ANALYSIS

Lecture Notes by
M.K. Warby and J.E. Furter
Department of Mathematical Sciences
Brunel University, Uxbridge UK

September 29, 2004

Contents

Initial Plan

On the next few pages you will find the syllabus for MA2930. Michael Warby (MKW) will be starting, then Jacques Furter (JEF) will take over by Week 6 or 7. Later, in the second term (from January 2004) there will be another shift, MKW will take about 4 weeks and JEF will then take the remainder of the module.

There will be one assignment given out in the autumn, one assignment (the largest one) due for early February and the final assignment (a group work) to be handed back at the end of the lecturing period with the combined worth of these assignments being 40% of the total mark for the module. It is intended that the first assignment will be given out in week 4 (near the end of October) and will be due in at the beginning of week 8. The second assignment will be given out in week 10 and will be due in during week 19 which is in the first part of February. The final one will be given out in week 22 to be handed in week 26 (just before the Easter break).

In these notes there is an approximate teaching schedule for all the material. You should note on the timetable that all the hours for this module are shown as LEC. Some weeks, for instance in the first 2 weeks all 3 hours will be used as lectures. From week 3 onwards one of the 3 hours will often be used as an example class or a seminar although this is not fixed at this stage. Please do not hesitate to ask at the beginning of a lecture if you have problems with some exercises. We will hand out notes in the lectures and we also intend making the printed notes available from the 'home page area' MKW on the www. This address together with my name, room number, internal telephone number and email address are listed below. Further on are the coordinates of JEF.

Name: Dr. M. K. Warby

Room: M301 (Mathematics building)

Internal telephone: 65604 (The Brunel number is ++44 1895 274000)

Mathematics department Fax: ++44 1895 203303

Email: Mike.Warby@brunel.ac.uk

URL: http://www.brunel.ac.uk/~icstmkw/home.html

Our proposed teaching style will be to give out reasonably detailed printed notes. Our lecturing styles might somewhat differ but they are compatible. During the lectures we will summarize the main points on the white boards and using the overhead projector (OHP) and we will work through key material on the boards. Although we will generally follow reasonably closely the material covered in the printed notes we would hope that the material covered on the board will in itself be an adequate self-contained coverage of the module. During the lectures we might emphasise points that are found difficult by this year's cohort (or more precisely those who attend!). We hence urge people to write down everything put on the board so as to have a record of what was covered in the lectures and to get used to writing mathematical arguments correctly. We will usually, for example, summarize at the start of the lectures the main points covered previously which in itself should be useful and which will generally not be contained in the printed notes. We will also give summaries in the printed notes at the end of a topic of the main point to help you focus on the main points a student needs to know to get a decent mark. From our personal experience, the act of taking notes also helps maintain concentration during the lectures. The printed notes can then be read after the lectures to help reinforce and perhaps clarify what has been covered. The seminars and the start of each lecture can additionally be used if you need us to re-explain and perhaps clarify specific aspects which you cannot understand or which you find difficult. Questions and comments are useful for us to see where you stand. Do not hesitate to ask, often someone else has the same, or a similar, point to be clarified.

Name: Dr. J. E. Furter

Room: `M104-5` (Mathematics building)

Internal telephone: `65183`

Email: `mastjef@brunel.ac.uk`

URL: `http://www.brunel.ac.uk/~mastjef/`

Regarding the printed notes themselves, a few conventions should be noted. The **Axioms**, **Definitions**, **Lemmas**, **Propositions**, **Theorems** and **Corollaries** need to be known. It is important to remember *accurately* all the **Axioms** and **Definitions**. You may be asked these in the final examination. *The term being defined appears in* roman characters *in an italic environment.* There might be some other definitions of terms in bold in the text. They are not examinable but often still useful. The lemmas, propositions, theorems and corollaries will generally be important and the statements of these should be remembered. You may need to make use of them or quote them in an answer to a question.

The proof for most of these will be given in the printed notes and during the lectures. As in many books, the symbol □ will be used to identify the end of the proof. Attempting to understand a proof will generally help to understand the result. For instance it will help you to remember the main properties of vectors spaces and their transformations or find them again if you forgot them. **By working through proofs you can check all of the mathematics we develop during the lectures. You do not need to take any of our statements on trust, you can verify them yourselves, in your own time.** We remind you that

- **to prove** a statement you need to show that it is true for <u>all</u> possible cases (or an equivalent logical statement).
- **to disprove** a statement, a single counter-example or a contradiction is enough,
- the use of letters for numbers in proofs allows us to use arguments which do not involve particular values, hence the generality of that argument.

- Similarly, some examples will also involve letters to be able to calculate *many* examples at the same time. It will illustrate that the particular values are not that important. It is the *structure* of the problem which is fundamental to understand its behaviour (like the vector product of two non aligned vectors is always perpendicular to both vectors), or to get the values leading to a *particular* problem like, when is the determinant of a matrix equal to 0?

You will not be expected to re-produce all proofs in an examination. The proofs of 'standard' results that you could be asked in an examination will be written as **PROOF** in the printed notes. The exercise sheets and the examination papers of the last few years give some indication as to what you might be asked. It is intended that they will become available on-line. Further information will be given out during the second part of the year.

THIS TEXT IS NOT A SUBSTITUTE FOR COMING TO THE CONTACT HOURS, NOR COMING TO THEM A SUBSTITUTE FOR READING BEFOREHAND THESE NOTES.

MANY EXAMPLES IN THE TEXT WILL ONLY BE WORKED OUT DURING THE CONTACT HOURS.

2004/2005

July

Su	Mo	Tu	We	Th	Fr	Sa
				1	2	3
4	5	6	7	8	9	10
11	12	13	14	15	16	17
18	19	20	21	22	23	24
25	26	27	28	29	30	31

August

Su	Mo	Tu	We	Th	Fr	Sa
1	2	3	4	5	6	7
8	9	10	11	12	13	14
15	16	17	18	19	20	21
22	23	24	25	26	27	28
29	30	31				

September

Su	Mo	Tu	We	Th	Fr	Sa
			1	2	3	4
5	6	7	8	9	10	11
12	13	14	15	16	17	18
19	20	21	22	23	24	25
26	27	28	29	30		

October

Su	Mo	Tu	We	Th	Fr	Sa
					1	2
3	4	5	6	7	8	9
10	11	12	13	14	15	16
17	18	19	20	21	22	23
24	25	26	27	28	29	30
31						

November

Su	Mo	Tu	We	Th	Fr	Sa
	1	2	3	4	5	6
7	8	9	10	11	12	13
14	15	16	17	18	19	20
21	22	23	24	25	26	27
28	29	30				

December

Su	Mo	Tu	We	Th	Fr	Sa
			1	2	3	4
5	6	7	8	9	10	11
12	13	14	15	16	17	18
19	20	21	22	23	24	25
26	27	28	29	30	31	

January

Su	Mo	Tu	We	Th	Fr	Sa
						1
2	3	4	5	6	7	8
9	10	11	12	13	14	15
16	17	18	19	20	21	22
23	24	25	26	27	28	29
30	31					

February

Su	Mo	Tu	We	Th	Fr	Sa
		1	2	3	4	5
6	7	8	9	10	11	12
13	14	15	16	17	18	19
20	21	22	23	24	25	26
27	28					

March

Su	Mo	Tu	We	Th	Fr	Sa
		1	2	3	4	5
6	7	8	9	10	11	12
13	14	15	16	17	18	19
20	21	22	23	24	25	26
27	28	29	30	31		

April

Su	Mo	Tu	We	Th	Fr	Sa
					1	2
3	4	5	6	7	8	9
10	11	12	13	14	15	16
17	18	19	20	21	22	23
24	25	26	27	28	29	30

May

Su	Mo	Tu	We	Th	Fr	Sa
1	2	3	4	5	6	7
8	9	10	11	12	13	14
15	16	17	18	19	20	21
22	23	24	25	26	27	28
29	30	31				

June

Su	Mo	Tu	We	Th	Fr	Sa
			1	2	3	4
5	6	7	8	9	10	11
12	13	14	15	16	17	18
19	20	21	22	23	24	25
26	27	28	29	30		

27	Sep	--	12	Nov	Weeks 1--7	Lectures by MKW
15	Nov	--	17	Dec	Weeks 8--12	Lectures by JEF
22	Dec	--	3	Jan	Weeks 13--15	*Christmas break*
4	Jan	--	28	Jan	Weeks 16--19	Lectures by MKW
31	Jan	--	24	Mar	Weeks 20--26	Lectures by JEF
25	Mar	--	15	Apr	Weeks 27--29	*Easter break*
18	Apr	--	2	May	Weeks 30--31	Revision
3	May	--	20	May	Weeks 32--34	Exams.

Please note the above details could change.

Syllabus

Module Code:	`MA2930`.
Module Title:	Analysis.
Level:	2.
Lecturers:	Dr. M. K. Warby/ Dr. J. E. Furter.
Credit value:	20.
Prerequisites:	`MA1915` (Calculus), `MA1920` (Linear Algebra), `MA1972` (Discrete Mathematics, Probability and Statistics).
Formative feedback:	Participation in the seminars and questionnaires.
Method of assessment:	Two hour unseen written examination worth 60% and 4 assignments worth in total 40%.

Main aims of the module

(i) To increase the students' understanding of sequences of real numbers or functions, continuity and various aspects relating to sequences and series of functions involving specifically point-wise and uniform convergence.

(ii) To provide an introduction to the theory of inner product spaces and its application to linear problems in applied mathematics via the concepts of best approximation and orthogonality.

(iii) To enhance and further develop the precision of the students' mathematical work and the rigour of their mathematical arguments.

Main topics of study

SOME PROPERTIES OF THE REAL NUMBERS: supremum and infimum; open and closed intervals.

SEQUENCES AND SERIES OF REAL NUMBERS: the ϵ–N definition of a limit, combining convergent sequences; standard convergent, bounded and monotonic sequences, the monotone convergence theorem; geometric and harmonic series; the integral, comparison, ratio and root tests; absolute and conditional convergence.

CONTINUOUS FUNCTIONS OF A REAL VARIABLE: the ϵ–δ and the sequential definitions of continuity at a point; Lipschitz and differentiable functions; the Contraction Mapping Theorem for functions of a real variable.

SEQUENCES OF FUNCTIONS: the uniform norm, pointwise and uniform convergence of a sequence of functions; uniform convergence preserves continuity. Interchanging the limit and the integral and the limit and differentiation.

DEEPER ASPECTS OF CONVERGENCE: the nested interval property; subsequences and the Bolzano Weierstrass theorem; Cauchy sequences and the Cauchy criterion for convergence of sequences of real numbers; Cauchy's criterion for uniform convergence for sequences of functions.

SERIES OF FUNCTIONS: series of functions and the Weierstrass M-test; power series including the radius of convergence and Taylor's series.

EUCLIDEAN SPACES: real vector spaces with examples; inner products, length, angular measure, distance; orthogonality, Gram-Schmidt process; orthogonal projections, distance to subspace; least square approximation; bases in Euclidean spaces; Bessels's inequality, Parseval's equality.

FOURIER APPROXIMATION: definition as a best approximation; convergence.

BOUNDARY-VALUE PROBLEMS: definitions and examples; eigenvalues and eigenfunctions; symmetric linear transformation; eigenvalue problems for simple Sturm-Liouville systems.

ORTHOGONAL POLYNOMIALS: Legendre polynomials and their standard properties.

Learning outcomes for the module

On successful completion of this module, students will be able to:

1. use key results of analysis to show that a given sequence or series of number converge;
2. use key results of analysis to show that a given sequence or series of functions converge uniformly and to deduce properties of the limit function;
3. to analyse sequences and series of numbers and functions; calculate best approximations in various context of Euclidean spaces;
4. compute eigenvalues and eigenfunctions of elementary Sturm-Liouville problems;
5. write mathematical work with precision.

Teaching/ learning methods/strategies used to enable the achievement of learning outcomes

72 contact hours comprising lectures and example classes (seminars).

In lectures concepts are introduced, definitions are stated, results (theorems) and techniques are explained, often proved. All material is illustrated with practical examples. Example classes are contact hours where students can practice exercises, examples and ask questions in a supportive environment.

The University expects student learning hours for their studies to be a notional 35-hour week (over 31 weeks) including taught classes, private study, revision and assessment. In practice, students will spend time on each module that will vary from week to week, between modules and between students. In particular, attention should be paid to the assessment pattern (assignments, no final examination ...) and the delivery pattern (number of contact hours) that will have an impact on the organisation of the working weeks during the Academic Year.

Average learning hours per week: 6 of which 3 hours are average student/tutor contact hours per week.

Assessment methods which enable student to demonstrate the learning outcomes for the module

Continuous assessment consisting of written assignments (one group work).	40%
A two hour unseen written examination.	60%

LO5 will be assessed by the assignments, LO1–4 will be assessed by both the assignments and the final examination.

Lecture notes

A paper version of notes, exercises and solutions to exercises are distributed during the lectures and seminars. These are also available on the www using the URL

`http://www.brunel.ac.uk/~icstmkw/ma2930/`

Reading list

Core Texts:

There is no core text for this module but several texts cover most parts of the module. Several of the texts cover the material in a more rigorous way than will be covered in this module and thus you may find the texts difficult.

Supplementary Texts:

R. G. Bartle and D. R. Sherbert (2000). *Introduction to Real Analysis*, (third edition). Wiley, Chichester. `QA300.B294`

M. C. Reed (1998). *Fundamental ideas of Analysis.* Wiley, Chichester. `QA300.R44`

If you enter "mathematical analysis" or "real analysis" when searching on the library catalogue then you will find many books which cover much of the material in the first half of the module.

D.L. Kreider, R.G. Kuller, D.R. Ostberg and F.W. Perkins. *An Introduction to Linear Analysis.* Addison-Wesley (1966). `QA402.I55`

H. Anton and C. Rorres (1994) and (2000). *Elementary Linear Algebra with Applications*, (7-8th ed.). Wiley, Chichester. `QA184.A582`

There are also many other books on Functional/Real Analysis (in particular in `QA320` and `QA184`). Look for sections on Inner Products, Euclidean or Hilbert spaces or, directly, for the topics covered in this course.

DO NOT HESITATE TO ASK US FOR ADVICE ABOUT ANY BOOK OR CHAPTER IN A BOOK YOU WOULD LIKE TO WORK ON. WE WILL BE PLEASED TO HELP.

MA2930 Analysis, 2004–2005 Teaching Schedule

Lecture number (approx)	Topics
1–2	Introduction and overview of the module.
2	Revision of the notation. The number systems $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$ and $\mathbb{R}$ and set notation such as $\{$, $\}$, $\cup$ and $\cap$.
3	Some properties of $\mathbb{R}$. Supremum and infimum. Open and closed intervals.
4–6	Sequences in $\mathbb{R}$. ϵ–N definition of a limit. Combining convergent sequences. Standard convergent sequences such as (r^n) when $\lvert r\rvert < 1$ and $(1/n^p)$ when $p \ge 0$. Bounded sequences. Monotonic sequences and the Monotone Convergence Theorem.
7–9	Series in $\mathbb{R}$. The geometric series and the harmonic series. Absolute and conditional convergence. The comparison test. The ratio test and the root test. The integral test.
10–12	Continuous functions on $\mathbb{R}$. The ϵ–δ definition of continuity at a point. The sequential definition of continuity at a point. Lipschitz functions and differentiable functions.
12–13	The Contraction Mapping Theorem in $\mathbb{R}$.
14–19	Real vector spaces. The axioms of a vector space. Subspaces and subspace test. Spans and linear dependence
20–23	Euclidean spaces. The axioms of a scalar product. Orthogonality. Orthonormal bases. The Gram-Schmidt process.
24–27	Best approximation. Orthogonal projections. Least square approximation.
28–29	Fourier Theory. Fourier approximation. Bases in Euclidean spaces. Bessel's inequality. Parseval's equality.

30–32	Sequences of functions. The uniform norm. Pointwise and uniform convergence of a sequence of functions. Uniform convergence preserves continuity. Interchanging the limit and the integral and interchanging the limit and differentiation.
33–35	Deeper aspects of convergence. The nested interval property. Subsequences and the Bolzano-Weierstrass theorem. Cauchy sequences and the Cauchy criterion for convergence of sequences of real numbers. Cauchy's criterion for uniform convergence for sequences of functions.
36–38	Series of functions. Series of functions and the Weierstrass M-test. Power series including the radius of convergence and Taylor's series.
39–40	Convergence of Fourier series. Pointwise and uniform convergence. Gibb's phenomenon.
41–46	Symmetric operators. Definition. Eigenvalues and eigenvectors. Separation of variable. Simple Sturm-Liouville boundary value problems.
47–50	Orthogonal polynomials. Legendre polynomials and their standard properties. Generating functions.

Chapter 0

An Introduction to the Module

0.1 Analysis

0.1.1 Limits

One of the key things that you do in analysis in some form is 'to take a limit'. As examples, you should have already met the following:

Defining a derivative

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

One way of defining a definite integral

$$\int_a^b f(x)\,dx := \lim_{n\to\infty} h \sum_{i=1}^{n} f(x_i), \quad \text{where } h = \frac{b-a}{n}, \quad x_i = a + (i - 1/2)h.$$

The above is concerned with finding the area under the curve of $y = f(x)$ from $x = a$ to $x = b$.

Defining continuity at a point

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

Defining the convergence of a sequence

A sequence (x_n) in $\mathbb{R}$ converges to x if for every $\epsilon > 0$ there exists $N = N(\epsilon)$ such that

$$|x_n - x| < \epsilon \quad \text{for all } n \geq N.$$

From earlier calculus modules and from study before you came to university, you should know that the differentiation operation and the integration operation (defined above respectively in a way which corresponds geometrically to determining the tangent to a curve and to determining the area under a curve) are actually the inverse of each other. Specifically, under certain assumptions,

$$f(x) = \frac{d}{dx}\left(\int_a^x f(t)\,dt\right) \quad \text{and} \quad f(x) = f(a) + \int_a^x f'(t)\,dt \ .$$

If you study the syllabus we re-visit all of these limiting situations in some form by considering further aspects relating to sequences of real numbers, we will briefly consider continuity of functions and towards the end of the module we will also consider sequences of functions. In this topic we will consider, for example, when it is valid to differentiate and integrate series of functions term by term. You will have probably have met most of these already in some form with the difference now being that things will be justified more rigorously.

0.1.2 Convergence

In numerical methods at Level 1 you meet fixed point iterations of the form

$$x_0 \text{ given}, \quad x_{n+1} = \phi(x_n), \quad n = 0, 1, 2, \cdots$$

as a way of approximately solving the equation

$$x = \phi(x)$$

which is in general a nonlinear equation. The emphasis in numerical methods is in generating a fixed point iteration which generates an approximation of sufficient accuracy with the minimum of computational effort. The emphasis in this module is slightly different in that we consider the conditions which are sufficient to imply that fixed points exist and that the sequence given above converges. The key result in this context is the Contraction Mapping Theorem.

0.1.3 Sequences and series of functions

In the case of sequences of functions (f_n) we will see that there are different types of convergence and the properties that the limit function f that we obtain will depend on what type of convergence that we have. For example, consider the sequence

$$f_n : (-1, \infty) \to \mathbb{R}, \quad f_n(x) := \frac{x^n}{1 + x^n} .$$

For each fixed x we consider the limit as $n \to \infty$ and to do this we need to separately consider the 3 cases $x \in (-1, 1)$, $x = 1$ and $x > 1$. If $x \in (-1, 1)$ then $x^n \to 0$. If $x = 1$ then $f_n(1) = 1/2$ for all n and if $x > 1$ then

$$f_n(x) := \frac{x^n}{1 + x^n} = \frac{1}{(1/x)^n + 1} \to \frac{1}{0 + 1} = 1 \quad \text{as } n \to \infty.$$

Thus the 'pointwise limit' function is the function f, $f : (-1, \infty) \to \mathbb{R}$ defined by

$$f(x) := \lim_{n\to\infty} f_n(x) = \begin{cases} 0 & \text{if } x \in (-1, 1), \\ 1/2 & \text{if } x = 1, \\ 1 & \text{if } x > 1. \end{cases}$$

Each function f_n is continuous but the limit function f is discontinuous at $x = 1$. Pointwise convergence is thus not sufficient to give a continuous limit. A condition which is sufficient for this is known as 'uniform convergence'.

When we have uniform convergence it is valid to perform many of the techniques that you may already be using. For example, if $|x| < 1$ then we have the geometric series

$$\frac{1}{1-x} = 1 + x + x^2 + x^3 + \cdots = \sum_{k=0}^{\infty} x^k$$

and integrating we have

$$\begin{aligned} -\ln(1-x) &= \int_0^x \frac{dt}{1-t} = \int_0^x \left(\sum_{k=0}^{\infty} t^k \right) dt \\ &= \sum_{k=0}^{\infty} \left(\int_0^x t^k \right) dt \\ &= \sum_{k=0}^{\infty} \left(\frac{x^{k+1}}{k+1} \right) \\ &= x + x^2/2 + x^3/3 + x^4/4 + \cdots . \end{aligned}$$

The interchange of the integral and the infinite sum is justified in this case because the series converges uniformly in any closed interval of the form $[-r, r] = \{x : \ |x| \le r\}$ for all $0 \le r < 1$. The knowledge of this and other results will thus help lead to a greater understanding of what techniques work and why they work.

0.1.4 Links with other modules

`MA2930` is only one module in your degree programme. The material we cover relates to methods in `MA2915`.

The earlier example with $f_n(x) = x^n/(1+x^n)$ illustrated how, roughly speaking, a sequence of 'smooth functions' can converge to a 'less smooth' limit function. Fourier[1] series and Fourier approximation, which are covered in different ways in both this module and the module `MA2915` (Linear and Numerical Methods) is a topic which is full of examples of this type. It can be shown that the periodic function

$$f : \mathbb{R} \to \mathbb{R}, \quad f(x) = |x| \quad \text{for } |x| \le \pi, \quad f(x + 2\pi) = f(x)$$

whose graph has a 'saw tooth' appearance is actually equal to its Fourier Series

$$\frac{\pi}{2} - \frac{4}{\pi} \left(\cos x + \frac{1}{3^2} \cos 3x + \frac{1}{5^2} \cos 5x + \cdots + \frac{1}{(2k-1)^2} \cos(2k-1)x + \cdots \right).$$

As a technical point, which will be covered in the module, the infinite series representation actually means the limit of the partial sums, i.e. it is the limit of the sequence $(f_n(x))$ where

$$f_n(x) = \frac{\pi}{2} - \frac{4}{\pi} \left(\cos x + \frac{1}{3^2} \cos 3x + \frac{1}{5^2} \cos 5x + \cdots + \frac{1}{(2n-1)^2} \cos(2n-1)x \right).$$

The graphs of f_1, f_2 and f_3 are shown in Figure 0.1 to attempt to illustrate what is happening. Each function f_n is infinitely differentiable but the limit function is only continuous with the

[1] J.B.J. Fourier (1768-1830) was a French mathematician that developed Fourier series in his analysis of solutions of the heat equation.

derivative not existing at $x = 0, \pm\pi, \pm 2\pi, \cdots, \pm k\pi, \cdots$. To attempt to understand this further observe that

$$f_n'(x) = \frac{4}{\pi}\left(\sin x + \frac{1}{3}\sin 3x + \frac{1}{5}\sin 5x + \cdots + \frac{1}{(2n-1)}\sin(2n-1)x\right).$$

The graphs of f_1', f_2' and f_4' are shown in Figure 0.2. It can be shown, although this will not be done in this module, that $(f_n'(x))$ does converge pointwise to the discontinuous 2π-periodic function g where

$$g : \mathbb{R} \to \mathbb{R}, \quad g(x) = \begin{cases} -1, & -\pi < x < 0, \\ 0, & x = 0, \\ 1, & 0 < x < \pi, \\ 0, & x = \pi. \end{cases}$$

The previous examples illustrates how sequences and series of smooth functions can have a less smooth limit. The converse can also occur and, possibly without appreciating it, you would have already have met this before. In numerical methods you are often concerned with approximating smooth (complicated) functions by putting together simpler functions. For example, if you approximate an infinitely differentiable function $f : [a, b] \to \mathbb{R}$ by a continuous piecewise linear function f_n involving n line segments then in general the derivative of f_n will not exist at the points at which the different line segments join. Thus although the slope of f_n is generally discontinuous the sequence of approximations (f_n) can still converge to f if they are constructed in the right way. We illustrate this below when (f_n) is a sequence of linear interpolants to $\sin x$ in $[0, 2\pi]$ based on equally spaced points and show in Figure 0.3 and 0.4 the graphs of f_4 and f_8.

0.2 Function Spaces and Differential Equations

A second aspect of this module is to extend ideas we developed in Linear Algebra at Level 1 to spaces of functions. One fundamental use is to be able to solve differential equations, ordinary and partial. First we shall look at some results from Level 1 about vectors and matrices we are going to generalise. In a second part we give you a string of problems that can be tackled with the ideas we are going to develop in this module.

0.2.1 Recall: symmetric matrices

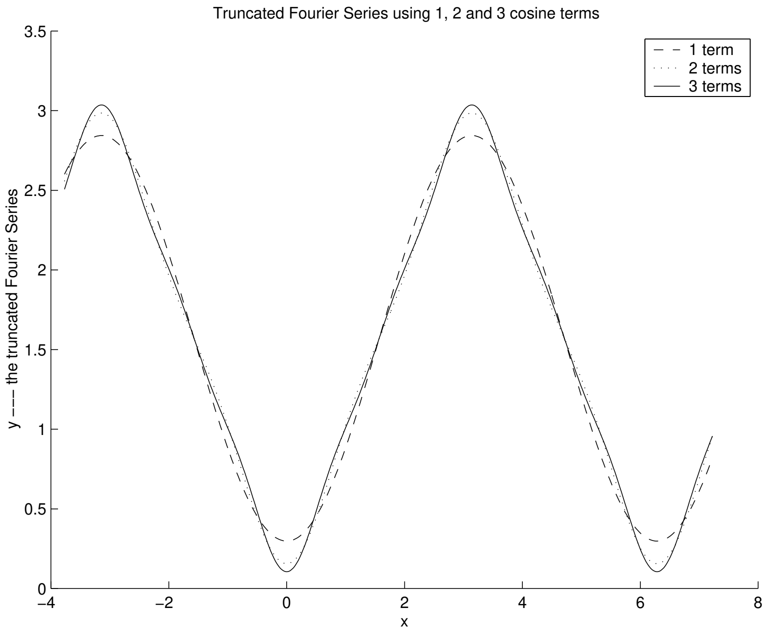
Let $\underline{v}, \underline{w} \in \mathbb{R}^n$ ($\underline{v} = (u_1 \ldots u_n)$, $\underline{w} = (w_1 \ldots w_n)$), we denote by $(\underline{v}, \underline{w})$ the dot product of the two vectors

$$(\underline{v}, \underline{w}) = \sum_{i=1}^{n} v_i w_i.$$

Let A be a real $n \times n$ *symmetric matrix*, that is $A = A^T$. We know that

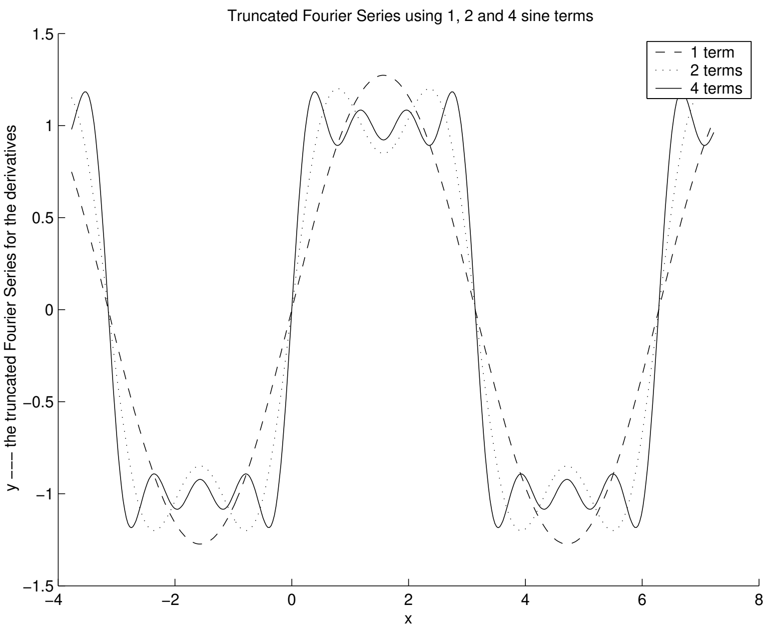
- all the *eigenvalues* of A ($A\underline{v} = \lambda\underline{v}$, $\underline{v} \neq \underline{0}$) are real numbers,
- there are n *eigenvectors* of A that are orthogonal to each other.

  More precisely, if $\{\underline{v}_i\}_{i=1}^n$ are the eigenvectors of A corresponding to the eigenvalues $\{\lambda_i\}_{i=1}^n \subset \mathbb{R}$ then $(\underline{v}_i, \underline{v}_j) = 0$ if $i \neq j$ (Recall that $(,)$ is the dot product). They are thus linearly independent and so form a *basis* of $\mathbb{R}^n$.



$$|x| = \frac{\pi}{2} - \frac{4}{\pi}\left(\cos x + \frac{1}{3^2}\cos 3x + \frac{1}{5^2}\cos 5x + \cdots + \frac{1}{(2k-1)^2}\cos(2k-1)x + \cdots\right).$$

Figure 0.1: Partial Fourier series sums for $f(x) := |x|$.



$$\frac{4}{\pi}\left(\sin x + \frac{1}{3}\sin 3x + \frac{1}{5}\sin 5x + \cdots + \frac{1}{(2n-1)}\sin(2n-1)x + \cdots\right).$$

Figure 0.2: Partial Fourier series sums for g.

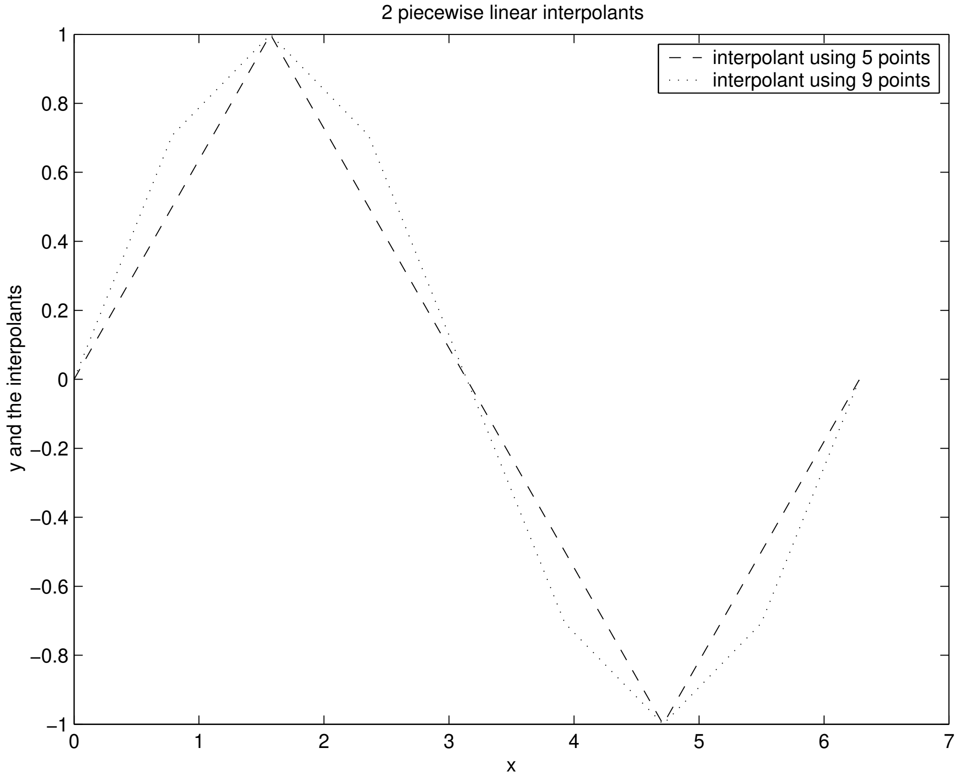


Figure 0.3: The piecewise linear functions using 5 and 9 points.

- the linear system of equation

$$A\underline{v} = \underline{b}$$

is easy to solve if we use the basis of eigenvectors of A.

Explicitly, suppose that the basis $\{\underline{v}_i\}_{i=1}^n$ of orthogonal eigenvectors has been chosen such that each vector is of length 1, $(\underline{v}_i, \underline{v}_i) = 1$ for all $1 \le i \le n$. Then

$$\begin{aligned} \underline{b} &= \sum_{i=1}^n (\underline{b}, \underline{v}_i)\, \underline{v}_i, \\ \underline{v} &= \sum_{i=1}^n (\lambda_i)^{-1} (\underline{b}, \underline{v}_i)\, \underline{v}_i, \end{aligned} \tag{1}$$

as it is readily verified by computing $A\underline{v}$.

Example

Let $A = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix}$. The eigenvalues of A are $\lambda_1 = 4$ and $\lambda_{2,3} = 1$ (twice, double eigenvalue). Corresponding orthogonal eigenvectors are $\underline{w}_1 = (1, 1, 1)$, $\underline{w}_2 = (1, -1, 0)$ and $\underline{w}_3 = (-1, -1, 2)$. We can normalise them and we get $\underline{v}_1 = \frac{1}{\sqrt{3}}(1, 1, 1)$, $\underline{v}_2 = \frac{1}{\sqrt{2}}(1, -1, 0)$ and $\underline{v}_3 = \frac{1}{\sqrt{6}}(-1, -1, 2)$. It is straightforward to check that they are orthogonal to each other and of length 1.

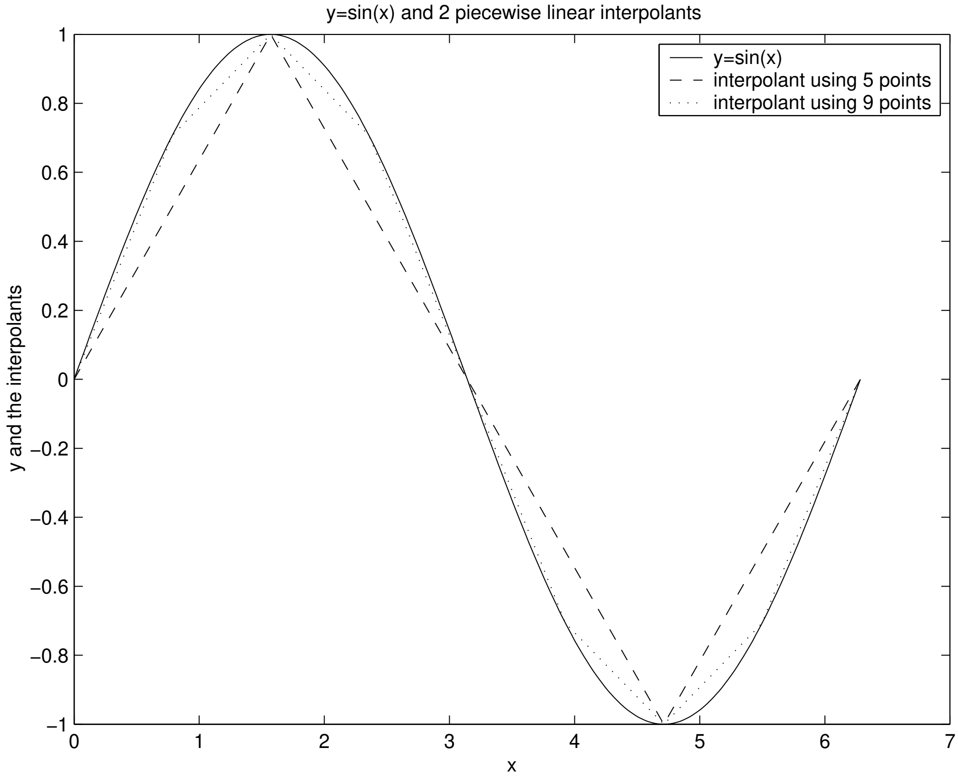


Figure 0.4: The piecewise linear interpolants and the function $\sin x$.

To solve $A\underline{v} = \underline{b} = (3, 4, 5)$, for instance, we only need to calculate

$$\begin{aligned}
\underline{v} &= (\lambda_1)^{-1}(\underline{b}, \underline{v}_1)\,\underline{v}_1 + (\lambda_2)^{-1}(\underline{b}, \underline{v}_2)\,\underline{v}_2 + (\lambda_3)^{-1}(\underline{b}, \underline{v}_3)\,\underline{v}_3 \\
&= \tfrac{1}{4}\tfrac{1}{3}\,(12)\,\underline{w}_1 + \tfrac{1}{2}\,(-1)\,\underline{w}_2 + \tfrac{1}{6}\,(3)\,\underline{w}_3 \\
&= \underline{w}_1 - \tfrac{1}{2}\,\underline{w}_2 + \tfrac{1}{2}\,\underline{w}_3 \\
&= (1, 1, 1) - \tfrac{1}{2}\,(1, -1, 0) + \tfrac{1}{2}\,(-1, -1, 2) \\
&= (0, 1, 2).
\end{aligned}$$

Note that (1) means that once we have the basis $\{\underline{v}_i\}_{i=1}^n$ we can solve $A\underline{v} = \underline{b}$ without using A^{-1} nor a Gaussian reduction.

0.2.2 Best approximation

Suppose that we have a line L in the plane and a point P situated outside L. Then there is a unique point on L which is at the shortest distance of P. It is given by the *foot* of the perpendicular to L from P. That foot of the perpendicular Q is called the **Best Approximation** to P in L.

This can be generalised for planes π (and subspaces in higher dimensions). Again, if we have a "nice" basis for π formed of orthogonal vectors of length 1, we can find a simple formula for the Best Approximation.

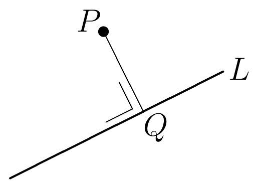


Figure 0.5: Best Approximation

Example

Let π be given by the plane $x + y + z = 0$. A basis for π is formed by the vectors $\underline{v}_1 = \frac{1}{\sqrt{2}}(1, 0, -1)$ and $\underline{v}_2 = \frac{1}{\sqrt{6}}(1, -2, 1)$. Let P be the point of coordinates $\underline{p} = (5, 2, 3)$, say, then the foot Q of the perpendicular to π through P is $\underline{OQ} = \underline{v}$ given by (cf. Chapter 2)

$$\underline{v} = (\underline{p}, \underline{v}_1)\, \underline{v}_1 + (\underline{p}, \underline{v}_2)\, \underline{v}_2 = (5/3, -4/3, -1/3).$$

0.2.3 Examples of linear problems

Although they appear in different parts of (real life) mathematics, the following problems have on thing in common:

THE USUAL WAYS TO HANDLE THEM USE WHAT IS COMMONLY REFERRED TO AS THE **LINEAR** *FRAMEWORKS AND TECHNIQUES*.

For instance,

(a) Solve $A\underline{x} = \underline{b}$ where A is a $n \times n$-matrix and $\underline{b} \in \mathbb{R}^n$.

(b) Suppose that we would like to determine experimentally the value of a constant c in a relation $y = cx$, where x, y are two measurable quantities. Usually we get a set of values $x_1 \ldots x_n$ for x and corresponding values $y_1 \ldots y_n$ for y, but the system $y_1 = cx_1, \ldots, y_n = cx_n$ cannot be solved for a single unknown c. So we abandon the idea to find an exact c and replace it with the idea to find the **best approximation** $\bar{c}$ which *minimises* the sum

$$\sum_{n=1}^{n} (cx_i - y_i)^2.$$

This method is called the method of **least squares** (cf. Statistics modules).

(c) Find the center of mass $\bar{x}$ of k particles, each of mass m_i, placed at a point $x_i \in \mathbb{R}^3$, $1 \le i \le k$. Let $m = \sum_{i=1}^{k} m_i$ be the total mass, the explicit formula is

$$m\bar{x} = \sum_{i=1}^{k} m_i x_i$$

(cf. Mechanics modules).

(d) Let $f : [0, 1] \to \mathbb{R}$ be some continuous function. Find a twice differentiable function (cf. Appendix A) $u : [0, 1] \to \mathbb{R}$ such that

$$\ddot{u}(t) + f(t)\, u(t) = 0$$

with the **boundary condition**

$$u(0) = u(1) = 0.$$

(e) Find a continuous function $u : \mathbb{R} \to \mathbb{R}$ such that

$$u(s) = 1 + \int_{-\infty}^{\infty} e^{-(t-s)^2} u(t)\, dt.$$

(f) Let $a < b$ and chose $n+1$ points $a \leq x_0 < x_1 \ldots < x_n \leq b$. Approximate a continuous function $f : [a, b] \to \mathbb{R}$ by a polynomial p_n of degree n such that

$$f(x_i) = p_n(x_i), \quad \text{for } 0 \leq i \leq n.$$

The x_i's are called **collocation points** (cf. MA1915).

(g) The equation of motion of a RCL-circuit is

$$L\ddot{q}(t) + R\dot{q}(t) + \frac{1}{C} q(t) = E \sin \omega t,$$

where the charge of the capacitor (depending on time) is $q(t)$, the current is $i(t) = \dot{q}(t)$ and the constants are the inductance L (coil of copper wire), the resistance R, the capacitance C and E is the voltage of an alternative source of current of frequency ω.

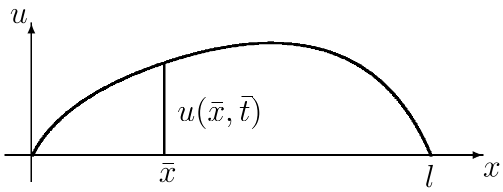


Figure 0.6: Profile of the elastic string at $t = \bar{t}$

(h) The simplest realistic equation of motion of an elastic string stretched along the x-axis from 0 to l, clamped at its end points $x = 0, l$, with initial position $u_o(x)$ and initial velocity $v(x)$ and then allowed to vibrate freely is the following **(hyperbolic) partial differential equation (PDE)** problem:

$$\frac{\partial^2 u}{\partial x^2}(x, t) = \frac{\rho}{K} \frac{\partial^2 u}{\partial t^2}(x, t)$$

with boundary conditions

$$u(0, t) = u(l, t) = 0$$

and **initial conditions**

$$\begin{aligned} u(x, 0) &= u_o(x), \\ \frac{\partial u}{\partial t}(x, 0) &= v(x), \end{aligned}$$

where $u(x, t)$ is the deformation of the string at $x \in [0, l]$ and time t, K is the horizontal tension and ρ the density of the string (both assumed to be constant).

(i) The equation governing the temperature distribution u in a slender rod of length L in which the rod is insulated excepted at 0 where its temperature losses are linearly dependent of its magnitude (of constant h) and the heat is being generated at a constant rate k per unit length, is another PDE problem, **parabolic** this time:

$$a^2 \frac{\partial u}{\partial t}(x,t) = \frac{\partial^2 u}{\partial x^2}(x,t) + k$$

with boundary conditions

$$\begin{aligned} u(L,t) &= 0, \\ \frac{\partial u}{\partial x}(0,t) &= -hu(0,t) \end{aligned}$$

and initial conditions

$$u(x,0) = u_o(x),$$

where u_o is the initial temperature distribution and a^2 is a constant depending on the density of the rod and the specific heat of its constitutive material.

(j) The electrical field V generated by a distribution of charges is described by Laplace's[2] equation

$$\Delta V = 0,$$

where

$$\Delta V(x,y,z) = \frac{\partial^2 V}{\partial x^2}(x,y,z) + \frac{\partial^2 V}{\partial y^2}(x,y,z) + \frac{\partial^2 V}{\partial z^2}(x,y,z)$$

is the **Laplacian operator**.

(k) The distribution $u(t,a)$ of an age-structured population (where t is time and a the different ages) satisfies

$$\frac{\partial u}{\partial t}(t,a) + \frac{\partial u}{\partial a}(t,a) = \mu(a)\, u(t,a)$$

where μ is the mortality at age a.

This is a **first order PDE**. The previous equations were second order (cf. the partial derivatives involved).

(l) The equation governing the density P of a population of plankton living in a domain $\Omega \subset \mathbb{R}^3$ of boundary $\partial\Omega$ is

$$\frac{\partial P}{\partial t}(x,t) + v_o \frac{\partial P}{\partial x}(x,t) = d^2 \frac{\partial^2 P}{\partial x^2}(x,t) + kP(x,t),$$

with initial conditions

$$P(x,0) = P_o(x)$$

and boundary conditions

$$P(x,t) = 0, \quad \forall x \in \partial\Omega,$$

where v_o is the velocity of the water current, d the diffusion coefficient of the plankton and k the growth coefficient of the plankton population.

The boundary condition means that the exterior of Ω is lethal for the species. This equation is again of second order but it has a **viscosity** term $\frac{\partial P}{\partial x}$.

[2] P.S. de Laplace (1749-1827) was a French mathematician that developed the transform carrying his name.

(m) The following problem involving partial derivatives of *fourth* order

$$\Delta^2 u(x,y) + \lambda \frac{\partial^2 u}{\partial x^2}(x,y) = 0$$

with boundary conditions

$$\begin{aligned} u(0,y) &= \frac{\partial^2 u}{\partial x^2}(0,y) = 0, \\ u(l_1,y) &= \frac{\partial^2 u}{\partial x^2}(l_1,y) = 0, \\ \frac{\partial u}{\partial y}(x,0) &= \frac{\partial^3 u}{\partial y^3}(x,0) = 0, \\ \frac{\partial u}{\partial y}(x,l_2) &= \frac{\partial^3 u}{\partial y^3}(x,l_2) = 0, \end{aligned}$$

where $\Delta u(x,y) = \frac{\partial^2 u}{\partial x^2}(x,y) + \frac{\partial^2 u}{\partial y^2}(x,y)$ is again the Laplacian operator, represents the behaviour of a concrete rectangular slab clamped on a pair of opposite edges at $x = 0, l_1$, simply supported along the other pair at $y = 0, l_2$, and subjected to a pressure λ on the clamped ends.

In this module we are not going to solve many of such problems, actually the only thorough treatment will be in the last assignment where we are going to deal with the heat equation on the sphere. But we are going to establish the bases of those linear techniques in some manageable generality. We refer to other modules dealing with Linear Algebra, Approximation/Splines, Waves, Electromagnetism, Elasticity, Integral Equations, Partial Differential Equations, Functional Equations etc. for a more thorough study of the previous equations.

Chapter 1

Techniques of Analysis

1.1 Preliminaries and Discussion

In this preliminary section we will introduce some of the notation to be used throughout the module together with some assumed knowledge about the real number system. This includes some results about real numbers that you may classify as 'elementary' and includes a discussion of the consequences of assuming that an interval is closed (as opposed to being open) in several different situations.

1.1.1 The number systems and some examples of set notation

We first define the natural numbers, the integers, the rational numbers and we then give our notation for real and complex numbers.

The set of natural numbers $\mathbb{N} := \{1, 2, 3, \cdots\}$.

The set of integers $\mathbb{Z} := \{0, \pm 1, \pm 2, \pm 3, \cdots\}$.

The set of rationals $\mathbb{Q} := \{m/n : \ m, n \in \mathbb{Z} \text{ and } n \neq 0\}$.

The set of real numbers is denoted by $\mathbb{R}$.

The set of complex numbers is $\mathbb{C} := \{x + iy : \ x, y \in \mathbb{R}\}$.

In the above the symbols := means that the left hand side is defined by the expression on the right. Also in the above, the sets have been defined by either listing the entries between the {} symbols or by expressing the set in the form

$$\{x : \ P(x)\}$$

where x is a typical element and $P(x)$ indicates the properties that it must satisfy for it to be in the set. When the universal set that x is from is not clear from the context or is not part of properties $P(x)$ then we will write instead

$$\{x \in S : \ P(x)\}$$

where $x \in S$ means that x *belongs to* the set S, i.e. x is a member or element of the set S. However, when the universal set is clear from the context (e.g. we are dealing with real numbers) then we will use the shorter notation $\{x : P(x)\}$ with $P(x)$ containing no statement about the type of the entry x as in a specification such as

$$\{x : 0 \le x < 1\}.$$

With the number systems introduced we now consider some set related notation and jargon. We have

$$\mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R} \subset \mathbb{C}$$

where $\subset$ means that the set to the left of the symbol is a *subset* of the set given on the right of the symbol. We also say that the set to the left of the symbol $\subset$ is *contained in* the set to the right of the symbol. In the above case the subsets are examples of *proper subsets* as there are integers which are not natural numbers (e.g. $0, -1, -2, \cdots$), there are rational numbers which are not integers (e.g. $1/2, 1/3, \cdots$) etc.. If A and B are sets then their **intersection** is

$$A \cap B := \{x : x \in A \text{ and } x \in B\}$$

and their **union** is

$$A \cup B := \{x : x \in A \text{ or } x \in B\}.$$

The mathematical meaning of the word 'or' here is what in the English language is referred to as the inclusive sense of the word. That is we mean that x belongs to A or x belongs to B or possibly x belongs to both A and B. In the English language this is sometimes written as x belongs to A and/or B.

The intersection and union notation can also be used when we have more than 2 sets. For example, if $A_1, \cdots, A_n$ are sets then we can define sets A and B by

$$\begin{aligned} A &:= \bigcup_{j=1}^{n} A_j = A_1 \cup A_2 \cup \cdots \cup A_n = \{x : x \in A_j \text{ for some } j,\ 1 \le j \le n\}. \\ B &:= \bigcap_{j=1}^{n} A_j = A_1 \cap A_2 \cap \cdots \cap A_n = \{x : x \in A_j \text{ for all } j,\ 1 \le j \le n\}. \end{aligned}$$

Indeed we can also use this notation when we have infinitely many sets (see later in the module when the Nested Interval Theorem is described).

Another common set operation is the **complement** operation which will be denoted by $\backslash$. This is used in the form

$$A \backslash B := \{x \in A : x \notin B\}$$

with the symbols $x \notin B$ meaning that x does not belong to B. Thus, for example, with this notation the irrational numbers are described by $\mathbb{R} \backslash \mathbb{Q}$. Finally, in this brief review of set terminology and notation, we use the symbol $\emptyset$ to denote the null set which is a set with no elements. Thus, for example, the set $\mathbb{Q} \backslash \mathbb{R} = \emptyset$ since $\mathbb{Q} \subset \mathbb{R}$.

Appendix: cartesian product

Let A and B be two sets, we can create a third set $A \times B$, the **cartesian product** of A and B. The cartesian product of A and B is the set of *ordered couples* (a, b) where $a \in A$ and $b \in B$. In formal notation

$$A \times B = \{\ (a, b)\ :\ a \in A,\ b \in B\ \}.$$

If $B = A$, we denote $A \times A$ by A^2, like with $\mathbb{R}^2$, $\mathbb{C}^2$, or more generally $\mathbb{R}^n$.

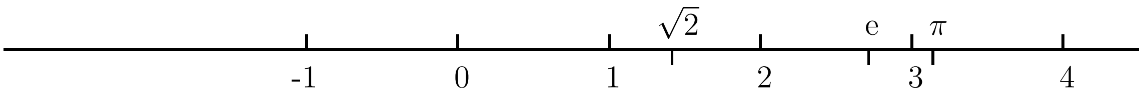
The cartesian product is often used to define *domains* of functions with several variables, like in the following examples:

- $f : \mathbb{R}^2 \to \mathbb{R}$ defined by $f(x, y) = \sin(x + y)$,
- it is also possible that the two variables do *not* belong to the same set, for instance (if $\mathbb{R}^*$ denotes $\mathbb{R} - \{0\}$) the formula $f(x, y) = \frac{x}{y}$ defines a function $f : \mathbb{R}\times\mathbb{R}^* \to \mathbb{R}$.

But the formula $f(x, y) = \ln(x^2+y^2-1)$ defines a function on the set $\{ (x, y) \ : \ x^2+y^2 > 1 \}$ that cannot be written as a cartesian product.

1.1.2 A characterisation of the real numbers $\mathbb{R}$

When $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$ and $\mathbb{C}$ were introduced you may have noted that $\mathbb{N}$, $\mathbb{Z}$ and $\mathbb{Q}$ were defined and that the complex numbers $\mathbb{C}$ was defined in terms of $\mathbb{R}$ assuming the existence of the imaginary unit $i = \sqrt{-1}$. However, we did not explicitly define what is meant by the real numbers which is in fact not a trivial matter. It is convenient to think of the real numbers as the points on a line as in the following figure.



We also have the concept of *order* with $x < y$ meaning that x is to the left of y on such a figure. We will assume the concept of order throughout this module (see the appendix thereafter for some more information on order relations). However we need something more precise in order to proceed in analysis when we cover topics such as limits. The characteristic property that we will use here is the so-called *completeness property* which is also sometimes known as the *supremum property* of $\mathbb{R}$. This is the property that we will take in this module as being an Axiom of the real numbers, that is it is a result which we need not prove. To state the property we need to first define what is meant by *bounded above* and we need to define what is meant by the *least upper bound* (also known as *supremum*).

Definition 1.1 *A set $S \subset \mathbb{R}$ is* bounded above *if it has an* upper bound*, i.e. if there exists $M \in \mathbb{R}$ such that*

$$x \le M \quad \textit{for all } x \in S \ .$$

Definition 1.2 *If S is bounded above then $u \in \mathbb{R}$ is a* least upper bound *or* supremum *of S if it is an upper bound of S and no number smaller than u is an upper bound. When we have a least upper bound we write*

$$u = \sup S.$$

Axiom 1.3 *(***The completeness property of** $\mathbb{R}$*) Every subset of $\mathbb{R}$ that is bounded above has a least upper bound in $\mathbb{R}$.*

This property thus guarantees the existence of a real number $\sup S$ which may or may not be a member of the set S. Also, it is a property which is not possessed by the rationals $\mathbb{Q}$. To illustrate these statements, note that if $S := (0, 1) := \{x : \ 0 < x < 1\}$ (an *open* interval) then $\sup S = 1$ which is not an element of S. If instead we take $S := [0, 1] := \{x : \ 0 \le x \le 1\}$ (a *closed* interval) then $\sup S = 1$ which is a member of S. In this case the set has a *maximum* element as well as a supremum and we have $\max S = \sup S = 1$. (We discuss other differences which result from having a closed interval instead of an open interval below.) If now we take

$$S := \{x \in \mathbb{Q} : \ x^2 < 2\}$$

then S is bounded above in $\mathbb{Q}$ (for example $3/2$ is an upper bound) but there is no least upper bound in $\mathbb{Q}$ as there is no rational number equal to $\sqrt{2}$.

In logical arguments that will be part of this module it is important to understand precisely what the term 'least upper bound' means. As has already been stated, it is an upper bound and nothing smaller is an upper bound. That is if $u = \sup S$ then $u - \epsilon$ is not an upper bound for any $\epsilon > 0$. Another way of saying this is that if we take any $\epsilon > 0$ then $u - \epsilon$ is not an upper bound and hence it is exceeded by at least one member of the set S. That is there exists $x \in S$ with $u - \epsilon < x \le u$. Thus although $\sup S$ may not be a member of S its existence is guaranteed when S is bounded above and we can get arbitrarily close to it within S.

The completeness property was expressed in terms of upper bounds and least upper bounds. A similar property holds if we had considered lower bounds and greatest lower bounds which are defined as follows.

Definition 1.4 *A set S is* bounded below *if there exists $L \in \mathbb{R}$ such that*

$$x \ge L \text{ for all } x \in S.$$

Definition 1.5 *If S is bounded below then $l \in \mathbb{R}$ is a* greatest lower bound *or* infimum *of S if it is a lower bound of S and no number greater than l is a lower bound. We write $l = \inf S$.*

As in the above, if the value $l \in S$ then the set S has a minimum and $l = \min S = \inf S$.

Appendix: order and equivalence relations

Relations between pairs of elements of a given set are important in Mathematics. In general a **relation** R on a set A in a subset of the cartesian product $R \subset A \times A$. Because we often have an explicit sentence linking x and y when they are related, if $(x, y) \in R$ we also denote it by xRy and we say that x is **related to** y. For instance, take $A = \mathbb{N}$ and xRy defined as x divides y. Clearly $(2, 4) \in R$ but $(4, 2) \notin R$, nor $(2, 3)$.

Obviously any subset can do, hence there are many possible relations on a given set (if A is finite, how many?). But two classes of relations are very important: the *equivalence* and *order* relations. Before we define them, we define properties that can hold for a given relation.

Definition 1.6

1. *A relation R is* reflexive *if xRx for all $x \in A$.*
2. *A relation R is* symmetric *if xRy implies that yRx also holds.*
3. *A relation R is* anti-symmetric *if when both xRy and yRx hold then $x = y$.*
4. *A relation R is* transitive *if every time xRy and yRz hold implies that xRz also holds.*

Note that $<$ is not reflexive because $a \not< a$ but $\le$ is because $a \le a$. Being a divisor is not symmetric, $(2, 4) \in R$ but $(4, 2) \notin R$. But having a common divisor is a symmetric relation. Being a divisor is transitive, but having a common divisor is not: $(4, 6)$ and $(6, 9)$ have common divisors, but $(4, 9)$ have not.

We can therefore define the two more important classes of relations.

Definition 1.7 *A relation R is a* (partial) order *if it is reflexive, antisymmetric and transitive.*

A relation R is an equivalence *relation if it is reflexive, symmetric and transitive.*

It is easy to check that $\leq$ is an order relation. Indeed it is our prototype for such relation. For equivalence relations, obviously $=$ is one but it is very restrictive. A more interesting one is on a set of sets, ARB if A has the same number of elements as B. Or, in $\mathbb{R}$, xRy if $[x] = [y]$ (if they having the same integer part $[x]$, the largest integer less or equal to x). Both are equivalence relation between objects that are not identical but have a common property.

1.1.3 Open and closed intervals

In the discussion following the completeness property of $\mathbb{R}$ which involved the supremum operation we illustrated the difference between the supremum and the maximum by considering bounded sets which were respectively open and closed. We consider here other differences which will be encountered during the module. First however we begin with the notation and the definition.

If $a, b \in \mathbb{R}$ with $a \leq b$ then we define

$$\begin{aligned} (a,b) &:= \{x \in \mathbb{R} : \; a < x < b\}, \quad \text{an open interval,} \\ [a,b] &:= \{x \in \mathbb{R} : \; a \leq x \leq b\}, \quad \text{a closed interval.} \end{aligned}$$

Thus the difference between an open and a closed interval is that an open interval does not include its two end points but a closed interval does. Hence the use of round or square brackets indicates whether or not the end value is included and using this convention we define

$$\begin{aligned} (a,b] &:= \{x \in \mathbb{R} : \; a < x \leq b\}, \\ [a,b) &:= \{x \in \mathbb{R} : \; a \leq x < b\}, \end{aligned}$$

to denote the intervals where exactly one of the end points is included. In these last two cases $(a, b]$ and $[a, b)$ are neither open or closed.

When open and closed intervals are first met it is easy to think that the mere inclusion or not of the end points is a minor technical point of little significance. Indeed, this is the case in many situations. For example, when considering the open interval $(0, 1)$ it was noted that it had a supremum but no maximum. The closed interval $[0, 1]$ has a maximum entry (which is of course also the supremum). There is a difference here but it is not very important in this case. Also, if you integrate a bounded function f from a to b then it makes no difference to the answer whether from 'a to b' means that the range of the integration is (a, b) or $[a, b]$. However, there are many other situations which we will meet in this module where the difference is significant. In the context of convergence there is a result which states that if a convergent sequence (x_n) is contained in $[a, b]$ then the limit $x \in [a, b]$. In fact we can have (x_n) contained in the open interval (a, b) but it may be the case that $x = a$ or $x = b$, i.e. the limit is at one of the end points. We can illustrate this with the simple example

$$x_n := \frac{1}{n} \in (0,2) \quad \text{for } n = 1, 2, \cdots$$

but

$$\lim_{n\to\infty} x_n = 0 \notin (0,2) \ .$$

Another situation where closure is significant is when we consider continuous functions $f : [a, b] \to \mathbb{R}$ where here we are assuming that f is continuous up to the end points $x = a$ and $x = b$. The inclusion of the end points enables us to deduce that the function is bounded and that the maximum and minimum values are attained at some point in $[a, b]$. That is, it can be shown that there exists $c, d \in [a, b]$ with

$$f(c) = \min\{f(x) : \ a \leq x \leq b\} \quad \text{and} \quad f(d) = \max\{f(x) : \ a \leq x \leq b\}.$$

If f is also differentiable then you will have probably already have done exercises involving determining such points by considering the function values $f(a)$ and $f(b)$ and by considering any intermediate points at which

$$f'(c) = 0$$

which characterises a point c at which we have a local maximum or minimum. When we only know that f is continuous on an open interval then it is no longer the case that f is necessarily bounded on the interval. For example, consider

$$f : (0, \infty) :\to \mathbb{R}, \quad f(x) = 1/x \ .$$

1.1.4 A consequences of the completeness property of $\mathbb{R}$

In the last subsection we reviewed some results which will be met later in the module which related to whether an interval is open or closed. We now return to some more ‘elementary’ results which are immediate consequences of the completeness property of $\mathbb{R}$. The results perhaps seem obvious though we will give proofs in each case to illustrate the sort of logical reasoning that will be part of this module. It is perhaps worth noting here that so-called ‘obvious’ results can be among the more difficult results to prove!

First a result that the set of natural numbers $\mathbb{N}$ is not bounded. In this we do need a characterising property of the natural numbers which is that if $m \in \mathbb{N}$ then $m + 1 \in \mathbb{N}$.

Theorem 1.8 (Archimedean Property) *If $x \in \mathbb{R}$ then there exists $n_x \in \mathbb{N}$ such that $x < n_x$.*

Proof. The method of the proof is proof by contradiction and thus we start with the superposition that the conclusion is false. That is we are assuming that x is not exceeded by any natural number or, in other words, that the set $\mathbb{N}$ is bounded above. But as $\mathbb{N} \subset \mathbb{R}$ the set must have a least upper bound $u \in \mathbb{R}$. u being a least upper bound implies that nothing smaller is an upper bound and thus the smaller number $u - 1$ is exceeded by at least one member $m \in \mathbb{N}$. Hence

$$u - 1 < m \quad \text{i.e. } u < m + 1.$$

But $m \in \mathbb{N}$ implies that $m + 1 \in \mathbb{N}$ and $m + 1$ exceeds the least upper bound u.This is our contradiction. □

A consequence of the Archimedean[1] property is that $1/n$ gets smaller and smaller the larger we make n. We make this precise in the following corollary.

[1] Archimedes (287BC-212BC) was a Greek mathematician and physicist who lived in the Greek founded city of Syracuse in Sicily. His work in physics and mechanics ranged from the theory (for instance on centre of gravity or hydrostatics with the Archimedes’ Principle for immersed bodies) to practical machines, even military one’s as it city was fighting off the Romans during the Second Punic War. He was actually eventually killed by Roman troops in unclear circumstances. But his main intellectual interest was in Mathematics, where he developed an early version of integration, wrote important treatises on spirals, geometry, number theory. What was particularly remarkable in his work is the way he used methods from one area to another at a time when mathematical theorisation was in its infancy.

Corollary 1.9 *Let $y \in \mathbb{R}$, $y > 0$. Then there exists $n \in \mathbb{N}$ with $0 < 1/n < y$.*

Proof. Take $x := 1/y$. This will be 'large' if y is 'small'. By the Archimedean principle it can be exceeded by some natural number n, i.e.

$$n > x = \frac{1}{y}, \quad \text{and thus } 0 < \frac{1}{n} < y. \qquad \square$$

Further results relating to the completeness property of $\mathbb{R}$ are considered later in this module.

1.1.5 The triangle inequality

As has already been stated, the concept of order in the case of real numbers is assumed throughout this module and this was required for us to be able to define an upper bound of a set of numbers. With a concept of order we have a concept of magnitude which is needed when we discuss convergence which is concerned with things getting closer and closer together (this will be defined precisely in the next chapter). In this module there are several occasions when we are not just concerned with the size of $x \in \mathbb{R}$ but we are concerned with the size of the combination of terms such as in expressions like $x + y$. The key result that we will use in dealing with this is the triangle inequality which you should have met before. For all $x, y \in \mathbb{R}$ we have

$$|x + y| \le |x| + |y| \ .$$

The result is also true in the case of complex numbers.

1.1.6 Some remarks about other properties of the number systems

In this section we have introduced the familiar number systems $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$ and $\mathbb{C}$, we have noted that $\mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R} \subset \mathbb{C}$ and we have characterised the real number system by the completion property which states that any set in $\mathbb{R}$ which is bounded above has a least upper bound or supremum. Given this property we have shown that as a consequence the natural numbers $\mathbb{N}$ are unbounded. These are sufficient for our subsequent study of sequences, limits, continuity and sequences of continuous functions which is central to this module.

There are a number of other results about the number systems that you will find in books on analysis which we will now briefly mention without any proofs. A simple first observation is that all the number systems have an infinite number of elements. We can list the entries of $\mathbb{N}$ and $\mathbb{Z}$ which technically means that they are *countable*. We can also list the rationals $\mathbb{Q}$ which might seem slightly surprising when you first meet this. If we restrict ourselves to the interval $(0, 1)$ then this can simply be done in the order $\{1/2, 1/3, 2/3, 1/4, 2/4, 3/4, 1/5, \cdots\}$, i.e. we consider all rationals with denominator 2, then 3, then 4 and so on. In this sense the rationals $\mathbb{Q}$ can be thought of as 'not such a large set' and we have already indicated that it fails to satisfy the completeness property. In another sense however $\mathbb{Q}$ can be considered as 'quite a large set'. Specifically, $\mathbb{Q}$ has the property that it is *dense* in $\mathbb{R}$. We will not give a technical meaning of this but it is sufficient in these comments to note that it implies that for all $x \in \mathbb{R}$ we can find a rational number $r \in \mathbb{Q}$ as close as we like to x. For example, if we truncate a decimal representation of x then we generate a rational approximation to x and the more terms that we take before truncating the closer is our rational approximation to x. The rationals are hence both countable and dense subset of $\mathbb{R}$. The question then is how much 'larger' is the set $\mathbb{R}$? The short answer to this is that in many senses it is 'much larger'. We have effectively stated in the completeness property that '$\mathbb{R}$ completes $\mathbb{Q}$' and it can be shown that $\mathbb{R}$ is not countable and this in turn implies that the set of irrational numbers $\mathbb{R}\backslash\mathbb{Q}$ is also not countable. Thus in this sense there are many more irrational numbers than there are rational numbers. To summarize these statements, we have that all the number systems are infinite with the reals $\mathbb{R}$ 'more infinite' than $\mathbb{N}$, $\mathbb{Z}$ and $\mathbb{Q}$.

In this module the complex numbers $\mathbb{C}$ and functions of a complex variable will not be considered further although many of the results of this module concerning sequences and limits, which will be done in $\mathbb{R}$, can often be done, possibly with minor modification, in $\mathbb{C}$. Complex numbers and functions of a complex variable are the topic of a module at level 3. However, a result about $\mathbb{C}$ worthy of noting here which shows 'how $\mathbb{C}$ extends $\mathbb{R}$' is known as the Fundamental Theorem of Algebra. One way of expressing this result is that any polynomial $p_n(z)$ of degree n can always be factored in the form

$$p_n(z) = C(z - z_1)(z - z_2)\cdots(z - z_n)$$

for some complex numbers C, $z_1, \cdots, z_n$, i.e. a polynomial of degree n always has n roots in the complex plane. Thus you can think of $\mathbb{R}$ as extending $\mathbb{Q}$ by including all the irrational numbers $(\sqrt{2}, \mathrm{e}, \pi, \cdots)$ and you can think of $\mathbb{C}$ as extending $\mathbb{R}$ by including all roots of polynomials. This is not the only reason for studying complex numbers or functions of a complex variable. The study of functions of a complex variable brings to light connections between some of the common functions of analysis in, for example, relations of the form

$$\exp iz = \cos z + i \sin z.$$

Note also

$$\mathrm{e}^{i\pi} = -1$$

which relates the real numbers e and π and the imaginary number i. Finally, power series such as

$$\sum_{n=1}^{\infty} a_n z^n$$

are best done in the complex number field where a radius of convergence means the radius of a disk of the form $\{z \in \mathbb{C} : \ |z| < R\}$ in the complex plane.

1.1.7 Summary

1. A set $S \subset \mathbb{R}$ is *bounded above* if it has an *upper bound*, i.e. if there exists $M \in \mathbb{R}$ such that

$$x \le M \quad \text{for all } x \in S \ .$$

If S is bounded above then $u \in \mathbb{R}$ is a *least upper bound* or *supremum* of S if it is an upper bound of S and no number smaller than u is an upper bound. When we have a least upper bound we write

$$u = \sup S.$$

A set S is *bounded below* if there exists $L \in \mathbb{R}$ such that

$$x \ge L \text{ for all } x \in S.$$

If S is bounded below then $l \in \mathbb{R}$ is a *greatest lower bound* or *infimum* of S if it is a lower bound of S and no number greater than l is a lower bound. We write $l = \inf S$.

2. (**The completeness property of** $\mathbb{R}$) Every subset of $\mathbb{R}$ that is bounded above has a least upper bound in $\mathbb{R}$.

3. (**Archimedean Property**) If $x \in \mathbb{R}$ then there exists $n_x \in \mathbb{N}$ such that $x < n_x$.

4. Let $y \in \mathbb{R}$, $y > 0$. Then there exists $n \in \mathbb{N}$ with $0 < 1/n < y$.

5. The triangle inequality: for all $x, y \in \mathbb{R}$ we have

$$|x + y| \leq |x| + |y| \ .$$

1.2 Sequences and Series in $\mathbb{R}$

A sequence in $\mathbb{R}$ is a function $x : \mathbb{N} \to \mathbb{R}$, i.e. for each natural number $n \in \mathbb{N}$ we get a real number $x(n)$ although in the context of sequences it is more common to denote the value at n by x_n. x_n is also referred to as the n**th term** or the n**th element**. The sequence of values $x_1, x_2, x_3, \cdots$ will usually be written as

$$(x_n) \quad \text{or} \quad (x_n : \ n \in \mathbb{N}) \quad \text{or} \quad (x_1, x_2, x_3, \cdots)$$

where, as in some texts, we use round brackets () to distinguish the sequence (x_n) from the set of values notation $\{x_n\}$ which assumes no specific ordering of the terms. That is, a set of real numbers is just a collection of values whereas a sequence is a collection of values in a specific order and our main interest is usually whether or not it converges which we next define.

1.2.1 The convergence of sequences

Definition 1.10 *The sequence (x_n)* converges *to x if for every $\epsilon > 0$ there exists a $N = N(\epsilon)$ such that*

$$|x_n - x| < \epsilon$$

for all $n \geq N$.

When a sequence (x_n) converges to x we write

$$\lim_{n\to\infty} x_n = x \quad \text{or} \quad x_n \to x \text{ as } n \to \infty.$$

In the above the interval $(x - \epsilon, x + \epsilon)$ is known as the ϵ**–neighbourhood** of x.

Definition 1.11 *A sequence which does not converge is said to* diverge*.*

It is important to have a good grasp of the above definition of convergence. It implies that however small we make the neighbourhood around x, all but a finite number of terms of the sequence will lie outside of the neighbourhood. To state this another way, the "tail of the sequence" will lie inside the neighbourhood. The smaller we make ϵ the larger $N = N(\epsilon)$ has to be but when we have convergence there is always a value $N(\epsilon)$ for every $\epsilon > 0$. The key phrase in the definition is the "for every $\epsilon > 0$" indicating that we can make the neighbourhood as small as we like. You should also note that the $N = N(\epsilon)$ term is not unique since if we have an $N(\epsilon)$ corresponding to a given $\epsilon > 0$ then any natural number larger that this N will also suffice to satisfy the definition.

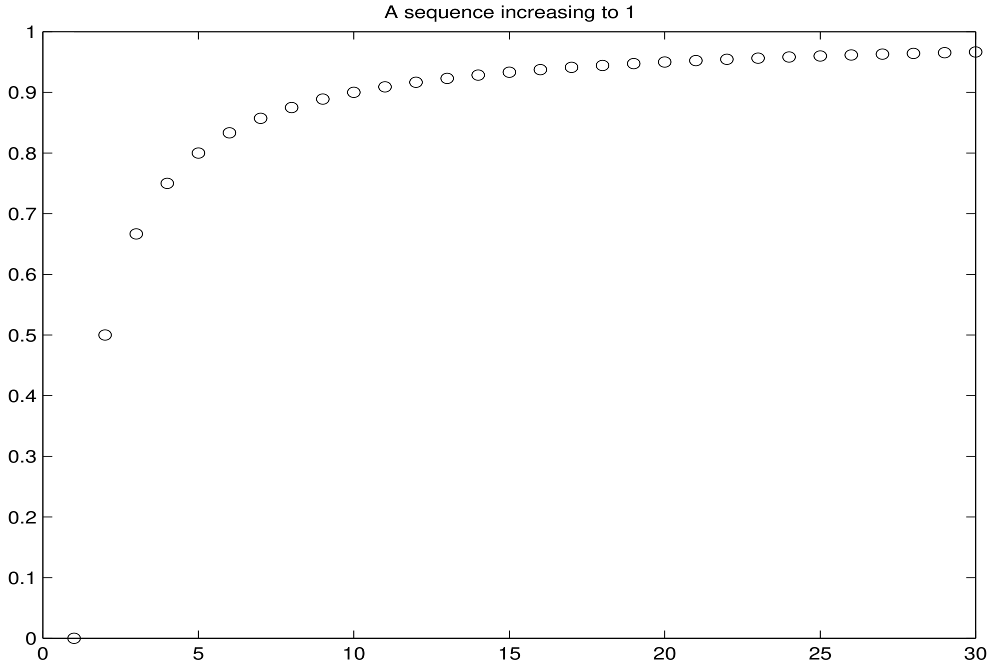


Figure 1.1: The first few values of a sequence increasing to 1

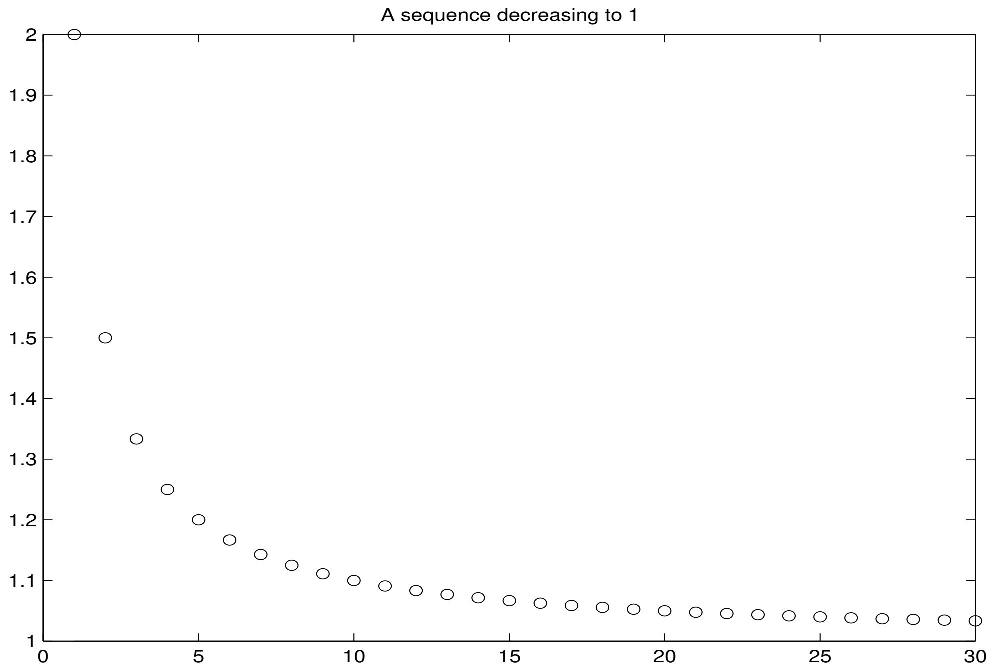


Figure 1.2: The first few values of a sequence decreasing to 1

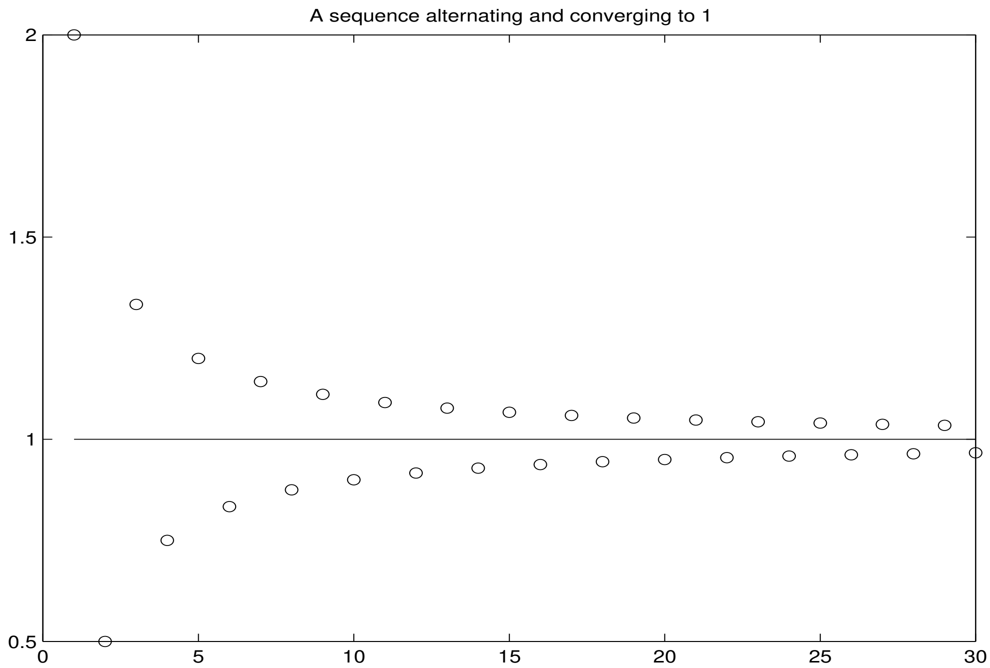


Figure 1.3: The first few values of a sequence alternating about the limit

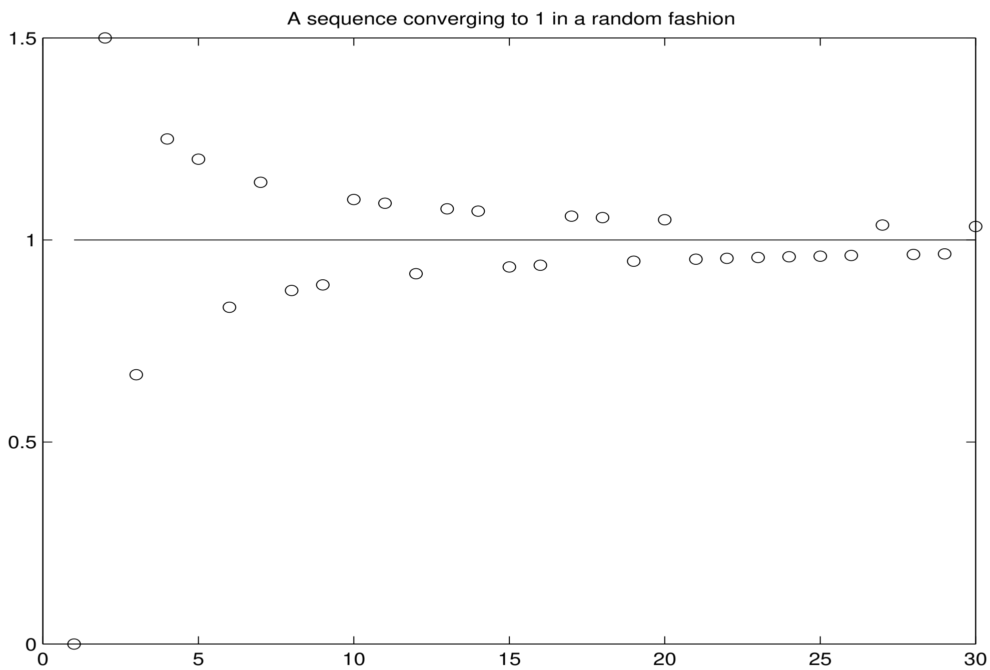


Figure 1.4: The first few values of a sequence converging in a random fashion to the limiting value

We show in Figures 1.1, 1.2, 1.3 and 1.4 what convergence means graphically in situations where the sequence increases to its limit, decreases to its limit, alternates about the limit and oscillates in no organised way about the limiting value.

Sequences were introduced and certain aspects covered in MA1915 (Calculus and Numerical Methods) and thus some of what follows will be a revision of that material with most results just stated and not formally proved. When no formal proof is given we shall instead make comments concerning their validity. The first thing to note is that the limit of a convergent sequence is unique.

Theorem 1.12 (Limit is unique) *If the sequence (x_n) converges then its limit is unique.*

PROOF. We use a proof by contradiction and thus we start by assuming that the converse is true. That is we assume that $x_n \to x$ and $x_n \to y$ with $x \neq y$. That is we have $|x-y| > 0$. Now, by the triangle inequality we have

$$|x-y| = |(x-x_n)+(x_n-y)| \le |x_n-x| + |x_n-y| \ .$$

We are given that $x_n \to x$ and $x_n \to y$ implies that for every $\epsilon > 0$ there exists N_x and N_y such that

$$\begin{aligned} |x_n - x| &< \epsilon \quad \text{for all } n \ge N_x, \\ |y - x_n| &< \epsilon \quad \text{for all } n \ge N_y. \end{aligned}$$

We want both of these to be true and thus take $n \ge \max\{N_x, N_y\}$ and we get

$$0 < |x-y| \le 2\epsilon \ .$$

This is true for every $\epsilon > 0$ and hence it is true for $\epsilon = |x-y|/4$ which immediately gives us a contradiction since we cannot have

$$0 < |x-y| \le \frac{2}{4}|x-y| \ . \qquad \square$$

Note the use of the triangle inequality in the above. This will be used many times in this module.

As has already been said, convergence implies that all but a finite number of terms of the sequence lie outside of a neighbourhood of this limit. A consequence of this is that the sequence is bounded which is formally defined as follows.

Definition 1.13 *A sequence (x_n) is* bounded *if there exists $M > 0$ such that $|x_n| \le M$ for all $n \in \mathbb{N}$.*

Note here that bounded means bounded above and bounded below. Also, note that by considering, for example, the sequence (x_n) with $x_n = (-1)^n$ a sequence which is bounded need not be convergent. The converse is however true which we next prove.

Theorem 1.14 (Convergence $\Longrightarrow$ bounded) *A convergent sequence is bounded.*

Proof. Let (x_n) be a convergent sequence and let x denote the limit. We use the definition of convergence and (arbitrarily) take $\epsilon = 1$. Then there exists $N = N(1)$ such that for all $n \geq N(1)$,

$$|x_n - x| < 1 \quad \text{for all } n \geq N(1).$$

Thus for $n \geq N(1)$

$$x - 1 \leq x_n \leq x + 1.$$

This bounds the tail of the sequence. If we also consider the terms before $N(1)$ then we get

$$|x_n| \leq M \quad \text{where } M = \max\{|x-1|, |x+1|, |x_1|, |x_2|, \cdots, |x_{N-1}|\}. \qquad \square$$

One way of describing the result just proved is that convergence means that the tail of the sequence is in a neighbourhood of the limit and the bound on all the terms in the sequence results because only a finite number of terms are outside of the tail. In the proof we took $\epsilon = 1$ in defining the tail. Any other value can be used.

As we have just shown, convergence implies bounded. From this it follows that not bounded implies not convergent which, in logic jargon, is the contrapositive statement.

The next thing to note is that sequences can be combined to give new sequences. We state this in the following theorem.

Theorem 1.15 *(***Combining convergent sequences***) Let (x_n) and (y_n) be convergent sequences in $\mathbb{R}$ with limits x and y respectively and let $\alpha \in \mathbb{R}$. Then*

(i) $\alpha x_n \to \alpha x$ as $n \to \infty$.

(ii) $x_n + y_n \to x + y$ as $n \to \infty$.

(iii) $x_n y_n \to xy$ as $n \to \infty$.

(iv) $x_n / y_n \to x/y$ as $n \to \infty$ provided $y_n \neq 0$ and $y \neq 0$.

Proof. We will only prove part (iii). The key to the proof is to re-write the difference $x_n y_n - xy$ in a form where we can compare x_n with x and y_n with y. We have

$$x_n y_n - xy = x_n(y_n - y) + y(x_n - x).$$

Then by considering the absolute value and using the triangle inequality we have

$$|x_n y_n - xy| \leq |x_n||y_n - x| + |y||x_n - x|.$$

Next we use a previous result that (x_n) being convergent implies that $|x_n|$ is bounded by a number $M > 0$. Hence

$$|x_n y_n - xy| \leq |x_n||y_n - y| + |y||x_n - x| \leq M|y_n - y| + |y||x_n - x| \leq K(|y_n - y| + |x_n - x|)$$

where $K = \max\{M, |y|\}$. We are now in a position to use the convergence property of both the sequences (x_n) and (y_n). Given $\epsilon > 0$ there exists N_x and N_y such that

$$|x_n - x| < \frac{\epsilon}{2K} \text{ for } n \geq N_x \quad \text{and} \quad |y_n - y| < \frac{\epsilon}{2K} \text{ for } n \geq N_y.$$

Then for $n \geq \max\{N_x, N_y\}$ we get

$$|x_n y_n - xy| < \epsilon. \qquad \square$$

We can also compare a given sequence with a possibly simpler sequence in an attempt to prove convergence. Specifically we have the following theorem.

Theorem 1.16 (Squeeze Theorem) *Suppose that (x_n), (y_n) and (z_n) are sequences of real numbers such that*

$$x_n \le y_n \le z_n \quad \textit{for all } n \in \mathbb{N}$$

and that

$$\lim_{n\to\infty} x_n = \lim_{n\to\infty} z_n.$$

Then the 'intermediate' sequence (y_n) is convergent and

$$\lim_{n\to\infty} x_n = \lim_{n\to\infty} y_n = \lim_{n\to\infty} z_n.$$

Proof. This is a direct proof. Let

$$w = \lim_{n\to\infty} x_n = \lim_{n\to\infty} z_n.$$

The convergence of (x_n) and (z_n) to w implies that given any $\epsilon > 0$ there exists $N \in \mathbb{N}$ such that

$$|x_n - w| < \epsilon \quad \text{and} \quad |z_n - w| < \epsilon \quad \text{for all } n \ge N.$$

Now the ordering of the elements

$$x_n \le y_n \le z_n \quad \Longrightarrow \quad x_n - w \le y_n - w \le z_n - w$$

and hence for $n \ge N$

$$-\epsilon < x_n - w \le y_n - w \le z_n - w < \epsilon, \quad \text{i.e.}|y_n - w| < \epsilon.$$

Thus $y_n \to w$ as $n \to \infty$. □

Some standard convergent sequences

We next state, with proofs in most cases, a few standard sequences with 'well known' limits. The proofs are not always trivial!

1. $x_n := 1/n \to 0$ as $n \to \infty$.

   This follows from the corollary to the Archimedean principle and it is a particular case of the sequence $x_n := 1/n^p$. For this more general sequence

   $x_n := 1/n^p \to 0$ as $n \to \infty$ for all (fixed) $p > 0$.

2. $x_n := r^n \to 0$ as $n \to \infty$ when $|r| < 1$.

   This is proved again later in these notes after we have described the Monotonic Convergence Theorem and indeed that proof is simpler than the one that follows.

   We can in fact deduce the convergence of this sequence to 0 by using only the Squeeze Theorem, that $1/n \to 0$ and the Binomial Theorem. From the Squeeze Theorem it is sufficient to show that $|r|^n \to 0$ and hence we can take $0 < r < 1$. Then we can write

   $$r = \frac{1}{1+a} \quad \text{where } a = \frac{1}{r} - 1 > 0\ .$$

   By the Binomial Theorem we have

   $$(1+a)^n = 1 + na + \text{positive terms} \ge 1 + na$$

giving

$$r^n = \frac{1}{(1+a)^n} \le \frac{1}{1+na} = \frac{1/n}{1/n+a} \to \frac{0}{0+a} = 0 \quad \text{as } n \to \infty.$$

The inequality $(1+a)^n \ge 1+na$ when $a \ge 0$ and $n \in \mathbb{N}$ is known as **Bernoulli's**[2] **inequality**.

3. $x_n := n^{1/n} \to 1$ as $n \to \infty$.

   The proof here is short but moderately difficult and, as in 2 above, it makes use of the Binomial expansion. We give below two slightly different ways in which this can be done.

   (i) Let $y_n = n^{1/n} - 1$ so that $n^{1/n} = 1 + y_n$ and $n = (1+y_n)^n$. Next note that $y_1 = 0$, $y_2 = 2^{1/2} - 1 > 0$, $y_3 = 3^{1/3} - 1 > 0$, $\cdots$, $y_n \ge 0$. The Binomial Theorem hence gives

   $$n = (1+y_n)^n = 1 + ny_n + \frac{n(n-1)}{2}y_n^2 + \text{positive terms} \ge 1 + \frac{n(n-1)}{2}y_n^2$$

   where the final bound is as a result of omitting some positive terms. Hence

   $$n - 1 \ge \frac{n(n-1)}{2}y_n^2 \quad \text{and} \quad y_n^2 \le \frac{2}{n} \to 0 \quad \text{as } n \to \infty.$$

   (ii) As a second way of proving the result consider again Bernoulli's inequality mentioned in the second example. We have

   $$0 < a, \quad (1+a)^n \ge 1 + na \quad \text{so that } 0 < a \le \frac{(1+a)^n - 1}{n} < \frac{(1+a)^n}{n}.$$

   If we take $a = n^{1/(2n)} - 1$ then $a \ge 0$ and

   $$0 < n^{\frac{1}{2n}} - 1 \le \frac{\left(n^{\frac{1}{2n}}\right)^n}{n} = \frac{1}{n^{\frac{1}{2}}}.$$

   Hence by the squeeze theorem $n^{1/(2n)} \to 1$ as $n \to \infty$ and thus $n^{1/n} = (n^{1/(2n)})^2 \to 1$ as $n \to \infty$.

4. If $c > 0$ then $x_n := c^{1/n} \to 1$ as $n \to \infty$. This follows from the previous result and the squeeze theorem.

   Suppose first that $c \ge 1$. As $c \ge 1$ is fixed then $1 \le c \le n$ for sufficiently large n and $c^{1/n} \le n^{1/n}$ because $x^{1/n}$ increases with x for $x \ge 1$. Thus for $n \ge c$

   $$0 \le c^{1/n} - 1 \le n^{1/n} - 1 \to 0 \quad \text{as } n \to \infty,$$

   i.e. $c^{1/n} \to 1$ as $n \to \infty$.

   If instead $0 < c < 1$ then $d = 1/c > 1$ and

   $$c^{1/n} = \frac{1}{d^{1/n}} \to \frac{1}{1} = 1 \quad \text{as } n \to \infty.$$

[2] The Bernouilli's were a three generation dynasty (fathers, brothers, sons, cousins etc.) of eight Swiss mathematicians that lived between 1654 and 1807.

Some examples

With the results concerning the convergence of a few standard sequences, with the theorem about combining convergent sequences and with the Squeeze Theorem, it is now possible to deduce the convergence and the limits of more complicated sequences. To illustrate this consider the following examples.

1. $x_n := \dfrac{1}{n^2+2} = \dfrac{1/n^2}{1+(2/n^2)} \to \dfrac{0}{1} = 0$ as $n \to \infty$.

2. $x_n := \dfrac{3n+1}{2n+5} = \dfrac{3+1/n}{2+5/n} \to \dfrac{3}{2}$ as $n \to \infty$.

3. $x_n := \dfrac{\sqrt{n}}{n+1} = \dfrac{1/\sqrt{n}}{1+1/\sqrt{n}} \to \dfrac{0}{1} = 0$ as $n \to \infty$.

4. $x_n := (2n)^{1/n} = 2^{1/n} n^{1/n} \to (1)(1) = 1$ as $n \to \infty$.

5. $x_n := \dfrac{\sin n}{n} \to 0$ as $n \to \infty$.

   This follows immediately from the Squeeze Theorem by noting that since $|\sin n| \le 1$ we have
   $$-\frac{1}{n} \le \frac{\sin n}{n} \le \frac{1}{n}.$$

6. $x_n := (\sin n)(\sin(1/n)) \to 0$ as $n \to \infty$.

   This follows from the Squeeze Theorem, the result $|\sin n| \le 1$ and the result that $|\sin x| \le |x|$ for all $x \in \mathbb{R}$.
   $$|x_n| = |(\sin n)(\sin(1/n))| \le |\sin(1/n)| \le 1/n \to 0 \text{ as } n \to \infty.$$

7. $x_n := \sqrt{n+1} - \sqrt{n}$.

   As written this is the difference of two divergent sequences and thus the result about combining convergent sequences does not apply. However we can re-express things in a form that we have convergent sequences as follows.

   Recall that
   $$a^2 - b^2 = (a-b)(a+b), \quad \text{which implies that } a - b = \frac{a^2-b^2}{a+b}.$$

   Hence with $a = \sqrt{n+1}$ and $b = \sqrt{n}$ we have
   $$\begin{aligned} x_n &= \frac{1}{\sqrt{n+1}+\sqrt{n}} = \frac{1}{\sqrt{n}\sqrt{1+1/n}+\sqrt{n}} \\ &= \left(\frac{1}{\sqrt{n}}\right)\left(\frac{1}{\sqrt{1+1/n}+1}\right) \to 0\left(\frac{1}{\sqrt{1+0}+1}\right) = 0 \text{ as } n \to \infty. \end{aligned}$$

8. $x_n := \dfrac{2n^2-3n}{n+1}$.

This is an example of a divergent sequence because the numerator grows at a faster rate than the denominator as $n \to \infty$. The sequence is unbounded. We can explain this properly by using the results covered about convergent sequences. We have

$$x_n = \frac{n^2(2-3/n)}{n(1+1/n)} = na_n \quad \text{where } a_n = \frac{2-3/n}{1+1/n}.$$

We have expressed x_n as the product of an unbounded term and a term which tends to a non-zero value as $a_n \to 2$ as $n \to \infty$. To explain why the sequence is unbounded we then can simply note that $a_n \to 2$ implies that $a_n > 1$ for all sufficiently large n and hence

$$x_n > n \quad \text{for all sufficiently large } n.$$

As x_n is unbounded it is divergent.

1.2.2 Convergent sequences and closed bounded intervals

Theorem 1.17 (Non-negative sequences) *If (x_n) is a convergent sequence of real numbers and if $x_n \ge 0$ for all $n \in \mathbb{N}$ then*

$$x = \lim_{n\to\infty} x_n \ge 0.$$

PROOF. We use a proof by contradiction. That is we start with the assumption that $x < 0$. But the elements of the tail of the sequence must be close to x and in particular this implies that for sufficiently large n, $x_n < 0$. In terms of ϵ's, we take $\epsilon = -x = |x| > 0$. The convergence means that there exists an $N \in \mathbb{N}$ such that

$$|x_n - x| < \epsilon, \quad \text{i.e. } x < x_n - x < -x \quad \Longrightarrow 2x < x_n < 0 \text{ for all } n > N.$$

This contradicts $x_n \ge 0$ for all $n \in \mathbb{N}$. □

If we replace the condition $x_n \ge 0$ by the condition $x_n \le 0$ then we get $x \le 0$ and if we consider a convergent sequence in a bounded interval $[a, b]$ then we get that the limit $x \in [a, b]$. This is an important property characterising a closed bounded interval. Being a closed bounded interval in $\mathbb{R}$ thus means that it contains its end points and it contains the limits of all convergent sequences contained within it. We summarize the result in the following corollary.

Corollary 1.18 Closed bounded intervals contain their limits *If (x_n) is a convergent sequence in the closed bounded interval $[a, b]$ then the limit $x \in [a, b]$.*

Remarks

In the definition of the convergence of a sequence and in the results involving composing sequences and in comparing sequences (the Squeeze Theorem) the limit of each sequence appeared in the statements. This is fine when the limit is known and the purpose of the analysis is just to verify that this is the case but there are many situations when there are no obvious candidate values to take. In the subsection that follows on monotone sequences and in a later section on Cauchy[3] sequences the results covered go, in a sense, some way towards remedying this situation by providing criterion for establishing the convergence of a sequence without any reference to the limit of the sequence. To put this another way, the results that follow will provide criterion for establishing the existence of the limit of sequences without explicitly giving us its value. This type of situation should be familiar to you from previous modules. For example, in the intermediate value theorem, Rolle's[4] Theorem, the Mean Value Theorem and Taylor's[5] Theorem with Remainder, you establish the existence of points in intervals but you can only usually determine what they are by numerical methods.

1.2.3 Monotone sequences

Let (x_n) be a sequence of real numbers. We start with some definitions of increasing, strictly increasing, decreasing, strictly decreasing, and monotone that we will use.

Definition 1.19 (x_n) *is* increasing *if* $x_1 \leq x_2 \leq \cdots \leq x_n \leq x_{n+1} \leq \cdots$.
(x_n) *is* strictly increasing *if* $x_1 < x_2 < \cdots \leq x_n < x_{n+1} < \cdots$.
(x_n) *is* decreasing *if* $x_1 \geq x_2 \geq \cdots \geq x_n \geq x_{n+1} \geq \cdots$.
(x_n) *is* strictly decreasing *if* $x_1 > x_2 > \cdots > x_n > x_{n+1} > \cdots$.
(x_n) *is* monotone *if it is either increasing or decreasing.*

As a comment here, some texts (a minority of texts from what I have seen) define increasing as $x_1 < x_2 < \cdots$ and decreasing as $x_1 > x_2 > \cdots$ which are defined in these notes as *strictly increasing* and *strictly decreasing.* This minor difference in terminology makes no difference in any of the results covered in this module.

The main result that we will consider about such sequences is as follows.

Theorem 1.20 (Monotone Convergence Theorem) *A monotone sequence converges if and only if it is bounded. Further we have the following.*

(i) If (x_n) is increasing then
$$\lim_{n\to\infty} x_n = \sup\{x_n : n \in \mathbb{N}\}.$$

(ii) If (x_n) is decreasing then
$$\lim_{n\to\infty} x_n = \inf\{x_n : n \in \mathbb{N}\}.$$

---

[3]A.L. Cauchy (1789-1857) was a French mathematician who made analysis what it is today via his work on convergence. The first analysis course mathematician are following today are still broadly based on his own at the École Polytechnique in 1821. Cauchy was also very productive in all areas of analysis and differential equations. His full life work covers 27 volumes.

[4]M. Rolle (1652-1719) was a French mathematician. Largely self-taught, he worked on analysis, geometry and arithmetic. He is best remembered for the theorem bearing his name. He also invented the notation $\sqrt[n]{x}$ and adopted the notion that if $a > b$ then $-b > -a$ in opposition to various senior mathematicians of his time.

[5]B. Taylor (1685-1731) was an English mathematician who worked on analysis, geometry and mechanics (vibrating string). He introduced the 'calculus of finite differences'. In that context, he invented integration by parts and gave his version of what is known as Taylor's Theorem, although the theorem was already known. He got embroiled in the disputes between English and continental mathematicians about following Newton's or Leibnitz's versions of calculus. He worked also on the foundation of descriptive and projective geometry but did not elaborate although he gave some deep results.

PROOF. We already know that a convergent sequence is bounded and hence we only need to prove the convergence given that the sequence is monotone. We consider here the case of an increasing sequence. The proof in the case of a decreasing sequence is similar.

Because (x_n) is bounded, the set $\{x_n : n \in \mathbb{N}\}$ has a least upper bound $u \in \mathbb{R}$ (Completeness property of $\mathbb{R}$). A least upper bound is such that nothing smaller is an upper bound. Hence given $\epsilon > 0$, $u - \epsilon$ is not an upper bound and consequently there exists x_N with $x_N > u - \epsilon$. But then the increasing property and the bounded above property requires that

$$u - \epsilon < x_N \le x_{N+1} \le x_{N+2} \le \cdots \le u.$$

Thus the tail of the sequence is within ϵ of u which implies that the sequence converges to u. □

Examples of monotone sequences

1. $x_n := \dfrac{1}{\sqrt{n}}$. This is a decreasing sequence bounded below by 0 and hence it converges by the Monotone Convergence Theorem. (Of course, the convergence of this sequence to 0 can be proved directly without the need to use the Monotone Convergence Theorem. Also, the result has already been covered when we stated that $1/n^p \to 0$ as $n \to \infty$ for all fixed $p > 0$.)

2. $x_n := r^n$, $0 < r < 1$ is also a decreasing sequence bounded below by 0 and hence it converges by the Monotone Convergence Theorem. To prove that the limit x is 0 note that $x_{n+1} = r x_n$ which, by letting $n \to \infty$ implies that

   $$x = rx, \quad \text{i.e. } x(1-r) = 0.$$

   Since $0 < r < 1$ we obtain that $x = 0$.

3. Later in this module when we consider the contraction mapping theorem we consider sequences of the form $x_{n+1} = \phi(x_n)$, $n = 0, 1, 2, \cdots$ for some given function ϕ and with some starting point x_0. Sequences of this type are often monotone. Consider for example the case of a function $\phi : \mathbb{R} \to \mathbb{R}$ which is such that $\phi'(x) \ge 0$ for all x and observe that by the Mean Value Theorem (a result you have met in Level 1)

   $$x_{n+1} - x_n = \phi(x_n) - \phi(x_{n-1}) = \phi'(\alpha_n)(x_n - x_{n-1})$$

   where α_n is some point between x_{n-1} and x_n. If the starting point x_0 is such that $x_1 = \phi(x_0) > x_0$ then this successively implies that $x_2 - x_1 \ge 0$, $x_3 - x_2 \ge 0$, $\cdots$, $x_{n+1} - x_n \ge 0$. That is $x_{n+1} \ge x_n \ge \cdots \ge x_2 \ge x_1 > x_0$, i.e. the sequence is increasing. Similar reasoning shows that if instead $x_1 < x_0$ then the sequence is decreasing.

4. Let $a_n > 0$ for all $n \in \mathbb{N}$ and let

   $$s_n := \sum_{k=1}^{n} a_k.$$

   (s_n) is the sequence of partial sums. This is an increasing sequence since $s_{n+1} - s_n = a_{n+1} > 0$ and hence it converges if and only if it is bounded above. When (s_n) converges we say that the *series* $\sum a_k$ converges. Series will be considered next.

1.2.4 Series in $\mathbb{R}$ – terminology and some basic results

Given a sequence $a_1, a_2, \cdots$ of real numbers we can form the sequence of partial sums

$$\begin{aligned} s_1 &= a_1, \\ s_2 &= a_1 + a_2 = s_1 + a_2, \\ \cdots & \\ s_n &= a_1 + a_2 + \cdots + a_n = s_{n-1} + a_n = \sum_{k=1}^{n} a_k \ . \end{aligned}$$

The sequence of partial sums is known as the **series** or **infinite series** generated by (a_n). We use in the case of series the notation $\sum a_n$ to denote the series generated by (a_n). The series is said to **converge** if the sequence of partial sums converges and when we have convergence we write

$$\sum_{k=1}^{\infty} a_k := \lim_{n\to\infty} \left(\sum_{k=1}^{n} a_k \right).$$

In the case of convergent sequences we refer to the limit of the sequence, in the case of convergent series we refer to the sum of the series. The value of the sum depends on all the terms (a_n) involved but the convergence or divergence of the series depends only on the entries in the tail. For example, if $\sum_{100}^{\infty} a_n$ converges then $\sum_{1}^{\infty} a_n$ also converges with the sums related by

$$\sum_{1}^{\infty} a_n = \left(\sum_{1}^{99} a_n \right) + \left(\sum_{100}^{\infty} a_n \right) \ .$$

Later in this module we will consider sequences and series of functions. In the case of a sequence of functions (f_n), all defined on a common domain, we will consider the function $\sum_1^\infty f_n$ and we will consider the value $\sum_1^\infty f_n(x)$ (when they exists) at each point x in the common domain. The aims then will be in determining whether or not such series converge and with the properties of the limit function. This in turn will require us to establish whether or not certain related series of numbers converge. We limit the consideration of series of numbers here to those results that are needed for the series of functions section at the end of the first part of this module.

As a series converges if and only if the sequence of partial sums converges, we can immediately apply all that we know about sequences to series. For example, if $\sum a_n$ and $\sum b_n$ converge then $\sum (a_n + b_n)$ and $\sum (c a_n)$ (for any constant c) also converge with

$$\begin{aligned} \sum_{1}^{\infty} (a_n + b_n) &= \sum_{1}^{\infty} a_n + \sum_{1}^{\infty} b_n \\ \sum_{1}^{\infty} (c a_n) &= c \sum_{1}^{\infty} a_n \ . \end{aligned}$$

As has already been stated, $\sum a_n$ converges means that (s_n) converges where $s_n := \sum_1^n a_k$. If $s = \lim s_n$ and we note that $a_n = s_n - s_{n-1}$ then

$$a_n = s_n - s_{n-1} \to s - s = 0 \quad \text{as } n \to \infty.$$

Thus if $\sum a_n$ converges then $a_n \to 0$ as $n \to \infty$, i.e. it is necessary for convergence of the series that the sequence (a_n) which generates the series converges to 0. We hence immediately deduce that if $a_n \not\to 0$ then the series diverges. The condition $a_n \to 0$ is necessary for convergence but it is not sufficient for convergence. The standard example to illustrate this is the harmonic series $\sum_1^\infty 1/n$ which diverges despite the fact that the terms $1/n \to 0$ as $n \to \infty$. We can show that this series diverges using the integral test given later. A slightly different way of doing this is to consider

$$t_n = \sum_{k=1}^{2^n} \frac{1}{k} = 1 + \frac{1}{2} + \left(\frac{1}{3} + \frac{1}{4}\right) + \left(\frac{1}{5} + \cdots + \frac{1}{8}\right) + \cdots + \left(\frac{1}{2^{n-1}+1} + \cdots + \frac{1}{2^n}\right).$$

Now for each bracketed term

$$\begin{aligned} \frac{1}{3} + \frac{1}{4} &> \frac{1}{4} + \frac{1}{4} = \frac{1}{2}, \\ \frac{1}{5} + \cdots + \frac{1}{8} &> \frac{1}{8} + \cdots + \frac{1}{8} = \frac{1}{2}, \\ &\cdots \\ \frac{1}{2^{n-1}+1} + \cdots + \frac{1}{2^n} &> \frac{1}{2^n} + \cdots + \frac{1}{2^n} = \frac{1}{2}, \end{aligned}$$

Thus

$$t_n \ge 1 + \frac{1}{2} + \cdots + \frac{1}{2} = 1 + \frac{n}{2}.$$

The right hand side is unbounded and thus the sequence of partial sums is unbounded by comparison.

A similar technique to the above can be used to show that the series $\sum 1/n^2$ converges. In this case let

$$\begin{aligned} t_n &= \sum_{k=1}^{2^{(n+1)}-1} \frac{1}{k^2} = 1 + \left(\frac{1}{2^2} + \frac{1}{3^2}\right) + \left(\frac{1}{4^2} + \cdots + \frac{1}{7^2}\right) + \cdots + \left(\frac{1}{2^{2n}} + \cdots + \frac{1}{(2^{(n+1)}-1)^2}\right) \\ &\le 1 + \frac{2}{2^2} + \frac{4}{4^2} + \cdots + \frac{2^n}{2^{2n}} \\ &= 1 + \frac{1}{2} + \frac{1}{2^2} + \cdots + \frac{1}{2^n} \\ &= \frac{1-(1/2)^{(n+1)}}{1-(1/2)} = 2 - \frac{1}{2^n} < 2 \end{aligned}$$

where in the above we have used the formula for summing a geometric series. The sequence of partial sums is hence increasing and bounded above by 2 and hence by the monotone convergence theorem it converges.

In the above two examples we had sequences of partial sums which are increasing as the terms of the series are non-negative. In the case of a series $\sum a_n$ in which entries $a_1, a_2, \cdots$ are both positive and negative it can be difficult to determine whether or not the series converges but we can deduce immediately that it converges when the related series of absolute values $\sum |a_n|$ converges. That is if $\sum |a_n|$ converges then $\sum a_n$ converges. This can be proved as follows. Let

$$s_n := \sum_1^n a_k \quad \text{and} \quad t_n := \sum_1^n |a_k|$$

and suppose that (t_n) is an increasing and bounded sequence (i.e. a convergent sequence). Also let

$$s_n^+ := \sum_{a_k>0,\ k\le n} a_k \quad \text{and} \quad s_n^- := \sum_{a_k<0,\ k\le n} a_k \ .$$

(s_n^+) involves adding the positive terms and hence it is increasing and bounded because

$$s_n^+ \le t_n \ .$$

(s_n^-) involves adding the negative terms and hence it is decreasing and bounded because

$$-s_n^- \le t_n, \quad s_n^- \ge -t_n \ .$$

Thus both (s_n^+) and (s_n^-) converge and because

$$s_n = s_n^+ + s_n^-$$

the sequence (s_n) also converges.

There is some terminology connected with this result.

Definition 1.21 $\sum a_n$ *is said to be* absolutely convergent *if the series* $\sum |a_n|$ *converges.* $\sum a_n$ *is said to be* conditionally convergent *if* $\sum a_n$ *converges but* $\sum |a_n|$ *diverges.*

In the above we have proved that if a series is absolutely convergent then it is convergent. That is, we can test for convergence by testing for absolute convergence. More specialised tests are usually required to deal with series which are conditionally convergent.

Given the above comments we restrict attention to non-negative sequences (a_n) and, as above, let the sequence of partial sums (s_n) be given by

$$s_n = a_1 + a_2 + \cdots + a_n = s_{n-1} + a_n = \sum_1^n a_k \ .$$

This is consequently an increasing sequence since $s_n - s_{n-1} = a_n \ge 0$. Increasing sequences converge if and only if they are bounded (Monotone Convergence Theorem) and hence we establish convergence or divergence by testing whether or not the sequence (s_n) is bounded. This is usually done by comparing the given series with a 'simpler' series which is known to converge or diverge. We make this precise in the following result.

Theorem 1.22 (Comparison test) *Let* (a_n) *and* (b_n) *be real sequences of non-negative numbers such that* $0 \le a_n \le b_n$ *for all* n*.*

(i) If $\sum b_n$ *converges then* $\sum a_n$ *converges.*

(ii) If $\sum a_n$ *diverges then* $\sum b_n$ *diverges.*

Proof. If $\sum b_n$ converges then the sequence of partial sums of $\sum a_n$ is increasing and bounded. Hence $\sum a_n$ converges.

If $\sum a_n$ diverges then the sequence of partial sums of (a_n) is unbounded. From the inequality the sequence of partial sums of (b_n) is also unbounded and hence $\sum b_n$ diverges. $\square$

1.2.5 Examples of series

(i) Let

$$a_n = \frac{n+3}{n^2+7} = \frac{1}{n}c_n, \quad \text{where} \quad c_n = \frac{1+(3/n)}{1+(7/n^2)} \to 1 \quad \text{as } n \to \infty.$$

We note that $a_n \to 0$ which is necessary for convergence but we also note that a_n is behaving like $1/n$ for large n and we know that $\sum 1/n$ diverges. We want to make use of this observation to prove that $\sum a_n$ also diverges with the minimum of effort by using the comparison test. Now as $c_n \to 1$ as $n \to \infty$ we know that there exists a N such that for $n \geq N$ we have $1/2 < c_n < 3/2$ for all $n \geq N$. Hence for $n \geq N$ we have

$$a_n = \frac{c_n}{n} \geq \frac{1}{2n}$$

and by the comparison test $\sum a_n$ diverges because $\sum 1/(2n)$ diverges.

(ii) Let now

$$a_n = \frac{n+3}{n^3+7} = \frac{c_n}{n^2} \quad \text{where} \quad c_n = \frac{1+3/n}{1+7/n^3} \to 1 \quad \text{as } n \to \infty.$$

Again we have $a_n \to 0$ with a_n behaving like $1/n^2$ and we know that $\sum 1/n^2$ converges. To prove that $\sum a_n$ converges we use the comparison test and to do this we need an upper bound on a_n. This can be obtained by similar reasoning to (i). As $c_n \to 1$ as $n \to \infty$ we know that there exists a N such that for $n \geq N$ we have $0 < c_n < 2$ for all $n \geq N$. Hence for $n \geq N$ we have

$$0 < a_n = \frac{c_n}{n^2} < \frac{2}{n^2}$$

and $\sum a_n$ converges by comparison with $\sum (2/n^2)$.

Note that in both (i) and (ii) it is not necessary here to struggle to work out a N such that the bounds hold for all $n \geq N$ although this is not too hard in these examples.

1.2.6 Comparing a series with a geometric series

One of the first series that you meet is the geometric series $\sum r^n$ which is a series for which you can give explicit representations for the partial sums. (This was used to show that $\sum 1/n^2$ is convergent.) If we define

$$s_n := 1 + r + r^2 + \cdots + r^{n-1} = \sum_{k=0}^{n-1} r^k$$

then

$$rs_n = r + r^2 + \cdots + r^n$$

and

$$(1-r)s_n = 1 - r^n \quad \text{giving} \quad s_n = \frac{1-r^n}{1-r}$$

provided $r \neq 1$. If $|r| < 1$ then $r^n \to 0$ as $n \to \infty$ and $s_n \to s$ where $s = 1/(1-r)$. Also,

$$|s_n - s| = \frac{|r|^n}{|1-r|}$$

and thus $s_n \to s$ at a geometric rate. As an example, when $r = 1/2$ we have

$$1 + \frac{1}{2} + \left(\frac{1}{2}\right)^2 + \left(\frac{1}{2}\right)^3 + \cdots + \left(\frac{1}{2}\right)^{n-1} = \frac{1-(1/2)^n}{1-(1/2)} = 2 - (1/2)^{n-1} \to 2 \quad \text{as } n \to \infty.$$

If $|r| > 1$ then the partial sums (s_n) and the terms in the series $\sum r^n$ are both unbounded and hence the series does not converge.

Based on this knowledge of the geometric series there are two named tests for convergence which involve comparing a given series with a geometric series. We first state both of them in their simplest form and give some examples before giving the proofs.

Theorem 1.23 (Ratio test) *Let (a_n) be a sequence of positive terms which are such that*

$$\lim_{n\to\infty} \frac{a_{n+1}}{a_n} = \rho\ .$$

(i) If $\rho < 1$ then the series converges.

(ii) If $\rho > 1$ then the series diverges.

Theorem 1.24 (Root test) *Let (a_n) be a sequence of positive terms which are such that*

$$\lim_{n\to\infty} a_n^{1/n} = \rho\ .$$

(i) If $\rho < 1$ then the series converges.

(ii) If $\rho > 1$ then the series diverges.

1.2.7 Examples of applying the ratio and root tests

The ratio and root tests are typically used in situations when we have a power series $\sum c_n x^n$. In this case we have $a_n = c_n x^n$ and

$$\frac{a_{n+1}}{a_n} = \frac{c_{n+1}}{c_n} x \quad \text{and} \quad |a_n|^{1/n} = |c_n|^{1/n}|x|.$$

We consider a few series of this type.

1. Consider $\sum x^n/n!$ which is the series defining $\exp(x)$ when we have $\sum_0^\infty x^n/n!$. We have not said here that x is positive and hence to apply the test we need to test for the absolute convergence of the series. We have

   $$a_n = \frac{x^n}{n!} \quad \text{giving} \quad \left|\frac{a_{n+1}}{a_n}\right| = \frac{|x|}{n+1} \to 0 \quad \text{as } n \to \infty.$$

   Thus the ratio test implies that the series is absolutely convergent for all x.

   The root test gives the same conclusion but is a bit harder to apply in this case. We have

   $$|a_n|^{1/n} = \frac{|x|}{(n!)^{1/n}}.$$

   To proceed we need to estimate $n!$. Now

   $$(2n)! = (2n)(2n-1)\cdots(n+1)(n)\cdots(3)(2) > (2n)(2n-1)\cdots(n+1) > n^n$$

where the first inequality is because we are omitting terms ≥ 1 and the second inequality is because we have the product of n terms each of which is greater than n. Similarly

$$\begin{aligned}(2n+1)! &= (2n+1)(2n)\cdots(n+1)(n)\cdots(3)(2)\\ &> (2n+1)(2n)\cdots(n+1) > (n+1)^{n+1} > (n+1/2)^{n+1/2}.\end{aligned}$$

Hence for all n (even and odd) we have

$$n! \geq \left(\frac{n}{2}\right)^{n/2} \quad \text{and} \quad (n!)^{1/n} \geq \left(\frac{n}{2}\right)^{1/2}$$

which in turn implies that

$$\left(\frac{1}{n!}\right)^{1/n} \leq \left(\frac{2}{n}\right)^{1/2} \to 0 \quad \text{as } n \to \infty.$$

To summarize, we have shown that $|a_n|^{1/n} \to 0$ as $n \to \infty$ for all x.

2. Consider $\sum x^n/n^2$, i.e. $a_n = x^n/n^2$.

$$\left|\frac{a_{n+1}}{a_n}\right| = |x|\left(\frac{n}{n+1}\right)^2 = |x|\left(\frac{1}{1+1/n}\right)^2 \to |x| \quad \text{as } n \to \infty.$$

   The root test gives the same conclusion since

$$|a_n|^{1/n} = |x|(1/n^{1/n})^2 \to |x| \quad \text{as } n \to \infty.$$

   The tests imply that the series converges if $|x| < 1$ and diverges if $|x| > 1$. The tests give no information when $|x| = 1$. From other tests it can be shown that the series converges absolutely for $|x| = 1$.

3. Consider $\sum nx^n$, i.e. $a_n = nx^n$.

$$\left|\frac{a_{n+1}}{a_n}\right| = |x|\left(\frac{n+1}{n}\right) = |x|\left(1+\frac{1}{n}\right) \to |x| \quad \text{as } n \to \infty.$$

   The root test gives the same conclusion since

$$|a_n|^{1/n} = |x|n^{1/n} \to |x| \quad \text{as } n \to \infty.$$

   The tests imply that the series converges if $|x| < 1$ and diverges if $|x| > 1$. The tests give no information when $|x| = 1$. However, when $|x| = 1$ we have $a_n = \pm n$ which does not tend to 0. Hence the series does not converge for any x satisfying $|x| = 1$.

4. Consider $\sum x^n/n$, i.e. $a_n = x^n/n$. We get similar conclusions to the last two examples in that the series converges absolutely when $|x| < 1$ and it diverges when $|x| > 1$. The ratio and root tests give no information when $|x| = 1$. In this case the series diverges when $x = 1$ (the harmonic series) but it can be shown to converge when $x = -1$ by a test known as the alternating series test.

5. Consider $\sum n!x^n$, i.e. $a_n = n!x^n$. In this case the ratio of successive terms is

$$\frac{a_{n+1}}{a_n} = (n+1)x$$

   which is unbounded for all $x \neq 0$. Thus the series diverges for all $x \neq 0$.

Proofs of the ratio and root tests

In the following we bound the series above by a convergent geometric series when $\rho < 1$ and we bound the series below by a divergent geometric series when $\rho > 1$.

Ratio test when $\rho < 1$
If $a_{n+1}/a_n \to \rho < 1$ then for any r in $\rho < r < 1$ the definition of convergence implies that there exists a N such that

$$\frac{a_{n+1}}{a_n} \le r \quad \text{for all } n \ge N.$$

If we consider the sequence (a_n) from this point onwards then

$$a_{N+k} = \left(\frac{a_{N+k}}{a_{N+k-1}}\right)\left(\frac{a_{N+k-1}}{a_{N+k-2}}\right)\cdots\left(\frac{a_{N+2}}{a_{N+1}}\right)\left(\frac{a_{N+1}}{a_N}\right)a_N \le a_N r^k \quad \text{for } k = 0, 1, 2, \cdots$$

since each of the bracketed terms is less than or equal to r. We know that the geometric series $\sum r^k$ converges when $0 \le r < 1$ and the comparison test then implies that $\sum_{k=0}^{\infty} a_{N+k} = \sum_N^{\infty} a_k$ also converges. As there are only a finite number of terms before $n = N$ we deduce that $\sum_1^{\infty} a_k$ converges.

Ratio test when $\rho > 1$
A similar reason to the above tells us that for any r satisfying $\rho > r > 1$ there exists an N such that

$$a_{N+k} = \left(\frac{a_{N+k}}{a_{N+k-1}}\right)\left(\frac{a_{N+k-1}}{a_{N+k-2}}\right)\cdots\left(\frac{a_{N+2}}{a_{N+1}}\right)\left(\frac{a_{N+1}}{a_N}\right)a_N \ge a_N r^k \quad \text{for } k = 0, 1, 2, \cdots .$$

By the comparison test the series diverges. In fact the terms become unbounded.

Root test when $\rho < 1$
Again, the definition of the limit implies that for any r satisfying $\rho < r < 1$ there exists a N such that

$$a_n^{1/n} \le r \quad \text{for all } n \ge N.$$

Then

$$a_n \le r^n \quad \text{for all } n \ge N.$$

By comparing with the convergent geometric series we deduce that the series converges.

Root test when $\rho > 1$
As in the previous cases the definition of the limit implies that for any r satisfying $\rho > r > 1$ there exists a N such that

$$a_n^{1/n} \ge r \quad \text{for all } n \ge N$$

and thus

$$a_n \ge r^n \quad \text{for all } n \ge N.$$

By the comparison test the series hence diverges.

Remarks

(i) The ratio test and the root test are often easy to apply as has been illustrated in the previous set of examples. When the ratio test can be used it can be shown that the root test can also be used and the same ρ is obtained, i.e. the same conclusion is reached. We have

$$\rho = \lim_{n\to\infty} a_{n+1}/a_n = \lim_{n\to\infty} a_n^{1/n} \ .$$

There are however situations when $(a_n^{1/n})$ converges but the sequence of ratios (a_{n+1}/a_n) does not and in this sense the ratio test is the stronger and the more widely applicable test.

(ii) The ratio test and the root test give no information when $\rho = 1$. Consider for example

$$a_n = \frac{1}{n} \quad \text{and} \quad b_n = \frac{1}{n^2}.$$

We have

$$\frac{a_{n+1}}{a_n} = \frac{n}{n+1} \to 1 \quad \text{and} \quad \frac{b_{n+1}}{b_n} = \left(\frac{n}{n+1}\right)^2 \to 1$$

with $\sum a_n = \sum 1/n$ diverging and with $\sum b_n = \sum 1/n^2$ converging. When $\rho = 1$ the only information that the test gives is that the series converges 'slower' than a geometric series or diverges 'slower' than a geometric series.

(iii) To use the ratio test you need the sequence (a_{n+1}/a_n) to converge and to use the root test you need the sequence $(a_n^{1/n})$ to converge. We state now, without proof, how you can remove this requirement in a stronger but more technical version of the root test.

The result involves considering the set of entries $\{a_m^{1/m}\}$. If this set is not bounded then the series diverges. If the set is bounded then we can define the two sequences

$$b_n := \inf\{a_m^{1/m} : \ m \ge n\} \quad \text{and} \quad c_n := \sup\{a_m^{1/m} : \ m \ge n\}.$$

As n increases the set $\{a_m^{1/m} : \ m \ge n\}$ diminishes and consequently (b_n) is an increasing sequence and (c_n) is a decreasing sequence and as they are both bounded they both converge with

$$\lim_{n\to\infty} b_n \le \lim_{n\to\infty} c_n \ .$$

When

$$\lim_{n\to\infty} \sup\{a_m^{1/m} : \ m \ge n\} < 1$$

we have convergence and when

$$\lim_{n\to\infty} \inf\{a_m^{1/m} : \ m \ge n\} > 1$$

we have divergence. The form of the root test given previously corresponds to when

$$\lim_{n\to\infty} \inf\{a_m^{1/m} : \ m \ge n\} = \lim_{n\to\infty} a_n^{1/n} = \lim_{n\to\infty} \sup\{a_m^{1/m} : \ m \ge n\}$$

which is the situation when $(a_n^{1/n})$ converges.

1.2.8 The integral test

We now consider a test which can be used for series of the type $\sum 1/n^p$ for fixed $p > 0$. Both the ratio test and the root test give no information. The test that we will consider applies to series of the form $\sum_{n=1}^{\infty} f(n)$ where $f : [1, \infty) \to [0, \infty)$ is a decreasing continuous function with $f(x) \to 0$ as $x \to \infty$. In the above case the function $f(x) := 1/x^p$ has this property.

Now if f is decreasing then for all $x \in [k-1, k]$ we have

$$f(k) \le f(x) \le f(k-1)$$

and by considering the area under the curves of $y = f(k)$, $y = f(x)$ and $y = f(k-1)$ when $k-1 \le x \le k$ we have

$$f(k) = \int_{k-1}^{k} f(k)\,dx \le \int_{k-1}^{k} f(x)\,dx \le \int_{k-1}^{k} f(k-1)\,dx = f(k-1),$$

see Figure 1.5. If we define

$$s_n := \sum_{k=1}^{n} f(k) \quad \text{and} \quad I_n := \int_1^n f(x)\,dx$$

then both (s_n) and (I_n) are increasing sequences and

$$s_n - f(1) = \sum_2^n f(k) \le \left(\int_1^2 + \int_2^3 + \cdots + \int_{n-1}^n \right) f(x)\,dx \le \sum_2^n f(k-1) = \sum_1^{n-1} f(k) = s_{n-1}\ .$$

That is

$$s_n - f(1) \le I_n \le s_{n-1}\ .$$

Hence if (I_n) converges then (s_n) converges and if (I_n) diverges then (s_n) diverges by the comparison test. We summarize these results in the following theorem.

Theorem 1.25 (Integral test) *If $f : [1, \infty) \to [0, \infty)$ is a decreasing continuous function then*

$$\sum_1^{\infty} f(n) \quad \text{and} \quad \int_1^{\infty} f(x)\,dx$$

either both converge or both diverge.

As an example of applying the test, consider again $f(x) := 1/x^p$. For $p \ne 1$ we have

$$I_n = \int_1^n f(x)\,dx = \int_1^n \frac{dx}{x^p} = \left[\frac{x^{-p+1}}{-p+1} \right]_1^n = \frac{1}{p-1} \left(1 - \frac{1}{n^{p-1}} \right)$$

and for $p = 1$ we have

$$I_n = \int_1^n \frac{dx}{x} = \ln n\ .$$

Hence if $p \le 1$ then the series diverges and if $p > 1$ then the series converges.

We can relate the limits of (s_n) and (I_n) in the convergent case. We have that if $s_n \to s$ and $I_n \to I$ as $n \to \infty$ then

$$s - f(1) \le I \le s$$

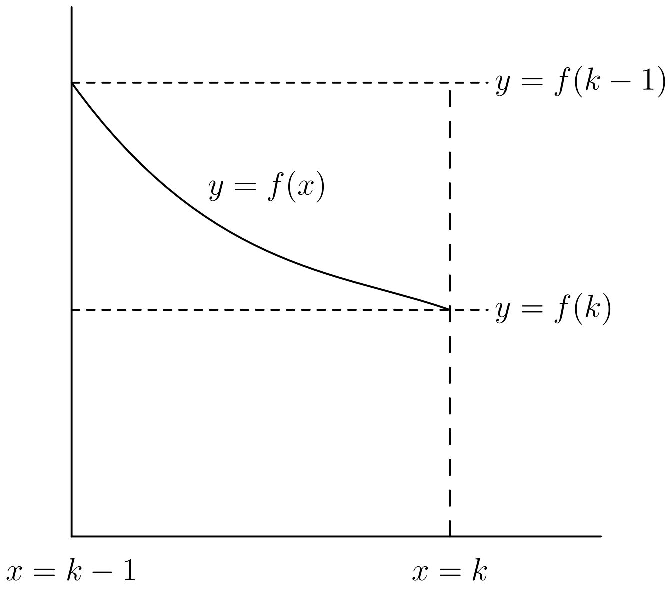


Figure 1.5: The area under the curve is less than the area under the line $y = f(k-1)$ and it is greater than the area under the line $y = f(k)$.

and thus the sum of the series s satisfies $s \in [I, I + f(1)]$.

We can also obtain information about the rate of convergence or divergence by considering again the inequality

$$f(k) \le \int_{k-1}^{k} f(x)\,dx \le f(k-1)\ .$$

We have

$$0 \le \left(f(k-1) - \int_{k-1}^{k} f(x)\,dx \right) \le f(k-1) - f(k)\ .$$

Summing from $k = 2$ to n gives

$$0 \le s_{n-1} - I_n \le f(1) - f(n) < f(1)$$

where

$$s_{n-1} - I_n = \left(f(1) - \int_{1}^{2} f(x)\,dx \right) + \cdots + \left(f(n-1) - \int_{n-1}^{n} f(x)\,dx \right)\ .$$

Since $(s_{n-1} - I_n)$ is increasing with n and bounded by $f(1)$ we know that it converges with

$$\lim_{n\to\infty} (s_{n-1} - I_n) = \gamma \le f(1)\ .$$

Thus the difference between the two sequences converges even in the case when the sequences themselves diverge. Thus when $f(x) := 1/x$ we obtain the result that

$$\lim_{n\to\infty} \left(\sum_{k=1}^{n-1} \frac{1}{k} - \ln n \right) = \gamma < 1.$$

In this case it can be shown that $\gamma = 0.577..$ which is known as Euler constant. Thus $\sum 1/n$ diverges as slowly as $\ln n$.

In the case when (s_n) converges to s we have

$$0 \le s_{n+j} - s_n = \sum_{k=n+1}^{n+j} f(k) \le \int_n^{n+j} f(x)\, dx$$

and letting $j \to \infty$ gives

$$0 \le s - s_n \le \int_n^{\infty} f(x)\, dx \ .$$

This gives a bound on how rapidly the sequence converges to s. In the case $f(x) := 1/x^p$, $p > 1$ we have

$$0 \le s - s_n \le \int_n^{\infty} \frac{dx}{x^p} \ .$$

Now

$$\int_n^{\infty} \frac{dx}{x^p} = \lim_{M\to\infty} \int_n^{M} \frac{dx}{x^p} = \frac{1}{p-1}\left(\frac{1}{n^{p-1}}\right).$$

When $p = 2$ we have

$$0 \le \sum_{n+1}^{\infty} \frac{1}{k^2} \le \frac{1}{n}.$$

Hence to obtain the sum accurate to 10^{-3} we need 10^3 terms and to obtain the sum accurate to 10^{-6} we need 10^6 terms. The series converges slowly.

1.2.9 Summary

We finish this section on sequences and series in $\mathbb{R}$ by summarizing the main definitions and results.

1. The sequence (x_n) *converges* to x if for every $\epsilon > 0$ there exists a $N = N(\epsilon)$ such that

   $$|x_n - x| < \epsilon$$

   for all $n \ge N$.

2. If the sequence (x_n) converges then its limit is unique.

3. A sequence (x_n) is *bounded* if there exists $M > 0$ such that $|x_n| \le M$ for all $n \in \mathbb{N}$.

4. A convergent sequence is bounded.

5. Let (x_n) and (y_n) be convergent sequences in $\mathbb{R}$ with limits x and y respectively and let $\alpha \in \mathbb{R}$. Then

   (i) $\alpha x_n \to \alpha x$ as $n \to \infty$.

   (ii) $x_n + y_n \to x + y$ as $n \to \infty$.

   (iii) $x_n y_n \to xy$ as $n \to \infty$.

   (iv) $x_n / y_n \to x/y$ as $n \to \infty$ provided $y_n \ne 0$ and $y \ne 0$.

6. Suppose that (x_n), (y_n) and (z_n) are sequences of real numbers such that

$$x_n \le y_n \le z_n \quad \text{for all } n \in \mathbb{N}$$

and that

$$\lim_{n\to\infty} x_n = \lim_{n\to\infty} z_n.$$

Then the 'intermediate' sequence (y_n) is convergent and

$$\lim_{n\to\infty} x_n = \lim_{n\to\infty} y_n = \lim_{n\to\infty} z_n.$$

7. $x_n := 1/n^p \to 0$ as $n \to \infty$ for all (fixed) $p > 0$.
$x_n := r^n \to 0$ as $n \to \infty$ when $|r| < 1$.
$x_n := n^{1/n} \to 1$ as $n \to \infty$.
If $c > 0$ then $x_n := c^{1/n} \to 1$ as $n \to \infty$.

8. If (x_n) is a convergent sequence in the closed bounded interval $[a, b]$ then the limit $x \in [a, b]$.

9. (x_n) is *increasing* if $x_1 \le x_2 \le \cdots \le x_n \le x_{n+1} \le \cdots$.
(x_n) is *strictly increasing* if $x_1 < x_2 < \cdots \le x_n < x_{n+1} < \cdots$.
(x_n) is *decreasing* if $x_1 \ge x_2 \ge \cdots \ge x_n \ge x_{n+1} \ge \cdots$.
(x_n) is *strictly decreasing* if $x_1 > x_2 > \cdots > x_n > x_{n+1} > \cdots$.
(x_n) is *monotone* if it is either increasing or decreasing.

10. A monotone sequence converges if and only if it is bounded. Further we have the following.

    (i) If (x_n) is increasing then

$$\lim_{n\to\infty} x_n = \sup\{x_n : \ n \in \mathbb{N}\}.$$

    (ii) If (x_n) is decreasing then

$$\lim_{n\to\infty} x_n = \inf\{x_n : \ n \in \mathbb{N}\}.$$

11. $\sum a_n$ converges if and only is the sequence of partial sums converges (definition).

12. $\sum a_n$ converges implies that $a_n \to 0$ as $n \to \infty$.

13. $\sum a_n$ converges absolutely if $\sum |a_n|$ converges. $\sum a_n$ converges conditionally if $\sum a_n$ converges but $\sum |a_n|$ diverges (definition).

14. If $\sum |a_n|$ converges then $\sum a_n$ converges.

15. Suppose $0 \le a_n \le b_n$ for all n. If $\sum b_n$ converges then $\sum a_n$ converges and if $\sum a_n$ diverges then $\sum b_n$ diverges (comparison test).

16. If $a_n > 0$ and $a_{n+1}/a_n \to \rho$ as $n \to \infty$ then $\sum a_n$ converges if $\rho < 1$ and diverges if $\rho > 1$ (ratio test). In the case $\rho < 1$ this can also be used to show that $a_n \to 0$.

17. If $a_n > 0$ and $a_n^{1/n} \to \rho$ as $n \to \infty$ then $\sum a_n$ converges if $\rho < 1$ and diverges if $\rho > 1$ (root test). In the case $\rho < 1$ this can also be used to show that $a_n \to 0$.

18. If $f : [1, \infty) \to [0, \infty)$ is a decreasing continuous function then

$$\sum_{k=1}^{\infty} f(k) \quad \text{and} \quad \int_1^{\infty} f(x)\, dx$$

both converge or both diverge (integral test).