{dx}$ in terms of z. To do so, differentiate both sides of the equation

$$y = \frac{1}{z}$$

with respect to t to obtain

$$\frac{dy}{dt} = -\frac{1}{z^2}\frac{dz}{dt}.$$

With this substitution, (3.28) becomes

$$-\frac{1}{z^2}\frac{dz}{dt} + \frac{2}{z} + \frac{t^2e^{2t}}{z^2} = 0.$$

Rearranging gives

$$\frac{dz}{dt} - 2z = t^2e^{2t},$$

which is a first order linear ODE in z. As usual, multiply through by the integrating factor e^{-2t} to obtain

$$\frac{d}{dt}(e^{-2t}z) = t^2.$$

Integrating gives

$$e^{-2t}z = \frac{t^3}{3} + C_0,$$

whereupon

$$z = e^{2t}\left(\frac{t^3}{3} + C_0\right),$$

where $C_0 \in \mathbb{R}$. Since $y = 1/z$, the solution to (3.28) is given by

$$y = \frac{3}{e^{2t}(t^3 + C)},$$

where $C = 3C_0 \in \mathbb{R}$. □

3.7 Modelling with first order ODEs

(Ref: SH10 §9.1, 9.2)

Many real-life problems can be analysed and solved by attempting to convert them into mathematics. In doing so, a number of assumptions have to be made and a theoretical framework set up which attempts to reflect what is happening in the real world. Such a framework is called a *mathematical model*. The reliability of that model can be judged by how well it predicts what actually happens in the real world.

To construct a mathematical model, one should

1. describe accurately the data we have,

2. decide exactly what information we want to extract from the model,

3. decide which variables in the model are dependent and which are independent, and

4. describe how the dependent variables change as the independent ones vary (which may lead to a differential equation).

Examples of mathematical modelling with differential equations have already been encountered (see Examples 3.3.4 and 3.4.3). The following two subsections provide further examples and discussion.

3.7.1 Mixing problems

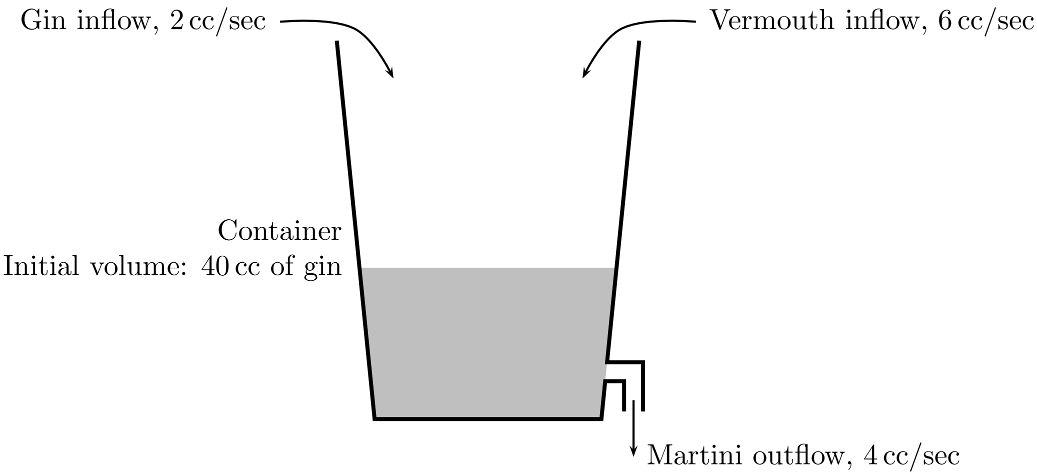
In this subsection we illustrate how to construct a mathematical model that leads to a differential equation.

Example 3.7.1. A martini drink is, in essence, a mixture of the two liquids *gin* and *vermouth*. James Blond insists that his martinis be prepared as follows. Initially, 40 cc of gin are placed in a large container. Then gin is poured into the container at a rate of 2 cc/sec and at the same time vermouth is poured in at a rate of 6 cc/sec. The mixture is constantly shaken (not stirred) and flows out at a rate of 4 cc/sec.

(a) Find an expression for the volume of vermouth in the container t seconds after the pouring commences.

(b) James likes his martini to have roughly two parts gin to three of vermouth. How many seconds should elapse before he stops pouring and inserts a cocktail glass in the outflow from the container?

Solution. (a) To begin, it is recommended that we place the relevant information on a diagram.



Next, it is important to identify what we want to find. In this case, we want a formula for the volume of Vermouth. So let $V(t)$ denote the volume of vermouth (in cc) in the container at time t.

Now we write down as much information about V as we can. We know that $V(0) = 0$. The other information given tells us how V changes with time. In particular,

$$\begin{aligned}\frac{dV}{dt} &= \text{ rate of change of } V \\ &= (\text{rate of inflow}) - (\text{rate of outflow}).\end{aligned} \tag{3.29}$$

Now the rate of inflow of vermouth is 6 cc/sec. To calculate the the rate of outflow, we note that the total volume of liquid in the container at time t is given by

$$40 + 2t + 6t - 4t = 40 + 4t.$$

Hence the *proportion* by volume of vermouth in the container at time t is

$$\frac{V(t)}{40+4t}. \tag{3.30}$$

Since the rate of outflow of liquid is 4 cc/sec, it follows that the rate of outflow of vermouth is

$$\frac{V(t)}{40+4t} \times 4\,\text{cc/sec}.$$

Following from (3.29), we obtain the initial value problem

$$\frac{dV}{dt} = 6 - \frac{V(t)}{10+t}, \qquad V(0) = 0. \tag{3.31}$$

Finally, we solve the IVP. The ODE is first order linear equation, whose solution (when $V(0) = 0$) is given by

$$V(t) = \frac{3t^2 + 60t}{10+t}. \tag{3.32}$$

The details for finding this solution are left to the reader as an exercise.

(b) The liquid will be two parts gin to three of vermouth exactly when three-fifths of the liquid is vermouth. Hence, using (3.30), we require that

$$\frac{V(t)}{40+4t} = \frac{3}{5}.$$

But $V(t)$ is given by (3.32), so we require that

$$\frac{3t^2+60t}{4(10+t)^2} = \frac{3}{5}.$$

By rearranging this equation, we obtain the quadratic equation

$$t^2 + 20t - 400 = 0.$$

The quadratic formula shows that

$$t = \sqrt{500} - 10 \approx 12.36,$$

where we have chosen the positive solution to the quadratic equation. Hence James should insert the cocktail glass 12.36 seconds after the mixing process begins. □

3.7.2 Population models

In this subsection we compare three different mathematical models for population growth.

Suppose that the city of Mathopolis initially has a population of 3,000,000 inhabitants and (initially) grows at a rate of 2% per annum. Can we predict what the population of the city will be in 10, 20 or 100 years time? In the next three examples, we examine different population models, discuss their accuracy and see what growth forecast each model gives for Mathopolis.

Example 3.7.2 (Population model 1)**.** In this model, we assume that the growth rate of the population remains constant. That is, we assume that the rate of change of population is proportional to the population, where the constant of proportionality r is the growth rate. In other words, we have the IVP

$$\frac{dP}{dt} = rP, \qquad P(0) = P_0,$$

where P is the population at time t, r is the growth rate and P_0 is the initial population. Solving this (separable) ODE gives the solution

$$P(t) = P_0 e^{rt}.$$

(This model goes back to Thomas Malthus' book *An Essay on the Principle of Population*, published in 1798.)

Application. For Mathopolis, $P_0 = 3,000,000$ and $r = 0.02$. Thus the population $P(t)$ of the city at time t is given by

$$P(t) = 3000000e^{0.02t}.$$

The table in Figure 3.1 shows the predicted population when t is 10, 20 and 100. The graph of P against t is also shown.

Criticisms of this model.

1. The model predicts that population will grow indefinitely. This ignores the fact that the resources and space needed to support such a population are finite.

2. The model ignores external factors (such as disease, natural disasters and wars) that have an effect on population size.

In light of the first criticism, we introduce a second model.

Example 3.7.3 (Population model 2)**.** In this model the rate of population growth is not proportional to the population. Instead, there is a critical population P_c which when exceeded causes the population P to decrease; otherwise the population increases. We try the IVP

$$\frac{dP}{dt} = k(P_c - P), \qquad P(0) = P_0, \tag{3.33}$$

where k is a positive constant. Hence the rate of change of P is positive if $P < P_c$ and negative if $P > P_c$.

To solve (3.33), we separate the variables and integrate:

$$\int \frac{dP}{P_c - P} = \int k\,dt.$$

In the case when $P < P_c$ we obtain

$$-\ln(P_c - P) = kt + C$$

and so

$$P = P_c - Ae^{-kt},$$

where $A = e^{-C}$. Now $P = P_0$ when $t = 0$ and so $A = P_m - P_0$. Hence

$$P(t) = P_c - (P_c - P_0)e^{-kt}.$$

(The case when $P > P_c$ gives the same solution, as can be easily verified by carefully working through the details.) Note that $P(t) \to P_c$ as $t \to \infty$.

Application. For Mathopolis, $P_0 = 3,000,000$ and $\frac{dP}{dt} = 0.02P_0 = 60,000$ when $t = 0$. Assume also that $P_c = 7,000,000$. (That is, due to available land, resources and other factors, one expects the city's maximum sustainable population size is 7 million.) By considering the differential equation (3.33) when $t = 0$, we conclude that

$$60000 = \frac{dP}{dt} = k(P_c - P_0) = k(7000000 - 3000000)$$

and hence that $k = 0.015$. Therefore the population $P(t)$ of Mathopolis at time t is given by

$$P(t) = 7000000 - 4000000e^{-0.015t}.$$

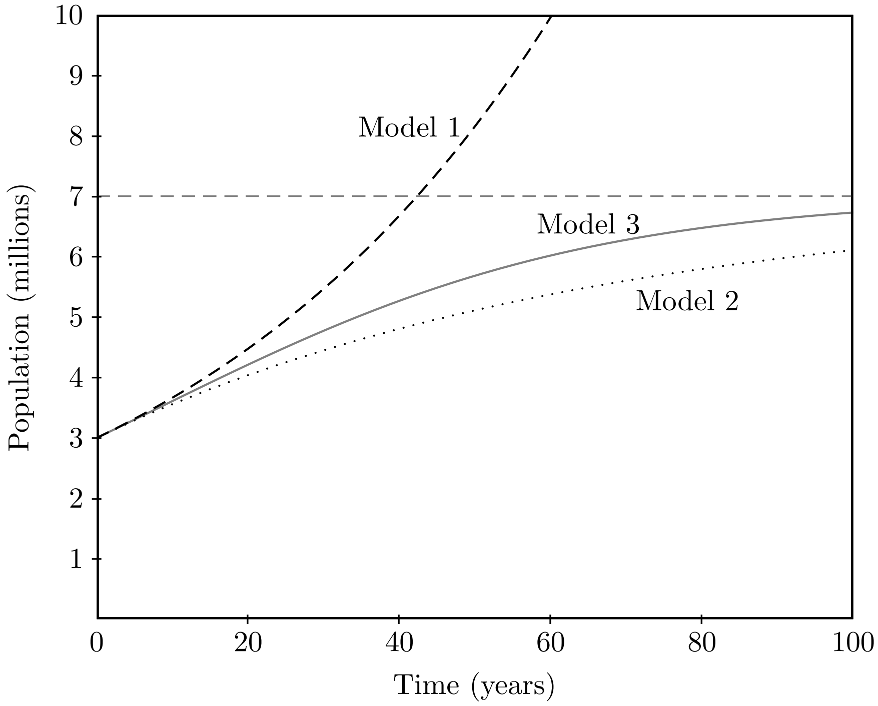
See Figure 3.1 for specific population projections under this model and a corresponding graph.

Criticisms of this model.

1. Observe that if P is close to 0 then $\frac{dP}{dt} \approx kP_c$, which means that when the population is very small the growth rate may be a large positive number. In fact, the rate of increase is most rapid for tiny populations!

	Initially	After 10 years	After 20 years	After 100 years
Model 1	3.00 million	3.66 million	4.48 million	22.17 million
Model 2	3.00 million	3.56 million	4.04 million	6.11 million
Model 3	3.00 million	3.61 million	4.21 million	6.73 million

(a) Table showing population forecast according to various models



(b) Graphs showing how models compare

Figure 3.1: Population growth for a city as projected by different models.

2. As with the first population model, external factors are ignored.

In light of the first criticism we introduce a third model.

Example 3.7.4 (Population model 3)**.** To overcome the first criticism of the previous model, we instead try the IVP

$$\frac{dP}{dt} = kP(P_c - P), \qquad P(0) = P_0, \tag{3.34}$$

where k is a positive constant and P_c is the critical population. Note that $\frac{dP}{dt}$ is small when P is small. (This model was first published by Pierre Verhulst in 1838 after he had read Thomas Malthus' *An Essay on the Principle of Population.*)

To solve (3.34), separation of variables and integration gives

$$\int \frac{dP}{P(P_c - P)} = \int k\,dt.$$

The integral on the left-hand side is evaluated by the method of partial fractions:

$$\frac{1}{P_c}\int \frac{1}{P} + \frac{1}{P_c - P}\,dP = \int k\,dt.$$

From here it is not difficult to show that

$$P(t) = \frac{P_c P_0}{P_0 + (P_m - P_0)e^{-kP_c t}}$$

(try this as an exercise). Note once again that $P(t) \to P_c$ as $t \to \infty$. The resulting curve (see, for example, the solid gray curve in Figure 3.1 (b)) is a called a *logistic curve.* The initial stage of growth is approximately exponential; then, as P approaches P_c the growth slows and approaches P_c asymptotically.

Application. Once again, for Mathopolis we have the data

$$P_0 = 3,000,000, \qquad P_c = 7,000,000 \qquad \text{and} \qquad \frac{dP}{dt} = 0.02P_0 = 60,000 \quad \text{when} \quad t = 0.$$

By considering the differential equation (3.34) when $t = 0$, we conclude that

$$60000 = \frac{dP}{dt} = kP_0(P_c - P_0)$$

and hence that $k = 5 \times 10^{-9}$. Therefore the population $P(t)$ of the city at time t is given by

$$P(t) = \frac{21000000}{3 + 4e^{-0.035t}}.$$

See Figure 3.1 for specific population projections under this model.

Criticisms of this model. As with the other population models, external factors are ignored. In particular, P_c may change due to factors such as technological advances or climate change.

Remark 3.7.5. Note that as the models introduced become more realistic, the mathematics needed to solve the corresponding differential equations is more sophisticated. In the case of modelling fluid flow (such as water flow in the pipes or air flow around an aeroplane wing), the set of (partial) differential equations which must be solved (known as the *Navier–Stokes equations*) raises difficulties that are beyond the grasp of current mathematical knowledge. For example, it is as yet unknown whether a solution to these equations always exists and whether (in the case that a solution exists) it is a 'smooth' solution. The Clay Mathematics Institute has listed these problems as one of the seven *Millennium problems* and carries prize money of US$1,000,000 for a correct solution.

3.8 Second order linear ODEs with constant coefficients

(Ref: SH10 §9.3, §19.4)

In this section we consider a special class of second order equations, known as *second order linear ODEs with constant coefficients.* An equation of this class has the form

$$\frac{d^2y}{dx^2} + a\frac{dy}{dx} + by = f(x), \tag{3.35}$$

where a and b are real numbers. Such equations naturally arise in modelling wave mechanics and prey-predator interaction. You will have seen in earlier calculus courses (and in Physics) that the second order equation $\dfrac{d^2x}{dt^2} + n^2x = 0$ is used to model simple harmonic motion.

3.8.1 The homogeneous case

To simplify our treatment of the second oder ODE (3.35), we first look at the case when $f(x) \equiv 0$ (which means that $f(x) = 0$ for all x).

Definition 3.8.1. A second order linear ODE with constant coefficients is said to be *homogeneous* if it is of the form

$$\frac{d^2y}{dx^2} + a\frac{dy}{dx} + by = 0, \tag{3.36}$$

where a and b are real numbers.

It turns out that we can always solve a homogeneous second order ODE with real coefficients. The first important observation towards proving this fact is given by the following lemma.

Lemma 3.8.2. *If y_1 and y_2 are two solutions to the differential equation (3.36) then any linear combination $Ay_1 + By_2$, where A and B are real numbers, is also a solution to (3.36).*

Proof. Suppose that y_1 and y_2 are two solutions to the differential equation (3.36). If $y = Ay_1 + By_2$, where A and B are real numbers, then

$$\begin{aligned}
y'' + ay' + by &= (Ay_1 + By_2)'' + a(Ay_1 + By_2)' + b(Ay_1 + By_2) \\
&= Ay_1'' + Ay_2'' + Aay_1' + Bay_2' + Aby_1 + Bby_2 \\
&= A(y_1'' + ay_1' + by_1) + B(y_2'' + ay_2' + by_2) \\
&= 0 + 0 \qquad\qquad \text{(since } y_1 \text{ and } y_2 \text{ are solutions)} \\
&= 0.
\end{aligned}$$

Hence y is also a solution to (3.36). □

Remark 3.8.3. It can also be shown that every second order ODE has at most two linearly independent solutions (this will be demonstrated in second year linear algebra courses). Hence if y_1 and y_2 are two *linearly independent* solutions of (3.36) then every solution y to (3.36) is of the form $y = Ay_1 + By_2$. In this context, y_1 and y_2 are linearly independent if and only if they are not constant multiples of each other.

In view of the above lemma and remark, to find a complete solution to (3.36), one only needs to find two linearly independent solutions. To look for a solution to (3.36), we try a function y that does not change too much when differentiated. (The idea is that, upon substitution, the terms on the left-hand side need to cancel each other out to give zero.) If $y = e^{\lambda x}$, where λ is a constant, then $y' = \lambda e^{\lambda x}$ and $y'' = \lambda^2 e^{\lambda x}$. When these are substituted into (3.36) we obtain

$$\lambda^2 e^{\lambda x} + a\lambda e^{\lambda x} + b e^{\lambda x} = 0.$$

Dividing by $e^{\lambda x}$ gives

$$\lambda^2 + a\lambda + b = 0, \tag{3.37}$$

Hence we have shown that

$$y = e^{\lambda x} \text{ is a solution to (3.36)} \qquad \text{if and only if} \qquad \lambda \text{ is a root of (3.37)}.$$

We give the quadratic equation (3.37) a special name.

Definition 3.8.4. The *characteristic equation* of the second order linear ODE

$$\frac{d^2y}{dx^2} + a\frac{dy}{dx} + by = 0,$$

is given by

$$\lambda^2 + a\lambda + b = 0. \tag{3.38}$$

The following example illustrates what we have learnt so far.

Example 3.8.5. Solve the second order homogeneous linear ODE

$$\frac{d^2y}{dx^2} - 5\frac{dy}{dx} + 6y = 0. \tag{3.39}$$

Solution. The characteristic equation associated to (3.39) is given by

$$\lambda^2 - 5\lambda + 6 = 0.$$

By solving the quadratic equation we find that $\lambda = 2, 3$. Hence $y_1 = e^{2x}$ and $y_2 = e^{3x}$ are solutions to (3.39). By Lemma 3.8.2 the linear combination y, given by

$$y = Ay_1 + By_2 = Ae^{2x} + Be^{3x}$$

where A and B are real numbers, is also a solution to (3.39). Moreover, since y_1 and y_2 are linearly independent, *every* solution is of this form (see Remark 3.8.3). ☐

Note that in the last example, the characteristic equation had two distinct real roots, thus leading to two linearly independent solutions to the homogeneous ODE. In general, there are three possibilities since all the coefficients we consider are *real* numbers. Either (i) the characteristic equation has two distinct real roots, (ii) the characteristic equation has a repeated real root, or (iii) the characteristic equation has two distinct complex roots (which are complex conjugates of each other).

Case (i): The characteristic equation has two distinct real roots λ_1 and λ_2. Then, as seen above, we obtain two linearly independent solutions $y_1 = e^{\lambda_1 x}$ and $y_2 = e^{\lambda_2 x}$. Hence the general solution is given by

$$y = Ae^{\lambda_1 x} + Be^{\lambda_2 x},$$

where A and B are real numbers.

Case (ii): The characteristic equation has a repeat real real root λ_1. Then one solution is given by $y_1 = e^{\lambda_1 x}$. Is there another independent solution? It turns out in this case that $y_2 = xe^{\lambda_1 x}$ also solves the homogeneous equation (3.36), as can be easily verified by substituting this into the left-hand side of (3.36). Hence the general solution is given by

$$y = Ae^{\lambda_1 x} + Bxe^{\lambda_1 x},$$

where A and B are real numbers.

Case (iii): The characteristic equation has two distinct complex roots $\alpha + \beta i$ and $\alpha - \beta i$, where α and β are real numbers and $\beta \neq 0$. Then we obtain two solutions

$$y_1 = e^{(\alpha+\beta i)x} \qquad \text{and} \qquad y_2 = e^{(\alpha-\beta i)x}.$$

Hence y is also a solution, where

$$y = Ce^{(\alpha+\beta i)x} + De^{(\alpha-\beta i)x},$$

and C and D are (complex) constants. The goal is to choose C and D so that the solution is real. Now Euler's formula

$$e^{i\theta} = \cos\theta + i\sin\theta$$

gives

$$\begin{aligned}
y &= Ce^{(\alpha+\beta i)x} + De^{(\alpha-\beta i)x} \\
&= e^{\alpha x}(Ce^{\beta i x} + De^{-\beta i x}) \\
&= e^{\alpha x}\big(C(\cos\beta x + i\sin\beta x) + D(\cos\beta x - i\sin\beta x)\big) \\
&= e^{\alpha x}\big((C+D)\cos\beta x + i(C-D)\sin\beta x\big) \\
&= e^{\alpha x}\big(A\cos\beta x + B\sin\beta x\big),
\end{aligned}$$

where $A = C + D$, $B = i(C - D)$. If we choose C and D to be complex conjugates of each other, but otherwise with arbitrary real and imaginary parts, then A and B will be real. It is easy to see that $e^{\alpha x}\cos(\beta x)$ and $e^{\alpha x}\sin(\beta x)$ are independent solutions; hence the general solution in this case is given by

$$y = e^{\alpha x}\big(A\cos\beta x + B\sin\beta x\big)$$

where A and B are real numbers.

We summarise our findings in the following theorem.

Theorem 3.8.6. *Consider the second order homogeneous ODE given by (3.36) and let λ_1 and λ_2 denote the roots of the corresponding characteristic equation (3.38).*

(i) If λ_1 and λ_2 are different real numbers then the solution to (3.36) is given by

$$y = Ae^{\lambda_1 x} + Be^{\lambda_2 x},$$

where $A, B \in \mathbb{R}$.

(ii) *If $\lambda_1 = \lambda_2$ then the solution to (3.36) is given by*

$$y = Ae^{\lambda_1 x} + Bxe^{\lambda_1 x},$$

where $A, B \in \mathbb{R}$.

(iii) *If $\lambda_1 = \alpha + \beta i$ and $\lambda_2 = \alpha - \beta i$, where $\alpha, \beta \in \mathbb{R}$ and $\beta \neq 0$, then the solution to (3.36) is given by*

$$y = e^{\alpha x}\big(A\cos(\beta x) + B\sin(\beta x)\big),$$

where $A, B \in \mathbb{R}$.

Example 3.8.7. Solve the following differential equations:

(a) $y'' - 6y' + 25 = 0$,

(b) $y'' + 4y' + 4y = 0$, with initial conditions $y(0) = 1$ and $y'(0) = 0$.

Solution. (a) The characteristic equation

$$\lambda^2 - 6\lambda + 25 = 0$$

has roots $3+4i$ and $3-4i$ (as determined by completing the square or using the quadratic formula). Hence the solution y to the ODE is given by

$$y = e^{3x}(A\cos 4x + B\sin 4x),$$

where A and B are real numbers.

(b) The left-hand side of the characteristic equation

$$\lambda^2 + 4\lambda + 4 = 0$$

is easily factorised to give

$$(\lambda + 2)^2 = 0.$$

Hence -2 is a repeated root and the solution y to the IVP is given by

$$y = Ae^{-2x} + Bxe^{-2x}, \tag{3.40}$$

where the constants A and B are to be determined by imposing initial conditions. Differentiation shows that

$$y' = -2Ae^{-2x} - 2Bxe^{-2x} + Be^{-2x}. \tag{3.41}$$

When $x = 0$, (3.40), (3.41) and the initial conditions imply that $A = 1$ and $B = 2$. Hence

$$y = e^{-2x} + 2xe^{-2x}$$

describes the solution y to the IVP. □

3.8.2 The non-homogeneous case

We return to solving the second order linear ODE (3.36) in the case when f is not identically zero. The main idea will be illustrated in the following example.

Example 3.8.8. Solve the equation

$$y'' - 5y' + 6y = 12x - 4. \tag{3.42}$$

Solution. Since the first and second derivatives of a polynomial are also polynomials, it seems likely that at least one particular solution y_P to the ODE is a polynomial. A little more thought shows that if y_P is polynomial that satisfies (3.42), then the degree of y_P is no greater than one. So we look for a particular solution y_P of the form

$$y_P = ax + b,$$

where a and b are real numbers whose values are to be determined. Now $y_P' = a$ and $y_P'' = 0$, so substituting y_P into (3.42) gives

$$0 - 5a + 6(ax + b) = 12x - 4.$$

By equating coefficients we find that $a = 2$ and $b = 1$. Hence one particular solution y_P to (3.39) is given by $y_P = 2x + 1$.

Are there any other solutions? The answer to this question is 'yes'. To find them, we consider the associated homogeneous ODE

$$y'' - 5y' + 6y = 0, \tag{3.43}$$

whose solution y_H is given by

$$y_H = Ae^{2x} + Be^{3x},$$

where A and B are real numbers (see Example 3.8.5). Now we will show that y, where $y = y_H + y_P$, is a solution to (3.42):

$$\begin{aligned}
y'' - 5y' + 6y &= (y_H + y_P)'' - 5(y_H + y_P)' + 6(y_H + y_P) \\
&= (y_H'' - 5y_H' + 6y_H) + (y_P'' - 5y_P' + 6y_P) && \text{(by linearity of differentiation)} \\
&= 0 + (12x - 4) && \text{(by the properties of } y_H \text{ and } y_P\text{)} \\
&= 12x - 4.
\end{aligned}$$

Hence y, given by

$$y = y_H + y_P = Ae^{2x} + Be^{3x} + 2x - 1$$

where A and B are real constants, also solves the ODE. In fact, this gives the general solution (the discussion in Subsection 3.8.4 explains why). □

Bearing in mind the above example, we now detail an algorithm for solving a second order ODE of the form (3.35).

1. Find the solution y_H to the corresponding homogeneous equation (3.36) (by first identifying the roots of the characteristic equation, as in Subsection 3.8.1).

2. Find a particular solution y_P to (3.35).

3. The general solution y to (3.35) is then given by $y = y_H + y_P$.

In general, it is best to perform Step 1 before Step 2; the reason for this will soon become obvious.

Step 1 has already been discussed in some detail, and Step 3 is easy. We need only devote some discussion to performing Step 2.

Example 3.8.9. Solve the ODE

$$y'' - 4y' + 5y = 20e^{-x}. \tag{3.44}$$

Proof. First, we solve the homogeneous equation

$$y'' - 4y' + 5y = 0.$$

The roots of the characteristic equation $\lambda^2 - 4\lambda + 5 = 0$ are $2 - i$ and $2 + i$ (these can be found using the quadratic formula). Hence the solution y_H of the homogeneous equation is given by

$$y_H = e^{2x}(A\cos x + B\sin x),$$

where A and B are real numbers.

Second, we find a particular solution y_P. Since the right hand side is an exponential, we try

$$y_P = ae^{-x},$$

where a is a constant to be determined. Note that $y_P' = -ae^{-x}$ and $y_P'' = ae^{-x}$. So substituting y_P into (3.44) gives

$$ae^{-x} + 4ae^{-x} + 5ae^{-x} = 20e^{-x}.$$

By comparing coefficients on each side we conclude that y_P is a particular solution if and only if $a = 2$. Hence our particular solution is given by $y_P = 2e^{-x}$.

Third, the general solution y to (3.44) is given by

$$\begin{aligned} y &= y_H + y_P \\ &= e^{2x}(A\cos x + B\sin x) + 2e^{-x}, \end{aligned}$$

where A and B are real numbers. □

Example 3.8.10. Solve the ODE

$$y'' - 3y' + 2y = 5e^{2x}. \tag{3.45}$$

Solution. First we solve the corresponding homogeneous ODE

$$y'' - 3y' + 2y = 0.$$

The characteristic equation $\lambda^2 - 3\lambda + 2 = 0$ has the solutions $\lambda = 1$ and $\lambda = 2$. Hence the solution y_H to the homogeneous equation is given by

$$y_H = Ae^{2x} + Be^{x}, \tag{3.46}$$

where A and B are real constants.

Second, we look for a particular solution y_P. Since the right-hand side of (3.45) is a multiple of e^{2x}, it seems natural to try $y_P = ae^{2x}$. However, note that ae^{2x} is a particular solution to the homogeneous equation. (To see this, simply set A as a and B as 0 in (3.46)). Hence substituting $y_P = ae^{2x}$ into (3.45) will produce 0 on the left-hand side, and consequently this guess for y_P will *not* work.

In this circumstance, the 'trick' for finding a particular solution y_P is to *multiply the old guess* ae^{2x} *by* x. That is, we now try

$$y_P = axe^{2x}.$$

(This guess for a particular solution is certainly not in the solution space for the homogeneous equation.) Now we find that

$$y_P' = ae^{2x}(2x+1) \qquad \text{and} \qquad y_P'' = 4ae^{2x}(x+1),$$

and hence substituting $y_P = axe^{2x}$ into (3.45) gives

$$4ae^{2x}(x+1) - 3ae^{2x}(2x+1) + 2axe^{2x} = 5e^{2x}.$$

This simplifies to

$$4ae^{2x} - 3ae^{2x} = 5e^{2x},$$

and we thereby deduce that $a = 5$. Therefore a particular solution y_P is given by $y_P = 5xe^{2x}$.

Finally, the general solution y to (3.45) is given by

$$\begin{aligned} y &= y_H + y_P \\ &= Ae^{2x} + Be^{x} + 5xe^{2x} \end{aligned}$$

where A and B are real constants. $\square$

Given a function f, the following table indicates which guess for y_P will always yield a particular solution for the nonhomogeneous ODE (3.35).

$f(x)$	Guess for particular solution y_P
$P(x)$ (a polynomial of degree n)	$Q(x)$ (a polynomial of degree n)
$P(x)e^{sx}$	$Q(x)e^{sx}$
$P(x)\cos(sx)$	$Q_1(x)\cos(sx) + Q_2(x)\sin(sx)$
$P(x)\sin(sx)$	$Q_1(x)\cos(sx) + Q_2(x)\sin(sx)$
$P(x)e^{sx}\cos(tx)$ or $P(x)e^{sx}\sin(tx)$	$Q_1(x)e^{sx}\cos(tx) + Q_2(x)e^{sx}\sin(tx)$

If any term of the guess for y_P is a solution to the homogeneous ODE, then multiply it by x.

If any term of the new guess is still a solution to the homogeneous ODE, then multiply by x again.

The next example illustrates the directive given in the last two rows of the table.

Example 3.8.11. Solve the ODE

$$y'' - 6y' + 9y = 8e^{3x}. \tag{3.47}$$

Solution. First, the characteristic equation

$$\lambda^2 - 6\lambda + 9 = 0$$

factorises as $(\lambda - 3)^2 = 0$ and thus has the repeated root 3. So the solution y_H to the corresponding homogeneous equation is given by

$$y_H = Ae^{3x} + Bxe^{3x}, \tag{3.48}$$

where A and B are real numbers.

Second, we search for a particular solution y_P. Our first guess $y_P = ae^{3x}$ will not work, as it is a solution to the corresponding homogeneous equation (to see why, simply set A as a and B as 0 in (3.48)). So we multiply by x to obtain a new guess $y_P = axe^{3x}$. However, this guess also solves the homogeneous equation. So once again, multiply by x to obtain a new guess $y_P = ax^2e^{3x}$. It is easy to see that y_P no longer lies in the solution space to the homogeneous equation. So this guess will work.

To determine the value of a, substitute $y_P = ax^2e^{3x}$ into (3.47). We leave it to the reader to verify that $a = 4$. Hence a particular solution y_P is given by $y_P = 4x^2e^{3x}$.

Finally, the general solution y is the sum of y_H and y_P, namely

$$y = Ae^{3x} + Bxe^{3x} + 4x^2e^{3x},$$

where A and B are real numbers. □

Remark 3.8.12. The previous example illustrates the importance of finding y_H before making a guess for y_P; without knowing y_H it is not possible to make a suitable guess for y_P. The next example emphasises this point.

Example 3.8.13. Consider the nonhomogeneous ODE

$$\frac{d^2y}{dt^2} + 6\frac{dy}{dt} + 13y = 5e^{3t}\cos(2t). \tag{3.49}$$

Write down the *form* of a particular solution y_P to this ODE. (You are *not* required to evaluate the undetermined coefficients appearing in the the form of y_P.)

Solution. The characteristic equation $\lambda^2 + 6\lambda + 13 = 0$ has roots $3 + 2i$ and $3 - 2i$ (these can be found by completing the square or using the quadratic formula). Hence the solution y_H to the corresponding homogeneous equation is given by

$$y_H = e^{3t}(A\cos 2t + B\sin 2t), \tag{3.50}$$

where $A, B \in \mathbb{R}$.

Since the right-hand side of (3.49) is a product of e^{3t} and $\cos 2t$, our initial guess for the particular solution y_P is given by

$$y_P = ae^{3t}\cos 2t + be^{3t}\sin 2t.$$

However, (3.50) shows that this guess solves the homogeneous equation. Instead, try

$$y_P = ate^{3t}\cos 2t + bte^{3t}\sin 2t.$$

This new guess for y_P does not lie in the solution space for the homogeneous equation; hence this gives the form of particular solution that we seek. □

Remark 3.8.14. One must take care in the use of the method of undetermined coefficients, and in particular the two rules at the end of the table given above.

Consider, for example, the differential equation

$$y'' - 6y' + 9y = x^2 e^{3x}.$$

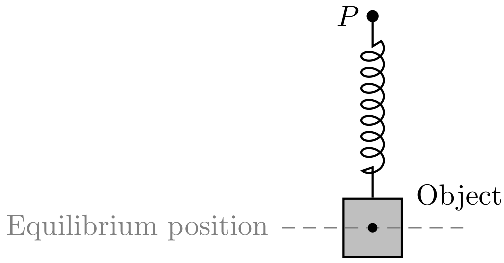
The homogeneous solution, from above, is $y_H = Ae^{3x} + Bxe^{3x}$. Now since x^2e^{3x} is **not** one of the homogeneous solutions, one would try the particular solution $y_P = (Cx^2 + Dx + E)e^{3x}$. However, since this contains terms which are part of the homogeneous solution, we need to multiply by x twice - thus $y_P = (Cx^4 + Dx^3 + Ex^2)e^{3x}$ - so that no term in this expression is a homogeneous solution. Substitution will reveal that D and E are zero, while $C = \frac{1}{12}$.

There is an easier method to solve such a problem using differential operators. This is left to more advanced courses in differential equations.

3.8.3 An application: vibrations and resonance

Many structures have a natural frequency of vibration. If an external agent causes them to vibrate at or near one of these frequencies then large oscillations build up and *resonance* occurs. This can cause disasters such as the collapse of bridges. The 'trick' used by singers breaking wine glasses is another example of this phenomenon. The first example of this subsection examines the behaviour of a natural oscillating system; the second example illustrates how introducing external vibrations into such a system causes resonance.

Example 3.8.15. Suppose that a spring is mounted to a (fixed) point P and that an object of mass m is suspended from the spring. Let x denote the (vertical) displacement from the equilibrium position (or resting position) of the object, taking x to be positive if it lies above the equilibrium position.



By using Newton's second law of motion, Hooke's law and making some simple assumptions, one can show that the system satisfies the differential equation

$$\frac{d^2x}{dt^2} + \omega^2 x = 0, \tag{3.51}$$

where ω is a positive constant that depends only on the mass of the object and the stiffness of the spring.

The object is pulled downwards from its equilibrium position by a distance of 4 units and then released from rest. Find $x(t)$ when $t \geq 0$.

Solution. We need to solve the differential equation (3.51) subject to the initial conditions $x(0) = -4$ and $x'(0) = 0$. The characteristic equation $\lambda^2 + \omega^2 = 0$ is easily solved, leading to the solution

$$x(t) = A\cos\omega t + B\sin\omega t$$

of (3.51), where A and B are constants. Before imposing initial conditions, we calculate $x'(t)$. This is given by

$$x'(t) = -A\omega \sin \omega t + B\omega \cos \omega t.$$

Now $x'(0) = 0$ implies that $0 = B\omega$, from which we conclude that $B = 0$ (since ω is positive). Finally, $x(0) = -4$ implies that $A = -4$. Hence the solution is given by

$$x(t) = -4 \cos \omega t.$$

This type of motion is known as *simple harmonic motion* and is graphed in Figure 3.2 (a). ☐

Example 3.8.16. Consider the same scenario as Example 3.8.15, except that now the point P vibrates up and down, such that its vertical displacement y is given by $y = 2\sin \Omega t$. In these circumstances, a simple physical argument shows that x obeys the differential equation

$$\frac{d^2x}{dt^2} + \omega^2 x = 2\sin \Omega t. \tag{3.52}$$

Describe the motion of the object, given that $x(0) = -4$ and $x'(0) = 0$.

Solution. We have already solved the homogeneous equation in Example 3.8.15. So we look for a particular solution x_P.

Case 1: Suppose that $\Omega \neq \omega$. Then we look for a particular solution of the form

$$x_P = C\cos \Omega t + D \sin \Omega t.$$

By substituting this into (3.52) we obtain

$$x_P = \frac{2}{\omega^2 - \Omega^2} \sin \Omega t.$$

By using the fact that $x = x_H + x_P$ and imposing initial conditions we find that

$$x(t) = -4\cos \omega t - \frac{\Omega}{\omega}\frac{2}{\omega^2 - \Omega^2} \sin \omega t + \frac{2}{\omega^2 - \Omega^2} \sin \Omega t.$$

This is simply another oscillating system. Its graph is shown in Figure 3.2 (b) (in the case when $\Omega = \frac{3}{2}\omega$).

Case 2: Suppose that $\Omega = \omega$. This time we look for a particular solution of the form

$$x_P = Ct\cos \omega t + Dt \sin \omega t.$$

By substituting this into (3.52) we obtain

$$x_P = -\frac{t}{\omega} \cos \omega t.$$

Using the fact that $x = x_H + x_P$ and imposing initial conditions we find that

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

Hence as t increases, the amplitude of the $\cos \omega t$ term grows without bound and the system becomes unstable. Its graph is shown in Figure 3.2 (c). ☐

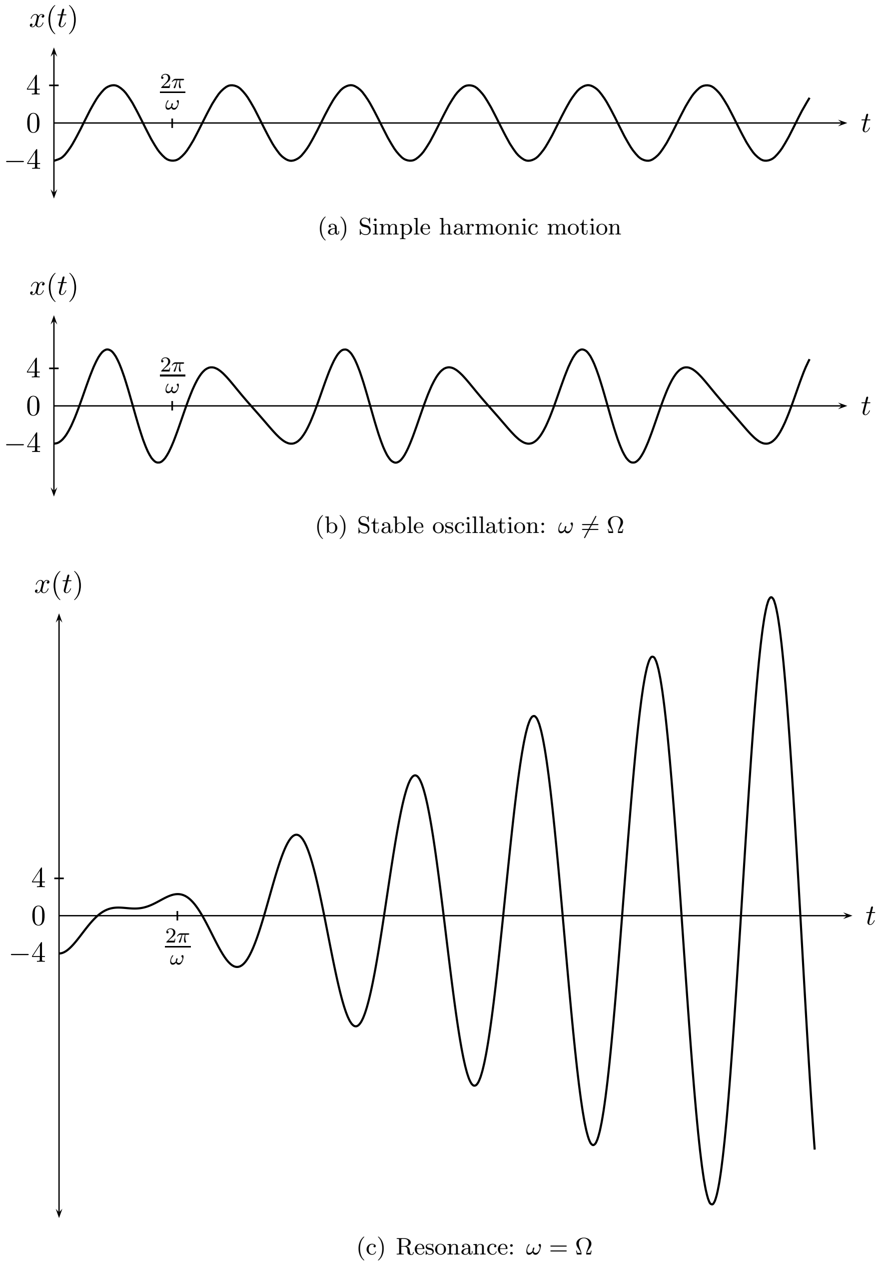


Figure 3.2: Different oscillating systems.

3.8.4 A connection with linear algebra

In Section 3.8, we have made two claims whose proofs have not been given yet; namely (A) that every second order homogeneous ODE has at most two linearly independent solutions and (B) that every solution y to a nonhomogeneous second order ODE can be written in the form $y_H + y_P$. Both of these assertions can be proved by using linear algebra. In this subsection we highlight the connection between second order linear ODEs and linear algebra, and explain why assertion (B) is true. The proof of assertion (A) is given in MATH2501. Students should not continue reading this subsection until Chapter 7 from MATH1231/1241 algebra has been completed.

Consider the second order ODE

$$y'' + ay' + by = f, \tag{3.53}$$

where a and b are real numbers and $f : \mathbb{R} \to \mathbb{R}$ is a function. Let V denote the vector space of all (infinitely) differentiable functions $y : \mathbb{R} \to \mathbb{R}$. We define the linear transformation $T : V \to V$ by the formula

$$T(y) = y'' + ay' + by.$$

Observe that

$$\begin{aligned} y_H \text{ is a solution to the homogeneous equation} \quad &\iff \quad y_H'' + ay_H' + by_H = 0 \\ &\iff \quad T(y_H) = 0 \\ &\iff \quad y_H \in \ker(T). \end{aligned}$$

Hence the general solution to the homogeneous equation *is* the kernel of T. Also observe that

$$\begin{aligned} y_P \text{ is a particular solution to the nonhomogeneous equation} \quad &\iff \quad y_P'' + ay_P' + by_P = f \\ &\iff \quad T(y_P) = f. \end{aligned}$$

Hence the ODE (3.53) has a particular solution if and only if f is in the image of T. Moreover, since

$$\begin{aligned} T(y_H + y_P) &= T(y_H) + T(y_P) && \text{(by the linearity of } T) \\ &= 0 + f \\ &= f, \end{aligned}$$

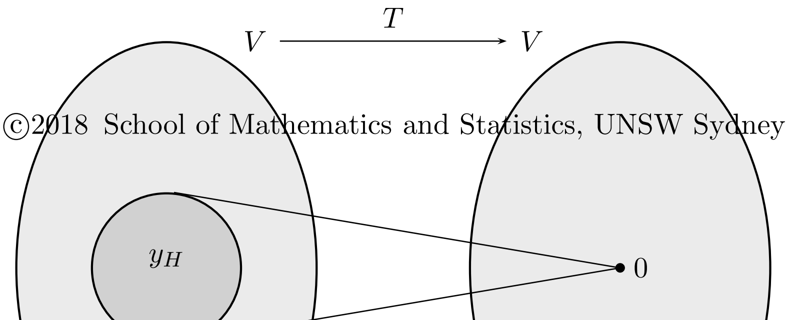
we conclude that $y_H + y_P$ is a also a solution to (3.53).

Finally, we prove assertion (B). Suppose that y is a solution to (3.53) and that y_P is some particular solution to (3.53). Then $T(y) = f$ and $T(y_P) = f$. Hence

$$\begin{aligned} T(y - y_P) &= T(y) - T(y_P) && \text{(by the linearity of } T) \\ &= f - f \\ &= 0. \end{aligned}$$

Therefore $y - y_P$ is in the kernel of T. But, as observed above, every function in the kernel of T is a solution y_H to the homogeneous equation. Hence $y - y_P = y_H$, where y_H is some solution to the homogeneous equation. It follows that $y = y_H + y_P$, thus completing the proof of assertion (B).

The kernel and image of T, and their connection to the homogeneous solution space and a particular solution for the differential equation, may be represented pictorially as shown below.



3.9 Maple notes

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

`dsolve(deqn, y(x));` solves the ordinary differential equation (or IVP) deqn for the function `y(x)`. For example,

```
> dsolve(diff(y(x), x$2) - y(x) = 1, y(x));
```

$$y(x) = -1 + \_C1 \exp(x) + \_C2 \exp(-x)$$

```
> dsolve({diff(v(t), t) + 2*t = 0, v(1)=5}, v(t));
```

$$v(t) = -t^2 + 6$$

Chapter 4

Taylor series

Polynomials are nice functions to work with. Their values can be easily evaluated using a finite number of additions and multiplications. They are easy to differentiate and integrate; moreover their derivatives and antiderivatives are also polynomials and thus have these properties. Some (but not all) of these properties are shared by a few other classes of functions. For example, the derivatives of exponential functions are exponentials, but in general exponentials cannot be evaluated using a finite number additions and multiplications. Many other useful functions share none of these properties.

If a function can be accurately approximated by a polynomial, then we can use the polynomial to approximate the values, derivatives and antiderivative of the function. This generalises an idea met in MATH1131, where we saw that a differentiable function can be locally approximated by a linear function (which is a degree 1 polynomial). In this chapter we will see that many n-times differentiable functions can be locally approximated by a polynomial of degree n. We will devote considerable time discussing how accurate such approximations are. Finally, we shall prove that some functions can not only be approximated by polynomials but are also equal to series consisting of infinitely many polynomial terms. Such series are known as *Taylor series*.

The ideas presented in this chapter have a long history, going back to Archimedes' *method of exhaustion*, which he used to approximate π. Beginning in the fourteenth century, a school of Indian mathematicians based in Kerala found accurate polynomial approximations to trigonometric functions. This allowed them to solve problems in astronomy. In the seventeenth century, the Scottish astronomer James Gregory independently employed similar techniques. However, it was not until 1715 that the English mathemtician Brook Taylor published a theorem which gave a general method for polynomial approximation, and which described precisely the errors involved. In the twenty-first century, computers and calculators regularly use algorithms to find the approximate value of functions; many of these modern computation techniques trace their roots back to Taylor's method.

4.1 Taylor polynomials

(Ref: SH10 §12.6, 12.7)

Recall from MATH1131 that the function $f : \mathbb{R} \to \mathbb{R}$, given by $f(x) = e^x$, is defined as the inverse of the function $\ln : (0, \infty) \to \mathbb{R}$, and that ln is defined in terms of an integral. How, then, does one evaluate $e^{0.1}$? (If the answer given is 'use a calculator,' then we merely ask the question, How does a calculator evaluate $e^{0.1}$?)

One approach is to suppose that $y = e^{0.1}$. Then we need to solve the equation $\ln y = 0.1$. By the definition of ln, this boils down to solving the integral equation

$$\int_1^y \frac{dt}{t} = 0.1.$$

From here we could guess an approximate value for y and use Riemann sums to check whether our guess is reasonable. Clearly this is an unsatisfactory approach to the problem.

Another method is to locally approximate the function f with a linear function. This technique was discussed in MATH1131 (see Chapter 4 of the MATH1131 calculus notes) and is based on the idea that the tangent lies close to the graph of f near the point of contact. Suppose, once again, that $f(x) = e^x$. To find an approximate value for $e^{0.1}$, we will approximate f using the tangent to the graph of f at 0. The tangent function p_1 at 0 is a polynomial of degree 1 and is given by

$$p_1(x) = 1 + x.$$

So when x is close to 0,

$$e^x \approx 1 + x$$

and hence $e^{0.1} \approx 1.1$.

The polynomial p_1 has the property that its value at 0 and gradient at 0 agree with the value and gradient of f at 0. That is,

$$p_1(0) = f(0) \qquad \text{and} \qquad p_1'(0) = f'(0)$$

If we want to improve our approximation, we could generalise this idea and look for a degree two polynomial p_2 such that the value, gradient and concavity of f and p_2 agree at 0. That is, if

$$p_2(x) = b_0 + b_1 x + b_2 x^2,$$

then we choose the coefficients b_0, b_1 and b_2 such that

$$p_2(0) = f(0), \qquad p_2'(0) = f'(0) \qquad \text{and} \qquad p_2''(0) = f''(0).$$

By calculating the first and second derivatives of f and p, we find that

$$b_0 = 1, \qquad b_1 = 1 \qquad \text{and} \qquad 2b_2 = 1.$$

Hence

$$p_2(x) = 1 + x + \frac{x^2}{2}.$$

One can see in Figure 4.1 that p_2 gives a better approximation to f *near* 0 than does p_1. Using p_2, we obtain the approximation

$$e^{0.1} = f(0.1) \approx p_2(0.1) = 1.105.$$

One can again improve this approximation by using a polynomial of degree 3. If

$$p_3(x) = c_0 + c_1 x + c_2 x^2 + c_3 x^3$$

then we can determine the unknown coefficients by solving the equations

$$p_3(0) = f(0), \qquad p_3'(0) = f'(0), \qquad p_3''(0) = f'(0) \qquad \text{and} \qquad p_3'''(0) = f'''(0).$$

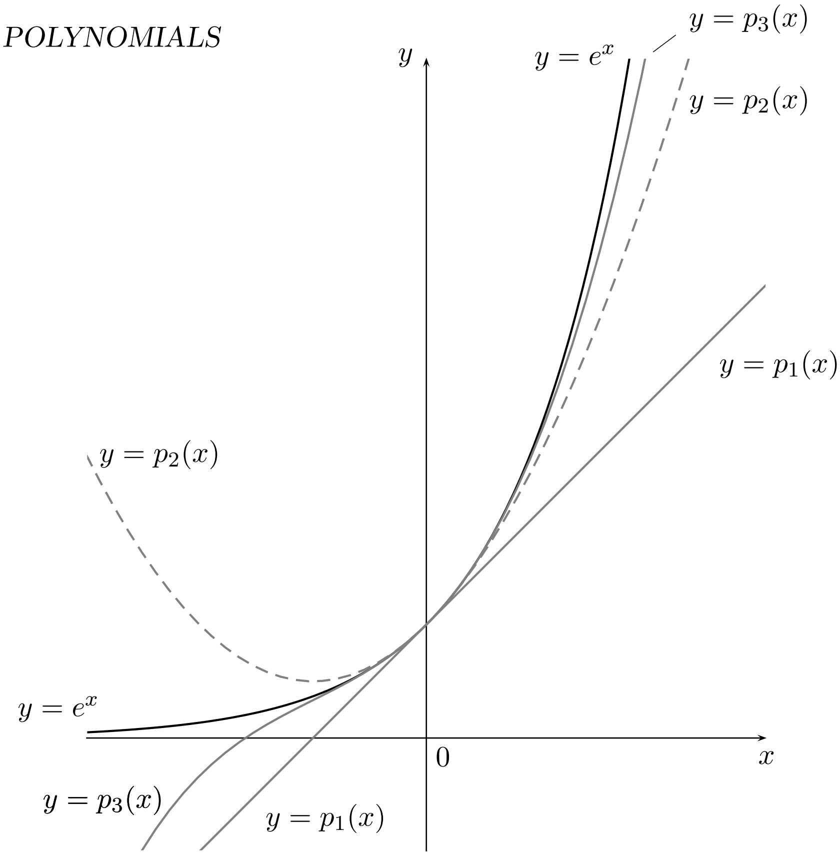


Figure 4.1: Polynomial approximations (in gray) for the exponential function (in black) about 0.

These equations imply that

$$c_0 = 1, \qquad c_1 = 1, \qquad 2c_2 = 1 \qquad \text{and} \qquad 6c_3 = 1$$

and hence

$$p_3(x) = 1 + x + \frac{x^2}{2} + \frac{x^3}{6}$$

(see Figure 4.1). Thus we have the approximation

$$e^{0.1} = f(0.1) \approx p_3(0.1) = 1.1051\dot{6}.$$

The table below compares the approximations to $e^{0.1}$ given by these first, second and third degree polynomials.

n	$p_n(x)$	$p_n(0.1)$	Error in approximation (3 s.f.)
1	$1 + x$	1.1	5.17×10^{-3}
2	$1 + x + \frac{x^2}{2}$	1.105	1.71×10^{-4}
3	$1 + x + \frac{x^2}{2} + \frac{x^3}{6}$	$1.1051\dot{6}$	4.25×10^{-6}

In fact, $e^{0.1}$ (rounded to five decimal places) is equal to 1.10517 and so the approximation $p_3(0.1)$ is accurate to four decimal places. Each of the polynomials p_1, p_2 and p_3 are called *Taylor polynomials* for f about 0.

Using the same technique, one can attempt to approximate any function f at 0 with an n degree polynomial p_n, provided that f is n-times differentiable at 0. Suppose that

$$p_n(x) = a_0 + a_1 x + a_2 x^2 + a_3 x^3 + \cdots + a_n x^n.$$

We require that

$$p_n(0) = f(0), \quad p_n'(0) = f'(0), \quad p_n''(0) = f''(0), \quad p_n^{(3)}(0) = f^{(3)}(0), \quad \ldots, \quad p_n^{(n)}(0) = f^{(n)}(0),$$

where $f^{(j)}$ denotes the jth derivative of f. By calculating the derivatives of p, one finds that

$$a_0 = f(0), \quad a_1 = f'(0), \quad 2!a_2 = f''(0), \quad 3!a_3 = f^{(3)}(0), \quad \ldots, \quad n!a_n = f^{(n)}(0).$$

Hence

$$p_n(x) = f(0) + f'(0)x + \frac{f''(0)}{2!}x^2 + \frac{f^{(3)}(0)}{3!}x^3 + \cdots + \frac{f^{(n)}(0)}{n!}x^n.$$

Definition 4.1.1. Suppose that f is n-times differentiable at 0. Then the *Taylor polynomial* p_n *of degree* n *for* f *about* 0 is given by

$$p_n(x) = f(0) + f'(0)x + \frac{f''(0)}{2!}x^2 + \frac{f^{(3)}(0)}{3!}x^3 + \cdots + \frac{f^{(n)}(0)}{n!}x^n.$$

We also call p_n the *nth Taylor polynomial for f about* 0.

Example 4.1.2. Suppose that $f(x) = e^x$ and $n \geq 0$. Find the Taylor polynomial of degree n for f about 0.

Solution. It is clear that

$$f'(x) = e^x, \quad f''(x) = e^x, \quad f''(x) = e^x, \quad \ldots, \quad f^{(n)}(x) = e^x.$$

Hence $f(0) = f'(0) = f''(0) = \cdots = f^{(n)}(0) = 1$ and

$$p_n(x) = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots + \frac{x^n}{n!}$$

gives the Taylor polynomial of degree n about 0. □

Example 4.1.3. Suppose that $f(x) = \sin x$. Find the Taylor polynomials for f up to degree seven about the point 0.

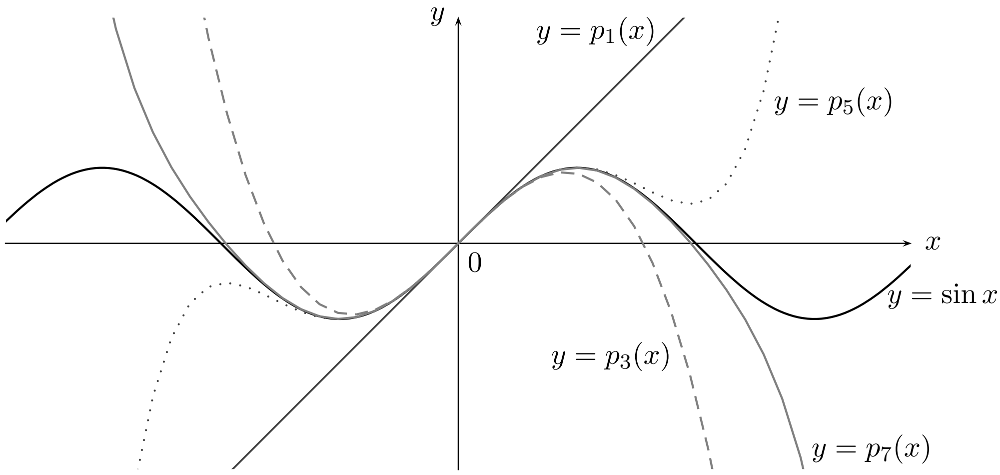
Solution. We have

$$\begin{aligned}
f(x) &= \sin x & f(0) &= 0 \\
f'(x) &= \cos x & f'(0) &= 1 \\
f''(x) &= -\sin x & f''(0) &= 0 \\
f^{(3)}(x) &= -\cos x & f^{(3)}(0) &= -1 \\
f^{(4)}(x) &= \sin x & f^{(4)}(0) &= 0 \\
&\vdots & &\vdots
\end{aligned}$$

and the pattern will repeat itself. Hence

$$\begin{aligned} p_1(x) &= x \\ p_3(x) &= x - \frac{x^3}{3!} \\ p_5(x) &= x - \frac{x^3}{3!} + \frac{x^5}{5!} \\ p_7(x) &= x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} \end{aligned}$$

are the Taylor polynomials for f up to degree seven about the point 0. Their graphs are compared to the sine function in the following diagram.



Note that the sine function is odd and that its Taylor polynomials about 0 are also odd. □

The diagram in Example 4.1.3 shows that the Taylor polynomials about 0 are good approximations for f near 0. However, the approximations get worse farther away from 0. If one wants to approximate a function f near a point a, where a is not close to 0, then one can try (i) using Taylor polynomials about 0 of higher degree (which may or may not give satisfactory results) or (ii) approximating f with *Taylor polynomials about* a. Method (ii) is particularly pertinent in the case when $f(x) = \ln x$; since f is not defined at 0, the Taylor polynomials for f about 0 do not exist.

Suppose f is n-times differentiable at a. To approximate f about a, we try a degree n polynomial p_n, given by

$$p_n(x) = a_0 + a_1(x-a) + a_2(x-a)^2 + \cdots + a_n(x-a)^n,$$

where

$$p_n(a) = f(a), \qquad p_n'(a) = f'(a), \qquad p_n''(a) = f''(a), \qquad \ldots, \qquad p_n^{(n)}(a) = f^{(n)}(a).$$

By calculating the derivatives of p_n at a, the unknown coefficients can be determined. Hence

$$p_n(x) = f(a) + f'(a)(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \frac{f^{(3)}(a)}{3!}(x-a)^3 + \cdots + \frac{f^{(n)}(a)}{n!}(x-a)^n.$$

This leads to the following definition.

Definition 4.1.4. Suppose that f is n-times differentiable at a. Then the *Taylor polynomial p_n of degree n for f about a* is given by

$$p_n(x) = f(a) + f'(a)(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \frac{f^{(3)}(a)}{3!}(x-a)^3 + \cdots + \frac{f^{(n)}(a)}{n!}(x-a)^n.$$

We also call p_n the *nth Taylor polynomial for f about a.*

The nth Taylor polynomial for f about a can be expressed using summation notation as

$$p_n(x) = \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(x-a)^k.$$

Here $0! = 1$ (by definition) and we use the convention of writing $f(a)$ as $f^{(0)}(a)$.

Example 4.1.5. Suppose that $f(x) = \ln x$. Find the Taylor polynomial of degree 5 for f about 1.

Solution. Since

$$\begin{aligned}
f(x) &= \ln x & f(1) &= 0 \\
f'(x) &= \frac{1}{x} & f'(1) &= 1 \\
f''(x) &= -\frac{1}{x^2} & f''(1) &= -1 \\
f^{(3)}(x) &= \frac{2!}{x^3} & f^{(3)}(1) &= 2! \\
f^{(4)}(x) &= -\frac{3!}{x^4} & f^{(4)}(1) &= -3! \\
f^{(5)}(x) &= \frac{4!}{x^5} & f^{(5)}(1) &= 4!,
\end{aligned}$$

we see that

$$\begin{aligned}
p_5(x) &= (x-1) + \frac{-1}{2!}(x-1)^2 + \frac{2!}{3!}(x-1)^3 + \frac{-3!}{4!}(x-1)^4 + \frac{4!}{5!}(x-1)^5 \\
&= (x-1) - \frac{(x-1)^2}{2} + \frac{(x-1)^3}{3} - \frac{(x-1)^4}{4} + \frac{(x-1)^5}{5}.
\end{aligned}$$

It is preferable to express such a polynomial in powers of $(x-1)$ rather than in powers of x. There is usually no need to expand each of the terms $(x-1)^k$. ☐

4.2 Taylor's theorem

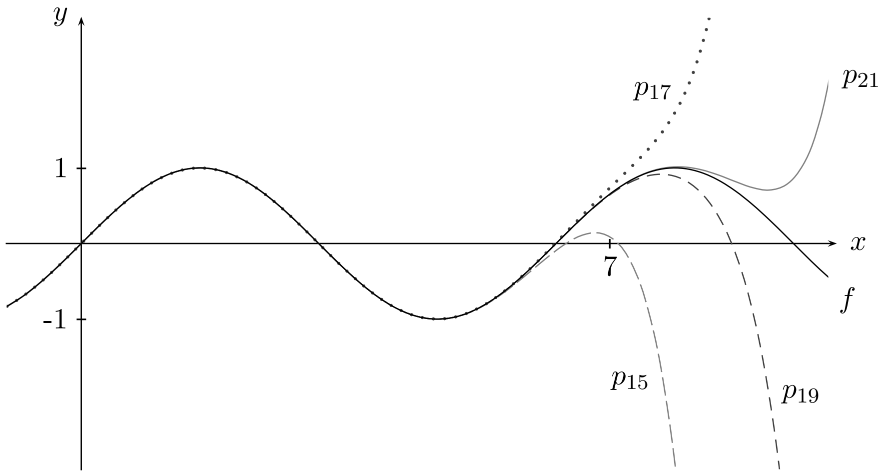
(Ref: SH10 §12.6, 12.7)

We saw in the last section that $e^{0.1}$ can be approximated using Taylor polynomials. Suppose that a calculator (with a 10 digit display) uses Taylor polynomials to calculate the (approximate) value of $e^{0.1}$. What degree Taylor polynomial should the calculator use so that the displayed value is accurate to 10 digits? Importantly, is there a way of answering this question without knowing in advance the decimal expansion of $e^{0.1}$?

In other words, we seek a method for determining how bad the error can get when approximating a function by one of its Taylor polynomials without knowing the precise values of the function. The functions f and g, given in the table below, further highlight the need for such a method.

Function	nth Taylor polynomial about 0
$f(x) = \sin x$	$p_n(x) = x - \dfrac{x^3}{3!} + \dfrac{x^5}{5!} - \dfrac{x^7}{7!} + \cdots \dfrac{x^n}{n!}$ (n odd)
$g(x) = \ln(1+x)$	$q_n(x) = x - \dfrac{x^2}{2} + \dfrac{x^3}{3} - \dfrac{x^5}{5} + \cdots \dfrac{x^n}{n}$

Suppose that we want to approximate $f(7)$. Although 7 is not close to 0, the graph below shows that the approximation $f(7) \approx p_n(7)$ is reasonable if $n = 19$ or $n = 21$.



By plotting more graphs, one would discover that the approximation $f(7) \approx p_n(7)$ seems to improve as n increases. We could conjecture that as $n \to \infty$ the error in the approximation approaches 0.

On the other hand, suppose that we want to approximate $g(1.5)$. Although 1.5 is closer to 0 than is 7, the Taylor polynomials do not provide a good approximation to the true value of $g(1.5)$, as Figure 4.2 shows. The Taylor polynomials approximate g very well on the interval $(-0.7, 0.7)$ and are not too bad on $(0.7, 1)$. However, to the right of 1 they should not be used to approximate the function. Moreover, while the higher degree Taylor polynomials give better approximations near 0, they give larger errors than lower degree polynomials to the right of 1. This surprising observation highlights the need for establishing a rigorous basis Taylor polynomial approximation. In particular, we must have some idea whether the error involved in each approximation is going to be large or small.

In light of the above discussion, we want an exact expression for the difference between a function and one of its Taylor polynomials. To obtain this expression, we use integration by parts.

Suppose that f has $n+1$ continuous derivatives on an open interval I containing 0. We are about to compute the nth Taylor polynomial of f about 0 in such a way that we keep track of the *difference* (also known as the *error* or *remainder*) between $f(x)$ and $p_n(x)$. Fix a number x in the interval I and note that

$$\int_0^x f'(t)\,dt = f(x) - f(0). \tag{4.1}$$

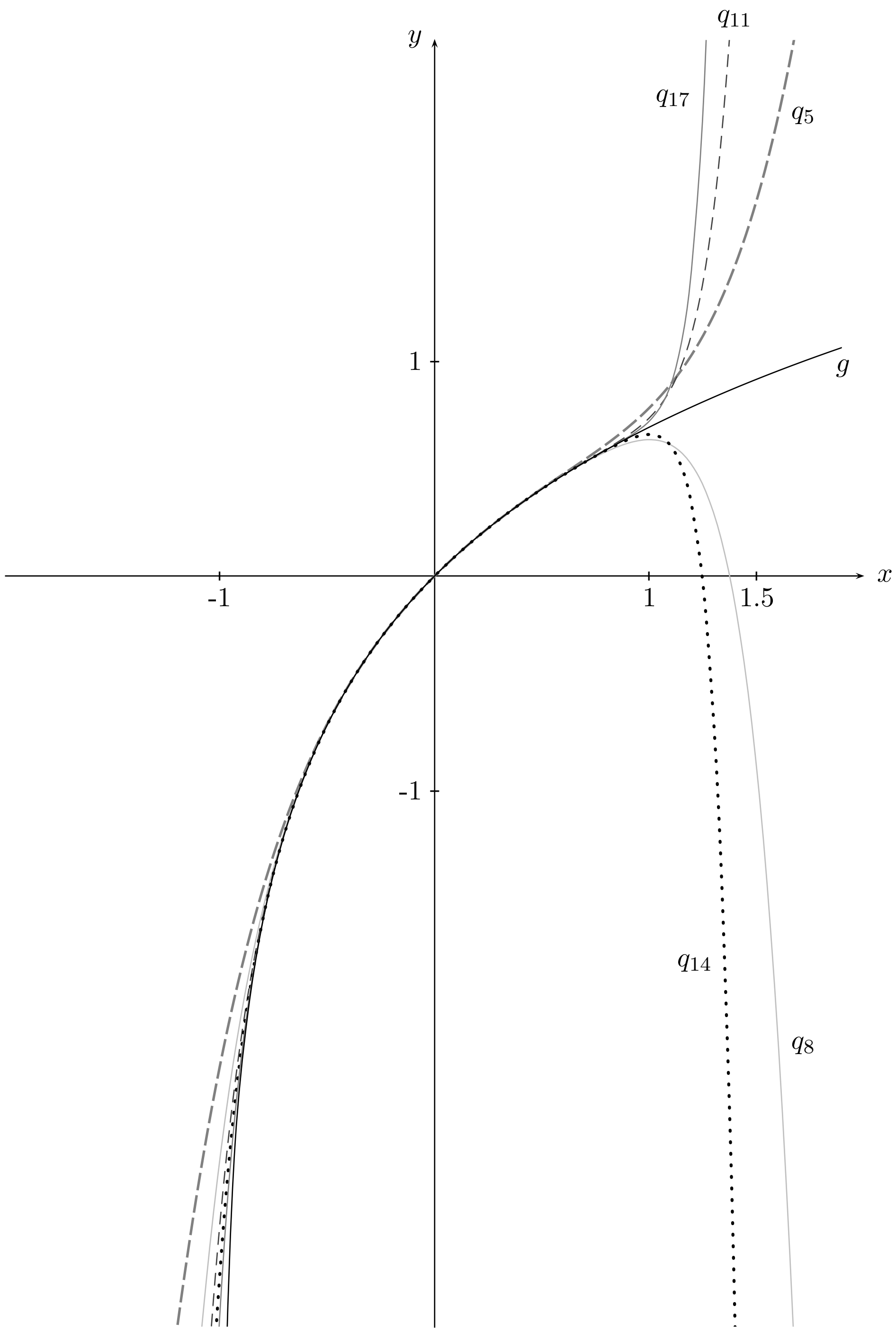


Figure 4.2: Taylor polynomials for $\ln(1 + x)$ about 0.

On the other hand, we can evaluate the same integral using integration by parts. If we set

$$u = f'(t) \qquad\qquad v = -(x-t)$$
$$\frac{du}{dt} = f''(t) \qquad\qquad \frac{dv}{dt} = 1$$

then integration by parts gives

$$\begin{aligned}\int_0^x f'(t)\,dt &= \Big[-f'(t)(x-t)\Big]_0^x + \int_0^x f''(t)(x-t)\,dt \\ &= f'(0)x + \int_0^x f''(t)(x-t)\,dt. \end{aligned} \tag{4.2}$$

Thus, from equations (4.1) and (4.2) we see that

$$f(x) = f(0) + f'(0)x + \int_0^x f''(t)(x-t)\,dt.$$

(Note that the first two terms on the right-hand side gives the Taylor polynomial of degree 1, and that the third term on the right-hand side is the error in the approximation $f(x) \approx p_1(x)$.) By applying integration by parts with

$$u = f''(t) \qquad\qquad v = -\tfrac{1}{2}(x-t)^2$$
$$\frac{du}{dt} = f'''(t) \qquad\qquad \frac{dv}{dt} = x-t,$$

we obtain

$$f(x) = f(0) + f'(0)x + \frac{f''(0)}{2!}x + \frac{1}{2!}\int_0^x f'''(t)(x-t)^2\,dt.$$

(The right-hand side is now the second Taylor polynomial of f with the associated error (also known as remainder)). If we continue integrating by parts then we obtain after n steps

$$f(x) = f(0) + f'(0)x + \frac{f''(0)}{2!}x + \frac{f^{(3)}(0)}{3!}x^3 + \ldots + \frac{f^{(n)}(0)}{n!}x^n + \frac{1}{n!}\int_0^x f^{(n+1)}(t)(x-t)^n\,dt.$$

That is,

$$f(x) = p_n(x) + R_{n+1}(x)$$

where p_n is the nth Taylor polynomial for f about 0 and

$$R_{n+1}(x) = \frac{1}{n!}\int_0^x f^{(n+1)}(t)(x-t)^n\,dt.$$

We call $R_{n+1}(x)$ the *remainder* term.

This argument may be easily adapted for Taylor polynomials about a point a, rather than 0. We thus have the following theorem.

Theorem 4.2.1 (Taylor's theorem). *Suppose that f has $n+1$ continuous derivatives on an open interval I containing a. Then for each x in I,*

$$f(x) = p_n(x) + R_{n+1}(x),$$

where the nth Taylor polynomial p_n about a is given by

$$p_n(x) = f(a) + f'(a)(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \frac{f^{(3)}(a)}{3!}(x-a)^3 + \cdots + \frac{f^{(n)}(a)}{n!}(x-a)^n$$

and the remainder $R_{n+1}(x)$ is given by

$$R_{n+1}(x) = \frac{1}{n!}\int_a^x f^{(n+1)}(t)(x-t)^n\,dt. \tag{4.3}$$

Taylor's theorem tells us that the error in the approximation $f(x) \approx p_n(x)$ is *exactly* equal to $R_{n+1}(x)$. However, the remainder $R_{n+1}(x)$, as given in the integral form (4.3), is usually very difficult to compute. A more convenient form is known as the *Lagrange formula for the remainder.*

Corollary 4.2.2 (Lagrange formula for the remainder)**.** *Suppose that f has $n+1$ continuous derivatives on an open interval I containing a. Then for each x in I,*

$$f(x) = p_n(x) + R_{n+1}(x),$$

where p_n is the nth Taylor polynomial for f about a and

$$R_{n+1}(x) = \frac{f^{(n+1)}(c)}{(n+1)!}(x-a)^{n+1} \tag{4.4}$$

for some real number c between a and x.

The proof of Corollary 4.2.2 is left as an exercise in the tutorial problems.

Remark 4.2.3. Taylor's theorem with the Lagrange formula for the remainder is a generalisation of the mean value theorem. To see this, suppose that f is once differentiable. Then

$$f(x) - p_0(x) = R_1(x),$$

which means that

$$f(x) - f(a) = f'(c)(x-a).$$

for some c between x and a. In other words,

$$\frac{f(x) - f(a)}{x-a} = f'(c)$$

for some c between x and a.

In most examples, it is difficult to find the exact value of the number c that appears in the Lagrange formula for the remainder. Instead, one uses the fact that c lies between x and a to find an upper bound for the remainder term. That is while we may not know *what* c is, we do know *where* c is.

Example 4.2.4. Suppose that $f(x) = \cos x$. By considering the second Taylor polynomial for f about 0, estimate $f(1/5)$ and find an upper bound for the error.

Solution. The second Taylor polynomial p_2 is given by

$$\begin{aligned} p_2(x) &= f(0) + f'(0)x + \frac{f''(0)}{2!}x^2 \\ &= \cos(0) - \sin(0)x - \frac{\cos(0)}{2!}x^2 \\ &= 1 - \frac{x^2}{2!}. \end{aligned}$$

Hence we use the estimate $\cos(\frac{1}{5}) \approx p_2(\frac{1}{5}) = \frac{49}{50}$. Using the Lagrange formula for the remainder, we calculate an upper bound for absolute error:

$$\begin{aligned} |\text{error}| &= |f(1/5) - p_2(1/5)| \\ &= |R_3(1/5)| && \text{(by Taylor's theorem)} \\ &= \left| \frac{f^{(3)}(c)}{3!}(1/5)^3 \right| && \text{(for some } c \text{ in } [0,1/5]) \\ &= \frac{\sin c}{6} \times \frac{1}{125} && \text{(for some } c \text{ in } [0,1/5]) \\ &\le \frac{1}{6} \times \frac{1}{125} && \text{(since } \sin c \le 1) \\ &= \frac{1}{750}. \end{aligned}$$

So an upper bound for the error is $\frac{1}{750}$, which is approximately 0.001333.

This upper bound is actually quite crude due to the estimate $\sin c \le 1$. If instead we use the inequality

$$\sin t < t \qquad \text{whenever } t > 0$$

(see Chapter 5 of MATH1131), then $\sin c < c < \frac{1}{5}$ and we obtain the upper bound

$$|\text{error}| < \frac{1}{3750}.$$

In summary, $\cos(\frac{1}{5}) \approx \frac{49}{50}$ and the absolute error in this estimate is less than $\frac{1}{3750}$ (that is, less than 0.0002667). □

Example 4.2.5. A calculator with a 10 digit display uses a Taylor polynomial about 0 to estimate $e^{0.1}$. What degree polynomial should be used to guarantee that $e^{0.1}$ is displayed accurately?

Solution. We consider the function f given by $f(x) = e^x$. By Taylor's theorem,

$$e^x = p_n(x) + R_{n+1}(x)$$

where

$$p_n(x) = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \ldots + \frac{x^n}{n!}$$

and

$$R_{n+1}(x) = \frac{e^c}{(n+1)!}x^{n+1}$$

for some c between 0 and x. Since the calculator will display the first ten digits of $p_n(0.1)$, it suffices if the error in the approximation $e^{0.1} \approx p_n(0.1)$ is less than 10^{-10}. That is, we need to find n large enough so that

$$|R_{n+1}(0.1)| < 10^{-10}. \tag{4.5}$$

Now

$$\begin{aligned}
|R_{n+1}(0.1)| &= \frac{e^c}{(n+1)!}(0.1)^{n+1} && \text{(for some } c \text{ in } [0, 0.1]) \\
&\le \frac{e^{0.1}}{(n+1)!}(0.1)^{n+1} && \text{(since exp is an increasing function)} \\
&< \frac{2}{(n+1)!}(0.1)^{n+1} && \text{(since } e^{0.1} < 3^{0.1} < 3^{1/2} < 2) \\
&= \frac{2}{(n+1)!}10^{-(n+1)} && \text{(since } 0.1 = 10^{-1}).
\end{aligned}$$

Now $\frac{2}{(n+1)!} < 1$ whenever $n \ge 1$, so we obtain the (very crude) estimate

$$|R_{n+1}(0.1)| < 10^{-(n+1)}$$

provided that $n \ge 1$. Thus if $n = 9$ then (4.5) is satisfied. That is, if the calculator displays the first 10 digits appearing in the decimal expansion of $p_9(0.1)$, then the number appearing on the display will be accurate.

However, we can do better than this. By trial and error we find that the first (positive) integer n such that

$$\frac{2}{(n+1)!}10^{-(n+1)} < 10^{-10}$$

is 6. That is, the calculator should display the first 10 digits appearing in the decimal expansion of $p_6(0.1)$.

(Note that we did not use the decimal expansion of $e^{0.1}$ to estimate the error.) □

4.2.1 Classifying stationary points

Taylor's theorem can also be applied to the problem of classifying the stationary points of differentiable functions. For example, the function f given by $f(x) = (x-3)^4$ has a stationary point at 3 (since $f'(3) = 0$). However, since $f''(3) = 0$, the second derivative test cannot be used to determine whether the stationary point is a maximum, minimum or horizontal point of inflexion. By using Taylor's theorem, one can deduce the following improvement on the second derivative test.

Corollary 4.2.6. *Suppose that f is n times differentiable at a and that $f'(a) = 0$. If*

$$f''(a) = f'''(a) = \ldots = f^{(k-1)}(a) = 0$$

but $f^{(k)}(a) \ne 0$, where $k \le n$, then

(i) a is a local minimum point if k is even and $f^{(k)}(a) > 0$;

(ii) a is a local maximum point if k is even and $f^{(k)}(a) < 0$;

(iii) a is an horizontal point of inflexion if k is odd.

Sketch proof. Suppose that f is n times differentiable at a, and that

$$f'(a) = 0 = f''(a) = f'''(a) = \ldots = f^{(k-1)}(a) = 0 \qquad \text{and} \qquad f^{(k)}(a) \ne 0.$$

We will make the additional assumptions that $f^{(k)}(x)$ exists for all x sufficiently close to a and that $f^{(k)}$ is continuous at a. (A proof of the general case is more complicated and will be omitted.)

Taylor's theorem (with Lagrange's remainder) implies that

$$\begin{aligned} f(x) &= f(a) + f'(a)(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \cdots + \frac{f^{(k-1)}(a)}{(k-1)!}(x-a)^{k-1} + \frac{f^{(k)}(c)}{k!}(x-a)^k \\ &= f(a) + 0 + 0 + \cdots + 0 + \frac{f^{(k)}(c)}{k!}(x-a)^k \\ &= f(a) + \frac{f^{(k)}(c)}{k!}(x-a)^k \end{aligned} \tag{4.6}$$

for some c between x and a.

We now prove case (i). If k is even then $(x-a)^k > 0$ whenever $x \neq a$. Since $f^{(k)}(a) > 0$ and $f^{(k)}$ is continuous at a, we conclude that $f^{(k)}(c) > 0$ whenever x (and hence c) is sufficiently close to a. By combining these inequalities with (4.6), we conclude that $f(x) \geq f(a)$ for all x sufficiently close to a. Hence a is a local minimum point for f.

Cases (ii) and (iii) are proved in a similar way. ☐

Example 4.2.7. Suppose that

$$f(x) = x^7 - 17x^6 + 101x^5 - 229x^4 + 3x^3 + 621x^2 - 297x - 567.$$

You are given that 3 is a stationary point of f. Classify this stationary point.

Solution. It is easy to check that

$$f'(3) = f''(3) = f'''(3) = 0 \qquad \text{and} \qquad f^{(4)}(3) = -1536 < 0.$$

By applying Corollary 4.2.6 (with k equal to 4), we conclude that 3 is a local maximum point for f. ☐

4.2.2 Some questions arising from Taylor's theorem

To set the agenda for the rest of the chapter, we return now to our solution of Example 4.2.5, where Taylor polynomials were used to estimate $e^{0.1}$. We saw that

$$e^{0.1} = 1 + (0.1) + \frac{(0.1)^2}{2!} + \frac{(0.1)^3}{3!} + \cdots \frac{(0.1)^n}{n!} + R_{n+1}(0.1), \tag{4.7}$$

where

$$0 < |R_{n+1}(0.1)| < \frac{2}{10^{n+1}(n+1)!}.$$

Since

$$\lim_{n\to\infty} \frac{2}{10^{n+1}(n+1)!} = 0,$$

it is seems reasonable to surmise that $|R_{n+1}(0.1)|$ (and hence $R_{n+1}(0.1)$) approaches 0 as $n \to \infty$. By letting n approach infinity in (4.7), this suggests that

$$e^{0.1} = 1 + (0.1) + \frac{(0.1)^2}{2!} + \frac{(0.1)^3}{3!} + \frac{(0.1)^4}{4!} + \frac{(0.1)^5}{5!} + \frac{(0.1)^6}{6!} + \cdots.$$

We therefore ask the following questions.

- What do we mean by $R_{n+1}(0.1) \to 0$ as $n \to \infty$? More generally, given a sequence of numbers $\{a_1, a_2, a_3, \ldots\}$, how do we determine the limiting behaviour of a_n as $n \to \infty$?

- What is meant by the infinite sum

$$1 + (0.1) + \frac{(0.1)^2}{2!} + \frac{(0.1)^3}{3!} + \frac{(0.1)^4}{4!} + \frac{(0.1)^5}{5!} + \frac{(0.1)^6}{6!} + \cdots?$$

  Does the sum converge to a real number or (since we are adding infinitely many positive numbers) does the sum diverge to infinity? If the sum does converge to a real number then is that number $e^{0.1}$? More generally, given an infinite sum (also called a *series*) $a_1 + a_2 + a_3 + \cdots$, how can we determine whether the series converges to a real number or else diverges?

- Each of these questions may be framed in a much larger context. Given a function f which is infinitely differentiable at a, Taylor's theorem gives

$$f(x) = f(a)+f'(a)(x-a)+\frac{f''(a)}{2!}(x-a)^2+\frac{f^{(3)}(a)}{3!}(x-a)^3+\cdots+\frac{f^{(n)}(a)}{n!}(x-a)^n+R_{n+1}(x).$$

  For what values of x will $\lim_{n\to\infty} R_{n+1}(x) = 0$? For what values of x will the infinite series

$$f(a) + f'(a)(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \frac{f^{(3)}(a)}{3!}(x-a)^3 + \cdots$$

  converge to a finite number? If the series does converge for some x, is it always true that

$$f(x) = f(a) + f'(a)(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \frac{f^{(3)}(a)}{3!}(x-a)^3 + \cdots?$$

The example of $\ln(x+1)$ given at the beginning of this section suggests that we cannot always answer 'yes' to each of these questions. To appreciate a comprehensive answer to the questions posed, we must spend some time studying the limiting behaviour of sequences and series. We do so in the next three sections.

4.3 Sequences

(Ref: SH10 §11.6–11.4)

A *sequence* is a real-valued function defined on (a subset of) the natural numbers. For example, the function $f : \mathbb{N} \to \mathbb{R}$, given by $f(n) = n^2$, is a sequence. Sequences are usually denoted using a and b, rather than f and g. In the case of sequences, subscript notation is traditionally preferred over function notation. Thus

$$a(n) = n^2 \qquad \forall n \in \mathbb{N}$$

and

$$a_n = n^2 \qquad \text{whenever } n = 0, 1, 2, 3, \ldots$$

describe the same sequence, but the second notation is used more frequently. Another way of describing the same sequence is

$$\{a_n\} = \{0, 1, 4, 16, 25, \ldots\},$$

or more precisely,

$$\{a_n\}_{n=0}^{\infty} = \{n^2\}_{n=0}^{\infty}.$$

In each case, the number a_n is called the *nth term of the sequence.*

Example 4.3.1. The sequence

$$\left\{\frac{1}{2}, \frac{2}{5}, \frac{3}{10}, \frac{4}{17}, \frac{5}{26}, \ldots\right\}$$

is described by the rule $a_n = \dfrac{n}{n^2+1}$ whenever $n \geq 1$.

Example 4.3.2. The Fibonacci sequence

$$\{1, 1, 2, 3, 5, 8, 13, 21, 34, 55, \ldots\}$$

is described *recursively* by the rule

$$a_n = \begin{cases} 1 & \text{if } n = 1 \text{ or } n = 2 \\ a_{n-1} + a_{n-2} & \text{if } n \geq 3. \end{cases}$$

Example 4.3.3. Suppose that f is $m+1$ times differentiable at a. Then for each x in $\mathbb{R}$, the remainder term $R_{n+1}(x)$ from Taylor's theorem gives rise to the (finite) sequence $\{R_{n+1}(x)\}_{n=1}^{m}$.

Since sequences are a class of functions, we can add, subtract, multiply and divide any two sequences that share a common domain. For example,

$$\{n^2\}_{n\in\mathbb{N}}\{\sqrt{n}\}_{n\in\mathbb{N}} = \{n^2\sqrt{n}\}_{n\in\mathbb{N}}.$$

If the domain (in this case $\mathbb{N}$) is understood, then one can simply write $\{n^2\}\{\sqrt{n}\} = \{n^2\sqrt{n}\}$.

4.3.1 Describing the limiting behaviour of sequences

Suppose that $\{a_n\}$ is a sequence. Our primary objective is to describe the behaviour of a_n as $n \to \infty$. There are two main types of behaviour. Either

(a) a_n approaches some finite number L, in which case we say that the sequence $\{a_n\}$ is *convergent* and write $\lim\limits_{n\to\infty} a_n = L$; or

(b) the sequence $\{a_n\}$ is not convergent, in which case we say that $\{a_n\}$ is *divergent*.

Divergent sequences can be further classified according to the list below.

(i) If $a_n \to \infty$ as $n \to \infty$ (that is, a_n grows without bound) then we say that the sequence $\{a_n\}$ *diverges to infinity*.

(ii) If $a_n \to -\infty$ as $n \to \infty$ then we say that the sequence $\{a_n\}$ *diverges to negative infinity*.

(iii) If $\{a_n\}$ has no limit as $n \to \infty$ but remains bounded then we say that $\{a_n\}$ *boundedly divergent*.

(iv) If $\{a_n\}$ exhibits none of the above behaviour then we say that $\{a_n\}$ *unboundedly divergent*.

There are rigorous definitions for each of these cases. We will see one of them later in Remark 4.3.5.

Example 4.3.4. Describe the behaviour of each sequence $\{a_n\}$ as $n \to \infty$.

(a) $a_n = n^3$

(b) $a_n = \sin(n\pi/2)$

(c) $a_n = \frac{3n^2}{n^2+4n+3}$

(d) $a_n = (-1)^n 2^n$

Solution. (a) Since $a_n \to \infty$ as $n \to \infty$, the sequence $\{a_n\}$ diverges to infinity.

(b) Since

$$\{a_n\}_{n=0}^{\infty} = \{0, 1, 0, -1, 0, 1, 0, -1, 0, \ldots\}$$

we see that $\{a_n\}$ is bounded but does not have a limit. Hence $\{a_n\}$ is boundedly divergent.

(c) Since

$$\frac{3n^2}{n^2+4n+3} = \frac{3}{1+4/n+3/n^2} \to \frac{3}{1+0+0}$$

as $n \to \infty$, we conclude that $\{a_n\}$ converges to 3.

(d) Since

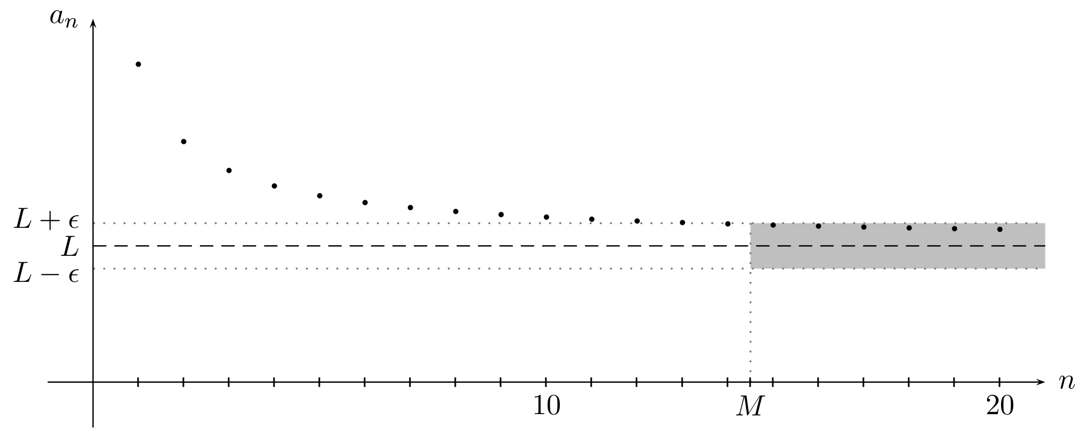
$$\{a_n\}_{n\in\mathbb{N}} = \{1, -2, 4, -8, 16, -32, 64, -128, \ldots\}$$

we see that $\{a_n\}$ is not bounded. Moreover, since the even terms of the sequence approach infinity while the odd terms approach negative infinity, we conclude that the sequence $\{a_n\}$ is unboundedly divergent. $\square$

Remark 4.3.5. The formal definition for $\lim_{n\to\infty} a_n = L$ is similar to that for $\lim_{x\to\infty} f(x) = L$ (see the MATH1131 calculus notes) and is given as follows.

> Suppose that $\{a_n\}_{n=0}^{\infty}$ is a sequence of real numbers and that $L \in \mathbb{R}$. We write
>
> $$\lim_{n\to\infty} a_n = L$$
>
> if, for every positive number ϵ, there is a number M such that $|a_n - L| < \epsilon$ whenever $n > M$.

This definition may be interpreted geometrically.



Given *any* small number ϵ, there is a point M such that the distance between the a_n and the limit L is less than ϵ for every n past M. In other words, for every small band about the limit L, there is a point M such the sequence always lies in the band past M.

4.3.2 Techniques for calculating limits of sequences

Many of the rules and techniques given in MATH1131 for calculating limits of functions also apply for limits of sequences.

The first proposition of this subsection shows that limits behave well under the standard arithmetic operations.

Proposition 4.3.6. *Suppose that* $\lim\limits_{n\to\infty} a_n$ *and* $\lim\limits_{n\to\infty} b_n$ *exist. Then*

(i) $\lim\limits_{n\to\infty}(a_n+b_n) = \lim\limits_{n\to\infty} a_n + \lim\limits_{n\to\infty} b_n$*;*

(ii) $\lim\limits_{n\to\infty}(a_nb_n) = \lim\limits_{n\to\infty} a_n \times \lim\limits_{n\to\infty} b_n$*;*

(iii) $\lim\limits_{n\to\infty}\dfrac{a_n}{b_n} = \dfrac{\lim\limits_{n\to\infty} a_n}{\lim\limits_{n\to\infty} b_n}$*, provided that* $\lim\limits_{n\to\infty} b_n \neq 0$ *and* $b_n \neq 0$ *for any* n*; and*

(iv) $\lim\limits_{n\to\infty}(\alpha a_n) = \alpha \lim\limits_{n\to\infty} a_n$ *for every real number* α*.*

Example 4.3.7. Suppose that $a_n = \sqrt{n^2+4n} - n$. Determine the limiting behaviour of a_n as $n\to\infty$.

Solution. We proceed in the same manner as if we were asked to calculate $\lim\limits_{x\to\infty}(\sqrt{x^2+4x}-x)$. Now

$$\begin{aligned}
a_n &= \sqrt{n^2+4n}-n \\
&= \frac{(\sqrt{n^2+4n}-n)(\sqrt{n^2+4n}+n)}{(\sqrt{n^2+4n}+n)} && \text{(multiplying top and bottom by the ‘conjugate’)} \\
&= \frac{n^2+4n-n^2}{\sqrt{n^2+4n}+n} && \text{(difference of two squares)} \\
&= \frac{4n}{\sqrt{n^2+4n}+n} \\
&= \frac{4}{\sqrt{1+4/n}+1} && \text{(dividing top and bottom by } n\text{).}
\end{aligned}$$

Clearly $4/n \to 0$ as $n\to\infty$ and so by applying the limit rules of the above proposition we find that $a_n \to 2$ as $n\to\infty$. □

The next proposition can be used when a function is composed with a sequence.

Proposition 4.3.8. *Suppose that* $\lim\limits_{n\to\infty} a_n = a$ *and that* f *is continuous at* a*. Then*

$$\lim_{n\to\infty} f(a_n) = f(a).$$

This proposition is easy to remember if f is continuous everywhere; it amounts to saying that the function and limit can be swapped, as shown below:

$$\lim_{n\to\infty} f(a_n) = f\left(\lim_{n\to\infty} a_n\right).$$

Example 4.3.9. Find $\lim\limits_{n\to\infty} \sin\left(\dfrac{\pi n^2}{4n^2+1}\right)$, if it exists.

Solution. Note that the sine function is continuous everywhere. Therefore

$$\begin{aligned}
\lim_{n\to\infty} \sin\left(\frac{\pi n^2}{4n^2+1}\right) &= \sin\left(\lim_{n\to\infty} \frac{\pi n^2}{4n^2+1}\right) \\
&= \sin\left(\lim_{n\to\infty} \frac{\pi}{4+1/n^2}\right) \\
&= \sin\left(\frac{\pi}{4}\right) \\
&= \frac{1}{\sqrt{2}}.
\end{aligned}$$

□

The following rule allows the use of l'Hôpital's rule when calculating limits of sequences.

Proposition 4.3.10. *Suppose that $\{a_n\}$ is a sequence and f is a function defined on some interval (b, ∞). If $a_n = f(n)$ for all n sufficiently large and $\lim_{x\to\infty} f(x)$ exists then*

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

Note that this proposition only works for limits which are finite. If $a_n = f(n)$ and $f(x)$ diverges as $x \to \infty$, one cannot say that a_n diverges. (Consider, for example, the case when $a_n = 0$ for all n and $f(x) = x\sin(\pi x)$ for all x.)

The next example shows how Proposition 4.3.10 is applied in conjunction with l'Hôpital's rule.

Example 4.3.11. Suppose that $a_n = (1 + 1/n)^n$. Determine the limiting behaviour of a_n as $n \to \infty$.

Solution. Suppose that $f(x) = (1 + 1/x)^x$ whenever $x > 0$. Then $a_n = f(n)$ whenever $n > 0$. Hence

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

The method for calculating the limit of f is standard: we first take the logarithm of f (to remove the power) and then rearrange the resulting function so that we can apply l'Hôpital's rule. That is,

$$\begin{aligned}
\lim_{n\to\infty} a_n &= \lim_{x\to\infty}\left(1+\frac{1}{x}\right)^x \\
&= \lim_{x\to\infty}\exp\left\{\ln\left(1+\frac{1}{x}\right)^x\right\} && \text{(since ln and exp are inverses)} \\
&= \exp\left\{\lim_{x\to\infty}\ln\left(1+\frac{1}{x}\right)^x\right\} && \text{(since exp is continuous everywhere)} \\
&= \exp\left\{\lim_{x\to\infty} x\ln\left(1+\frac{1}{x}\right)\right\} && \text{(by the log law)} \\
&= \exp\left\{\lim_{x\to\infty}\frac{\ln(1+1/x)}{1/x}\right\} && \text{(to prepare for l'Hôpital's rule)} \\
&= \exp\left\{\lim_{x\to\infty}\frac{\left(\frac{-1/x^2}{1+1/x}\right)}{-1/x^2}\right\} && \text{(by l'Hôpital's rule)} \\
&= \exp\left\{\lim_{x\to\infty}\frac{1}{1+1/x}\right\} && \text{(simplifying the fraction)} \\
&= \exp\{1\} \\
&= e.
\end{aligned}$$

In summary, a_n converges to e as $n \to \infty$. □

Remark 4.3.12. The limit $\lim_{n\to\infty}\left(1+\frac{1}{n}\right)^n = e$ is a standard result and should be familiar to students.

Finally, we present a version of the pinching theorem for sequences.

Proposition 4.3.13 (The pinching theorem for sequences)**.** *Suppose that $\{a_n\}$, $\{b_n\}$ and $\{c_n\}$ are sequences and that for some positive integer N the inequality*

$$a_n \le b_n \le c_n$$

is satisfied whenever $n > N$. If $\lim_{n\to\infty} a_n = \lim_{n\to\infty} c_n = L$ then $\lim_{n\to\infty} b_n = L$.

The following example is important because sequences involving both factorials and powers arise frequently in applications of Taylor's theorem.

Example 4.3.14. Suppose that $a_n = \dfrac{n!}{n^n}$. Discuss the limiting behaviour of a_n as $n \to \infty$.

Solution. Note that

$$\begin{aligned}
a_n &= \frac{1}{n}\cdot\frac{2}{n}\cdot\frac{3}{n}\cdots\frac{n}{n} \\
&\le \frac{1}{n}\cdot\frac{n}{n}\cdot\frac{n}{n}\cdots\frac{n}{n} \\
&= \frac{1}{n}
\end{aligned}$$

whenever $n \ge 1$. On the other hand, a_n is always positive. Thus

$$0 \le a_n \le \frac{1}{n}.$$

As $n \to \infty$ we conclude that $a_n \to 0$ by the pinching theorem. □

A similar technique was used in Lemma 2.2.5 to prove that

$$\lim_{n\to\infty} \frac{c^n}{n!} = 0$$

whenever $c > 0$.

Remark 4.3.15. It is helpful to have a good intuition of the *order of growth of sequences*. The following table compares the growth of various sequences as $n \to \infty$. The lower down on the table the sequence appears, the faster it grows at infinity.

a_n	growth rate as $n \to \infty$
1	constant: does not grow
$\ln n$	grows slowly
n^k, where $k > 0$	growth rate is faster for larger k
c^n, where $c > 1$	growth rate is faster for larger c
$n!$	grows rapidly
n^n	grows very rapidly

For example, the ordering in the table reflects the fact that

$$\lim_{n\to\infty} \frac{e^n}{n!} = 0 \qquad \text{while} \qquad \lim_{n\to\infty} \frac{n!}{n^n} = 0.$$

The final theorem is of great theoretical importance and will be used in later sections. We begin with a definition.

Definition 4.3.16. A sequence $\{a_n\}_{n=0}^{\infty}$ of real numbers is said to be

(a) *increasing* if $a_n < a_{n+1}$ for each natural number n,

(b) *nondecreasing* if $a_n \le a_{n+1}$ for each natural number n,

(c) *decreasing* if $a_n > a_{n+1}$ for each natural number n, and

(d) *nonincreasing* if $a_n \ge a_{n+1}$ for each natural number n.

If any of these four properties holds then the sequence is said to be *monotonic*.

Theorem 4.3.17. *If $\{a_n\}_{n=0}^{\infty}$ is a bounded monotonic sequence of real numbers then it is converges to some real number L.*

This theorem is proved using a property that distinguishes the real numbers from the rational numbers. Given a bounded monotonic sequence of rational numbers, it is not true, in general, that the sequence converges to a rational number.

4.3.3 Suprema and infima [X]

In this subsection we summarise some more advanced ideas for students studying MATH1241.

Definition 4.3.18. Suppose that $\{a_n\}_{n=0}^{\infty}$ is a sequence of real numbers.

(a) We say that M is an *upper bound* for $\{a_n\}_{n=0}^{\infty}$ if $a_n \le M$ for every natural number n.

(b) We say that M is a *lower bound* for $\{a_n\}_{n=0}^{\infty}$ if $a_n \ge M$ for every natural number n.

(c) We say that K is *the least upper bound* for $\{a_n\}_{n=0}^{\infty}$ if K is an upper bound for $\{a_n\}_{n=0}^{\infty}$ and $K \le M$ whenever M is an upper bound for $\{a_n\}_{n=0}^{\infty}$.

(d) We say that K is *the greatest lower bound* for $\{a_n\}_{n=0}^{\infty}$ if K is a lower bound for $\{a_n\}_{n=0}^{\infty}$ and $K \ge M$ whenever M is a lower bound for $\{a_n\}_{n=0}^{\infty}$.

Example 4.3.19. Find the greatest lower bound and least upper bound for the sequence $\{a_n\}_{n=1}^{\infty}$, where $a_n = \dfrac{(-1)^n n}{n+1}$. Prove your answer.

Solution. We write out the first few terms of the sequence to get a feel for what is happening:

$$\left\{-\frac{1}{2}, \frac{2}{3}, -\frac{3}{4}, \frac{4}{5}, -\frac{5}{6}, \frac{6}{7}, -\frac{7}{8}, \dots\right\}.$$

It is clear that the odd terms approach -1 (from above) while the even terms approach 1 (from below).

We will prove that 1 is the least upper bound. First, it is clear that $|a_n| = \frac{n}{n+1} < 1$ for every positive integer n. So 1 is an upper bound. Suppose now that K is also an upper bound but that $K < 1$. Hence $K = 1 - \epsilon$ for some positive number ϵ, while

$$\frac{n}{n+1} < K$$

for every positive integer n. Therefore

$$\frac{n}{n+1} < 1 - \epsilon$$

and hence

$$1 - \frac{1}{n+1} < 1 - \epsilon$$

for every positive integer n. But rearranging this inequality gives

$$n < \frac{1}{\epsilon} - 1$$

for every positive integer n, which gives a contradiction since the set of positive integers has no upper bound. Hence no such K exists. We conclude that 1 is the least upper bound.

Using a similar technique, one can show that -1 is the greatest lower bound for the sequence.

Note that the sequence has neither a maximum nor minimum value. ☐

The fact that every bounded monotonic sequence of real numbers has a limit in $\mathbb{R}$ (see Theorem 4.3.17) follows from one of the axioms of the real number system. This axiom is called the *least upper bound axiom* and may be stated as

> 'Every nonempty set of real numbers that has an upper bound has a *least* upper bound.'

Note that this axiom is not true for the rational number system.

To prove Theorem 4.3.17 in the case when $\{a_n\}_{n=0}^{\infty}$ is a bounded increasing sequence of real numbers, we note that the values of the sequence forms a bounded nonempty set of real numbers. By the least upper bound axiom, it therefore has an upper bound, which we denote by L. Using the definition of the limit (see Remark 4.3.5), one can now show that $\lim_{n\to\infty} a_n = L$. The proof of the other cases is similar.

We now introduce some alternate terminology and new notation for least upper bound and greatest lower bound.

Definition 4.3.20. Suppose that $\{a_n\}_{n=0}^{\infty}$ is a sequence of real numbers.

(a) If $\{a_n\}_{n=0}^{\infty}$ has a least upper bound M, then M is also called the *supremum* of $\{a_n\}_{n=0}^{\infty}$ and is denoted by

$$\sup_{n\geq 0} a_n \qquad \text{or} \qquad \sup\{a_n : n \geq 0\}.$$

(b) If $\{a_n\}_{n=0}^{\infty}$ has a greatest lower bound M, then M is also called the *infimum* of $\{a_n\}_{n=0}^{\infty}$ and is denoted by

$$\inf_{n\geq 0} a_n \qquad \text{or} \qquad \inf\{a_n : n \geq 0\}.$$

The plural for supremum and infimum is *suprema* and *infima*.

4.4 Infinite series

(Ref: SH10 §12.1, 12.2)

At the end of Section 4.2, we asked the question What is meant by the infinite sum

$$1 + (0.1) + \frac{(0.1)^2}{2!} + \frac{(0.1)^3}{3!} + \frac{(0.1)^4}{4!} + \frac{(0.1)^5}{5!} + \frac{(0.1)^6}{6!} + \cdots?$$

In this section, we will develop a framework that gives meaning to this infinite sum by using existing notions for convergence (and divergence) of sequences. The key is to recognise that

- the terms of the above series form a sequence $\{a_k\}_{k=0}^{\infty}$, where
$$a_k = \frac{(0.1)^k}{k!};$$
and
- if s_n denotes the sum of the first n terms in the series (so that
$$s_n = 1 + (0.1) + \frac{(0.1)^2}{2!} + \frac{(0.1)^3}{3!} + \cdots + \frac{(0.1)^n}{n!}$$
whenever $n \geq 0$) then $\{s_n\}_{n=0}^{\infty}$ is also a sequence.

Thus questions concerning the meaning of an infinite series can be reduced to studying the limiting behaviour of $\{s_n\}_{n=0}^{\infty}$. (Naturally, the limiting behaviour of $\{s_n\}$ depends on limiting properties of the sequence $\{a_k\}$; we will pay more attention to this aspect of the theory in Section 4.5.)

Definition 4.4.1. Suppose that $\{a_k\}_{k=0}^{\infty}$ is a sequence of real numbers. For each natural number n, let s_n denote the nth partial sum given by

$$s_n = a_0 + a_1 + a_2 + \cdots + a_n = \sum_{k=0}^{n} a_k.$$

If the sequence $\{s_n\}_{n=0}^{\infty}$ of partial sums converges to a number L then we say that the infinite series $\sum_{k=0}^{\infty} a_k$ *converges to* L and we write

$$\sum_{k=0}^{\infty} a_k = L.$$

In this case we also say that the series is *summable.* If the sequence $\{s_n\}_{n=0}^{\infty}$ of partial sums diverges then we say that the infinite series $\displaystyle\sum_{k=0}^{\infty} a_k$ *diverges.*

The following example illustrates this definition.

Example 4.4.2 (Geometric series)**.** Suppose that $r \in \mathbb{R}$ and consider the *geometric series*

$$\sum_{k=0}^{\infty} r^k = 1 + r + r^2 + r^3 + r^4 + \cdots .$$

Determine the values of r for which the series (a) converges and (b) diverges.

Solution. If $r \neq 1$ then s_n is given by the formula

$$s_n = \frac{1 - r^{n+1}}{1 - r}$$

(as is taught in high school). In the case when $r = 1$, we simply have

$$s_n = \underbrace{1 + 1 + 1 + \cdots + 1}_{n \text{ times}} = n.$$

In summary,

$$s_n = \begin{cases} \frac{1-r^{n+1}}{1-r} & \text{if } r \neq 1 \\ n & \text{if } r = 1. \end{cases}$$

To determine the convergence (or otherwise) of the infinite series $\sum_{k=0}^{\infty} r^k$, we simply determine the convergence of $\{s_n\}_{n=0}^{\infty}$. We break this up into four cases.

- If $|r| < 1$ then $r^{n+1} \to 0$ as $n \to \infty$ and so

  $$\lim_{n\to\infty} s_n = \lim_{n\to\infty} \frac{1 - r^{n+1}}{1 - r} = \frac{1}{1 - r}.$$

  Since the sequence of partial sums converges, so does the series and thus

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

- If $|r| > 1$ then r^{n+1} diverges (either to infinity if $r > 1$ or unboundedly if $r < -1$). The sequence $\{s_n\}$ of partial sums therefore diverges and consequently the series $\sum_{k=0}^{\infty} r^k$ also diverges.
- If $r = 1$ then $s_n = n \to \infty$ as $n \to \infty$. Consequently the series diverges.
- If $r = -1$ then it is easily seen that $\{s_n\}_{n=0}^{\infty} = \{1, 0, 1, 0, 1, \ldots\}$. Hence $\{s_n\}$ is boundedly divergent and the series

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

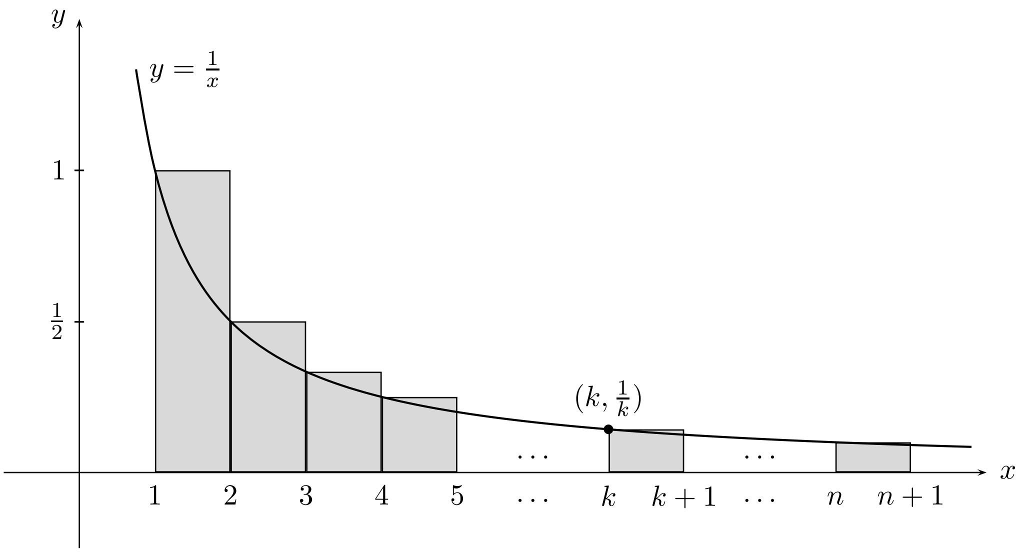
also diverges.

$\square$

The next example illustrates the technique of comparing a series with an integral.

Example 4.4.3 (The harmonic series)**.** Show that the *harmonic series* $\sum_{k=1}^{\infty} \frac{1}{k}$ diverges.

Solution. This proof uses results from integration. Consider the diagram below.



It is clear that the area of the kth rectangle is $1/k$, which is also equal to the kth term of the series. Moreover, the area under the rectangles on the interval $[1, n+1]$ is greater than the area under the curve under the same interval. From these two observations it follows that

$$\begin{aligned} s_n &= \sum_{k=1}^{n} \frac{1}{k} \\ &\geq \int_1^{n+1} \frac{1}{x}\,dx \\ &= \Big[\ln x\Big]_1^{n+1} \\ &= \ln(n+1). \end{aligned}$$

Now $\ln(n+1) \to \infty$ as $n \to \infty$, so we conclude that $s_n \to \infty$ as $n \to \infty$. Hence the series diverges. $\square$

Using the technique illustrated in the previous example, one can show that

- $\displaystyle\sum_{k=1}^{\infty} \frac{1}{k^2}$ converges,
- $\displaystyle\sum_{k=1}^{\infty} \frac{1}{\sqrt{k}}$ diverges, and
- $\displaystyle\sum_{k=2}^{\infty} \frac{1}{k \ln k}$ diverges.

All one needs to do is draw a diagram and compare the series with an appropriate improper integral.

Since a convergent infinite series is the limit of a sequence (the sequence of partial sums), many results for sequences can be interpreted as results for series. The following proposition illustrates this point.

Proposition 4.4.4. *Suppose that* $\displaystyle\sum_{k=0}^{\infty} a_k$ *and* $\displaystyle\sum_{k=0}^{\infty} b_k$ *are two summable series. Then*

(i) $\displaystyle\sum_{k=0}^{\infty}(a_k + b_k) = \sum_{k=0}^{\infty} a_k + \sum_{k=0}^{\infty} b_k$*; and*

(ii) $\displaystyle\sum_{k=0}^{\infty}(\alpha a_k) = \alpha \sum_{k=0}^{\infty} a_k$ *for every real number* α.

Proof. Let s_n and t_n denote the partial sums of $\sum_{k=0}^{\infty} a_k$ and $\sum_{k=0}^{\infty} b_k$ respectively. Now apply Proposition 4.3.6 (i) and (iv) to the sequences $\{s_n\}$ and $\{t_n\}$. $\square$

Remark 4.4.5. While all the terms of a convergent series contribute to the value of the series, the convergence (or otherwise) of any series only depends on the 'tail' of the series. That is, the first hundred, thousand or even billion terms of the series are irrelevant to the question of whether the series converges. More precisely, given any positive integer N,

$$\sum_{k=0}^{\infty} a_k \quad \text{converges} \qquad \text{if and only if} \qquad \sum_{k=N}^{\infty} a_k \quad \text{converges.}$$

With this in mind, all of the theorems presented in the next section are just as true for series of the form

$$\sum_{k=1}^{\infty} a_k, \qquad \sum_{k=50}^{\infty} a_k \qquad \text{or} \qquad \sum_{k=2000}^{\infty} a_k$$

as they are for series of the form $\displaystyle\sum_{k=0}^{\infty} a_k$. When the starting point for the series does not matter, one sometimes simply writes

$$\sum^{\infty} a_k \qquad \text{or} \qquad \sum a_k.$$

4.5 Tests for series convergence

(Ref: SH10 §12.3–12.5)

To determine whether a series

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

converges or diverges, mathematicians have developed some simple tests. Typically, these tests examine the behaviour of the *sequence* $\{a_k\}$ and thereby deduce the convergence (or otherwise) of the corresponding *series*. In this section we introduce three such tests: the *kth term test*, the *ratio test* and the *alternating series test*.

Remark 4.5.1 (Warning)**.** Care must be taken through this section not to confuse sequences and series. For example, suppose that $a_k = \frac{1}{k}$. While the *sequence* $\{a_k\}$ converges, the infinite *series* $\sum a_k$ does not.

4.5.1 Some preliminary results on series summation

The next two results are fundamental to the study of infinite series. They will later be used to establish some simple tests for convergence or divergence.

Lemma 4.5.2. *Suppose that $\{a_k\}_{k=0}^{\infty}$ is a sequence of positive numbers and let s_n denote the partial sum given by*

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

If $\{s_n\}_{n=0}^{\infty}$ is a bounded sequence then the infinite series $\sum_{k=0}^{\infty} a_k$ is convergent.

Proof. For any natural number n,

$$s_{n+1} = s_n + a_{n+1} > s_n,$$

since a_{n+1} is positive. Hence $\{s_n\}_{n=0}^{\infty}$ is a bounded increasing sequence and hence has a limit L (see Theorem 4.3.17). Therefore

$$\sum_{k=0}^{\infty} a_k = L$$

and the series converges. ☐

4.5.2 The kth term divergence test

The next test we introduce is a simple test for divergence. *One should always use this test first when trying to decide whether a series converges.*

Theorem 4.5.3 (The kth term test for divergence.)**.** *If $a_k \not\to 0$ as $k \to \infty$ then $\sum_{k=0}^{\infty} a_k$ diverges.*

Before proving the theorem, we give an example.

Example 4.5.4. Determine whether the series $\displaystyle\sum_{k=1}^{\infty} \frac{k^2+2k}{\sqrt{k^4+2}}$ converges.

Solution. If $a_k = \dfrac{k^2+2k}{\sqrt{k^4+2}}$ then

$$\begin{aligned}\lim_{k\to\infty} a_k &= \lim_{k\to\infty} \frac{k^2+2k}{\sqrt{k^4+2}} \\ &= \lim_{k\to\infty} \frac{1+2/k}{\sqrt{1+2/k^4}} && \text{(by dividing top and bottom by } k^2\text{)} \\ &= 1.\end{aligned}$$

Since $a_k \not\to 0$ as $k \to \infty$, the series diverges by the kth term test. □

Remark 4.5.5. The kth term test is not a test for convergence. For example, consider the series $\sum_{k=1}^{\infty} \frac{1}{k}$. In this case, $\lim_{k\to\infty} 1/k = 0$ but the series diverges (see Example 4.4.3).

The kth term divergence test is equivalent to the following theorem, whose proof we give below.

Theorem 4.5.6. *If the series $\displaystyle\sum_{k=0}^{\infty} a_k$ converges then $a_k \to 0$ as $k \to \infty$.*

Proof. [**H**] Suppose that $\displaystyle\sum_{k=0}^{\infty} a_k$ converges to the real number L and let s_n denote the nth partial sum of the series. Then

$$\begin{aligned}s_n - s_{n-1} &= (a_0 + a_1 + \cdots + a_{n-1} + a_n) - (a_0 + a_1 + \cdots + a_{n-1}) \\ &= a_n. \qquad (4.8)\end{aligned}$$

Now $\lim_{n\to\infty} s_n = \lim_{n\to\infty} s_{n-1} = L$ and so

$$\begin{aligned}\lim_{n\to\infty} a_n &= \lim_{n\to\infty} (s_n - s_{n-1}) && \text{(by (4.8))} \\ &= \lim_{n\to\infty} s_n - \lim_{n\to\infty} s_{n-1} \\ &= L - L \\ &= 0,\end{aligned}$$

thus completing the proof. □

4.5.3 The integral test

The idea from Example 4.4.3, where we bounded a sum by an integral, can be applied more generally to produce a test for either convergence or divergence.

Theorem 4.5.7 (The integral test)**.** *Suppose that $\sum a_k$ is an infinite series with positive terms. Suppose $f(x)$ is a positive integrable function decreasing on $[1,\infty)$ such that for each positive integer k, $f(k) = a_k$.*

(i) If $\displaystyle\int_1^{\infty} f(x)\,dx$ converges then so does $\displaystyle\sum_{k=1}^{\infty} a_k$.

(ii) If $\displaystyle\int_1^{\infty} f(x)\,dx$ diverges then so does $\displaystyle\sum_{k=1}^{\infty} a_k$.

The proof is similar to that given in Example 4.4.3.

Example 4.5.8. Determine whether or not the following series converge.

(a) $\displaystyle\sum_{k=1}^{\infty} \frac{1}{k^2}$ (b) $\displaystyle\sum_{k=1}^{\infty} \frac{k}{2k^2+1}$ (c) $\displaystyle\sum_{k=2}^{\infty} \frac{1}{k(\log k)^2}$

Proof. (a) Consider the improper integral, $\displaystyle\int_1^{\infty} \frac{1}{x^2}\,dx$.

$$\int_1^{\infty} \frac{1}{x^2}\,dx = \lim_{N\to\infty} \int_1^N \frac{1}{x^2}\,dx = \lim_{N\to\infty} \left[-\frac{1}{x}\right]_1^N = 1.$$

Since the improper integral converges, so does the series $\displaystyle\sum_{k=1}^{\infty} \frac{1}{k^2}$.

(Note: This is a famous series, first summed by Euler. Its value is remarkably $\frac{\pi^2}{6}$. This will be proven in later courses, but a proof of it appeared in the NSW Extension 2 paper, 2010.)

(b) Consider the improper integral, $\displaystyle\int_1^{\infty} \frac{x}{2x^2+1}\,dx$.

$$\int_1^{\infty} \frac{x}{2x^2+1}\,dx = \lim_{N\to\infty} \int_1^N \frac{x}{2x^2+1}\,dx = \lim_{N\to\infty} \left[\frac{1}{4}\log(2x^2+1)\right]_1^N \to \infty.$$

Since the improper integral diverges, so does the series $\displaystyle\sum_{k=1}^{\infty} \frac{k}{2k^2+1}$.

(c) Consider the improper integral, $\displaystyle\int_2^{\infty} \frac{1}{x(\log x)^2}\,dx$.

$$\int_2^{\infty} \frac{1}{x(\log x)^2}\,dx = \lim_{N\to\infty} \int_2^N \frac{1}{x(\log x)^2}\,dx = \lim_{N\to\infty} \left[-\frac{1}{\log x}\right]_2^N = \frac{1}{\log 2}.$$

Since the improper integral converges, so does the series $\displaystyle\sum_{k=1}^{\infty} \frac{1}{k(\log k)^2}$. ☐

4.5.4 The comparison test

The integral test was basically a comparison between each term of a given series and the area of a corresponding rectangle. This idea can also be applied to the terms of two series. If each term of a given series is less than each term of another series - whose convergence is easy to determine, then we can conclude the given series also converges. A similar test can be found for divergence.

Theorem 4.5.9 (The comparison test). *Suppose that $\{a_k\}_{k=0}^{\infty}$ and $\{b_k\}_{k=0}^{\infty}$ are two positive sequences such that $a_k \le b_k$ for every natural number k.*

(i) If $\sum_{k=0}^{\infty} b_k$ converges then $\sum_{k=0}^{\infty} a_k$ also converges.

(ii) If $\sum_{k=0}^{\infty} a_k$ diverges then $\sum_{k=0}^{\infty} b_k$ also diverges.

The comparison test is often used in conjunction with series of the following type.

Proposition 4.5.10 (Convergence and divergence of p-series)**.** *The series*

$$\sum_{k=1}^{\infty} \frac{1}{k^p}$$

converges if $p > 1$ and diverges if $p \leq 1$.

Proof. This theorem may be easily proved by the integral test, or by adapting the proof in Example 4.4.3. The details are left to the reader. $\square$

Example 4.5.11. Determine whether or not the following series converge.

$$\text{(a)} \sum_{k=1}^{\infty} \frac{k}{k^3+1} \qquad \text{(b)} \sum_{k=2}^{\infty} \frac{1}{\sqrt{k^2-1}} \qquad \text{(c)} \sum_{k=1}^{\infty} \frac{1}{\sqrt{k^2+1}}$$

Typically, one needs some intuition as to whether the series will converge or diverge before the comparsion test is used to construct a rigorous solution.

Solution. (a) By considering the dominant term (as $k \to \infty$), we see that

$$\frac{k}{k^3+1} \approx \frac{k}{k^3} = \frac{1}{k^2}$$

whenever k is a large positive integer. Since $\sum \frac{1}{k^2}$ converges (p-series when $p = 2$), this suggests that $\sum_{k=1}^{\infty} \frac{k}{k^3+1}$ also converges.

To prove that $\sum_{k=1}^{\infty} \frac{k}{k^3+1}$ converges, we use part (i) of the comparison test. Note that

$$0 \leq \frac{k}{k^3+1} \leq \frac{k}{k^3} = \frac{1}{k^2}$$

whenever $k \geq 1$. So suppose that $a_k = \frac{k}{k^3+1}$ and $b_k = \frac{1}{k^2}$. We have shown that $0 \leq a_k \leq b_k$. Since $\sum b_k$ converges, it follows from the comparison test that $\sum a_k$ converges.

(b) Looking at the dominant terms, $\displaystyle\sum_{k=2}^{\infty} \frac{1}{\sqrt{k^2-1}} \approx \sum_{k=2}^{\infty} \frac{1}{\sqrt{k^2}} = \sum_{k=2}^{\infty} \frac{1}{k}$.

Since this series diverges (p-series when $p = 1$), this suggests that $\displaystyle\sum_{k=2}^{\infty} \frac{1}{\sqrt{k^2-1}}$ also diverges.

To prove that $\displaystyle\sum_{k=2}^{\infty} \frac{1}{\sqrt{k^2-1}}$ diverges, we use part (ii) of the comparison test.

Note that $\dfrac{1}{\sqrt{k^2-1}} \geq \dfrac{1}{\sqrt{k^2}} = \dfrac{1}{k}$, whenever $k \geq 2$. So suppose that $a_k = \frac{1}{\sqrt{k^2-1}}$ and $b_k = \frac{1}{k}$. We have shown that $a_k \geq b_k$.

Since $\sum b_k$ diverges, it follows from the comparison test that $\sum a_k$ diverges.

(c) The same analysis as in (b) suggests that this series also diverges. However, the inequality $\dfrac{1}{\sqrt{k^2+1}} \geq \dfrac{1}{\sqrt{k^2}} = \dfrac{1}{k}$ is *false*. To overcome this hurdle, we introduce a 'fudge factor' to obtain the inequality we want. Thus, $\dfrac{1}{\sqrt{k^2+1}} \geq \dfrac{1}{2\sqrt{k^2}} = \dfrac{1}{2k}$ for all sufficiently large k (in fact $k \geq 1$ will do-you should check this!). So suppose that $a_k = \frac{1}{\sqrt{k^2+1}}$ and $b_k = \frac{1}{2k}$. We have shown that $a_k \geq b_k$.

Since $\sum b_k$ diverges, it follows from the comparison test that $\sum a_k$ diverges. $\square$

4.5.5 [X] The limit form of the comparison test

This form of the comparison test is extremely useful and allows us to rely on our intuition without having to work with inequalities. The disadvantage is that it does not always work in quite the same way as the straight comparison test does.

Proposition 4.5.12. *Suppose a_n, b_n are sequences with positive terms and suppose $\lim_{n\to\infty} \frac{a_n}{b_n}$ is finite and* **not** *zero, then $\sum^{\infty} a_n$ converges if and only if $\sum^{\infty} b_n$ converges.*

Proof. Suppose $\lim_{n\to\infty} \frac{a_n}{b_n} = K > 0$. For any given $\epsilon > 0$, we have

$$\left|\frac{a_k}{b_k} - K\right| < \epsilon$$

for all sufficiently large k. For such k we have

$$K - \epsilon < \frac{a_k}{b_k} < K + \epsilon \Rightarrow (K - \epsilon)b_k < a_k < (K + \epsilon)b_k.$$

Thus, from the last inequality, if $\sum a_k$ converges then $\sum (K - \epsilon)b_k$ converges and hence $\sum b_k$ does also; if $\sum b_k$ converges then $\sum (K + \epsilon)b_k$ converges and hence $\sum a_k$ does also. $\square$

Example 4.5.13. Discuss the convergence of $\displaystyle\sum_{k=5}^{\infty} \frac{k^2}{k^4 + 3}$

Proof. For large k the summand $a_k = \frac{k^2}{k^4+3}$ is roughly $b_k = \frac{1}{k^2}$.

Now $\lim_{n\to\infty} \frac{a_n}{b_n} = 1$. Hence, since $\displaystyle\sum_{k=5}^{\infty} \frac{1}{k^2}$ converges (by p-series with $p = 2$), so does $\displaystyle\sum_{k=5}^{\infty} \frac{k^2}{k^4 + 3}$. $\square$

Example 4.5.14. Discuss the convergence of $\displaystyle\sum_{k=1}^{\infty} \sin(\frac{1}{k})$

Proof. Since $\sin x \approx x$ for small x, we try comparing $a_k = \sin(\frac{1}{k})$, (whose terms are positive), with $b_k = \frac{1}{k}$.

Now $\lim_{n\to\infty} \frac{a_n}{b_n} = \frac{\sin(\frac{1}{k})}{\frac{1}{k}} = 1$. Hence, since $\displaystyle\sum_{k=1}^{\infty} \frac{1}{k}$ diverges (by p-series with $p = 1$), so does $\displaystyle\sum_{k=1}^{\infty} \sin(\frac{1}{k})$. $\square$

Remark 4.5.15. The series $\sum \frac{1}{\log n}$ diverges, since, for $n \geq 1$, we have $\log n < n$ giving $\frac{1}{\log n} > \frac{1}{n}$ and so the comparison test may be applied. On other hand, is we try to use the limit comparison test with $a_n = \frac{1}{\log n}$ and $b_n = \frac{1}{n}$, then $\lim_{n\to\infty} \frac{a_n}{b_n} = \lim_{n\to\infty} \frac{n}{\log n} \to \infty$ and so the test fails.

4.5.6 The ratio test

We now introduce a simple convergence and divergence test known as the ratio test. It is important to note that this test can only be applied to series whose terms are positive.

Theorem 4.5.16 (The ratio test)**.** *Suppose that $\sum a_k$ is an infinite series with positive terms and that*

$$\lim_{k\to\infty} \frac{a_{k+1}}{a_k} = r.$$

(i) If $r < 1$ then $\sum a_k$ converges.

(ii) If $r > 1$ then $\sum a_k$ diverges.

Remark 4.5.17. The ratio test does not specify what happens if $r = 1$. In this case, the test is inconclusive; the series may converge or diverge.

The reason the ratio test works is that the tail of any series $\sum a_k$ with 'ratio' r given by

$$r = \lim_{k\to\infty} \frac{a_{k+1}}{a_k}$$

behaves like a geometric series with common ratio r. When $r < 1$, the geometric series is convergent and hence $\sum a_k$ also converges. Similarly, When $r > 1$, the geometric series is divergent and therefore $\sum a_k$ also diverges. Of course, these assertions need to be proved.

Before we see a proof of the ratio test, we shall see how it is applied. As seen in the examples below, the ratio test is particularly useful when $k!$ or kth powers appear in each term a_k.

Example 4.5.18. Determine whether or not the following series converge.

(a) $\displaystyle\sum_{k=1}^{\infty} \frac{1}{k!}$ (b) $\displaystyle\sum_{k=1}^{\infty} \frac{k^k}{k!}$ (c) $\displaystyle\sum_{k=1}^{\infty} \frac{1}{k}$

Solution. (a) Suppose that $a_k = \dfrac{1}{k!}$. Then

$$r = \lim_{k\to\infty} \frac{a_{k+1}}{a_k} = \lim_{k\to\infty} \frac{k!}{(k+1)!} = \lim_{k\to\infty} \frac{1}{k+1} = 0.$$

Since $r < 1$, the series converges by the ratio test.

(b) Suppose that $a_k = \dfrac{k^k}{k!}$. Then

$$\begin{aligned}
r &= \lim_{k\to\infty} \frac{a_{k+1}}{a_k} \\
&= \lim_{k\to\infty} \frac{(k+1)^{k+1}}{(k+1)!} \cdot \frac{k!}{k^k} \\
&= \lim_{k\to\infty} \frac{(k+1)^k(k+1)}{k!(k+1)} \cdot \frac{k!}{k^k} \\
&= \lim_{k\to\infty} \frac{(k+1)^k}{k^k} \\
&= \lim_{k\to\infty} \left(\frac{k+1}{k}\right)^k \\
&= \lim_{k\to\infty} \left(1 + \frac{1}{k}\right)^k \\
&= e
\end{aligned}$$

(see Example 4.3.11 for a calculation of this well known limit). Since $r = e > 1$, the series diverges by the ratio test.

(c) Since $r = \displaystyle\lim_{k\to\infty} \frac{k}{k+1} = 1$, we cannot say from the ratio test whether or not the series converges. It can be shown using another method that the series diverges (see Example 4.4.3). $\square$

Sketch proof of the ratio test. [**X**] (i) Suppose that $r < 1$ and choose R such that $r < R < 1$. Since $\displaystyle\lim_{k\to\infty} \frac{a_{k+1}}{a_k} = r$, the terms of the sequence $\left\{\frac{a_{k+1}}{a_k}\right\}$ eventually gets so close to r that they must also be less than R. More precisely, there is an integer N such that

$$\frac{a_{k+1}}{a_k} < R \qquad \text{whenever } k \geq N. \tag{4.9}$$

Our goal from here is to show that the tail $\sum_{k=N}^{\infty} a_k$ of the series can be bounded above by a convergent geometric series with common ratio R. From (4.9) we see that

$$a_{N+1} < Ra_N, \qquad a_{N+2} < Ra_{N+1} < R^2 a_N, \qquad a_{N+3} < Ra_{N+2} < R^2 a_{N+1} < R^3 a_N,$$

and more generally that

$$a_{N+j} < R^j a_N \qquad \text{whenever } j \geq 0. \tag{4.10}$$

Hence

$$\begin{aligned}
\sum_{k=N}^{\infty} a_k &= \sum_{j=0}^{\infty} a_{N+j} \\
&< \sum_{j=0}^{\infty} R^j a_N && \text{(by inequality (4.10))} \\
&= a_N \sum_{j=0}^{\infty} R^j && \text{(which is a geometric series)} \\
&= \frac{a_N}{1-R} && \text{(since } R < 1\text{).}
\end{aligned}$$

Since the tail of the series $\sum a_k$ converges, the series itself must converge.

(ii) If $r > 1$, then

$$\lim_{k\to\infty} \frac{a_{k+1}}{a_k} > 1$$

and so

$$\frac{a_k + 1}{a_k} > 1$$

for all k sufficiently large. That is, $a_{k+1} > a_k$ for all k sufficiently large, which means that the positive sequence $\{a_k\}$ eventually becomes an increasing sequence. Hence $a_k \not\to 0$ as $k \to \infty$. It follows from the kth term test that the series is divergent. $\square$

4.5.7 Leibniz' test for alternating series

So far we have given a convergence test for series all of whose terms are positive. If the terms are all negative, then we simply multiply the series by -1 to obtain series whose terms are positive. However, if the series has a mixture of positive and negative terms, we cannot apply this trick. In the next two subsections, we deal with series whose terms have mixed signs. The simplest case is when the sign alternates from term to term.

Definition 4.5.19. If $\{a_k\}_{k=0}^{\infty}$ is a sequence of positive real numbers, then the series

$$a_0 - a_1 + a_2 - a_3 + a_4 - a_5 + a_6 - a_7 + a_8 - a_9 + \cdots$$

is called an *alternating series.*

An alternating series is often written in the form

$$\sum_{k=0}^{\infty}(-1)^k a_k.$$

The following theorem, proved by Leibniz in the early eighteenth century, is a simple test for the convergence of alternating series.

Theorem 4.5.20 (Alternating series test)**.** *Suppose that $\{a_k\}_{k=0}^{\infty}$ is a sequence of real numbers satisfying the following properties:*

(a) $a_k \geq 0$*;*

(b) $a_k \geq a_{k+1}$ *for all k (that is, the sequence is nonincreasing); and*

(c) $\lim_{k\to\infty} a_k = 0$*.*

Then the alternating series $\sum_{k=0}^{\infty}(-1)^k a_k$ *converges.*

Before proving the theorem, we give and example and state a corollary.

Example 4.5.21. Determine whether the series $\displaystyle\sum_{k=2}^{\infty}\frac{(-1)^k k}{k^2+1}$ is summable.

Solution. Since this is an alternating series, one naturally tries the alternating series test. Suppose that $a_k = \dfrac{k}{k^2+1}$. We need to check that $\{a_k\}$ satisfies hypotheses (a), (b) and (c) of the alternating series test.

It is clear that (a) and (c) hold. To prove (b), consider the function f given by

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

Now

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

and hence $f'(x) < 0$ whenever $x > 1$. That is, f is decreasing on the interval $(1,\infty)$. Since $f(k) = a_k$ whenever $k \geq 2$, it follows that $\{a_k\}_{k=2}^{\infty}$ is a decreasing sequence.

We now apply the alternating series test and deduce that the series converges. □

If the hypotheses of the alternating sequence test are satisfied, then not only do we know that the series converges to some limit L, but we can also approximate L with any partial sum s_n and obtain an upper bound for the corresponding error.

Corollary 4.5.22. *Suppose that $\{a_k\}_{k=0}^{\infty}$ is a sequence of numbers satisfying properties (a), (b) and (c) of the alternating series test. Denote the value of the convergent series $\sum_{k=0}^{\infty}(-1)^k a_k$ by L and the nth partial sum of the same series by s_n. Then*

$$|s_n - L| \le a_{n+1} \tag{4.11}$$

for every natural number n.

In effect, the corollary says that if you chop the series off after the nth term, the error in approximation will be less than the $(n+1)$st term. The proof is given at the end of this subsection.

Example 4.5.23. Estimate the value of the series $\displaystyle\sum_{k=0}^{\infty} \frac{(-1)^k}{k^2+1}$ such that the error is less than $\frac{1}{100}$.

Solution. It is easy to verify that the sequence $\{a_k\}$, where $a_k = \frac{1}{k^2+1}$, satisfies the hypotheses of the alternating series test. Hence the infinite series converges. Denote the value of the series by L. We will use the estimate $s_n \approx L$, where n is chosen such that the absolute error is less than $\frac{1}{100}$. Now

$$\begin{aligned} \text{absolute error} &= |s_n - L| \\ &\le a_{n+1} && \text{(by Corollary 4.5.22)} \\ &= \frac{1}{(n+1)^2+1}. \end{aligned}$$

So it is enough to guarantee that

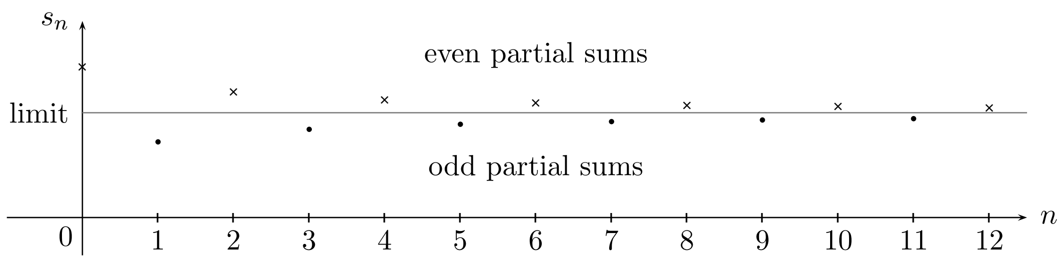
$$\frac{1}{(n+1)^2+1} < \frac{1}{100}.$$

Clearly the smallest positive integer n that satisfies this inequality is 9. So

$$L \approx s_9 = \sum_{k=0}^{9} \frac{(-1)^k}{k^2+1} = 1 - \frac{1}{5} + \frac{1}{10} - \frac{1}{17} + \cdots - \frac{1}{82}$$

and the error in this approximation is less than $\frac{1}{100}$. (Using Maple, one finds that $s_9 = 0.6305785114$ (correct to 10 decimal places) and so $L \approx 0.63$.) $\square$

We conclude this subsection by proving the alternating series test and its corollary. The following diagram illustrates the typical behaviour of the partial sums of a series that satisfies hypotheses (a), (b) and (c) of the alternating series test. It will be helpful to bear this diagram in mind when reading the proofs.



Proof of Theorem 4.5.20. [**X**] Let s_n denote the nth partial sum of the series and suppose that properties (a), (b) and (c) hold. The proof proceeds in three steps.

Step 1. We will prove that the sequence $\{s_{2n}\}_{n=0}^{\infty}$ of even partial sums is bounded above by 0. Now

$$s_{2n} = (a_0 - a_1) + (a_2 - a_3) + (a_4 - a_5) + \cdots + (a_{2n-2} - a_{2n-1}) + a_{2n}. \tag{4.12}$$

By property (b) we see that

$$a_0 - a_1 \geq 0, \qquad a_2 - a_3 \geq 0, \qquad , \ldots, \qquad a_{2n-2} - a_{2n-1} \geq 0$$

and by property (a) it is evident that $a_{2n} \geq 0$. It follows from (4.12) that $s_{2n} \geq 0$ for every n.

Step 2. We will prove that the sequence $\{s_{2n}\}_{n=0}^{\infty}$ of even partial sums is nonincreasing. Now

$$\begin{aligned} s_{2n} - s_{2n+2} &= (a_0 - a_1 + \ldots + a_{2n}) - (a_0 - a_1 + \ldots + a_{2n} - a_{2n+1} + a_{2n+2}) \\ &= a_{2n+1} - a_{2n+2} \\ &\geq 0 \end{aligned}$$

since $a_{2n+1} \geq a_{2n+2}$ by property (b). Thus $s_{2n} \geq s_{2n+2}$ for all n, which means that $\{s_{2n}\}_{n=0}^{\infty}$ is nonincreasing.

Step 3. From Steps 1 and 2, we conclude that $\{s_{2n}\}_{n=0}^{\infty}$ is a bounded monotonic sequence and hence convergent. Call the limit of this sequence L. If we can show that the sequence $\{s_{2n+1}\}_{n=0}^{\infty}$ of odd partial sums also converges to L, then we can conclude that $\{s_n\}$ converges. Now

$$s_{2n+1} = s_{2n} + a_{2n+1} \to L + 0$$

as $n \to \infty$ by property (c). Hence $\{s_n\}$ converges. □

Proof of Corollary 4.5.22. [**X**] Suppose that $\{a_k\}$ satisfies the hypotheses (a), (b) and (c) of the alternating series test. Since the infinite series converges, $\lim_{n\to\infty} s_n = L$ for some real number L. If n is odd then

$$s_{n+2} = s_n + a_{n+1} - a_{n+2} \geq s_n,$$

and so the odd partial sums increase towards L from below. If n is even then

$$s_{n+2} = s_n - a_{n+1} + a_{n+2} \leq s_n,$$

and so the even partial sums decrease towards L from above. Hence if n is odd then

$$s_n \leq L \leq s_{n+1} = s_n + a_{n+1},$$

while if n is even then

$$s_n - a_{n+1} = s_{n+1} \leq L \leq s_n.$$

That is,

$$\text{either} \qquad s_n \leq L \leq s_n + a_{n+1} \qquad \text{or} \qquad s_n - a_{n+1} \leq L \leq s_n.$$

Both cases imply (4.11). □

4.5.8 Absolute and conditional convergence

Consider the series

$$1+\frac{1}{2!}-\frac{1}{3!}+\frac{1}{4!}+\frac{1}{5!}-\frac{1}{6!}+\frac{1}{7!}+\frac{1}{8!}-\frac{1}{9!}+\cdots. \tag{4.13}$$

Since

$$\lim_{k\to\infty}|a_k| = \lim_{k\to\infty}\frac{1}{k!} = 0,$$

the kth term test for divergence does not apply. One cannot apply the ratio test (since not all the terms positive) and clearly this is not an alternating series. Is there a way of determining whether the series is summable? The theorem given below proves very helpful in this instance. First we give a definition.

Definition 4.5.24. A series $\sum_{k=0}^{\infty} a_k$ is said to be *absolutely convergent* if the series

$$\sum_{k=0}^{\infty}|a_k|$$

is convergent.

Theorem 4.5.25. *If a series is absolutely convergent then it converges.*

Proof. [**H**] Suppose that the series $\sum_{k=0}^{\infty} a_k$ converges absolutely. For each natural number k,

$$-|a_k| \le a_k \le |a_k|$$

and hence

$$0 \le a_k + |a_k| \le 2|a_k|. \tag{4.14}$$

Since $\sum|a_k|$ converges, it follows that $2\sum|a_k|$ converges and hence $\sum 2|a_k|$ converges. By the comparison test, we deduce from (4.14) that $\sum(a_k+|a_k|)$ converges. Now

$$a_k = (a_k+|a_k|) - |a_k|.$$

Since the sum of the terms on the right-hand side converges (by Proposition 4.4.4), the sum of the terms on the left must also converge. That is, $\sum_{k=0}^{\infty} a_k$ converges, thus completing the proof. ☐

Example 4.5.26. Determine whether or not the series given by (4.13) is convergent.

Solution. Let a_k denote the kth term of the series. By Theorem 4.5.25, it is enough to show that the series converges absolutely. Now

$$\sum_{k=1}^{\infty}|a_k| = \sum_{k=1}^{\infty}\frac{1}{k!}$$

and this series converges by the ratio test (see Example 4.5.18 (a)). ☐

Not every convergent series is absolutely convergent. For example, the alternating series

$$1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - \frac{1}{6} + \cdots \tag{4.15}$$

converges (by the alternating series test) but the corresponding absolute series

$$1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \frac{1}{5} + \frac{1}{6} + \cdots$$

diverges (see Example 4.4.3). In this situation, we say that the series (4.15) is *conditionally convergent.*

Definition 4.5.27. A series $\sum_{k=0}^{\infty} a_k$ is said to be *conditionally convergent* if it converges but the series

$$\sum_{k=0}^{\infty} |a_k|$$

diverges.

The distinction between conditionally and absolutely convergent series is brought into bold relief when considering rearrangements of series.

Definition 4.5.28. A *rearrangement* of a series $\sum a_k$ is a series that has exactly the same terms but that is summed in a different order.

For example,

$$1 + \frac{1}{3} + \frac{1}{5} - \frac{1}{2} + \frac{1}{7} + \frac{1}{9} - \frac{1}{4} + \frac{1}{11} + \cdots$$

and

$$-\frac{1}{2} + 1 - \frac{1}{4} + \frac{1}{3} - \frac{1}{6} + \frac{1}{5} - \frac{1}{8} + \frac{1}{7} - \cdots$$

are both rearrangements of

$$1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - \frac{1}{6} + \frac{1}{7} - \frac{1}{8} + \cdots .$$

If the series is finite, then every rearrangement has the same value (since addition of real numbers is commutative). However, if the series is infinite, then the value (if it exists) of each rearrangement is determined by a limit of partial sums, and one cannot appeal to commutativity of addition, as in the finite case. In fact, some rather surprising phenomena occur with rearrangements of conditionally convergent series.

Theorem 4.5.29. *Suppose that* $\sum_{k=0}^{\infty} a_k$ *is an infinite series.*

(i) *If $\sum a_k$ converges absolutely, then every rearrangement of the series converges absolutely and all rearrangements have the same limit as $\sum a_k$.*

(ii) *If $\sum a_k$ converges conditionally, then given any real number L, the series has a rearrangement that converges to L. Moreover, every conditionally convergent series has a rearrangement that diverges to ∞, and another rearrangement that diverges to $-\infty$.*

This theorem was published in 1867 by Riemann. One of the tutorial problems illustrates that the conditionally convergent series

$$1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - \frac{1}{6} + \frac{1}{7} - \cdots$$

can be rearranged to sum to different real numbers. The moral of the story is that one should not rearrange a conditionally convergent series to determine its value.

4.6 Taylor series

(Ref: SH10 §12.7)

At the end of Section 4.2 we posed the question, When is it true that

$$f(x) = f(a) + f'(a)(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \frac{f^{(3)}(a)}{3!}(x-a)^3 + \cdots ?$$

Having studied sequences and series of real numbers, we now have the tools to deal with this and related questions. First, we give the above series expansion a special name.

Definition 4.6.1. Suppose that a function f has derivatives of all orders at a. Then the series

$$f(a) + f'(a)(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \frac{f^{(3)}(a)}{3!}(x-a)^3 + \cdots ,$$

which may also be written as

$$\sum_{k=0}^{\infty} \frac{f^{(k)}(a)}{k!}(x-a)^k,$$

is called the *Taylor series for f about a*. In the case when $a = 0$, the series is also called the *Maclaurin series for f*.

Next, we need to define what we mean by the convergence (or divergence) of a Taylor series.

Definition 4.6.2. Suppose that I is an interval and that f has derivatives of all orders at some point a. We say that

(a) the Taylor series for f about a *converges on* I if the series
$$\sum_{k=0}^{\infty} \frac{f^{(k)}(a)}{k!}(x-a)^k$$
converges for each point x in I;

(b) the Taylor series for f about a *converges to* f *on* I if for each x in I, x lies in the domain of f and
$$f(x) = \sum_{k=0}^{\infty} \frac{f^{(k)}(a)}{k!}(x-a)^k;$$
and

(c) the Taylor series for f at a *diverges on* I if the series
$$\sum_{k=0}^{\infty} \frac{f^{(k)}(a)}{k!}(x-a)^k$$
diverges for each point x in I.

Thus the question asked at the beginning of this section may be rephrased as,

For what intervals I will the Taylor series of a function f converge to f?

The following corollary to Taylor's theorem helps answer this question.

Corollary 4.6.3. *Suppose that f has derivatives of all orders at a and that x lies in the domain of f. Let $R_{n+1}(x)$ denote the remainder term of Theorem 4.2.1 (or its equivalent form as given in Corollary 4.2.2). If*
$$\lim_{n\to\infty} R_{n+1}(x) = 0 \tag{4.16}$$
then
$$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!}(x-a)^n.$$

Proof. Let p_n denote the nth Taylor polynomial for f about a. Taylor's theorem implies that
$$f(x) = p_n(x) + R_{n+1}(x). \tag{4.17}$$

Note that $p_n(x)$ is the nth partial sum of the series
$$\sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!}(x-a)^n.$$

So we only have to show that
$$\lim_{n\to\infty} p_n(x) = f(x). \tag{4.18}$$

Now

$$\begin{aligned}|p_n(x) - f(x)| &= |R_{n+1}(x)| && \text{(by (4.17))}\\ &\to 0\end{aligned}$$

as $n \to \infty$ by (4.16). That is, the distance between $p_n(x)$ and $f(x)$ can be made as small as we like. Hence (4.18) follows. □

Example 4.6.4. Suppose that $x \in \mathbb{R}$. Show that

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \cdots .$$

Solution. Suppose that $f(t) = e^t$ and fix x in $\mathbb{R}$. The sum of the terms up to (and including) $\frac{x^n}{n!}$ is equal to $p_n(x)$, where p_n is the nth Taylor polynomial for f about 0. So by Corollary 4.6.3, we only need to show that $\lim_{n\to\infty} R_{n+1}(x) = 0$. Now

$$f^{(n)}(t) = e^t \qquad \forall t \in \mathbb{R},$$

and so by the Lagrange formula for the remainder,

$$R_{n+1}(x) = \frac{e^c}{(n+1)!}x^{n+1}$$

for some c between 0 and x. Now $e^c \le e^{|c|} \le e^{|x|}$. If $M = e^{|x|}$ then

$$\begin{aligned}0 &\le |R_{n+1}(x)|\\ &= \frac{e^c}{(n+1)!}|x|^{n+1}\\ &\le \frac{M|x|^{n+1}}{(n+1)!}\\ &\to 0\end{aligned}$$

as $n \to \infty$ by Lemma 2.2.5. Therefore

$$\lim_{n\to\infty} R_{n+1}(x) = 0$$

by the pinching theorem for sequences. □

We have just proved that the Taylor series about 0 for the exponential function converges to the exponential function on $\mathbb{R}$. We say that the exponential function is *represented* by its Taylor series about 0 on $\mathbb{R}$. Some other convergent Taylor series representations are given in the next theorem.

Theorem 4.6.5. *The following formulae hold whenever x lies in the given interval.*

$$\begin{aligned}
\frac{1}{1-x} &= 1 + x + x^2 + x^3 + x^4 + \ldots && x \in (-1,1) \\
e^x &= 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \cdots && x \in \mathbb{R} \\
\sin x &= x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \cdots && x \in \mathbb{R} \\
\cos x &= 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \cdots && x \in \mathbb{R} \\
\sinh x &= x + \frac{x^3}{3!} + \frac{x^5}{5!} + \frac{x^7}{7!} + \cdots && x \in \mathbb{R} \\
\cosh x &= 1 + \frac{x^2}{2!} + \frac{x^4}{4!} + \frac{x^6}{6!} + \cdots && x \in \mathbb{R} \\
\ln(1+x) &= x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \cdots && x \in (-1,1] \\
\tan^{-1} x &= x - \frac{x^3}{3} + \frac{x^5}{5} - \frac{x^7}{7} + \cdots && x \in [-1,1]
\end{aligned}$$

Moreover, if x lies outside the given interval then the corresponding Maclaurin series diverges.

Most of these formulae can be proved by showing that the Lagrange formula for the remainder tends to 0 as $n \to \infty$. However, sometimes one must resort to using the integral form of the remainder (as given by Theorem 4.2.1). One of the tutorial problems illustrates its use. In Section 4.8, we introduce tools that provide an alternate approach to deriving some of these expansions.

Remark 4.6.6. The Taylor series expansions for $\sin x$ and $\ln(1+x)$ given by Theorem 4.6.5 explain the phenomena discussed at the beginning of Section 4.2. In particular, the Taylor series for $\sin x$ converges for all x in $\mathbb{R}$, which explains why $\sin(7)$ could be approximated by Taylor polynomials of sufficiently high degree. On the other hand, we cannot use Taylor polynomials to approximate $\ln(1+x)$ when $x > 1$ (as suggested by Figure 4.2) because the Taylor series *diverges* when $x > 1$. This explains why higher order Taylor polynomials give worse approximations for $\ln(1+x)$ when $x > 1$.

Remark 4.6.7. The Maclaurin series given by Theorem 4.6.5 can be used to obtain beautiful series expansions for some irrational numbers. By substituting particular values for x into an appropriate Maclaurin series, one finds that

$$\begin{aligned}
e &= 1 + \frac{1}{1!} + \frac{1}{2!} + \frac{1}{3!} + \frac{1}{4!} + \cdots \\
\ln 2 &= 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - \cdots \\
\frac{\pi}{4} &= 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \cdots .
\end{aligned}$$

Unfortunately, the last two series converge too slowly to be of high computational value.

Remark 4.6.8. If f equals its Taylor series on an interval I, then the corresponding Taylor polynomial p_n can be used to approximate f on I. However, it is important to appreciate that some Taylor series (such as that for e^x) converge much more quickly to the function than do others (such as that for $\ln(1+x)$, which converges slowly). If the series converges very slowly then the approximation $f \approx p_n$ is only accurate when n is very large.

Remark 4.6.9. If a Taylor series converges to a function f on an interval I, then (obviously) the Taylor series converges on I. However, the converse is not true. That is, if the Taylor series converges on I, then one cannot conclude that the Taylor series converges *to* f on I. In the tutorial problems we give one example of a function f whose whose Maclaurin series converges on the entire real line but only converges to f at the origin.

4.7 Power series

(Ref: SH10 §12.8)

A Maclaurin series is a series of the form

$$a_0 + a_1x + a_2x^2 + a_3x^3 + a_4x^4 + a_5x^5 + \cdots, \tag{4.19}$$

where each coefficient a_k is given by

$$a_k = \frac{f^{(k)}(0)}{k!}$$

for some function f that is infinitely differentiable at 0. For the remainder of this chapter, we study more general series of the form (4.19), where each coefficient a_k is not necessarily a Taylor coefficient. Such a series is called a *power series*.

Definition 4.7.1. Suppose that $\{a_k\}_{k=0}^{\infty}$ is a sequence of real numbers and that $a \in \mathbb{R}$. A series of the form

$$\sum_{k=0}^{\infty} a_k x^k$$

is called a *power series in powers of* x. A series of the form

$$\sum_{k=0}^{\infty} a_k (x-a)^k$$

is called a *power series in powers of* $x - a$.

Thus a Maclaurin series is a power series in powers of x, while a Taylor series about a is a power series in powers of $x - a$. For the last two sections of this chapter, we discuss the convergence, addition, multiplication, integration and differentiation of power series. Hence whatever is said about power series also applies to Maclaurin and Taylor series.

In this section, we focus on the convergence and divergence of power series.

Definition 4.7.2. Suppose that $\{a_k\}_{k=0}^{\infty}$ is a sequence of real numbers, I is an interval and a is a real number. We say that a power series $\displaystyle\sum_{k=0}^{\infty} a_k(x-a)^k$ converges

(a) at a real number c if the series $\displaystyle\sum_{k=0}^{\infty} a_k(c-a)^k$ converges;

(b) on the interval I if the series $\displaystyle\sum_{k=0}^{\infty} a_k(x-a)^k$ converges for each x in I.

We say that a power series $\displaystyle\sum_{k=0}^{\infty} a_k(x-a)^k$ diverges

(a) at a real number c if the series $\displaystyle\sum_{k=0}^{\infty} a_k(c-a)^k$ diverges;

(b) on the interval I if the series $\displaystyle\sum_{k=0}^{\infty} a_k(x-a)^k$ diverges for each x in I.

Using the ratio test, one can often determine for what values of x a power series converges absolutely.

Example 4.7.3. Find an interval I such that the power series $\displaystyle\sum_{k=0}^{\infty} \frac{kx^k}{3^k}$ converges on I.

Solution. We first find an interval I on which the series converges absolutely. To do so, we apply the ratio test to the (absolute) series

$$\sum_{k=0}^{\infty} \left|\frac{kx^k}{3^k}\right|.$$

Now

$$\begin{aligned} r &= \lim_{k\to\infty} \left|\frac{(k+1)x^{k+1}}{3^{k+1}}\right| \cdot \left|\frac{3^k}{kx^k}\right| \\ &= \lim_{k\to\infty} \frac{(k+1)|x|}{3k} \\ &= \frac{|x|}{3}. \end{aligned}$$

To conclude from the ratio test that the series converges, we require that $r < 1$, which means that $|x| < 3$. So the series converges absolutely whenever $-3 < x < 3$. Hence the series $\displaystyle\sum_{k=0}^{\infty} \frac{kx^k}{3^k}$ converges on the interval $(-3, 3)$. □

Remark 4.7.4. In the previous example, one can also conclude that the series diverges when $|x| > 3$. To see this, fix x in $\mathbb{R}$ such that $|x| > 3$. Let b_k denote the kth term

$$\frac{kx^k}{3^k}$$

of the series. Now, by the same calculation as before,

$$\lim_{k\to\infty}\left|\frac{b_{k+1}}{b_k}\right| = \frac{|x|}{3}.$$

Since $|x| > 3$ we deduce that $\lim\limits_{k\to\infty} \frac{|b_{k+1}|}{|b_k|} > 1$. This shows that $\frac{|b_{k+1}|}{|b_k|} > 1$ whenever k is sufficiently large. Rearranging implies that

$$|b_{k+1}| > |b_k|$$

for all sufficiently large k and hence the tail of the sequence $\{|b_k|\}$ is increasing. Thus $|b_k| \not\to 0$ as $k \to \infty$. We conclude that $b_k \not\to 0$ as $k \to \infty$ and hence $\sum b_k$ diverges by the kth term test for divergence.

4.7.1 Radius of Convergence

In the previous example, the power series, in powers of x, converged for $|x| < 3$. We call the number 3 the **radius of convergence** of the power series. It is half the length of the interval of convergence.

Definition 4.7.5. If a power series of the form $\sum\limits_{k=0}^{\infty} a_k(x-a)^k$ converges at all points in some interval $(-R+a, R+a)$, or equivalently, for $|z-a| < R$, then the number R is called the *radius of convergence* for the power series. The corresponding interval $(-R+a, R+a)$ is called the *open interval of convergence* for the power series.
If the power series converges for all real x, we say that the radius of convergence is infinite.

Notes: 1. The term 'radius' is used since, when x is replaced by the complex variable z, the open interval is replaced by an open disc, $|z-a| < R$, in the Argand plane. The number R then is the radius of this open disc.

2. It is easiest to find the interval of convergence first, using the ratio test, and then write down the radius of convergence. It can be shown that $R = \lim\limits_{n\to\infty}\left|\frac{a_k}{a_{k+1}}\right|$, provided this limit exists. There are, however, power series which have a radius of convergence, but for which this limit does not exist.

By generalising the solution to Example 4.7.3 and the argument in Remark 4.7.4, one obtains the following theorem.

Theorem 4.7.6. *Suppose that $\{a_k\}_{k=0}^{\infty}$ is a sequence of real numbers such that*

$$\lim_{k\to\infty}\left|\frac{a_k}{a_{k+1}}\right| = R$$

for some real number R. Then the power series $\sum\limits_{k=0}^{\infty} a_k(x-a)^k$

(i) *converges absolutely whenever* $|x-a|<R$, *and*

(ii) *diverges whenever* $|x-a|>R$.

Proof of Theorem 4.7.6. [**H**] The proof of (i) is similar to the solution of Example 4.7.3. We apply the ratio test to the series

$$\sum_{k=0}^{\infty} |a_k(x-a)^k|.$$

Now

$$r = \lim_{k\to\infty} \frac{|a_{k+1}(x-a)^{k+1}|}{|a_k(x-a)^k|} = \lim_{k\to\infty} \left|\frac{a_{k+1}}{a_k}\right| |x-a|.$$

We have convergence whenever $r<1$, which corresponds to the condition that

$$\lim_{k\to\infty} \left|\frac{a_{k+1}}{a_k}\right| |x-a| < 1.$$

By rearranging we find that the series converges absolutely whenever

$$|x-a| < \frac{1}{\lim\limits_{k\to\infty} \left|\frac{a_{k+1}}{a_k}\right|} = \lim_{k\to\infty} \left|\frac{a_k}{a_{k+1}}\right| = R,$$

where R is the limit given in the theorem.

The proof of (ii) is a simple modification of the argument given in Remark 4.7.4. $\square$

Note that we cannot tell from the theorem whether or not the power series converges at the endpoints $a+R$ or $a-R$. Sometimes the power series will converge at one endpoint but not at the other. Other times it will converge at both endpoints or diverge at both endpoints.

Example 4.7.7. Find the largest open interval on which the power series

$$\sum_{k=0}^{\infty} \frac{(5x+2)^k}{k^2+1}$$

will converge.

Solution. We apply the ratio test to the series $\sum_{k=0}^{\infty} \left|\frac{(5x+2)^k}{k^2+1}\right|$. Now

$$\begin{aligned} r &= \lim_{k\to\infty} \left|\frac{(5x+2)^{k+1}}{(k+1)^2+1}\right| \left|\frac{k^2+1}{(5x+2)^k}\right| \\ &= \lim_{k\to\infty} \frac{k^2+1}{(k+1)^2+1}|5x+2| \\ &= |5x+2|. \end{aligned}$$

We require that $r<1$ and so that $|5x+2|<1$. Hence

$$-1 < 5x+2 < 1$$

or in other words,

$$-\frac{3}{5} < x < -\frac{1}{5}.$$

So the largest open interval of convergence is $(-\frac{3}{5}, -\frac{1}{5})$.

Hence the radius of convergence is $\frac{1}{5}$. $\square$

4.7.2 Convergence of power series at endpoints [X]

Students studying MATH1241 are also required to determine whether a power series converges at the endpoints of its interval of convergence. As the next example illustrates, one deduces the convergence at each endpoint by substituting the endpoint into the power series and determining whether the resulting series of real numbers converges.

Example 4.7.8. Find the interval of convergence (including endpoints, if appropriate) for the power series $\displaystyle\sum_{k=2}^{\infty} \frac{x^k}{\ln k}$.

Solution. First we find the open interval of convergence. Now

$$\begin{aligned} r &= \lim_{k\to\infty} \left| \frac{x^{k+1}}{\ln(k+1)} \cdot \frac{\ln k}{x^k} \right| \\ &= \lim_{k\to\infty} \frac{\ln k}{\ln(k+1)} |x| \\ &= \lim_{k\to\infty} \frac{1/k}{1/(k+1)} |x| && \text{(by l'Hôpital's rule)} \\ &= \lim_{x\to\infty} \frac{k+1}{k} |x| \\ &= |x|. \end{aligned}$$

The series converges absolutely whenever $r = |x| < 1$. Hence the largest open interval of convergence is $(-1, 1)$ and the series diverges whenever $|x| > 1$.

Now we determine whether the series converges at the endpoints 1 and -1. When $x = 1$, the series becomes

$$\sum_{k=2}^{\infty} \frac{1}{\ln k},$$

which diverges by comparison with the harmonic series $\sum \frac{1}{k}$. When $x = -1$, the series becomes

$$\sum_{k=2}^{\infty} \frac{(-1)^k}{\ln k},$$

which is alternating and converges (conditionally) by the alternating series test.

Hence the interval of convergence for the power series is $[-1, 1)$. (Note that the largest interval on which the power series is *absolutely* convergent is $(-1, 1)$.) □

4.8 Manipulation of power series

(Ref: SH10 §12.9)

In this section we investigate what sense (if any) can be made of adding, multiplying, differentiating and integrating power series. Since differentiation and integration are operations applied to functions, it is most natural to approach this investigation by viewing a power series as a function.

Suppose that a power series $\displaystyle\sum_{n=0}^{\infty} a_k x^k$ converges in the interval $(-R, R)$, where R is its radius of convergence. Then one can define a function $f : (-R, R) \to \mathbb{R}$ by the formula

$$f(x) = \sum_{k=0}^{\infty} a_k x^k \qquad \text{whenever } |x| < R.$$

Thus the value of f at each point x is a convergent sum of real numbers. Sometimes it is possible to find a *closed* form for f, but other times we must approximate each value $f(x)$ by using partial sums.

Example 4.8.1. Suppose that f is given by the rule

$$f(x) = \sum_{k=0}^{\infty} x^k.$$

By using, say, the ratio test, we see that the series converges whenever $|x| < 1$ and diverges when $|x| > 1$. Hence natural domain for f is $(-1, 1)$.

In fact, by summing the geometric series, we find that $f(x) = \dfrac{1}{1-x}$ whenever $|x| < 1$. This is the *closed form* of $f(x)$.

Example 4.8.2. Suppose that f is defined by the rule

$$f(x) = \sum_{k=0}^{\infty} \frac{x^k}{k!}.$$

We instantly recognise this series as the Maclaurin series for the exponential function. Since this series converges on $\mathbb{R}$, the maximal domain of f is $\mathbb{R}$. The closed form of $f(x)$ is given by $f(x) = e^x$ whenever $x \in \mathbb{R}$.

Example 4.8.3. Suppose that f is defined by the rule

$$f(x) = \sum_{k=1}^{\infty} \frac{x^k}{k^2}.$$

By the ratio test, we find that the series converges whenever $|x| < 1$ and diverges whenever $|x| > 1$. Therefore we take the domain of f to be $(-1, 1)$. (Students in MATH1141 will note that the power series also converges when $x = 1$ and $x = -1$. So the domain could be extended to $[-1, 1]$.)

There seems to be no obvious closed form for $f(x)$ whenever $|x| < 1$. How, then, does one evaluate $f(-\frac{1}{2})$? Note that

$$\begin{aligned} f(-\tfrac{1}{2}) &= \sum_{k=1}^{\infty} \frac{(-1)^k}{2^k\, k^2} \\ &= -\frac{1}{2^1\, 1^2} + \frac{1}{2^2\, 2^2} - \frac{1}{2^3\, 3^3} + \frac{1}{2^4\, 4^2} - \frac{1}{2^5\, 5^2} + \cdots \end{aligned}$$

and that the right-hand side is an alternating series. Thus

$$f(-\tfrac{1}{2}) \approx -\frac{1}{2^1\, 1^2} + \frac{1}{2^2\, 2^2} - \frac{1}{2^3\, 3^3} + \frac{1}{2^4\, 4^2} - \frac{1}{2^5\, 5^2} + \cdots + \frac{1}{2^{10}\, 10^2},$$

where the absolute error in this approximation is less than $\dfrac{1}{2^{11}\, 11^2}$ by Corollary 4.5.22.

It turns out that power series are very well behaved as functions defined on their interval of convergence. Given two power series with the same interval of convergence, you can add, subtract and multiply them together in the 'natural' way. Power series are also differentiable and integrable, and their derivatives and antiderivatives can also be expressed as power series in the 'natural' way. The following theorems articulate the precise details. Their proofs are given later in Subsection 4.8.1.

Theorem 4.8.4. *Suppose that the functions $f : I \to \mathbb{R}$ and $g : I \to \mathbb{R}$ are defined by*

$$f(x) = \sum_{k=0}^{\infty} a_k (x-a)^k \qquad \text{and} \qquad g(x) = \sum_{k=0}^{\infty} b_k (x-a)^k,$$

where both power series converge on the interval I. Then, whenever $x \in I$,

$$(f+g)(x) = \sum_{k=0}^{\infty} (a_k + b_k)(x-a)^k$$

and

$$(fg)(x) = \sum_{k=0}^{\infty} c_k (x-a)^k, \tag{4.20}$$

where

$$c_k = \sum_{j=0}^{k} a_j b_{k-j}.$$

Remark 4.8.5. The product formula (4.20) says that

$$\begin{aligned}(fg)(x) = a_0b_0 + (a_0b_1 + a_1b_0)(x-a) + (a_0b_2 + a_1b_1 + a_2b_0)(x-a)^2 \\ + (a_0b_3 + a_1b_2 + a_2b_1 + a_3b_0)(x-a)^3 + \cdots .\end{aligned}$$

This is the natural generalisation of polynomial multiplication.

Theorem 4.8.6. *Suppose that $f : I \to \mathbb{R}$ is defined by*

$$f(x) = \sum_{k=0}^{\infty} a_k (x-a)^k$$

whenever $x \in I$, where I denotes the open interval of convergence for the power series. Then

(i) f is differentiable on I and

$$f'(x) = \sum_{k=1}^{\infty} k a_k (x-a)^{k-1}$$

whenever $x \in I$; and

(ii) f is integrable on I and an antiderivative F for f is given by

$$F(x) = \sum_{k=0}^{\infty} \frac{a_k}{k+1}(x-a)^{k+1} + C$$

whenever $x \in I$, where C is a constant.

Remark 4.8.7. This theorem says that a power series can be differentiated and integrated 'term by term' inside its open interval of convergence. Another way of saying this is that

$$\frac{d}{dx}\left(\sum_{k=1}^{\infty} a_k(x-a)^k\right) = \sum_{k=1}^{\infty}\left(\frac{d}{dx}a_k(x-a)^k\right)$$

and

$$\int \left(\sum_{k=1}^{\infty} a_k(x-a)^k\right)\,dx = \sum_{k=1}^{\infty}\left(\int a_k(x-a)^k\,dx\right).$$

Students should note that one cannot always swap infinite summation with differentiation (or with integration). For example, if

$$f(x) = \sum_{k=1}^{\infty} \frac{\sin(2^k x)}{2^k}$$

then f is a sum of differentiable functions but f is not differentiable anywhere!

The following corollary follows from Theorem 4.8.6.

Corollary 4.8.8. *Suppose that $f : I \to \mathbb{R}$ is defined by*

$$f(x) = \sum_{k=0}^{\infty} a_k(x-a)^k$$

whenever $x \in I$, where I denotes the open interval of convergence for the power series. Then f is continuous on I and has derivatives of all orders on I.

Proof. By Theorem 4.8.6 (i), f is differentiable on I and is therefore continuous on I.

We now prove that f has derivatives of all orders on I. Suppose that n is *any* natural number. It suffices to show that f is n-times differentiable on I. By Theorem 4.8.6 (i), f is once differentiable on I and its derivative f' has a power series expansion that converges on I. Now apply Theorem 4.8.6 (i) to f'. We conclude that f' is differentiable on I and its derivativef'' has a power series expansion that converges on I. Now apply Theorem 4.8.6 (i) to f''. Continuing in this way, after n steps we conclude that $f^{(n-1)}$ is differentiable on I with derivative $f^{(n)}$. Hence f is n-times differentiable on I, thus completing the proof. $\square$

Remark 4.8.9. Suppose that a function $f : I \to \mathbb{R}$ is defined by

$$f(x) = \sum_{k=0}^{\infty} a_k(x-a)^k,$$

where I is the open interval of convergence for the power series. By the corollary, f has derivatives of all orders at a and therefore has a Taylor series about a. One can easily show that the Taylor series for f about a converges on I to f. Thus the Taylor series for f about a is equal to the power series that defines f.

Theorems 4.8.4 and 4.8.6 allow us to find valid Taylor series expansions of functions without having to derive the Taylor coefficients and verify that the remainder term from Taylor's theorem vanishes.

Example 4.8.10. Given the Taylor expansions

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \frac{x^5}{5!} + \frac{x^6}{6!} + \cdots \qquad (x \in \mathbb{R}) \tag{4.21}$$

and

$$\frac{1}{1-x} = 1 + x + x^2 + x^3 + x^4 + x^5 + \cdots \qquad (|x| < 1), \tag{4.22}$$

find Taylor expansions for each function f, making sure that you state the interval of convergence.

(a) $f(x) = \cosh x$ (b) $f(x) = \sinh x$ (c) $f(x) = \tan^{-1}(x)$

Solution. (a) First note that $\cosh x = \frac{1}{2}(e^x + e^{-x})$. So we aim to add the Maclaurin series for e^x and e^{-x}. By replacing x with $-x$ in (4.21), we find that

$$e^{-x} = 1 - x + \frac{x^2}{2!} - \frac{x^3}{3!} + \frac{x^4}{4!} - \frac{x^5}{5!} + \frac{x^6}{6!} - \cdots \qquad (x \in \mathbb{R}).$$

So

$$e^x + e^{-x} = 2 + 2\frac{x^2}{2!} + 2\frac{x^4}{4!} + 2\frac{x^6}{6!} + \cdots \qquad (x \in \mathbb{R})$$

by Theorem 4.8.4. Hence

$$\begin{aligned} \cosh x &= \frac{1}{2}(e^x + e^{-x}) \\ &= 1 + \frac{x^2}{2!} + \frac{x^4}{4!} + \frac{x^6}{6!} + \cdots \end{aligned}$$

whenever $x \in \mathbb{R}$.

(b) By differentiating both sides of the expansion

$$\cosh x = 1 + \frac{x^2}{2!} + \frac{x^4}{4!} + \frac{x^6}{6!} + \frac{x^8}{8!} + \cdots \qquad (x \in \mathbb{R})$$

we find that

$$\sinh x = x + \frac{x^3}{3!} + \frac{x^5}{5!} + \frac{x^7}{7!} + \cdots$$

whenever $x \in \mathbb{R}$.

(c) If $|x| < 1$ then $|-x^2| < 1$. So we can replace x with $-x^2$ in (4.22) to obtain the convergent expansion

$$\frac{1}{1+x^2} = 1 - x^2 + x^4 - x^6 + x^8 - x^{10} + \cdots$$

whenever $|x| < 1$. By integrating both sides this of identity, we find that

$$\tan^{-1} x = x - \frac{x^3}{3} + \frac{x^5}{5} - \frac{x^7}{7} + \frac{x^9}{9} - \frac{x^{11}}{11} + \cdots$$

whenever $|x| < 1$. □

The function f, given by $f(x) = e^{-x^2}$, has no antiderivative among the elementary functions. In MATH1131, we used Riemann sums to estimate the area underneath the graph of f. As the next example shows, the use of Taylor series provides a more efficient approach to the same problem.

Example 4.8.11. Suppose that $f(x) = e^{-x^2}$. By using the Maclaurin expansion for f, estimate $\int_0^1 f(x)\,dx$ and give an upper bound for the absolute error.

Solution. We begin with the Maclaurin expansion

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \frac{x^5}{5!} + \cdots,$$

which is valid for all x in $\mathbb{R}$. By replacing x with $-x^2$ we find that

$$e^{-x^2} = 1 - x^2 + \frac{x^4}{2!} - \frac{x^6}{3!} + \frac{x^8}{4!} - \frac{x^{10}}{5!} + \cdots$$

for all real numbers x. By integrating both sides of this equation on the interval $[0, 1]$, we find that

$$\begin{aligned}\int_0^1 e^{-x^2}\,dx &= \left[x - \frac{x^3}{3} + \frac{x^5}{5(2!)} - \frac{x^7}{7(3!)} + \frac{x^9}{9(4!)} - \frac{x^{11}}{11(5!)} + \cdots\right]_0^1 \\ &= 1 - \frac{1}{3} + \frac{1}{5(2!)} - \frac{1}{7(3!)} + \frac{1}{9(4!)} - \frac{1}{11(5!)} + \cdots.\end{aligned}$$

Hence the integral is expressed as alternating series of the form $\sum(-1)^k a_k$, where $\{a_k\}$ is a positive decreasing sequence. If we estimate the series using a partial sum, then Corollary 4.5.22 gives an upper bound for the absolute error. For example, the absolute error in the approximation

$$\int_0^1 e^{-x^2}\,dx \approx 1 - \frac{1}{3} + \frac{1}{5(2!)} - \frac{1}{7(3!)}$$

is no greater than $\frac{1}{9(4!)}$. One can evaluate this partial sum numerically to obtain

$$\int_0^1 e^{-x^2}\,dx \approx 0.7429$$

with an error no greater than 0.005. □

4.8.1 Proof of theorems in Section 4.8 [X]

In this subsection, we prove Theorems 4.8.4 for the case where $a = 0$. It is not hard to adapt the presented proofs to the general case.

Proof of Theorem 4.8.4. We shall only prove the product formula (4.20) when $a = 0$. Fix x in I and define the partial sums $s_n(x)$ and $t_n(x)$ by

$$s_n(x) = \sum_{k=0}^{n} a_k x^k \qquad \text{and} \qquad t_n(x) = \sum_{k=0}^{n} b_k x^k.$$

Then $f(x) = \lim_{n\to\infty} s_n(x)$ and $g(x) = \lim_{n\to\infty} t_n(x)$. So using Proposition 4.3.6, we find that the sequence $\{s_n(x)t_n(x)\}$ converges and

$$(f.g)(x) = f(x).g(x) = \lim_{n\to\infty} s_n(x) \times \lim_{n\to\infty} t_n(x) = \lim_{n\to\infty} \big(s_n(x)t_n(x)\big).$$

But

$$\begin{aligned} s_n(x)t_n(x) &= \big(a_0 + a_1x + a_2x^2 + \cdots a_nx^n\big) \times \big(b_0 + a_1x + b_2x^2 + \cdots + b_nx^n\big) \\ &= a_0b_0 + (a_0b_1 + a_1b_0)x + (a_0b_2 + a_1b_1 + a_2b_0)x^2 \\ &\quad + \cdots + (a_0b_n + a_1b_{n-1} + \cdots + a_{n-1}b_1 + a_nb_0)x^n. \end{aligned}$$

As $n \to \infty$, one obtains (4.20). □

We move now to the proof of Theorem 4.8.6, which shall be broken into two parts. First we prove the differentiation result, which is difficult and uses the mean value theorem. After this, the integration result can be easily deduced from the first part.

Proof of Theorem 4.8.6 (i). Suppose that $f : (-R, R) \to \mathbb{R}$ is defined by

$$f(x) = \sum_{k=0}^{\infty} a_k x^k$$

whenever $x \in (-R, R)$, where R is the radius of convergence for the power series. It can be easily shown (see the tutorial problems) that the radius of convergence for the power series

$$\sum_{k=1}^{\infty} k a_k x^{k-1}$$

is also R. So define the function $g : (-R, R) \to \mathbb{R}$ by the formula

$$g(x) = \sum_{k=1}^{\infty} k a_k x^{k-1}.$$

Fix, now, a number x in $(-R, R)$. Our task is to show that $f'(x) = g(x)$, or in other words, that

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

Now if $x + h \in (-R, R)$ and $h \neq 0$ then

$$\begin{aligned} \left| g(x) - \frac{f(x+h) - f(x)}{h} \right| &= \left| \sum_{k=1}^{\infty} k a_k x^{k-1} - \sum_{k=0}^{\infty} \frac{a_k(x+h)^k - a_k x^k}{h} \right| \\ &= \left| \sum_{k=1}^{\infty} k a_k x^{k-1} - \sum_{k=1}^{\infty} a_k \left(\frac{(x+h)^k - x^k}{h} \right) \right|. \end{aligned}$$

By the mean value theorem,

$$\frac{(x+h)^k - x^k}{h} = k\, c_k^{k-1}$$

for some real number c_k between x and $x + h$. Hence

$$\begin{aligned} \left| g(x) - \frac{f(x+h) - f(x)}{h} \right| &= \left| \sum_{k=1}^{\infty} k a_k x^{k-1} - \sum_{k=1}^{\infty} k a_k c_k^{k-1} \right| \\ &= \left| \sum_{k=1}^{\infty} k a_k \big(x^{k-1} - c_k^{k-1}\big) \right| \\ &= \left| \sum_{k=2}^{\infty} k a_k \big(x^{k-1} - c_k^{k-1}\big) \right|. \end{aligned}$$

Again by the mean value theorem,

$$\frac{x^{k-1} - c_k^{k-1}}{x - c_k} = (k-1)d_{k-1}^{k-2}$$

for some real number d_{k-1} between x and c_k. Hence

$$\left|x^{k-1} - c_k^{k-1}\right| = |x - c_k| \left|(k-1)d_{k-1}^{k-2}\right|.$$

Now $|x - c_k| < |h|$ and $|d_{k-1}| < M$, where $M = \max\{|x|, |x+h|\}$. So

$$\left|x^{k-1} - c_k^{k-1}\right| \leq |h| \left|(k-1)M^{k-2}\right|.$$

Thus

$$\left|g(x) - \frac{f(x+h) - f(x)}{h}\right| \leq |h| \sum_{k=2}^{\infty} \left|k(k-1)a_k M^{k-2}\right|.$$

One can show using the ratio test that the series on the right-hand side converges and hence

$$\lim_{h \to 0} |h| \sum_{k=2}^{\infty} \left|k(k-1)a_k M^{k-2}\right| = 0.$$

Therefore

$$\lim_{h \to 0} \left|g(x) - \frac{f(x+h) - f(x)}{h}\right| = 0$$

(by the pinching theorem for limits) and we conclude that

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

Hence $f'(x) = g(x)$ for all x in $(-R, R)$. □

Now that we have proved that a power series is differentiable inside its open interval of convergence, it is relatively easy to prove that it is integrable inside this interval.

Proof of Theorem 4.8.6 (ii). Suppose that $f : (-R, R) \to \mathbb{R}$ is defined by

$$f(x) = \sum_{k=0}^{\infty} a_k x^k$$

whenever $x \in (-R, R)$, where R is the radius of convergence for the power series. It can be easily shown (see the tutorial problems) that the radius of convergence for the power series

$$\sum_{k=0}^{\infty} \frac{a_k}{k+1} x^{k+1}$$

is also R. So define the function $F : (-R, R) \to \mathbb{R}$ by the formula

$$F(x) = \sum_{k=0}^{\infty} \frac{a_k}{k+1} x^{k+1}.$$

We now apply the differentiability theorem (Theorem 4.8.6 (i)) to F. In particular,

$$\begin{aligned} F'(x) &= \sum_{k=0}^{\infty} \frac{d}{dx}\left(\frac{a_k}{k+1}x^{k+1}\right) \\ &= \sum_{k=0}^{\infty} a_k x^k \\ &= f(x) \end{aligned}$$

whenever $x \in (-R, R)$. Hence F is an antiderivative for f on $(-R, R)$ and hence f is integrable on $(-R, R)$. □

4.9 Maple notes

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

`sum(f(k) , k=m..n);` computes the sum of `f(k)` as `k` runs from `m` to `n`. For example,

```
> sum(k^2, k=1..4);
```

$$30$$

```
> sum(k^2, k=1..n);
```

$$\frac{1}{3}(n+1)^3 - \frac{1}{2}(n+1)^2 + \frac{1}{6}n + \frac{1}{6}$$

```
> sum(1/k^2, k=1..infinity);
```

$$\frac{1}{6}\pi^2$$

`?powseries` will give information about the MAPLE package for manipulating formal power series.

`taylor(expr, x=a, k);` computes the Taylor series for `expr` about `x=a`, up to the term of order `k`.

`convert(taylor(expr, x=a, k), polynom);` computes the Taylor polynomial of order `k-1` for `expr` about `x=a`.

`coeftayl(expr, x=a, k);` computes the kth coefficient in the Taylor series expansion of `expr` about `x=a`.

For example,

```
> taylor(sin(x) ,x=0,8);
```

$$x - \frac{1}{6}x^3 + \frac{1}{120}x^5 - \frac{1}{5040}x^7 + O(x^8)$$

```
> convert(%,polynom);
```

$$x - \frac{1}{6}x^3 + \frac{1}{120}x^5 - \frac{1}{5040}x^7$$

```
> coeftayl(sin(x),x=0,11);
```

$$-\frac{1}{39916800}$$

Chapter 5

Averages, arc length, speed and surface area

In this chapter we look at the application of calculus to the following problems:

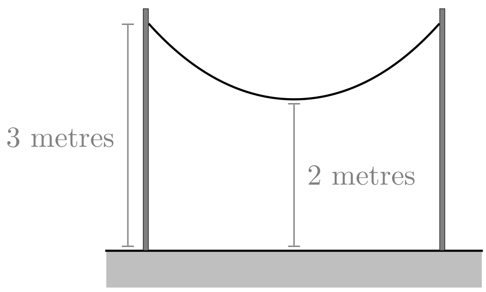
- finding the average height of a cable, or the average temperature over a certain time interval;
- finding the length of a curve;
- finding the speed of a particle that travels along a curve in the plane; and
- finding the surface area of certain solids.

Each of these applications involves integration, either through approximating a quantity with a Riemann sum, or through approximating the rate of change of a quantity and thereafter applying the fundamental theorem of calculus.

5.1 The average value of a function

(Ref: SH10 §5.9)

A cable is suspended between two poles as shown.

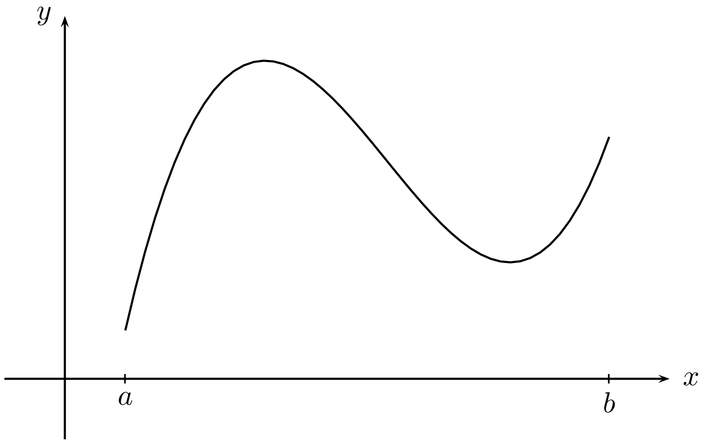


What is the average height of the cable above the ground? Clearly it must be somewhere between 2 and 3 metres. To obtain a precise answer, recall from MATH1131 that any suspended cable is the graph of a function f of the form

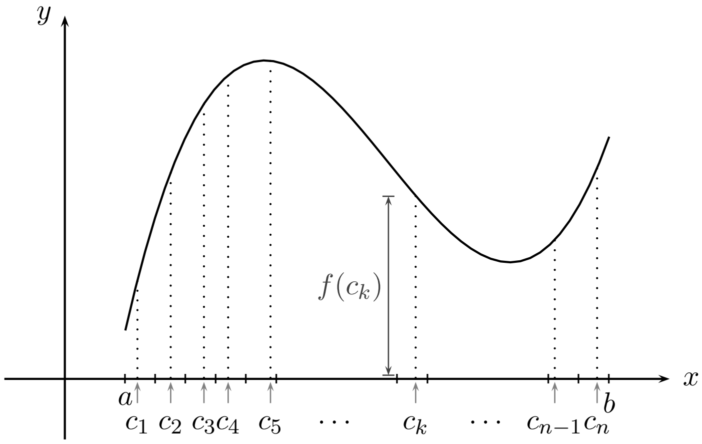
$$f(x) = \frac{1}{c}\cosh(cx)$$

over some interval $[a, b]$, where c is a constant that depends on the tension in and mass of the cable, and where the coordinate system is suitably chosen. Thus we rephrase the question as, 'What is the average value of f on the interval $[a, b]$?'

To answer this question, we need a suitable definition for the average value of a function f over an interval $[a, b]$. To motivate such a definition, consider the function f whose graph is shown below.



One way to proceed is the following. Divide the interval $[a, b]$ into n subintervals of equal length. We sample the height of the graph in the kth subinterval by choosing a point c_k in that subinterval and calculating $f(c_k)$.



The average (that is, arithmetic mean) a_n of these sampled heights is given by

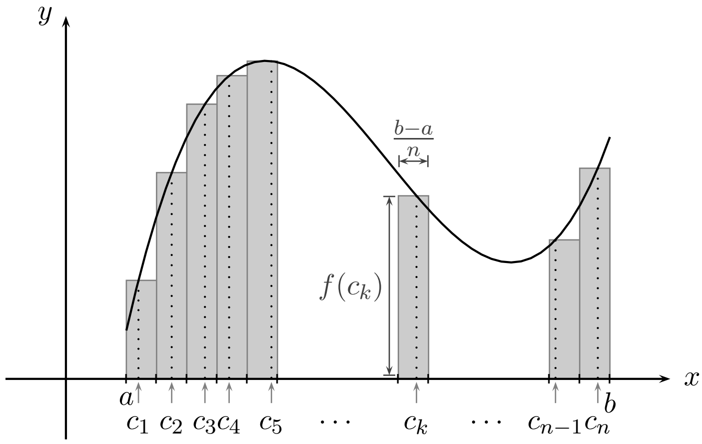
$$a_n = \frac{1}{n}\big(f(c_1) + f(c_2) + f(c_3) + \cdots + f(c_n)\big).$$

As the number n of subintervals increases, a_n should get closer to what we intuitively understand by 'the average height of the graph.'

Now, by multiplying and dividing a_n by $(b-a)$, we find that

$$\begin{aligned} a_n &= \frac{1}{n}\sum_{k=1}^{n} f(c_k) \\ &= \frac{1}{n}\sum_{k=1}^{n} f(c_k)\left(\frac{b-a}{b-a}\right) \\ &= \frac{1}{b-a}\sum_{k=1}^{n} f(c_k)\left(\frac{b-a}{n}\right). \end{aligned} \tag{5.1}$$

Note that $f(c_k)$ and $\dfrac{b-a}{n}$ is the height and width of the kth rectangle in the following diagram.



So the sum in (5.1) is a Riemann sum. If f is Riemann integrable then

$$\lim_{n\to\infty} a_n = \frac{1}{b-a}\lim_{n\to\infty}\sum_{k=1}^{n} f(c_k)\left(\frac{b-a}{n}\right) = \frac{1}{b-a}\int_a^b f(x)\,dx.$$

This leads to the following definition.

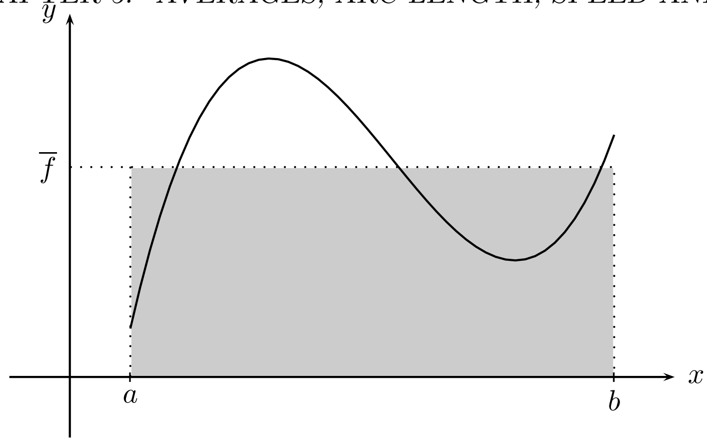
Definition 5.1.1. Suppose that f is integrable on a closed interval $[a, b]$. Then the *average value* $\overline{f}$ of f on $[a, b]$ is defined by the formula

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

Remark 5.1.2. By rearranging this formula, we see that $\overline{f}$ is the unique constant such that

$$(b-a)\overline{f} = \int_a^b f(x)\,dx.$$

Interpreted geometrically, $\overline{f}$ is the unique y-value such that the area of the shaded rectangle is equal to the area under the graph of f.



We now answer the question posed at the beginning of this section.

Example 5.1.3 (The average height of a suspended cable)**.** The suspended cable illustrated at the beginning of this section is a curve given by the equation

$$y = 2\cosh(x/2), \qquad -a \le x \le a,$$

where the x-axis runs along the ground, the y-axis passes through the vertex of the curve and $a = 2\cosh^{-1}(3/2)$. Find, to the nearest centimetre, the average height of the cable above the ground.

Note: The fact that the x-axis runs along the ground is special to this example. This will not necessarily be the case for any given suspended cable.

Solution. (a) Suppose that $f(x) = 2\cosh(x/2)$. The average value $\overline{f}$ of f on the interval $[-a, a]$ is given by

$$\begin{aligned}
\overline{f} &= \frac{1}{a+a}\int_{-a}^{a} 2\cosh(x/2)\,dx \\
&= \frac{2}{2a}\int_{0}^{a} 2\cosh(x/2)\,dx && \text{(since cosh is even)} \\
&= \frac{1}{a}\Big[4\sinh(x/2)\Big]_0^a \\
&= \frac{4}{a}\sinh(a/2) \\
&= \frac{2}{\cosh^{-1}(3/2)}\sinh\left(\cosh^{-1}(3/2)\right).
\end{aligned}$$

Now $\cosh^2(t) - \sinh^2(t) = 1$ and so

$$\sinh\left(\cosh^{-1}(3/2)\right) = \sqrt{\cosh^2\left(\cosh^{-1}(3/2)\right) - 1} = \frac{\sqrt{5}}{2}.$$

Hence

$$\overline{f} = \frac{2}{\cosh^{-1}(3/2)} \times \frac{\sqrt{5}}{2}.$$

By using a calculator, we find that $\overline{f} \approx 2.32$. So the average height of the cable above the ground is approximately 2.32 metres. □

We saw in MATH1131 that every continuous function f defined on a closed interval $[a, b]$ attains its maximum and minimum values. (This result is called the maximum-minimum theorem; see Chapter 2 in the MATH1131 calculus notes). The next theorem says that such a function also attains its average value.

Theorem 5.1.4 (The mean value theorem for integrals)**.** *Suppose that f is continuous on $[a, b]$. Then there is a number c in (a, b) such that*

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

Restated, the conclusion of the mean value theorem for integrals says that there exists a point c in $[a, b]$ such that $f(c) = \overline{f}$, where $\overline{f}$ is the average value of f on $[a, b]$.

Proof. Define $F : [a, b] \to \mathbb{R}$ by the formula

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

By the fundamental theorem of calculus, F is continuous on $[a, b]$, differentiable on (a, b) and $F'(x) = f(x)$. By the mean value theorem, there exists $c \in [a, b]$ such that

$$\frac{F(b) - F(a)}{b-a} = F'(c). \tag{5.2}$$

But

$$F(a) = 0, \qquad F(b) = \int_a^b f(t)\,dt \qquad \text{and} \qquad F'(c) = f(c).$$

Hence (5.2) implies that

$$\frac{1}{b-a}\int_a^b f(t)\,dt = f(c)$$

as required. ☐

Remark 5.1.5. A more general version of the mean value theorem for integrals is given in the tutorial problems for Chapter 4 and is used to prove the Lagrange formula for the remainder in Taylor's theorem.

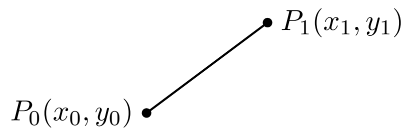
5.2 The arc length of a curve

(Ref: SH10 §10.7)

Suppose that $P_0(x_0, y_0)$ and $P_1(x_1, y_1)$ are two points in $\mathbb{R}^2$. The distance between P_0 and P_1 is given by

$$\mathrm{dist}(P_0, P_1) = \sqrt{(x_1 - x_0)^2 + (y_1 - y_0)^2}.$$

Suppose that the points P_0 and P_1 are the endpoints of a straight line segment P_0P_1, as shown.



Then we define the *length* of the line segment P_0P_1 to be the distance between P_0 and P_1. In other words,

$$\boxed{\text{length}(P_0P_1) = \sqrt{(x_1 - x_0)^2 + (y_1 - y_0)^2}.}$$

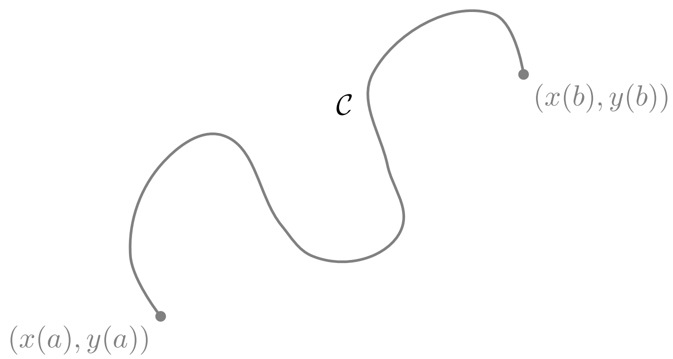
The line segment P_0P_1 is a special example of a curve in $\mathbb{R}^2$. In this section we study the lengths of curves that are not necessarily straight line segments. We begin by presenting an intuitive derivation of a formula that gives the arc length of a curve. (A rigorous approach to proving the validity of such a formula would involve giving a formal definition for arc length and considering the limits of some technically difficult Riemann sums; we shall not delve into this here.) The subsections following this derivation present variations of this formula and some examples.

5.2.1 An intuitive derivation of the arc length formula

Suppose that $\mathcal{C}$ is a curve in $\mathbb{R}^2$. The goal is to give an heuristic derivation of a formula for the arc length of $\mathcal{C}$. We make the assumption that $\mathcal{C}$ can be expressed in parametric form as

$$\mathcal{C} = \big\{(x(t), y(t)) \in \mathbb{R}^2 : a \le t \le b\big\},$$

where x and y are differentiable functions of t.

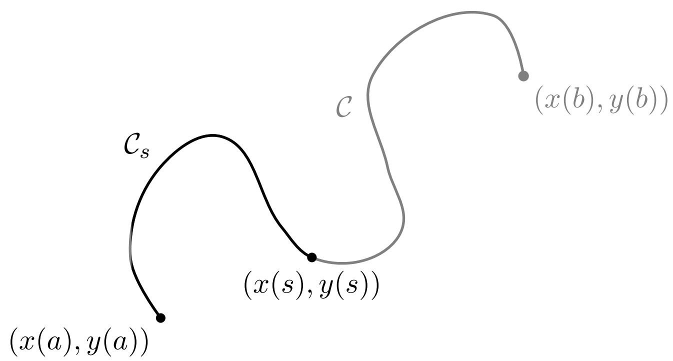


We also assume that the parametrisation is chosen so that the path traversed by the moving point $(x(t), y(t))$ does not retrace its steps (either forwards or backwards).

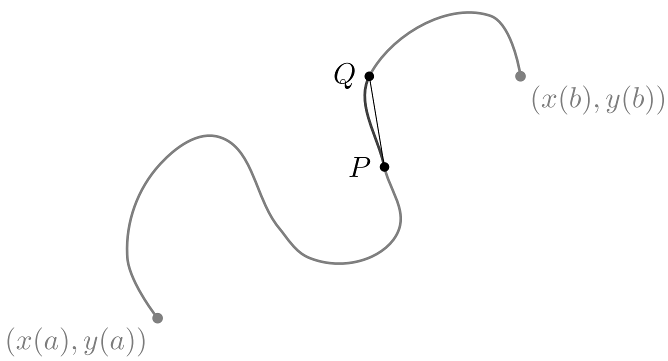
When $a \le t \le b$, let $\ell(s)$ denote the arc length of the curve $\mathcal{C}_s$, given by

$$\mathcal{C}_s = \big\{(x(t), y(t)) \in \mathbb{R}^2 : a \le t \le s\big\}.$$

The segment $\mathcal{C}_s$ corresponding to the arc length $\ell(s)$ is illustrated in black in the diagram below.



The idea is to take a small segment of the curve and approximate its length with the length of a secant. Suppose that $a < t < b$ and that h is a small real nonzero number. Consider the points $P(x(t), y(t))$ and $Q(x(t+h), y(t+h))$.



The length of the arc from P to Q is approximately equal to the length of secant PQ. That is

$$\ell(t+h) - \ell(t) \approx \sqrt{[x(t+h) - x(t)]^2 + [y(t+h) - y(t)]^2},$$

where the length of the secant is calculated using the distance formula. If we divide both sides by h then

$$\frac{\ell(t+h) - \ell(t)}{h} \approx \sqrt{\left[\frac{x(t+h) - x(t)}{h}\right]^2 + \left[\frac{y(t+h) - y(t)}{h}\right]^2}.$$

This approximation gets better as h gets smaller. If we make the assumption that ℓ is a differentiable function of s then, by taking the limit as h approaches zero, one obtains

$$\ell'(t) = \sqrt{[x'(t)]^2 + [y'(t)]^2}.$$

Hence

$$\ell(s) = \int_a^s \sqrt{[x'(t)]^2 + [y'(t)]^2}\, dt + K$$

for some constant of integration K, by the fundamental theorem of calculus. To evaluate K, note that $\ell(a) = 0$. So if $s = a$ then

$$0 = \ell(a) = \int_a^a \sqrt{[x'(t)]^2 + [y'(t)]^2}\, dt + K = 0 + K = K.$$

Hence $K = 0$ and thus

$$\ell(s) = \int_a^s \sqrt{[x'(t)]^2 + [y'(t)]^2}\, dt.$$

Finally, the length of the entire curve is $\ell(b)$. So the arc length of $\mathcal{C}$ is

$$\int_a^b \sqrt{[x'(t)]^2 + [y'(t)]^2}\, dt.$$

Our findings are summarised at the beginning of the next subsection.

5.2.2 Arc length for a parametrised curve

Suppose that a curve $\mathcal{C}$ can be expressed in parametric form as

$$\mathcal{C} = \{(x(t), y(t)) \in \mathbb{R}^2 : a \le t \le b\},$$

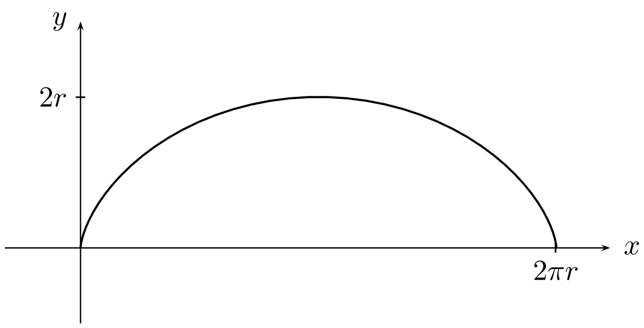
where x and y are differentiable functions of t. Then its arc length ℓ is given by the formula

$$\boxed{\ell = \int_a^b \sqrt{[x'(t)]^2 + [y'(t)]^2}\, dt} \tag{5.3}$$

The next example shows how this formula is applied. We consider the cycloid, which is closely related to the so-called 'curve of fastest descent' (see Section 7.2 of the MATH1131 calculus course notes).

Example 5.2.1 (The arc length of a cycloid)**.** Find the arc length of one arch of the cycloid

$$x(t) = r(t - \sin t), \qquad y(t) = r(1 - \cos t), \qquad 0 \le t \le 2\pi.$$



Solution. We begin by calculating the derivatives:

$$x'(t) = r(1 - \cos t), \qquad y'(t) = r \sin t.$$

Hence

$$\begin{aligned}[x'(t)]^2 + [y'(t)]^2 &= r^2(1 - 2\cos t + \cos^2 t) + r^2 \sin^2 t \\ &= 2r^2(1 - \cos t),\end{aligned}$$

since $\cos^2 t + \sin^2 t = 1$. We want to substitute this into formula (5.3). Before doing so, it is best to express $1 - \cos t$ as a square. Now $1 - \cos 2\theta = 2\sin^2\theta$ and so

$$\begin{aligned}[x'(t)]^2 + [y'(t)]^2 &= 2r^2(1 - \cos t) \\ &= 4r^2 \sin^2(t/2).\end{aligned}$$

Hence (5.3) gives

$$\begin{aligned}\ell &= \int_0^{2\pi} \sqrt{[x'(t)]^2 + [y'(t)]^2}\, dt \\ &= \int_0^{2\pi} \sqrt{4r^2 \sin^2(t/2)}\, dt.\end{aligned}$$

At this point we should be careful with taking the square root, since $\sqrt{a^2} = a$ only when $a \geq 0$. Now $\sin(t/2)$ is positive whenever $0 < t < 2\pi$, and so taking the squareroot in the 'naïve' way causes no problems. Therefore

$$\begin{aligned} \ell &= \int_0^{2\pi} 2r\sin(t/2)\,dt \\ &= 2r\Big[-2\cos(t/2)\Big]_0^{2\pi} \\ &= 8r. \end{aligned}$$

So the arc length of one arch of the cycloid is $8r$ units. ☐

Remark 5.2.2. For parametrisations of closed curves, one should be careful with the limits of integration. For example, a circle of radius r and centre $(0,0)$ may be parametrised as

$$x(t) = r\cos t, \qquad y(t) = r\sin t, \qquad 0 \leq t \leq 2\pi.$$

Hence

$$\ell = \int_0^{2\pi} \sqrt{[-r\sin t]^2 + [r\cos t]^2}\,dt = \int_0^{2\pi} r\,dt = 2\pi r,$$

which shows that the circumference of the circle is $2\pi r$, as expected.

On the other hand, if we use the parametrisation

$$x(t) = r\cos 2t, \qquad y(t) = r\sin 2t$$

then

$$\int_0^{2\pi} \sqrt{[-2r\sin 2t]^2 + [2r\cos 2t]^2}\,dt = \int_0^{2\pi} 2r\,dt = 4\pi r,$$

which is *not* the circumference of the circle. The reason for this is that, as t varies from 0 to 2π, the point $(x(t), y(t))$ moves around the circle twice! To find the arc length using this parametrisation, one should instead integrate from 0 to π.

5.2.3 Arc length for the graph of a function

Suppose that f is a function of one variable. To find the arc length of the graph of f on the interval $[a, b]$, we parametrise the curve

$$y = f(x)$$

by

$$x(t) = t, \qquad y(t) = f(t), \qquad a \leq t \leq b.$$

Now

$$x'(t) = 1 \qquad \text{and} \qquad y'(t) = f'(t).$$

By using the arc length formula (5.3), one finds that

$$\ell = \int_a^b \sqrt{1 + [f'(t)]^2}\,dt$$

We usually write the variable of integration as x.

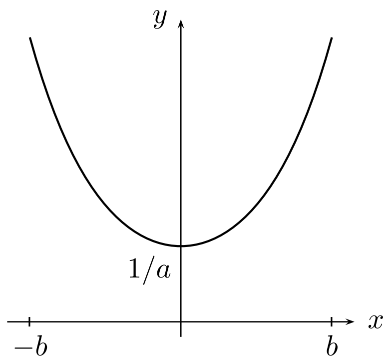
In summary, the arc length ℓ of the graph of a function f on the interval $[a, b]$ is given by

$$\boxed{\ell = \int_a^b \sqrt{1 + [f'(x)]^2}\, dx.} \tag{5.4}$$

The use of this formula is illustrated below. We remind readers that a catenary is the shape of a hanging cable and is described using the hyperbolic cosine function. See Chapter 10 of the MATH1131 calculus notes for further details.

Example 5.2.3 (The arc length of a catenary)**.** Find the arc length of a catenary whose graph is given by

$$y = \frac{1}{a}\cosh(ax), \qquad x \in [-b, b].$$



Solution. Suppose that $f(x) = \frac{1}{a}\cosh(ax)$. By formula (5.4),

$$\begin{aligned}
\ell &= \int_{-b}^{b} \sqrt{1 + [f'(x)]^2}\, dx \\
&= \int_{-b}^{b} \sqrt{1 + \sinh^2(ax)}\, dx \\
&= \int_{-b}^{b} \sqrt{\cosh^2(ax)}\, dx && \text{(since } \cosh^2 t - \sinh^2 t = 1\text{)} \\
&= \int_{-b}^{b} \cosh(ax)\, dx && \text{(since cosh is always positive)} \\
&= 2\int_{0}^{b} \cosh(ax)\, dx && \text{(since cosh is even)} \\
&= 2\left[\frac{1}{a}\sinh ax\right]_0^b \\
&= \frac{2}{a}\sinh(ab) \\
&= \frac{1}{a}\left(e^{ab} - e^{-ab}\right).
\end{aligned}$$

So the arc length of the catenary is $\frac{1}{a}\left(e^{ab} - e^{-ab}\right)$ units. □

5.2.4 Arc length for a polar curve

Suppose that a curve is described using polar coordinates by

$$r = f(\theta), \qquad \theta_0 \le \theta \le \theta_1.$$

Since

$$x = r\cos\theta = f(\theta)\cos\theta \qquad \text{and} \qquad y = r\sin\theta = f(\theta)\sin\theta,$$

we have a parametrisation for the curve in terms of θ. Now

$$x'(\theta) = -f(\theta)\sin\theta + f'(\theta)\cos\theta, \qquad \text{while} \qquad y'(\theta) = f(\theta)\cos\theta + f'(\theta)\sin\theta.$$

Hence

$$\begin{aligned}[x'(\theta)]^2 + [y'(\theta)]^2 &= [f(\theta)]^2\sin^2\theta - 2f(\theta)f'(\theta)\sin\theta\cos\theta + [f'(\theta)]^2\cos^2\theta \\ &\quad + [f(\theta)]^2\cos^2\theta + 2f(\theta)f'(\theta)\sin\theta\cos\theta + [f'(\theta)]^2\sin^2\theta \\ &= [f(\theta)]^2 + [f'(\theta)]^2,\end{aligned}$$

where we have used the fact that $\sin^2\theta + \cos^2\theta = 1$. So by using the parametric form (5.3) for arc length, we find that the arc length ℓ of the polar curve is given by

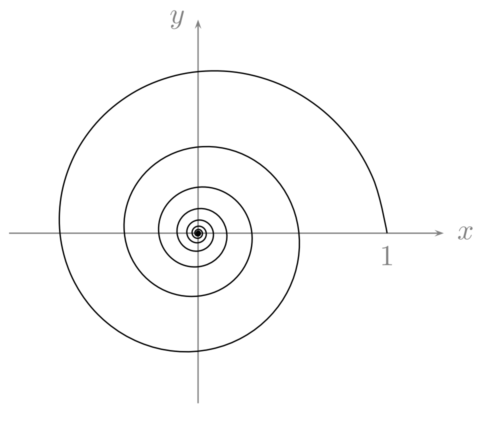
$$\ell = \int_{\theta_0}^{\theta_1} \sqrt{[f(\theta)]^2 + [f'(\theta)]^2}\, d\theta.$$

We usually write $f(\theta)$ as r and $f'(\theta)$ as $\frac{dr}{d\theta}$.

In summary, the arc length ℓ of a polar curve is given by

$$\boxed{\ell = \int_{\theta_0}^{\theta_1} \sqrt{r^2 + \left(\frac{dr}{d\theta}\right)^2}\, d\theta.}$$

Example 5.2.4. The spiral graphed below is given by the polar equation $r = e^{-\theta/10}$, where $\theta \ge 0$. Is the total arc length finite? Explain.



Solution. We begin by calculating the arc length for the segment of the curve when $0 \le \theta \le \theta_1$. Now

$$r = e^{-\theta/10} \qquad \text{and} \qquad \frac{dr}{d\theta} = -\frac{1}{10}e^{-\theta/10}.$$

Hence

$$\begin{aligned}
\ell &= \int_0^{\theta_1} \sqrt{\left(e^{-\theta/10}\right)^2 + \left(-\tfrac{1}{10}e^{-\theta/10}\right)^2}\, d\theta \\
&= \int_0^{\theta_1} \sqrt{\left(1+\tfrac{1}{100}\right) e^{-2\theta/10}}\, d\theta \\
&= \frac{\sqrt{101}}{10} \int_0^{\theta_1} e^{-\theta/10}\, d\theta \\
&= \sqrt{101}\left(1 - e^{-\theta_1/10}\right).
\end{aligned}$$

Now as $\theta_1 \to \infty$, $\ell \to \sqrt{101}$. Hence the total arc length is finite and equals $\sqrt{101}$ units. □

5.3 The speed of a moving particle

(Ref: SH10 §10.7)

In Chapter 4 of the MATH1131 calculus notes, we discussed the speed of a particle that moves along a straight line. Now we consider the speed of a particle that moves along a curve in the plane.

Suppose that a particle P is moving in the plane and that its position at time t is given by $(x(t), y(t))$. The distance $s(t)$ that the particle has travelled from time zero to any later time t is given by the formula

$$s(t) = \int_0^t \sqrt{[x'(u)]^2 + [y'(u)]^2}\, du$$

(which is simply the arc length formula for the path that P traverses in this time interval). By definition, the speed of P is the rate of change of its distance with respect to time. So if $v(t)$ denotes the speed of P at time t then

$$v(t) = s'(t) = \sqrt{[x'(t)]^2 + [y'(t)]^2}$$

by the fundamental theorem of calculus.

In summary, the speed $v(t)$ of a particle P at time t is given by

$$\boxed{v(t) = \sqrt{[x'(t)]^2 + [y'(t)]^2},}$$

where the functions x and y give the position $(x(t), y(t))$ of P at time t.

Example 5.3.1. A stone is thrown horizontally from the deck of the Sydney Harbour Bridge at 20 metres per second. Its position $(x(t), y(t))$ exactly t seconds after the stone is thrown is given by

$$x(t) = 20t, \qquad y(t) = 50 - 5t^2, \qquad 0 \le t \le \sqrt{10},$$

where $y(t)$ is the height above the water (see Example 7.2.1 in the MATH1131 calculus notes). Find the speed of the stone an instant before it hits the water.

Solution. We have

$$x'(t) = 20 \qquad \text{and} \qquad y'(t) = -10t.$$

So the speed $v(t)$ of the stone at time t is given by

$$v(t) = \sqrt{20^2 + 100t^2}$$

whenever $0 < t < \sqrt{10}$. Now the stone hits the water when $y(t) = 0$, which is precisely when $t = \sqrt{10}$. Hence the speed of the stone *an instant before* it hits the water is given by

$$\lim_{t \to (\sqrt{10})^-} v(t) = \sqrt{20^2 + 100(\sqrt{10})^2} = 10\sqrt{14} \approx 37.42.$$

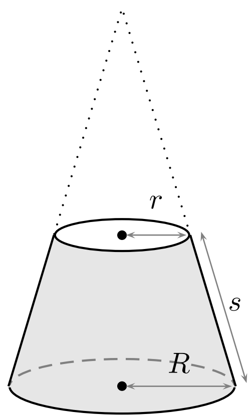
So the speed of the stone just before it hits the water is approximately 37.42 metres per second. ☐

5.4 Surface area

(Ref: SH10 §10.8)

The problem of finding the surface area of a surface or solid in $\mathbb{R}^3$ is not easy. In this section, we focus on finding the surface area for a surface (or solid) that is formed by rotating a curve about one of the axes. In the first subsection we derive, intuitively, a formula for the area of such a surface. In the second subsection, the relevant formulae are summarised and examples given.

The formulae presented all rely on the formula for the surface area of the frustum of a right circular cone.



Given a frustum of slant height s and radii r and R, the surface area A of the 'curved surface' is given by

$$\boxed{A = \pi(r + R)s.} \tag{5.5}$$

This formula may be proved using elementary methods and is left as an exercise in the tutorial problems.

5.4.1 An heuristic derivation for the surface area of a surface of revolution

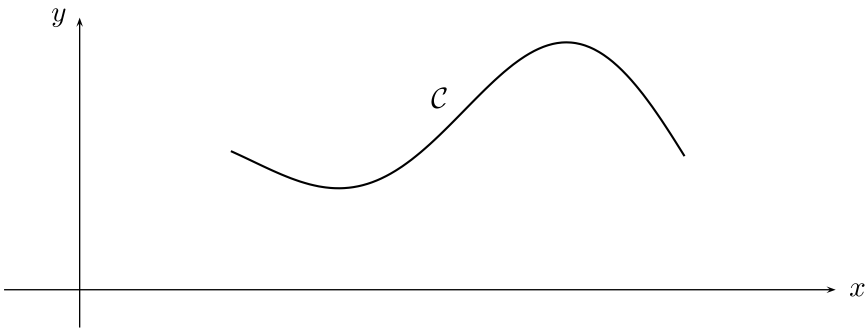
Suppose that a curve $\mathcal{C}$ has parametrisation given by

$$\mathcal{C} = \big\{(x(t), y(t)) \in \mathbb{R}^2 : a \le t \le b\big\}.$$

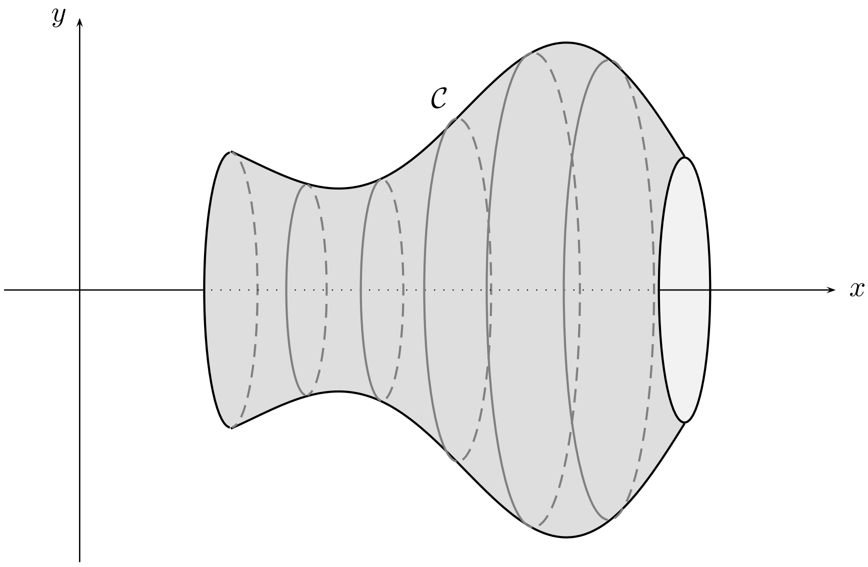
We will assume that

- the curve lies in the upper half-plane (more precisely, $x(t) \ge 0$ and if the curve meets the x-axis then it does so at only a finite number of points); and
- the curve $\mathcal{C}$ is *simple*: if $(x(t_0), y(t_0)) = (x(t_1), y(t_1))$ then $t_0 = t_1$ (that is, the curve does not intersect itself).

An example of such a curve is shown below.



If we rotate the curve about the axis, then a *surface of revolution* is formed.



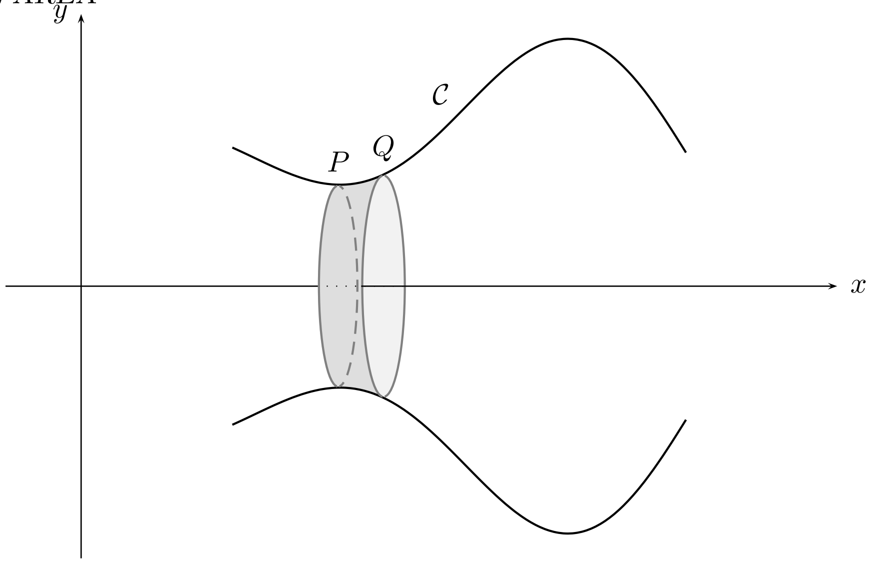
The goal of this subsection is to derive a formula for the area of this surface.

Let $A(s)$ denote the area of the surface formed when the curve segment

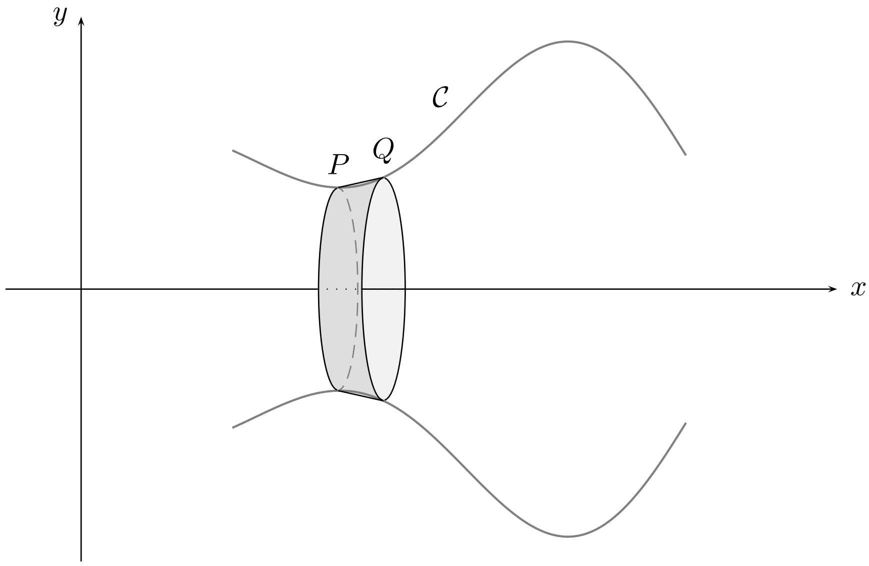
$$\left\{(x(t), y(t)) \in \mathbb{R}^2 : a \leq t \leq s\right\}$$

is rotated about the x-axis. We make the assumption that A is a differentiable function.

Our immediate goal is to compute the derivative of A. Fix t in (a, b) and suppose that h is a small nonzero real number. Consider the points $P(x(t), y(t))$ and $Q(x(t+h), y(t+h))$ and note that $A(t+h) - A(t)$ is the area of the surface formed by rotating the curve segment from P to Q about the x-axis.



Since h is small, this surface area is approximately equal to the area of the surface formed when the secant PQ is rotated about the x axis.



This area may be calculated using formula (5.5), where

$$\text{slant height } = \sqrt{[x(t+h)-x(t)]^2+[y(t+h)-y(t)]^2}.$$

Hence

$$A(t+h)-A(h) \approx \pi\big(y(t+h)+y(t)\big)\sqrt{[x(t+h)-x(t)]^2+[y(t+h)-y(t)]^2},$$

and dividing both sides by h gives

$$\frac{A(t+h)-A(h)}{h} \approx \pi\big(y(t+h)+y(t)\big)\sqrt{\left[\frac{x(t+h)-x(t)}{h}\right]^2+\left[\frac{y(t+h)-y(t)}{h}\right]^2}.$$

Note that approximation improves as h gets smaller. Moreover, since y is differentiable at t, it follows that y is continuous at t and so $y(t+h) \to y(t)$ as $h \to 0$. By taking the limit as h

approaches zero, we obtain

$$\begin{aligned} A'(t) &= \pi\big(y(t)+y(t)\big)\sqrt{[x'(t)]^2+[y'(t)]^2} \\ &= 2\pi y(t)\sqrt{[x'(t)]^2+[y'(t)]^2}. \end{aligned}$$

This gives an expression for the derivative of A. By applying the fundamental theorem of calculus, we find that

$$A(s) = \int_a^s 2\pi y(t)\sqrt{[x'(t)]^2+[y'(t)]^2}\,dt + K$$

for some constant of integration K. To evaluate K, note that $A(a) = 0$. So the substitution $s = a$ yields

$$0 = A(a) = \int_a^a 2\pi y(t)\sqrt{[x'(t)]^2+[y'(t)]^2}\,dt + K = 0 + K = K.$$

Thus $K = 0$ and hence

$$A(s) = \int_a^s 2\pi y(t)\sqrt{[x'(t)]^2+[y'(t)]^2}\,dt.$$

Finally, the substitution $s = b$ yields

$$A(b) = \int_a^b 2\pi y(t)\sqrt{[x'(t)]^2+[y'(t)]^2}\,dt,$$

which is a formula for the area of the surface of revolution formed by rotating $\mathcal{C}$ about the x-axis.

5.4.2 Surface area formulae and examples

Assume that a curve $\mathcal{C}$ lies in the upper-half plane and is simple (see the assumptions stated at the beginning of Subsection 5.4.1). We present formulae for the area of the surface of revolution about the x-axis when $\mathcal{C}$ is described either parametrically, as the graph of a function or using polar coordinates. In each case we assume that the appropriate derivatives exist.

If $\mathcal{C}$ is described parametrically by

$$\mathcal{C} = \big\{(x(t), y(t)) \in \mathbb{R}^2 : a \le t \le b\big\},$$

then the area A of the surface of revolution about the x-axis is given by

$$\boxed{A = \int_a^b 2\pi y(t)\sqrt{[x'(t)]^2+[y'(t)]^2}\,dt.} \tag{5.6}$$

If $\mathcal{C}$ is the graph

$$y = f(x), \qquad x \in [a, b],$$

of a function f on $[a, b]$ then the area A of the surface of revolution about the x-axis is given by

$$\boxed{A = \int_a^b 2\pi f(x)\sqrt{1+[f'(x)]^2}\,dx.} \tag{5.7}$$

If $\mathcal{C}$ is described using polar coordinates by

$$r = f(\theta), \qquad \theta_0 \le \theta \le \theta_1,$$

then the area A of the surface of revolution about the x-axis is given by

$$\boxed{A = \int_{\theta_0}^{\theta_1} 2\pi r\sin\theta\sqrt{r^2+\left(\frac{dr}{d\theta}\right)^2}\,d\theta.} \tag{5.8}$$

Remark 5.4.1. Formula (5.6) was derived heuristically in Subsection 5.4.1. Formula (5.7) may be easily deduced from (5.6) by using the following parameterisation of the graph of f:

$$x(t) = t, \qquad y(t) = f(t), \qquad a \le t \le b.$$

Formula (5.8) may be easily deduced from (5.6) by using the following parameterisation of the polar curve:

$$x(\theta) = f(\theta)\cos\theta, \qquad y(\theta) = f(\theta)\sin\theta, \qquad \theta_0 \le \theta \le \theta_1.$$

Remark 5.4.2. In parametric form, the formula for the area A of surface of revolution about the y-axis is given by

$$A = \int_a^b 2\pi x(t)\sqrt{[x'(t)]^2 + [y'(t)]^2}\, dt.$$

Other versions of this formula may be easily deduced using the parametrisations given in the previous remark.

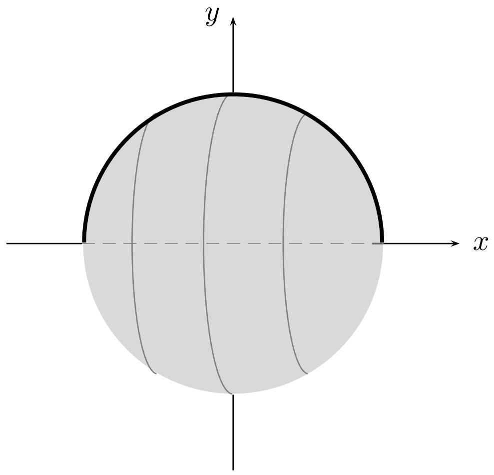
Remark 5.4.3. Each of these formulae only give the area of the *surface* of revolution. To find the surface area of the *solid* of revolution, one must also add the surface area contributed by any circular 'caps' appearing at each end of the surface of revolution. See, for example, Example 5.4.5.

Example 5.4.4. Find the surface area of a sphere of radius r.

Solution. A sphere of radius r is formed by rotating the curve

$$x(t) = r\cos t, \qquad y(t) = r\sin t, \qquad 0 \le t \le \pi$$

about the x-axis.



By using formula (5.6), we find that the surface area A is given by

$$\begin{aligned} A &= \int_0^\pi 2\pi r\sin t\sqrt{[-r\sin t]^2 + [r\cos t]^2}\, dt \\ &= \int_0^\pi 2\pi r\sin t\sqrt{r^2(\sin^2 t + \cos^2 t)}\, dt \\ &= \int_0^\pi 2\pi r^2 \sin t\, dt \\ &= 2\pi r^2\Big[-\cos t\Big]_0^\pi \\ &= 4\pi r^2. \end{aligned}$$

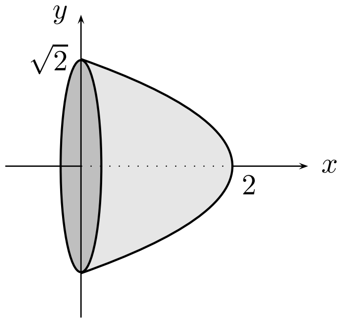
So the surface area of a sphere of radius r is $4\pi r^2$. $\square$

Example 5.4.5. A solid $\mathcal{S}$ is formed by rotating the curve given by

$$y = \sqrt{2-x}, \qquad x \in [0,2],$$

about the x-axis. Find the surface area of the solid, making sure that the every face of the solid is accounted for.

Solution. The solid $\mathcal{S}$ is drawn below.



It has two faces: the truncated paraboloid (shaded in lighter gray) and the circular cap (shaded in darker gray). The area A_1 of the truncated paraboloid is given by

$$\begin{aligned}
A_1 &= \int_0^2 2\pi f(x)\sqrt{1+[f'(x)]^2}\,dx \qquad\qquad \text{(where } f(x)=\sqrt{2-x}\text{)}\\
&= \int_0^2 2\pi\sqrt{2-x}\left(1+\left(\frac{-1}{2\sqrt{2-x}}\right)^2\right)^{1/2} dx\\
&= \int_0^2 2\pi\sqrt{2-x}\left(1+\frac{1}{4(2-x)}\right)^{1/2} dx\\
&= \pi\int_0^2 \sqrt{4(2-x)+1}\,dx\\
&= \pi\int_0^2 \sqrt{9-4x}\,dx\\
&= \pi\left[\frac{(9-4x)^{3/2}}{-6}\right]_0^2\\
&= \frac{13\pi}{3}.
\end{aligned}$$

The area A_2 of the circular cap is given by

$$A_2 = \pi r^2 = \pi(\sqrt{2})^2 = 2\pi.$$

Hence the total surface area of the solid $\mathcal{S}$ is given by

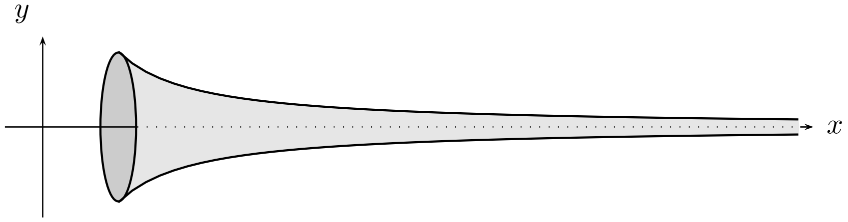
$$A_1 + A_2 = \frac{19\pi}{3},$$

which is approximately 19.9 square units. □

We end with a simple but interesting example. Recall from high school that the volume V of the solid formed when the graph of a function $f:[a,b]\to[0,\infty)$ is rotated about the x-axis is given by

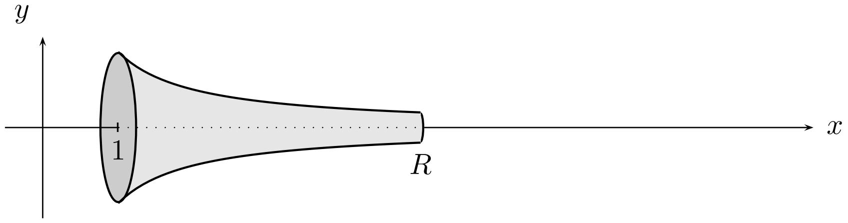
$$V = \int_a^b \pi[f(x)]^2\,dx. \tag{5.9}$$

Example 5.4.6 (Gabriel's horn)**.** Suppose that the function $f : [1, \infty) \to \mathbb{R}$ is defined by $f(x) = \frac{1}{x}$. Rotate the graph of f about the x-axis, as shown.



The surface of revolution formed is known as *Gabriel's horn* (after the biblical figure Gabriel) or *Torricelli's trumpet* (after the mathematician and philosopher Evangelista Torricelli, who was a pupil of Galileo). Even though the length of the horn is infinite, it may still be possible to calculate its surface area and the volume of the corresponding solid.

We begin with the volume. First, consider the volume V_R of the truncated solid shown below.



Using formula (5.9),

$$\begin{aligned} V_R &= \int_1^R \frac{\pi}{x^2}\, dx \\ &= \pi \left[-\frac{1}{x} \right]_1^R \\ &= \pi \left(1 - \frac{1}{R} \right). \end{aligned}$$

Now $\lim_{R\to\infty} V_R = \pi$, and so the volume of the solid of revolution is finite and equal to π cubic units.

We now examine the surface area. The area A of the surface of revolution for the truncated curve is given by

$$A = 2\pi \int_1^R \frac{1}{x} \sqrt{1 + \frac{1}{x^4}}\, dx.$$

Finding an antiderivative for the integrand looks difficult. However, we are mainly interested in whether or not the improper integral

$$\int_1^\infty \frac{1}{x} \sqrt{1 + \frac{1}{x^4}}\, dx$$

converges. Note that

$$\frac{1}{x} \sqrt{1 + \frac{1}{x^4}} > \frac{1}{x} \sqrt{1 + 0} = \frac{1}{x}$$

whenever $x \geq 1$. Since the improper integral

$$\int_1^\infty \frac{1}{x}\, dx$$

diverges, so too does the integral

$$\int_1^\infty \frac{1}{x}\sqrt{1+\frac{1}{x^4}}\,dx$$

by the comparison test for integrals (see Chapter 8 in the MATH1131 calculus notes). Hence the surface area of Gabriel's horn is infinite.

The fact that Gabriel's horn has finite volume (π cubic units) and infinite surface area leads to the following paradox. To paint the outside surface of the horn, one requires an infinite amount of paint since the surface area is infinite. However, to paint the inside surface of the horn, one only needs at most π cubic units of paint. Simply fill the horn with paint and then remove whatever paint is not touching the surface. This is sometimes called the *painter's paradox*.

Question: How is the paradox resolved?

Index