ANALYTIC MATHEMATICS (I)

CM0188

Antonia J. Jones

15 December 2003
UNIVERSITY OF WALES, CARDIFF (Cardiff University)
DEPARTMENT OF COMPUTER SCIENCE (COMSC)

COURSE: CM0188
MODULE: Mathematics for Computer Science
LECTURER: Antonia J. Jones, COMSC
DATED: Originally 15 August 1997
LAST REVISED: 15 December 2003
ACCESS: Lecturer (extn 4653, room N2.15).

Overhead slides are posted on:

http://www.cs.cf.ac.uk/user/Antonia.J.Jones/lectures

electronically as pdf Acrobat files. It is not normally necessary for students attending the course to print this file as complete sets of printed slides will be issued.

©2003 Antonia J. Jones. Permission is hereby granted to any web surfer for downloading, printing and use of this material for personal study only. Copyright permission is explicitly withheld for modification, re-circulation or publication by any other means, or commercial exploitation in any manner whatsoever, of this file or the material therein

Recommended text:


There are multiple copies of this book (First Edition) available on short loan from the library. Nevertheless purchase of the book is very strongly recommended. It covers a great deal more than the mathematics needed for just this module and will be useful at all stages in the degree course. The complete first year course (i.e. this semester plus last semester) will cover the material of Chapters 1, 2, 3, bits of the set theory from Chapter 6, Chapters 7, 8, and maxima/minima and Taylor's theorem from Chapter 9 of the book.

The work of last semester represents a review of what (ideally) you should have become familiar with at A-level, although some new terminology was introduced. Please read the file Numbers.pdf in the Lectures directory on my website. The work of this semester represents the new material that is vital for many subsequent modules of the degree course. In order to help this
process along it is essential that you attend the tutorials and use the coursebook to check things you are unsure about and to form the basis of independent work.

- The material in these notes represents the important new material of the course and where possible you should look ahead and try to become familiar with what is going to happen beforehand.

Whilst we have made every effort to help you along, there is a limit to how many examples and how much detail one can give in 20 lectures or so - the notes/slides contain the essential points but are quite condensed. You must be prepared to exploit the tutorials and pursue independent study using the book.

**Mathematica**

Almost everything we learn in this course can be illustrated using Mathematica, a programming system for doing mathematics. In the PC LAB every machine has Mathematica installed and there are example programs illustrating various aspects of this course. Just as a calculator does arithmetic for you so can Mathematica do symbolic calculation and display many results graphically (most of the graphs in these slides were produced using Mathematica). An ability to use Mathematica will be useful throughout the whole degree, so it is really worth making an effort to understand the sample programs.

Other useful references:


CONTENTS

Review ................................................................................. 6
Exponential and log functions ................................................. 7
Geometric progression (series) .................................................. 9
Complex numbers - a quick review ......................................... 11
Facts about complex numbers ................................................. 13
The modulus of a complex number .......................................... 14
The complex plane (Argand diagram) and roots of unity .......... 16
Rotations in the complex plane ................................................. 18
Proof of De Moivre’s theorem ................................................... 21

I Sequences and series ................................................................ 22
Sequences - definition .......................................................... 22
Summation notation ............................................................ 23
Product notation ............................................................... 23
Sequences of reals ............................................................. 24
Definition of convergence of a sequence .................................. 28
General principle of convergence ............................................. 31
Properties of limits ............................................................. 32
Proof of theorem 5 ............................................................ 33
Bounded monotonic sequences .............................................. 36
Infinite series ................................................................... 38
Exercises on geometric progressions ...................................... 41
* A population growth model ................................................. 42
Properties of series ............................................................. 45
\(\sum 1/n\) diverges ........................................................... 46
If \(\sum a_n\) converges then \(\lim a_n = 0\) ............................. 47
Alternating series theorem ................................................... 48
Absolute convergence ......................................................... 50
Products of series ............................................................... 51
Tests for convergence .......................................................... 52
The comparison test ........................................................... 53
The ratio test ................................................................... 55
Alternative definition of cos and sin ....................................... 56
The integral test ............................................................... 58
Power series ...................................................................... 63
The exponential power series ................................................. 64
Properties of \(\exp(x)\) .......................................................... 65
Radius of convergence ......................................................... 66
* Continuity of power series ............................................... 67
* Differentiation of power series ........................................... 68
Alternative definition of \(\log(x)\) ............................................. 69
Law of powers (revisited) ...................................................... 70
Complex power series ........................................................ 71
Euler’s theorem ............................................................... 72
Properties of cos and sin ...................................................... 74

II Calculus: Continuity, Differentiation and Integration .............. 76
Limits ............................................................................ 76
Continuous functions .......................................................... 79
Definition of continuity ....................................................... 80
Big O, Little o and - ......................................................... 83
LIST OF FIGURES

Figure 1 The Exponential and Natural Logarithm functions. ............................................... 7
Figure 2 The Argand diagram and the sixth roots of unity. ............................................... 16
Figure 3 Convergence to $e$. .................................................................................. 25
Figure 4 Convergence to $\sqrt{2}$. ........................................................................ 26
Figure 5 Convergence to $e$. .................................................................................. 38
Figure 6 Population growth for small $r$. ........................................................ 43
Figure 7 The amazing results for larger $r$. .......................................................... 44
Figure 8 Example of Alternating Series theorem - convergence to $\log_2$. ................. 49
Figure 9 Comparison of the integral and the sum. ................................................... 58
Figure 10 Graph of $\sin(1/x)$ near zero. ........................................................................ 81
Figure 11 The slope of a function. ............................................................................. 85
Figure 12 Forming upper and lower sums. ........................................................... 100
Figure 13 The surface $\sqrt{x^2 + y^2}$ plotted using the Mathematica Plot3D command. ... 121
Figure 14 The graph of a differentiable function whose mixed partial derivatives at the origin do not have the same sign. ................................................................. 122
Figure 15 Counter example to the Only-Extreme-Point-In-Town test. ............................ 124
Figure 16 Forming upper and lower sums. ........................................................... 135
Figure 17 Velocity against time. ................................................................................ 138
Review

Last semester we sketched the structure of numbers (see the supplementary file Numbers.pdf)

\[ \mathbb{N} \subseteq \mathbb{Z} \subseteq \mathbb{Q} \subseteq \mathbb{R} \subseteq \mathbb{C} \]

The integers \( \mathbb{Z} \) form a ring and \( \mathbb{Q}, \mathbb{R}, \mathbb{C} \) form fields. The distinguishing characteristic of a field is the existence of a multiplicative inverse for every non-zero element.

There are other fields. In particular there are finite fields \( \mathbb{F}(p) \), one for each prime \( p \).

\( \mathbb{R} \) is constructed from \( \mathbb{Q} \) using topological considerations (essentially \( \mathbb{R} \) plugs all the ‘holes’ in \( \mathbb{Q} \) - technically we say \( \mathbb{R} \) is ‘complete’).

The complex numbers have the surprising property that they are algebraically closed, i.e. every polynomial equation of degree \( n \) with coefficients in \( \mathbb{C} \) has exactly \( n \) solutions in \( \mathbb{C} \).
Exponential and log functions.

Functions of the type $y = a^x$ ($a > 0$) are called exponential functions.

One particular exponential function has the important property that its slope at every point is equal to the value of the function (i.e. the function is its own derivative). We write this function as

$$y = e^x = \text{Exp}(x)$$

where $e \approx 2.7182818...$ is a particularly important number. We shall see how to compute $e$ later.
The vital properties of $\text{Exp}$ are as follows

\[ e^x e^y = \text{Exp}(x) \cdot \text{Exp}(y) = \text{Exp}(x + y) = e^{x+y} \]

\[ e^0 = \text{Exp}(0) = 1 \]

From the graph we can see that $\text{Exp}(x)$ is a steadily increasing function of $x$ and so we can define an inverse function. This function is called $\log_e(x)$ (or $\ln(x)$ or often just $\log(x)$), the *natural logarithm* of $x$.

Notice that $\log_e(x)$ ($x \in \mathbb{R}$) is only defined if $x > 0$.

The vital properties of $\log_e(x)$ are

\[ y = \text{Exp}(x) \text{ if and only if } \log_e(y) = x \]

\[ \log_e(ab) = \log_e(a) + \log_e(b) \]

\[ \log_e(1) = 0 \]

It is interesting to observe that if one has a way of computing the log function (and reversing the process) then the problem of multiplying two numbers is reduced to an addition. (This was the basis of doing arithmetic using logarithms before electronic calculators came along.)
Geometric progression (series).

Suppose we want to sum (as a closed formula) the expression

\[ S_n = a + ar + ar^2 + ... + ar^n \]

(where \(a\) and \(r\) are real numbers) involving \(n+1\) terms. One way to do this is as follows.

Multiply the equation both sides by \(r\) to obtain

\[ rS_n = ar + ar^2 + ... + ar^n + ar^{n+1} \]

Hence, subtracting we have

\[ S_n - rS_n = a - ar^{n+1} \]

and so

\[ (1 - r)S_n = a(1 - r^{n+1}) \]

Now what we should like to do is to divide both sides by \(1 - r\). But if \(r = 1\) this would be a first degree mathematical felony (roughly equivalent to murder). In that case we can’t divide by \(1 - r\) because it is zero. On the other hand if \(r = 1\) the answer is obviously \(S_n = (n + 1)a\) so there is no difficulty in finding the sum. If \(r \neq 1\) we can avoid death by lethal injection and divide by \(1 - r\). We then obtain

\[ S_n = \frac{a(1 - r^{n+1})}{(1 - r)} \]
This expression for $S_n$ is known as the sum of a geometric progression and we can use it to illustrate a number of issues raised later in the course. Notice the formula is the sum of $n + 1$ terms (not $n$) because we have a first term $a$, which we can consider as being multiplied by $r$ to the zeroth power.
Complex numbers - a quick review

We create an ‘algebraic extension’ of the reals by adjoining the $\sqrt{-1}$.

We can do this formally as follows.

**Definition.** A complex number is an ordered pair $(a, b)$ of real numbers for which addition $\oplus$ and multiplication are defined by

**Addition rule:** $(a, b) \oplus (c, d) \equiv (a + c, b + d)$

**Multiplication rule:**

$$(a, b) \otimes (c, d) \equiv (a*c - b*d, a*d + b*c)$$

Here $*$ and $+$ are ordinary multiplication and addition between real numbers. We denote the set of complex numbers by $\mathbb{C}$.

**Exercise 1.** Addition and multiplication so defined are commutative, associative and multiplication is distributive over addition.

**Exercise 2.** $\mathbb{C}$ is a field. (Check that the axioms for a field are satisfied).
We can identify the reals $\mathbb{R}$ with the subset $\{(a, 0) \in \mathbb{C} : a \in \mathbb{R}\}$, we could write this identification as $(a, 0) \in \mathbb{C} \leftrightarrow a \in \mathbb{R}$.

Now we note that $(0, 1) \otimes (0, 1) \equiv (-1, 0)$ and that $(-1, 0) \in \mathbb{C}$ has been identified with $-1 \in \mathbb{R}$. Hence in the complex numbers the equation $x^2 = -1$ is soluble (its solution is $x = (0, 1) \leftrightarrow i$).

We write $i = (0, 1)$ and $(x, y) \in \mathbb{C}$ as $x + iy$ and drop all the circles around things.

Now here is a very remarkable fact:

**Theorem 1** (Gauss). Every polynomial equation of degree $n \geq 1$

$$a_n x^n + a_{n-1} x^{n-1} + \ldots + a_0 = 0 \quad (a_i \in \mathbb{C})$$

has exactly $n$ solutions in $\mathbb{C}$ (some of these solutions may be the same one repeated).

Thus having adjoined just one element $i$ to the reals we have got a field which is algebraically closed, i.e. the solutions of all polynomial equations with coefficients in $\mathbb{C}$ also lie in $\mathbb{C}$.
Facts about complex numbers.

E.g The quadratic equation $ax^2 + bx + c = 0$ ($a \neq 0$, $b$, $c \in \mathbb{C}$) has exactly two solutions in $\mathbb{C}$ and these are given by

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

There are general formulae to solve cubic and quartic polynomial equations in terms of their coefficients but it is a remarkable theorem (due to Galois) that there are no such general simple algebraic formulae for polynomial equations of degree $n \geq 5$.

Now there is a price to be paid for all these nice properties of $\mathbb{C}$. It is:

- $\mathbb{C}$ is not ordered.

We cannot order $\mathbb{C}$ with an order inherited from $\mathbb{R}$.

This is essentially because $\mathbb{R}$ is a one dimensional linear set and $\mathbb{C}$ is basically a two dimensional object.
The modulus of a complex number

We can extend the absolute value function \(| . |\) from the reals to the complex numbers by

**Definition.** If \( z = x + iy \) then we define

\[
|z| = \sqrt{x^2 + y^2}
\]

Then \( |z| \geq 0 \ \forall \ z \in \mathbb{C} \), and this definition is also consistent with the earlier one \( |x| = \max\{x, -x\} \) for \( x \in \mathbb{R} \).

If \( z = x + iy \) and \( \bar{z} = x - iy \) then \( \bar{z} \) is called the *complex conjugate* of \( z \).

**Note.** \( z^*\bar{z} = (x + iy)(x - iy) = x^2 + y^2 = |z|^2 \).

**Exercise 3.** Using the \( \bar{z} \) notation prove that with this definition

\[
|z + w| \leq |z| + |w| \quad \forall \ z, w \in \mathbb{C}
\]

(see p 146 for solution).

Notice that if a quadratic equation with real coefficients has two complex roots then one root is the complex conjugate of the other.
In general the complex zeros of any polynomial with real coefficients can be grouped into complex conjugate pairs.
The complex plane (Argand diagram) and roots of unity

Figure 2 The Argand diagram and the sixth roots of unity.

The figure (top) shows how complex numbers can be regarded as points in a plane (the complex plane or Argand diagram).
Notice that any complex number $z$ can be written as

$$z = |z|(\cos \theta + i \sin \theta)$$

for some $\theta$, where $-\pi < \theta \leq \pi$. The angle $\theta$ is called the *argument* of $z$, written $\theta = \text{Arg}(z)$ and here

$$|\cos \theta + i\sin \theta| = \sqrt{(\cos^2 \theta + \sin^2 \theta)} = 1$$

The figure (bottom) also shows the complex sixth roots of unity, i.e. the solutions of the equation $z^n = 1$ for $n = 6$.

If we write

$$\omega = \frac{1}{2} + \frac{\sqrt{3}}{2} i = \cos \left(\frac{2\pi}{6}\right) + i \sin \left(\frac{2\pi}{6}\right)$$

then we observe that the zeros of $z^6 - 1$ can be written as $\omega, \omega^2, ..., \omega^5, \omega^6 = 1$. For this reason $\omega$ is called a *primitive* sixth root of unity (its powers generate all the others). Observe that $\omega^2$ is not a primitive sixth root of unity.

We also notice that

$$1 + \omega + \omega^2 + ... + \omega^5 = 0$$

These observations generalise to $n$th roots of unity for any positive integer $n > 1$. 
Rotations in the complex plane

If the point \((x, y)\) is rotated about the origin through angle \(\theta\) the resulting point is given by the equations

\[
x' = x\cos\theta - y\sin\theta
\]
\[
y' = x\sin\theta + y\cos\theta
\]

Written in matrix form these equations become

\[
\begin{pmatrix}
x' \\
y'
\end{pmatrix} =
\begin{pmatrix}
\cos\theta & -\sin\theta \\
\sin\theta & \cos\theta
\end{pmatrix}
\begin{pmatrix}
x \\
y
\end{pmatrix}
\]

If we imagine the point \((x, y)\) a point \(x + iy\) in the complex plane then there is an easy way to remember these formulae. We simply observe that multiplying \(x + iy\) by \(\cos\theta + isin\theta\) is equivalent to a rotation through \(\theta\). Thus

\[
x' + iy' = (\cos\theta + isin\theta)(x + iy)
\]
\[
= (x\cos\theta - y\sin\theta) + i(x\sin\theta + y\cos\theta)
\]

Remark: In three dimensions we need \(3 \times 3\) matrices to describe a rotation about some axis (needed in many computer graphics applications). Unfortunately, there is no easy way (as above) to remember these equations which are quite complicated (see Mathematica file Rots3D.nb).
Exercise 4. Use the same idea to deduce the addition formulae for \( \cos(A + B) \) and \( \sin(A + B) \). Starting with the complex number \( x + iy \) perform first a rotation through angle \( A \) and second a further rotation through angle \( B \).

Solution. Rotate through angle \( A \)

\[
z' = (\cos A + i \sin A)(x + iy)
= (x \cos A - y \sin A) + i(x \sin A + y \cos A)
\]

Rotate through angle \( B \)

\[
z'' = (\cos B + i \sin B)(x' + iy')
= (\cos B + i \sin B)(\cos A + i y \sin A)(x + iy)
= ((\cos B \cos A - \sin B \sin A) + i(\cos B \sin A + \sin B \cos A))(x + iy)
\]

Hence \( \cos(A + B) = \cos A \cos B - \sin A \sin B \) and similarly \( \sin(A + B) = \sin A \cos B + \cos A \sin B \).

Alternatively using matrix multiplication. Rotation through \( A \) followed by rotation through \( B \) is given by

\[
\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \cos B & -\sin B \\ \sin B & \cos B \end{pmatrix} \begin{pmatrix} \cos A & -\sin A \\ \sin A & \cos A \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}
= \begin{pmatrix} \cos B \cos A - \sin B \sin A & -\cos B \sin A - \sin B \cos A \\ \sin B \cos A + \cos B \sin A & -\sin B \sin A + \cos B \cos A \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}
\]

which is equivalent to the previous formulae.
Exercise. 5 Prove by induction on $n$ that for any positive integer $n$ we have

Theorem 2 (De Moivre). For all positive integers $n$ and any real $\theta$

$$(\cos\theta + i \sin\theta)^n = \cos n\theta + i \sin n\theta$$

Hint. You will need to assume the addition formulae for $\cos$ and $\sin$ (which we just proved), i.e.

$$\cos(\theta + \varphi) = \cos \theta \cos \varphi - \sin \theta \sin \varphi$$

$$\sin(\theta + \varphi) = \cos \theta \sin \varphi + \sin \theta \cos \varphi$$
Proof of De Moivre’s theorem

Base case. The theorem is obviously true for $n = 1$.

Induction step. Suppose that for some particular integer $k$ we have

$$(\cos \theta + i \sin \theta)^k = \cos k \theta + i \sin k \theta$$

Then multiplying both sides by $\cos \theta + i \sin \theta$ we have

$$(\cos \theta + i \sin \theta)^{k+1} = (\cos \theta + i \sin \theta)(\cos k \theta + i \sin k \theta)$$

$$= (\cos \theta \cos k \theta - \sin \theta \sin k \theta) + i (\cos \theta \sin k \theta + \sin \theta \cos k \theta)$$

$$= \cos(\theta + k \theta) + i \sin(\theta + k \theta) \quad \text{(by the addn formulae)}$$

$$= \cos(k + 1) \theta + i \sin(k + 1) \theta$$

Thus IF the result is true for $n = k$ THEN it is true for $n = k + 1$ and the theorem follows by the axiom of mathematical induction.
I Sequences and series

Sequences - definition

We can exploit the natural ordering of \( \mathbb{N} \) to impose an order on any infinite sequence of numbers. We make the

**Definition.** A *sequence* is a (countable) set which can be put into 1-1 correspondence with \( \mathbb{N} \). The sequence is ordered using the order inherited from its 1-1 correspondence with \( \mathbb{N} \).

If the sequence can be put into 1-1 correspondence with a finite subset of \( \mathbb{N} \) then it is called a *finite* sequence, otherwise it is called an *infinite* sequence.

We are mostly interested in infinite sequences.

We usually talk of the sequence

\[(s_1, s_2, s_3, ... ) = (s_n : n \in \mathbb{N})\]

or simply the sequence \( s_n \) \((n = 1, 2, 3, ...).

E.g. The sequence of squares \( s_n = n^2 \) \((n = 1, 2, 3, ...).\)
Summation notation Suppose $a_1, a_2, ..., a_n$ are numbers to be added together. We write

$$\sum_{k=1}^{n} a_k = a_1 + a_2 + ... + a_n$$

Examples.

$$\sum_{k=1}^{n} k = \frac{1}{2}n(n+1)$$

$$\sum_{k=1}^{n} k^2 = \frac{1}{6}n(n + 1)(2n + 1)$$

Product notation. Suppose $a_1, a_2, ..., a_n$ are numbers to be multiplied together. We write

$$\prod_{k=1}^{n} a_k = a_1 \cdot a_2 \cdot ... \cdot a_n$$

For more examples see page 25 of the coursebook.
Sequences of reals

We saw earlier that

**Definition.** A *sequence* is a countable set with an order inherited from its 1-1 correspondence with $\mathbb{N}$.

We now concentrate on sequences of real or complex numbers, i.e. (usually infinite) ordered lists of the form

$$a_1, a_2, a_3, \ldots,$$

where $a_n \in \mathbb{R}$ or $\mathbb{C}$ (all $n$).

Here are some examples:

$$a_n = n \text{ (all } n \in \mathbb{N})$$

i.e. The (infinite) sequence 1, 2, 3, ...

$$a_n = 1/n \text{ (all } n \in \mathbb{N})$$

i.e. The (infinite) sequence 1/1, 1/2, 1/3, 1/4,...
\[ a_n = (1 + \frac{1}{n})^n, \text{ i.e. The (infinite) sequence} \]

<table>
<thead>
<tr>
<th>n</th>
<th>(a_n)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2.</td>
</tr>
<tr>
<td>2</td>
<td>2.25</td>
</tr>
<tr>
<td>3</td>
<td>2.37037037037037</td>
</tr>
<tr>
<td>4</td>
<td>2.44140625</td>
</tr>
<tr>
<td>5</td>
<td>2.48832</td>
</tr>
<tr>
<td>6</td>
<td>2.521626371742113</td>
</tr>
<tr>
<td>7</td>
<td>2.546499697040713</td>
</tr>
<tr>
<td>8</td>
<td>2.565784513950348</td>
</tr>
<tr>
<td>9</td>
<td>2.581174791713197</td>
</tr>
<tr>
<td>10</td>
<td>2.5937424601</td>
</tr>
<tr>
<td>...</td>
<td></td>
</tr>
</tbody>
</table>

In fact this sequence approaches \(e = 2.7182818284\ldots\) as \(n\) becomes very large (but it does so rather slowly).

Figure 3 Convergence to \(e\).
Now consider the sequence defined by $a_1 = 1$ and

$$a_n = \frac{a_{n-1}^2 + 2}{2a_{n-1}} \quad (n \geq 2)$$

This gives

<table>
<thead>
<tr>
<th>$n$</th>
<th>$a_n$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>1.5</td>
</tr>
<tr>
<td>3</td>
<td>1.4166666666666667</td>
</tr>
<tr>
<td>4</td>
<td>1.41421568627451</td>
</tr>
<tr>
<td>5</td>
<td>1.41421356237469</td>
</tr>
<tr>
<td>6</td>
<td>1.414213562373095</td>
</tr>
<tr>
<td>7</td>
<td>1.414213562373095</td>
</tr>
<tr>
<td>8</td>
<td>1.414213562373095</td>
</tr>
<tr>
<td>9</td>
<td>1.414213562373095</td>
</tr>
<tr>
<td>10</td>
<td>1.414213562373095</td>
</tr>
<tr>
<td>...</td>
<td></td>
</tr>
</tbody>
</table>

Figure 4 Convergence to $\sqrt{2}$.

Exercise 6. Suppose for a moment that we could prove that this sequence $a_n$ approached some number $l$ (say) as $n$ became very large. What is $l$ and why?
Solution

IF $a_n$ approaches $l$ as $n$ becomes large THEN $l$ must satisfy the approximate equality

$$l \approx \frac{l^2 + 2}{2l}$$

i.e.

$$2l^2 \approx l^2 + 2$$

i.e.

$$l^2 \approx 2 \quad \text{or} \quad l \approx \sqrt{2}$$

where we choose the positive square root because the sequence consists of positive terms.

Note this says that IF $l$ exists THEN it must be $\sqrt{2}$, but by itself this does not prove that $l$ exists.

*Note.* One way to prove that $l$ exists would be to show that for $n \geq 2$ the sequence $a_n$ is monotonic decreasing and bounded below. We can then use a standard theorem (Theorem 6) to infer that $l$ exists.

Regardless of the existence question this example is quite interesting to us: it shows a way of constructing an algorithm which computes $\sqrt{2}$ very rapidly. (Later we give more powerful methods for computing powers of numbers.)
Definition of convergence of a sequence.

Some sequences (like the last two above) have the property that as \( n \) becomes very large the values \( a_n \) become eventually as close as we like to some fixed real number.

We can formalise this by playing the \((\varepsilon, N)\)-game. You give me a (small) \( \varepsilon > 0 \) and I have to find an \( N = N(\varepsilon) \) such that

\[
|a_n - l| < \varepsilon \quad \text{for all} \quad n > N = N(\varepsilon)
\]

If no matter how small (but always positive) your \( \varepsilon \) is I can always find an \( N \) which satisfies this condition then I win the \((\varepsilon, N)\)-game and we call \( l \) the limit of the sequence \( a_n \) as \( n \) tends to infinity.

**Definition.** We say the sequence \( a_n \to l \) as \( n \to \infty \) if for every \( \varepsilon > 0 \) there exists \( N = N(\varepsilon) \) such that

\[
|a_n - l| < \varepsilon \quad \forall \quad n > N
\]

If this is the case we say the sequence \( a_n \) is convergent to the limit \( l \) and write

\[
\lim_{n \to \infty} a_n = l
\]
NOTE: The above definition is probably the most important single idea that you will meet in mathematical analysis.

Example. The sequence \( a_n = (-1)^n/n \to 0 \) as \( n \to \infty \). If you give me \( \varepsilon > 0 \) then I will choose \( N = \text{Floor}[1/\varepsilon] + 1 \) so that

\[
|a_n - 0| = \left| \frac{(-1)^n}{n} \right| < \frac{1}{N} < \varepsilon
\]

whenever \( n > N \). (Floor\([x]\) is the biggest integer less than or equal to \( x \))

Thus I win the \((\varepsilon, N)\)-game and we can write

\[
\lim_{n \to \infty} \frac{(-1)^n}{n} = 0
\]

The \((\varepsilon, N)\)-game captures the idea that we are interested in the behaviour of \( a_n \) as \( n \) becomes very large.
Exercises. 7 Which of the following sequences converge and to what limit? For those that do, show how to win the $(\epsilon, N)$-game.

(i) $a_n = (n-1)/(n+1) \ (n = 1, 2, 3, \ldots)$

(ii) $a_n = \sin(1/n) \ (n = 1, 2, 3, \ldots)$

(iii) $a_n = 1 + (-1)^n/(2^n) \ (n = 1, 2, 3, \ldots)$

One way to work on these problems is to use Mathematica to plot the sequences for $1 \leq n \leq 1000$ (say). See the Mathematica file Sequences.nb for examples. This should give you a clear idea of what is happening.
General principle of convergence.

You will observe that one problem with the definition of a limit for a sequence is that one needs to know what the limit is before one can construct a proof via the \((\varepsilon, N)\)-game that the sequence is convergent (i.e. the limit exists).

This difficulty is not as bad as it might appear.

Firstly, in many cases it is fairly clear what the limit ought to be. Secondly, it is possible to derive a criterion for convergence which does not depend on knowing the limit. In fact

**Theorem 3 (General Principle of Convergence - GPC).** The sequence \(a_n\) \((n = 1, 2, 3, \ldots)\) is convergent if and only if for every \(\varepsilon > 0\) there exists \(N = N(\varepsilon)\) such that

\[
|a_n - a_m| < \varepsilon \quad \forall \quad n, m > N
\]

Sequences which satisfy this condition are called Cauchy sequences. So another way to state the theorem is that a sequence is convergent if and only if it is a Cauchy sequence. We do not go into the proof, but merely remark that the theorem says that a sequence is convergent if and only if for sufficiently large \(n\) all subsequent terms are arbitrarily close together. This criterion of convergence is very useful for proving certain types of theorem (e.g. Theorem 11).
Properties of limits.

Some basic consequences of the definition of a limit whose derivations are straightforward:

**Theorem 4.** If \( a_n \ (n = 1, 2, 3, \ldots) \) is a convergent sequence the limit is unique.

*Note.* This justifies speaking of the limit of a sequence (i.e. there can only be one).

**Proof.** Suppose \( l_1 \neq l_2 \) both satisfy the definition of convergence. Then for every \( \varepsilon > 0 \) there exists \( N_1(\varepsilon) \) and \( N_2(\varepsilon) \) such that both
\[
|a_n - l_1| < \varepsilon \quad \forall \ n > N_1
\]
\[
|a_n - l_2| < \varepsilon \quad \forall \ n > N_2
\]
hold \( \forall \ n > N = \max\{N_1(\varepsilon), N_2(\varepsilon)\} \). Thus
\[
|l_1 - l_2| = |(l_1 - a_n) + (a_n - l_2)|
\]
\[
\leq |l_1 - a_n| + |a_n - l_2|
\]
\[
\leq \varepsilon + \varepsilon = 2\varepsilon \quad \forall \ n > N
\]

Hence by choosing \( \varepsilon > 0 \) smaller than half the distance between \( l_1 \) and \( l_2 \) we obtain a contradiction. Hence \( l_1 = l_2 \).
Theorem 5. If \( \lim a_n = r \) and \( \lim b_n = s \) then

(i) \( \lim (a_n + b_n) = r + s \)
(ii) \( \lim a_n b_n = rs \)
(iii) If \( s \neq 0 \) then \( \lim \frac{a_n}{b_n} = \frac{r}{s} \)

Proof of theorem 5.

We prove (i) and (iii) and leave (ii) as an exercise for the reader.

(i) By the triangle inequality we have

\[
|a_n + b_n - (r + s)| \leq |a_n - r| + |b_n - s|
\]

Since \( a_n \to r \) as \( n \to \infty \) by definition given \( \varepsilon > 0 \) \( \exists N_1 = N_1(\varepsilon) \) such that \( |a_n - r| < \varepsilon/2 \ \forall \ n > N_1 \).

Similarly, since \( b_n \to s \) as \( n \to \infty \) by definition \( \exists N_2 = N_1(\varepsilon) \) such that \( |b_n - s| < \varepsilon/2 \ \forall \ n > N_2 \). Hence

\[
|(a_n + b_n) - (r + s)| \leq \varepsilon/2 + \varepsilon/2 = \varepsilon
\]

for all \( n > N = \max\{N_1, N_2\} \). i.e. by the definition \( a_n + b_n \to r + s \) as \( n \to \infty \).
(iii) Proving this is a little more difficult.

We start by observing that, since \( s \neq 0 \),

\[
\frac{|a_n - r|}{b_n} = \frac{|sa_n - rb_n|}{|b_n s|} \leq \frac{|sa_n - sr + sr - rb_n|}{|b_n s|}
\]

\[
\leq \frac{1}{|b_n| |s|} \left| s(a_n - r) + r(s - b_n) \right|
\]

\[
\leq \frac{1}{|b_n| |s|} \left( |s| |a_n - r| + |r| |b_n - s| \right)
\]

\[
\leq \frac{1}{|b_n|} |a_n - r| + \frac{1}{|b_n|} |r| |b_n - s|
\]

Now plainly we can make the two terms on the right arbitrarily small as \( n \to \infty \) provided that \( 1/|b_n| \) can be kept bounded, i.e. provided \( |b_n| \) can be bounded away from zero as \( n \to \infty \).

Now \( b_n \to s \) and \( s \neq 0 \), so taking \( \varepsilon = |s|/2 \) in the definition of convergence \( \exists N_0 = N_0(s) \) such that

\[
|b_n - s| < \frac{|s|}{2} \quad \forall \ n > N_0(s)
\]
Thus
\[
\left| \frac{a_n - r}{b_n} - \frac{r}{s} \right| \leq \frac{2}{|s|} |a_n - r| + \frac{2|r|}{|s|^2} |b_n - s|
\]

Now given \( \varepsilon > 0 \) choose \( N_1 = N_1(\varepsilon, s) \) so that

\[
|a_n - r| < \varepsilon |s|/4
\]

and \( N_2 = N_2(\varepsilon, r, s) \) so that

\[
|b_n - s| < \varepsilon |s|^2/(4|r|)
\]

Then
\[
\left| \frac{a_n - r}{b_n} - \frac{r}{s} \right| < \frac{2}{|s|} \frac{\varepsilon |s|}{4} + \frac{2|r|}{|s|^2} \frac{\varepsilon |s|^2}{4|r|} = \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon
\]

provided \( n > N = \max\{N_0, N_1, N_2\} \).

These proofs are quite technical - you will not be asked to construct such proofs but you should try to see how the definition of convergence allows us to prove these results.
Bounded monotonic sequences

**Definition.** A sequence $a_n$ $(n = 1, 2, 3, ...)$ is said to be *monotonic increasing* (or decreasing) if $a_n \leq a_{n+1}$ (or $a_n \geq a_{n+1}$) for all $n \geq 1$.

**Theorem 6.** A bounded monotonic sequence is convergent.

**Proof.** Let $a_n$ be a bounded monotonic sequence. For definiteness suppose the sequence is monotonic increasing and bounded above.

Then there exists a *least* (there is a technical point here which we have skipped over) upper bound $U \in \mathbb{R}$ for the set of points defined by the sequence, i.e. $a_n \leq U$ for all $n \geq 1$ and $U$ is the least such number.

We assert that $a_n \to U$ as $n \to \infty$. For given $\varepsilon > 0$, $U - \varepsilon$ is *not* an upper bound (because $U$ is the least upper bound). Hence $\exists N = N(\varepsilon) \in \mathbb{N}$ such that

$$U - \varepsilon \leq a_N \leq U$$
But $a_N \leq a_n$ for all $n \geq N$ (because the sequence is monotonic increasing). Hence

$$U - \varepsilon \leq a_N \leq a_n \leq U \quad \forall \quad n \geq N$$

But this implies (from the definition of convergence) that

$$\lim_{n \to \infty} a_n = U$$

and the theorem is proved.

**Example.** We return to the sequence $a_n = (1 + 1/n)^n$.

**Thought.** Looking at the table, or Figure 4, for this sequence suggests that it is monotonic increasing. If we could prove that it is bounded above we could infer that it has a limit.
Infinite series

We now seek to assign a meaning to such expressions as

\[ 1 + \frac{1}{1!} + \frac{1}{2!} + \frac{1}{3!} + \ldots \]

which is an example of an infinite series. Let us calculate

\[ S_n = 1 + \frac{1}{1!} + \frac{1}{2!} + \ldots + \frac{1}{n!} \]

\[
\begin{array}{c|c}
 n & S_n \\
--- & --- \\
 1 & 2.000000000000000 \\
 2 & 2.500000000000000 \\
 3 & 2.666666666666666 \\
 4 & 2.708333333333333 \\
 5 & 2.716666666666667 \\
 6 & 2.718055555555555 \\
 7 & 2.718253968253968 \\
 8 & 2.71827876984127 \\
 9 & 2.718281525573192 \\
 10 & 2.718281801146384 \\
\end{array}
\]

Figure 5 Convergence to $e$. 
In fact this sequence approaches

\[ e = 2.7182818284... \]

as \( n \) becomes very large (notice that it does so rather faster than our previous sequence which generated \( e \) as a limit).

This suggests how we can assign a meaning to the idea of the sum of an infinite series.

Let

\[ \sum_{n=1}^{\infty} a_n = a_1 + a_2 + a_3 + \ldots \]

be an infinite series and define

\[ s_n = a_1 + a_2 + \ldots + a_n \quad (n = 1, 2, 3, \ldots) \]

We call \((s_n)\) the sequence of \(n\)-th partial sums of the series \( \sum a_n \).

We can use this idea to define the sum of an infinite series.

**Definition.** We say that the infinite series \( \sum a_n \) converges to the sum \( s \) if and only if the sequence \( s_n \) \((n = 1, 2, 3, \ldots)\) of partial sums of the series converges to \( s \).
If this is the case we write

\[ s = \sum_{n=1}^{\infty} a_n \]

**Examples.**

\[ \sum_{n=1}^{\infty} \frac{1}{2^n} = \frac{1}{2} + \frac{1}{2^2} + \frac{1}{2^3} + \ldots = 1 \]

\[ \sum_{n=1}^{\infty} \frac{1}{n^2} = 1 + \frac{1}{2^2} + \frac{1}{3^2} + \ldots = \zeta(2) = \frac{\pi^2}{6} \]

\[ \sum_{n=1}^{\infty} \frac{1}{n} = 1 + \frac{1}{2} + \frac{1}{3} + \ldots = ? \]

\[ \sum_{n=1}^{\infty} \frac{(-1)^n}{n} = 1 - \frac{1}{2} + \frac{1}{3} - \ldots = ? \]

The first of these is fairly easily established (for example by induction show \( s_n = 1 - 1/(2^n) \)). It is an example of a Geometric Series. The remaining three pose various problems.
Exercises on geometric progressions

Let $a, r \in \mathbb{R}$. Consider the finite series

$$S_n = a + ar + ar^2 + ... + ar^n$$

Show that

$$S_n = \frac{a(1 - r^{n+1})}{1 - r} \quad (r \neq 1)$$

(we did this already earlier). Now consider what happens as $n \to \infty$. Under what circumstances does the (infinite) series converge/diverge?

(i) Trivial. The Sun is $93 \times 10^6$ miles away. An extremely large piece of paper is $1/1000$ inch thick. We proceed to fold the paper, doubling the thickness with each fold. How many times would we have to fold the paper to reach from Earth to the Sun?
(ii) *Deeper.* A population growth model. We normalise the size of the population to [0, 1].

Let \( x_0 \in (0, 1) \) be the initial population and \( x_n \) its size after \( n \) years. The growth rate \( R \) is the relative increase per year

\[
R = \frac{x_{n+1} - x_n}{x_n}
\]

If this is a constant (\( r \) say) the dynamic law is

\[
x_{n+1} = (1 + r)x_n
\]

After \( n \) years the population size is

\[
x_n = (1 + r)^n x_0
\]

To limit this exponential growth Verhulst assumed the rate \( R \) to vary with the population size: he postulated the size dependent rate \( R \) to be proportional to \( 1 - x_n \).
We write $R = r(1 - x_n)$. Hence if $x_n << 1$ and $r$ small the population may still increase until it stops growing when $x_n \approx 1$.

The new dynamical rule then becomes

$$x_n = (1 + r)x_n - rx_n^2$$

*For what growth rates can populations attain stability?*
**Exercise.** (*Mathematica* file pops.nb)

Suppose \( x_0 = 0.1 \) and use *Mathematica* to simulate population growth for \( r \) in the range 0.5 to 3.

For \( 1.9 \leq r \leq 3 \) plot the \( r \) value horizontally and plot iterations of the population for \( 120 \leq n \leq 1000 \) vertically (or more if time permits - this program is going to take a while to give a detailed picture - it also requires an awful lot of memory).

![Graph](image.png)

Figure 7 The amazing results for larger \( r \).

Comment: The population is following the period doubling route to chaos.
In examining questions such as these we have to ask two questions:

1. Does the series converge - can we prove it (one way or the other)?

2. If it does converge what can we say about the sum? The sum may exist but it may not have some nice expression like $\pi^2/6$. Even if it has this may be quite hard to prove.

**Properties of series**

The basic properties of series follow from those of the sequences of partial sums. Thus

**Theorem 7.** If $\sum a_n \ (n = 1, 2, 3, \ldots)$ is a convergent series the sum is unique.

**Proof.** This is immediate from Theorem 4.

**Theorem 8.** If $\sum a_n = r$ and $\sum b_n = s$ then $\sum (a_n + b_n) = r + s$.

**Proof.** This is immediate from Theorem 5(i).

In forming products we have to be more careful (ratios don’t work at all).
Let us return to the infinite series $\sum a_n$, where $a_n = 1/n$, i.e.

$$\sum_{n=1}^{\infty} \frac{1}{n} = 1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \frac{1}{5} + \ldots$$

Consider

$$a_3 + a_4 = \frac{1}{3} + \frac{1}{4} > \frac{1}{4} + \frac{1}{4} = \frac{1}{2}$$

$$a_5 + a_6 + a_7 + a_8 = \frac{1}{5} + \frac{1}{6} + \frac{1}{7} + \frac{1}{8} > \frac{1}{8} + \frac{1}{8} + \frac{1}{8} + \frac{1}{8} = \frac{4}{8} = \frac{1}{2}$$

$$\sum_{i = 2^{k-1} + 1}^{2^k} a_i > \frac{2^{k-1}}{2^k} = \frac{1}{2}$$

From this we can conclude that

$$s_{2^k} > \frac{k + 2}{2} \quad \forall \quad k \geq 1$$

i.e. we can make $s_n$ ($n = 2^k$) as large as we like by taking $n$ sufficiently large.

In such a case we say the series diverges to infinity.

Notice in this case $a_n \to 0$ even though $\sum a_n$ diverges.
If $\sum a_n$ converges then $\lim a_n = 0$

**Theorem 9.** If $\sum a_n$ converges then $\lim a_n = 0$.

**Proof.** Let $\sum a_n = s$ and let $s_n$ denote the $n$th partial sum of the series. By Theorem 4 we have

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

which proves the result. [1]

*Note.* The previous example shows *the converse is not true*, i.e.

$$\lim a_n = 0 \text{ DOES NOT imply } \sum a_n \text{ converges.}$$

The following theorem is a kind of converse.
Alternating series theorem.

Theorem 10. (Alternating series). Suppose for all $n \geq 1$ \(a_n \geq a_{n+1} \geq 0\) and \(\lim a_n = 0\), then \(\sum a_n\) may or may not converge, but

\[
\sum_{n=1}^{\infty} (-1)^{n-1}a_n = a_1 - a_2 + a_3 - a_4 + \ldots
\]

is convergent.

Example. This theorem implies that the series

\[
\sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n} = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \ldots
\]

is convergent. But what is the sum? (We can use the power series for \(\log(1 + x)\) on page 95 to show the sum is \(\log 2\).)

Proof of the theorem. Let \(s_n\) be the sequence of partial sums of the alternating series. We observe that, because \(a_n \geq a_{n+1} \geq 0\) for all \(n \geq 1\),

\[
s_2 \leq s_4 \leq s_6 \leq \ldots \leq s_5 \leq s_3 \leq s_1
\]

Thus \(s_2, s_4, s_6, \ldots\) is a monotonic increasing sequence bounded above and hence has a least upper bound (lub) \(s\) (say) \(\in \mathbb{R}\).
Similarly, $s_1, s_3, s_5, \ldots$ is a monotonic decreasing sequence bounded below and hence has a greatest lower bound (glb) $t$ (say) $\in \mathbb{R}$.

Thus for all $k \in \mathbb{N}$

$$s_{2k} \leq s \leq t \leq s_{2k+1}$$

But, since $\lim a_n = 0$,

$$s_{2k+1} - s_{2k} = a_{2k+1} \to 0 \text{ as } k \to \infty$$

Hence $s = t$ and $\lim s_n = s$ as required.

\[ s = 1 - 1/2 + 1/3 - \ldots \pm 1/n \]

Figure 8 Example of Alternating Series theorem - convergence to $\log_e 2$.
Absolute convergence

Definition. We say the series $\Sigma a_n$ is *absolutely convergent* if the series $\Sigma |a_n|$ is convergent.

This is a useful concept, especially when dealing with power series. The most useful theorem in this connection is:

**Theorem 11.** If a series is absolutely convergent then it is convergent.

**Proof.** Let $S_n$ be the partial sums of $\sum |a_n|$ and $s_n$ be the partial sums of $\sum a_n$ then for $n > m$ we have, by the triangle inequality,

$$|s_n - s_m| = |a_n + \ldots + a_{m+1}|$$

$$\leq |a_n| + \ldots + |a_{m+1}|$$

$$\leq |S_n - S_m|$$

Since $\sum |a_n|$ converges, by the General Principle of Convergence there exist $N = N(\varepsilon)$ so that the last term is less than $\varepsilon$ for all $n > m \geq N$. But then the inequality implies (again by the GPC) that $\sum a_n$ is convergent.
**Products of series**

The theorem on absolute convergence is often stated as ‘an absolutely convergent series is convergent’. It is worth remarking that the converse is false (example?).

Using this theorem we can prove the following useful result.

**Theorem 12** (Product of series). If

\[ s = \sum_{n=0}^{\infty} a_n \]

\[ t = \sum_{n=0}^{\infty} b_n \]

are absolutely convergent series and we define

\[ c_n = \sum_{i+j=n} a_i b_j \quad (0 \leq i, j \leq n) \]

Then \( \sum c_n \) is absolutely convergent and

\[ \sum_{n=0}^{\infty} c_n = \left( \sum_{n=0}^{\infty} a_n \right) \left( \sum_{n=0}^{\infty} b_n \right) = s.t \]
Tests for convergence

To help us with the problem of whether a given series converges there are a number of useful tests. We shall look at three:

*The comparison test.*

*The ratio test.*

and *The integral test.* (Even though we haven’t yet discussed integration.)

In general any test of convergence *may* tell us whether a given series converges or diverges, but sometimes *will not* tell us anything.

- Consequently the converse of any test of convergence is *invariably false* - we shall illustrate this with examples later.
The comparison test

This is the simplest and most obvious test for convergence/divergence. It relies on comparing the series for which the convergence/divergence is in question with a series for which the answer is already known.

Theorem 13. If \( a_n \geq 0 \) and \( b_n \geq 0 \) and there is an absolute constant \( K > 0 \) and an integer \( N \) such that

\[
0 \leq a_n \leq Kb_n \quad \forall \quad n > N
\]

then if \( \sum b_n \) is convergent so is \( \sum a_n \).

Proof. Since \( b_n \geq 0 \) and \( \sum b_n \) is convergent by the General Principle of Convergence (applied to the sequence of partial sums) \( \exists \ N_1 \) such that

\[
0 \leq \sum_{k=m}^{n} b_n < \varepsilon/K \quad \forall \quad n > m > N_1
\]

Thus whenever \( n > m > N_2 = \max\{N_1, N\} \) we have

\[
0 \leq \sum_{k=m}^{n} a_n \leq K \sum_{k=m}^{n} b_n < \varepsilon \quad \forall \quad n > m > N_2
\]

and again by the GPC it follows that \( \sum a_n \) is convergent.
**Example.** The series

$$\sum_{n = 2}^{\infty} \frac{1}{2^n \log n}$$

is convergent by comparison with the geometric series $\sum 2^{-n}$. To establish this we just need to observe that $1/\log n \leq 1/\log 2$ for all $n \geq 2$. 
The ratio test

**Theorem 14.** Suppose $a_n$ ($n = 1, 2, 3, ...$) is a sequence such that

$$\lim_{n \to \infty} \left| \frac{a_{n+1}}{a_n} \right| = \alpha$$

then

(i) If $\alpha < 1$ then the series is convergent.

(ii) If $\alpha > 1$ then the series is divergent.

(iii) If $\alpha = 1$ then we can say nothing (the series may or may not converge).

**Remark**: Notice the theorem refers to the limit of the ratio - not to just the ratio! (This is a very common mistake in exams.)

**Proof.** A proof can be constructed using the comparison test based on our knowledge of geometric series.

**Examples.**

1. $a_n = 1/(n!)$ then $a_{n+1}/a_n = 1/(n+1) \to 0 < 1$ as $n \to \infty$. Hence $\sum a_n$ converges.

2. $a_n = 1/n^2$ then $a_{n+1}/a_n = n^2/(n+1)^2 \to 1$ as $n \to \infty$. Hence the ratio test tells us nothing.
More examples.

**Alternative definition of cos and sin.** For all real \( \theta \), the functions \( \cos \theta \) and \( \sin \theta \) are defined by

\[
\cos \theta = 1 - \frac{\theta^2}{2!} + \frac{\theta^4}{4!} - \ldots = \sum_{n=0}^{\infty} (-1)^n \frac{\theta^{2n}}{(2n)!}
\]

\[
\sin \theta = \theta - \frac{\theta^3}{3!} + \frac{\theta^5}{5!} - \ldots = \sum_{n=0}^{\infty} (-1)^n \frac{\theta^{2n+1}}{(2n+1)!}
\]

For this definition to make sense the series have to converge. In fact both series converge for all real \( \theta \). For example, consider the series for \( \sin \theta \). Using the ratio test we have for \( \theta \neq 0 \) (the series obviously converges for \( \theta = 0 \))

\[
\left| \frac{a_{n+1}}{a_n} \right| = \frac{\theta^{2n+3}}{(2n+3)!} \cdot \frac{(2n+1)!}{\theta^{2n+1}} = \frac{\theta^2}{(2n+3)(2n+2)} \to 0 \quad \text{as} \quad n \to \infty
\]

Hence the series is absolutely convergent (hence convergent) for all real \( \theta \). Similarly for \( \cos \theta \).

Now \( \cos(0) = 1 \), and a little calculation based on the observation that \( \cos(2) \) is an alternating series shows that \( \cos(2) < -0.4 < 0 \).

If \( \cos \) is a continuous function (and we shall give reasons later why \( \cos \) as defined by this series is continuous at every point) this suggests that there is a zero of \( \cos \theta \) lying between 0 and 2.
Definition. (Of $\pi$) We define the least positive zero of $\cos \theta$ to be $\pi/2$.

Notice also that (from these definitions) $\cos \theta$ and $\sin \theta$ are even ($f(x) = f(-x)$ for all real $x$) and odd ($f(x) = -f(-x)$ for all real $x$) functions respectively.
The integral test

Figure 9 Comparison of the integral and the sum.

**Theorem 15.** Suppose that $f(x)$ is a real valued function such that $f(x)$ is monotonic decreasing, $f(x) \geq 0$ for all $x \geq 1$ and $\lim (x \to \infty) f(x) = 0$. Write $a_k = f(k)$ ($k \in \mathbb{N}$), then

$$\left\{ \sum_{k=1}^{\infty} a_k \text{ converges} \right\} \text{ if and only if } \left\{ \int_{1}^{\infty} f(x)dx \text{ converges} \right\}$$
The idea should be clear from Figure 9.

Because the function is required to be positive and monotonic decreasing the integral, which is the area under the curve, can be bracketed above by the upper sums and below by the lower sums.

In situations where we can actually find the indefinite integral of \( f(x) \) this test is easy to use and very useful. When using it \textit{one must be careful to verify that all the conditions are satisfied}.

\textit{Note}. Saying that the \textit{infinite} integral converges means

\[
\lim_{x \to \infty} \int_{x=1}^{x} f(x)dx \text{ exists}
\]

i.e. is some real number.

The reason why this test is useful is that it is often easy to see whether the (infinite) integral exists (or does not exist) in cases where the convergence (or divergence) of the series may not be clear.
Sketch proof. From Figure 9 it is clear that if \( f \) is positive and monotonic decreasing
\[
\sum_{k=2}^{N} a_k \leq \int_{1}^{N} f(x) dx \leq \sum_{k=1}^{N-1} a_k
\]
Thus the integral is bounded either side by sums of the series and must converge or diverge with the series.

Similarly, the series can be bounded on either side by integrals. So the series must converge or diverge with the integral.

(The condition \( \lim (x \to \infty) f(x) = 0 \) is needed to nail down this argument.)

Examples. To apply the test to \( \sum 1/n \) we use the function \( f(x) = 1/x (x \geq 1) \). Then \( f(x) \geq 0 \) is monotonic decreasing as \( x \) increases and \( \lim (x \to \infty) f(x) = 0 \). So the conditions of the test are satisfied.

Now
\[
\int_{x=1}^{X} \frac{dx}{x} = [\log x]_{1}^{X} = \log X \to \infty \text{ as } x \to \infty
\]
Hence the integral is divergent and so, by the theorem, the series must be divergent (as, in fact, we earlier proved).
**Exercise.** Do the series

\[
\sum_{n=2}^{\infty} \frac{1}{n \log n}
\]

\[
\sum_{n=2}^{\infty} \frac{1}{n^{1+\varepsilon}} \quad (\varepsilon > 0)
\]

converge or diverge?
Puzzle. Suppose I rotate the graph of $1/x$ about the $x$-axis to form a kind of infinite trumpet (from $x = 1$ to $\infty$). What is

(a) The volume of this trumpet?

(b) The area of this trumpet?

Is there anything that strikes you as odd about your conclusions (suppose I wanted to paint the surface of the trumpet - how much paint would I need?).
Power series

Definition. A power series in a variable $x$ is a series of the form

$$\sum_{n=0}^{\infty} a_n x^n$$

The main question we have to ask about a power series is:

- For what values of $x$ is the series convergent?

Examples.

$$1 + x + x^2 + x^3 + \ldots$$

is a power series. In fact it is a geometric series and we already know that it is convergent when $|x| < 1$ and divergent if $|x| > 1$.

$$1 + x + x^4 + x^9 + x^{16} + \ldots$$

where

$$a_n = \begin{cases} 
1, & \text{if } n \text{ is a square} \\
0, & \text{otherwise}
\end{cases}$$

is a power series. For what $x$ is it convergent? (Actually answering that is a bit tricky, so it is just an illustrative example).
The exponential power series

Almost certainly the most important power series of all is the exponential series, defined as

$$\text{Exp}(x) = 1 + \frac{x}{1!} + \frac{x^2}{2!} + \frac{x^3}{3!} + \ldots$$

$$= \sum_{n=0}^{\infty} \frac{x^n}{n!}$$

Since

$$\frac{|x|^{n+1} \cdot n!}{(n+1)! \cdot |x|^n} = \frac{1}{n+1} \cdot |x| \to 0 \quad \forall \ x \in \mathbb{R}$$

it follows from the ratio test that this series converges absolutely for all real $x$, and hence converges for all real $x$ by the theorem on absolute convergence.

- Thus $\text{Exp}(x)$ defined in this way is defined for all real $x$. 
Properties of Exp(x)

(i) Exp(0) = 1 (clear).

(ii) Exp(x + y) = Exp(x).Exp(y) for all x, y ∈ ℝ.

Sketch proof. The series for Exp(x) and Exp(y) are both absolutely convergent. Hence the product theorem for series applies. Now

\[ c_n = \sum_{i+j=n} \frac{x^i \cdot y^j}{i! \cdot j!} = \sum_{i+j=n} \frac{1}{(i!) (j!) \cdot x^i y^{n-i}} \]

\[ = \left( \sum_{i=0}^{n} \frac{n!}{(i!)(n-i)!} x^i y^{n-i} \right) \frac{1}{n!} = \left( \sum_{i=0}^{n} \binom{n}{i} x^i y^{n-i} \right) \frac{1}{n!} \]

\[ = \frac{(x + y)^n}{n!} \]

and this leads to the conclusion.  

Consequence. Putting y = -x we have

\[ \text{Exp}(-x) = \frac{1}{\text{Exp}(x)} \]

If x > 0 plainly Exp(x) > 0 but then if y = -x we also have Exp(-x) > 0. Hence Exp(x) > 0 ∀ x ∈ ℝ.

- The exponential function is positive for all real x.
Radius of convergence

Definition. If a power series is convergent for all \( x \) with \( |x| < R \) and divergent for \( x \) with \( |x| > R \) (notice we say nothing about what happens if \( |x| = R \)) then we call \( R \) the \textit{radius of convergence} of the power series.

Thus the radius of convergence is the greatest \( R \in \mathbb{R} \) such that the series converges inside the circle \( |x| = R \).

A circle when \( x \) is allowed to take complex values, the interval \([-R, R]\) if \( x \) is real. Thus it is called the \textit{radius} because for power series with complex terms the region of convergence is a \textit{circle} of radius \( R \) in the complex plane.
* Continuity of power series

Although we haven’t yet discussed the idea of a 
*continuous* function it is worth observing that:

- A power series defines a continuous function inside its radius of convergence.

One of the significant properties of real continuous functions is that if \( f \) is continuous and \( f(x_1) > 0 \) and \( f(x_2) < 0 \) there must be a point \( x \), between \( x_1 \) and \( x_2 \), where \( f(x) = 0 \).

This was the key property we used when we defined \( \pi \) as the number such that \( \pi/2 \) is the least positive zero of \( \cos \theta \).
* Differentiation of power series

Another important fact about power series which we shall not prove in detail is:

- Inside the radius of convergence a power series is differentiable term by term (in the sense that the resulting series is the derivative of the sum of the original series).

**Example.** If we assume that for integer \( n \)

\[
\frac{d}{dx}(x^n) = nx^{n-1}
\]

and apply term-by-term differentiation to the exponential series we obtain

\[
\frac{d}{dx}\text{Exp}(x) = 0 + 1 + 2\frac{x}{2!} + 3\frac{x^2}{3!} + ... = \text{Exp}(x)
\]

which establishes (given our assumptions) that:

- The derivative of the exponential function is itself.

This is a critical property of the exponential function. It is essentially the only function with this property.
Consequence. Since the exponential function is always positive (for real \( x \)) it follows that the derivative of the exponential function is always positive.

- The exponential function is a strictly monotonic increasing function of \( x \).

From this it follows that the exponential function takes \( \mathbb{R} \to \mathbb{R}^+ \) in a 1-1 and onto fashion. Consequently the exponential function has a well defined inverse function.

Alternative definition of \( \log_e \). We call the inverse function of \( \text{Exp} : \mathbb{R} \to \mathbb{R}^+ \) the natural logarithm function, written \( \log_e(x) \) (or often just \( \text{Log}(x) \) or \( \log x \)).

By definition the logarithm function has the property that

\[
\log(\exp(x)) = x \quad \forall x \in \mathbb{R}
\]

Notice in Figure 1 that the Log function is not defined for \( x \leq 0 \) (in fact it goes to \(-\infty\) as \( x \) approaches zero from above). This is because it is the inverse of \( \exp \), and \( \exp(x) > 0 \) for all \( x \).
Law of powers (revisited)

Now we know about the Exp function and Log we can see how the theorem $a^x \cdot a^y = a^{x+y} (a > 0)$, which rested on $\text{Exp}(x + y) = \text{Exp}(x) \cdot \text{Exp}(y)$ for all $x, y \in \mathbb{R}$, can be justified via the theory of power series and the definition of Exp and its inverse function Log.

In fact the definition of $a^x$ as $\text{Exp}(x \log a)$ can be extended to $a < 0$ and complex $a$ and $x$ provided we can extend the definition of Log to the whole complex plane.

Although one tricky points arises in the process of following this idea through (there is an ambiguity in Log which stems from the fact that $\text{Exp}(2\pi ni) = 1$ for any $n \in \mathbb{Z}$), with some care one can establish these functions over the whole complex plane except at $a = 0$ (where, whatever one does, the Log function will always have a singularity).
Complex power series

We note in passing that all the definitions of convergence and most of the various theorems we have proved about sequences and series are true (with almost no modification) for sequences and series of complex numbers rather than just reals.

The reason for this is that the only really important property of the reals not possessed by the complex numbers is ordering, and none of the arguments or proofs made any serious appeal to ordering.

The proofs all really depended on completeness (the real and complex numbers ‘have no holes’ i.e. every sequence which ‘ought’ to converge does converge) and the triangle inequality.
Euler’s theorem

For complex numbers $z = x + i y$ we defined

$$|z| = \sqrt{x^2 + y^2}$$

(i.e. the distance from the origin of the point $z$ in the complex plane). Now it is a fact that the triangle inequality is still true for complex numbers

$$|z_1 + z_2| \leq |z_1| + |z_2|$$

for all $z_1, z_2 \in \mathbb{C}$.

Hence all definitions and theorems carry through for complex numbers.

**Example.** The series for $\exp(x)$ can also be used to define $\exp(z)$ for all $z \in \mathbb{C}$.

Having $\exp(z)$ well defined for complex numbers leads to some extremely interesting conclusions. For example

**Theorem 16** (Euler). For all real $\theta$

$$\exp(i\theta) = \cos \theta + i\sin \theta$$
Proof. By definition

$$\text{Exp}(i\theta) = \sum_{n=0}^{\infty} \frac{(i\theta)^n}{n!}$$

$$= 1 + i\frac{\theta}{1!} - \frac{\theta^2}{2!} - i\frac{\theta^3}{3!} + \frac{\theta^4}{4!} + i\frac{\theta^5}{5!} - \frac{\theta^7}{7!} - \ldots$$

$$= \left(1 - \frac{\theta^2}{2!} + \frac{\theta^4}{4!} - \ldots\right) + i\left(\theta - \frac{\theta^3}{3!} + \frac{\theta^5}{5!} - \ldots\right)$$

$$= \cos \theta + i\sin \theta$$

For the purists we would need to justify the rearrangement of the series - in fact this is justified by the fact that all the series are absolutely convergent.

Euler’s theorem leads immediately to the following

**Theorem 17.** For all real (and complex) $\theta$ we have

$$\cos \theta = \frac{1}{2}(\text{Exp}(i\theta) + \text{Exp}(-i\theta))$$

$$\sin \theta = \frac{1}{2i}(\text{Exp}(i\theta) - \text{Exp}(-i\theta))$$
Properties of cos and sin

**Corollary.** (Addition formulae for cos and sin). For all real \( \theta \) and \( \varphi \) we have

\[
\begin{align*}
\cos(\theta + \varphi) &= \cos \theta \cos \varphi - \sin \theta \sin \varphi \\
\sin(\theta + \varphi) &= \cos \theta \sin \varphi + \sin \theta \cos \varphi
\end{align*}
\]

**Proof.** By the previous theorem

\[
\begin{align*}
\text{Exp}(i\theta) \cdot \text{Exp}(i\varphi) &= (\cos \theta + i \sin \theta)(\cos \varphi + i \sin \varphi) \\
&= (\cos \theta \cos \varphi - \sin \theta \sin \varphi) + i(\sin \theta \sin \varphi + \cos \theta \cos \varphi)
\end{align*}
\]

But using the fact that \( \text{Exp}(x) \cdot \text{Exp}(y) = \text{Exp}(x + y) \) and again applying the previous theorem we have

\[
\begin{align*}
\text{Exp}(i\theta) \cdot \text{Exp}(i\varphi) &= \text{Exp}(i(\theta + \varphi)) \\
&= \cos(\theta + \varphi) + i \sin(\theta + \varphi)
\end{align*}
\]

Now equating real and imaginary parts we obtain the result.

**Consequences.** Using the fact that \( \cos(\pi/2) = 0 \) (definition of \( \pi \)) we have

\[
\begin{align*}
\sin \pi &= \sin\left(\frac{\pi}{2} + \frac{\pi}{2}\right) = \sin\left(\frac{\pi}{2}\right)0 + 0.\sin\left(\frac{\pi}{2}\right) = 0
\end{align*}
\]

From which, in a similar fashion, it follows that \( \sin(2\pi) = 0 \).
Similarly, since sin is an odd function,

\[ 1 = \cos(0) = \cos\left(\frac{\pi}{2} - \frac{\pi}{2}\right) \]

\[ = \cos\left(\frac{\pi}{2}\right)\cos\left(-\frac{\pi}{2}\right) - \sin\left(\frac{\pi}{2}\right)\sin\left(-\frac{\pi}{2}\right) \]

\[ = \sin\left(\frac{\pi}{2}\right)\sin\left(\frac{\pi}{2}\right) \quad (\cos\frac{\pi}{2} = 0) \]

i.e. \( \sin^2(\pi/2) = 1 \). From which we deduce

\[ \cos\pi = \cos\left(\frac{\pi}{2} + \frac{\pi}{2}\right) \]

\[ = 0.0 - \sin\left(\frac{\pi}{2}\right)\sin\left(\frac{\pi}{2}\right) = -1 \]

So that \( \cos\pi = -1 \), from which it similarly follows that \( \sin(\pi) = 0 \).

We can now deduce the following rather amazing fact

\[ \text{Exp}(i\pi) = \cos\pi + i\sin\pi = -1 \]

From this via Euler’s theorem it easily follows that sin and cos are periodic functions with period \( 2\pi \).

**Exercise.** Prove (using results that we have established) that

\[ \cos^2\theta + \sin^2\theta = 1 \]

for all real \( \theta \).
Limits

Absolutely fundamental to the calculus is the notion of a limit. We have encountered this idea in the context of letting $n \to \infty$ in some sequence $a_n$. Now we shall consider a more general notion where a real (or complex) number $x \to x_0 \in \mathbb{R}$ and we are interested in the behaviour of some real (or complex) valued function $f$.

We can again formalise a definition in terms of an $(\varepsilon, \delta)$-game. You give me any $\varepsilon > 0$ and my object is to choose a $\delta > 0$ so that whenever $0 < |x - x_0| < \varepsilon$ then $|f(x) - l| < \varepsilon$. If I can do this no matter what $\varepsilon > 0$ you give me then $f(x)$ tends to $l$ as $x \to x_0$ and I win the $(\varepsilon, \delta)$-game. If you succeed in finding an $\varepsilon > 0$ for which there is no such $\delta > 0$ then you win and $f(x)$ does not tend to $l$.

**Definition.** We say $f(x) \to l$ ($f(x)$ tends to a limit $l$) as $x \to x_0$ if for every $\varepsilon > 0 \exists \delta > 0$ such that

$$|f(x) - l| < \varepsilon \ \forall \ x \text{ such that } 0 < |x - x_0| < \delta$$

*Note.* It is vital to observe that we do not consider the value of $f$ at $x = x_0$ (it may not be defined).
Note. The definition can easily be extended to the case where $x \to \infty$ or $f(x) \to \infty$, but if $f(x) \to \infty$ we do not say the limit exists we just say $f \to \infty$.

For the limit to exist it must actually be a real (or complex) number.

Examples.

$f(x) = 1/x$. $\lim_{x \to 2} f(x) = 1/2$ but $\lim_{x \to 0} f(x)$ does not exist. In fact as $x \to 0$ from below $f(x) \to -\infty$ and as $x \to 0$ from above $f(x) \to +\infty$.

$f(x) = x \sin(1/x)$ as $x \to 0$. This is slightly more tricky. One thing we know (we have already established that $\cos^2\theta + \sin^2\theta = 1$) is that
\[|\sin\theta| \leq 1 \quad \forall \quad \theta \in \mathbb{R}\]

Hence, provided $x \neq 0$,
\[\left|x \sin\left(\frac{1}{x}\right)\right| \leq |x| < \varepsilon\]

whenever $|x| < \varepsilon$. Thus I can win the $(\varepsilon, \delta)$-game by choosing $\delta = (1/2)\varepsilon$. 

77
More examples.

\[ \lim_{x \to \infty} x^d = \begin{cases} 
\infty, & \text{if } d > 0 \\
1, & \text{if } d = 0 \\
0, & \text{if } d < 0 
\end{cases} \]

\[ \lim_{x \to \infty} a^x = \begin{cases} 
\infty, & \text{if } a > 1 \\
1, & \text{if } a = 1 \\
0, & \text{if } 0 < a < 1 
\end{cases} \]

**Theorem 18.**

\[ \lim_{x \to x_0} (f(x) \pm g(x)) = \lim_{x \to x_0} f(x) \pm \lim_{x \to x_0} g(x) \]

**Proof.** EER. This is very similar to the proof of Theorem 5(i).
Continuous functions

The idea of continuously deforming one object (set) $X$ to another $Y$ is very intuitive. A formalisation of this notion requires two things: (1) The two sets $X$ and $Y$ must come equipped with a notion of nearness - there must be some sense in which we can say one point is ‘near to’ another; (2) The function $f: X \rightarrow Y$ must preserve the nearness properties of points, i.e. if $x_1$ and $x_2$ are close in $X$ then $f(x_1)$ and $f(x_2)$ should be close in $Y$.

An example of a continuous mapping would be to draw a curve on an elastic sheet. Now stretch the sheet in any way and the curve will distort, but this is a continuous transformation.

If we took a pair of scissors to the elastic sheet, cut it up into pieces, and glued the pieces back together this could easily result in our single curve becoming several bits of curve - such a transformation would not be continuous.

Again we can formalise the notion of continuity using an extension of our earlier notion of the $(\varepsilon, N)$-game.
Definition of continuity.

To establish continuity of a function at some point $x_0$ we play an $(\varepsilon, \delta)$-game. You give me any $\varepsilon > 0$ and my object is to choose a $\delta > 0$ so that whenever $0 < |x - x_0| < \delta$ then $|f(x) - f(x_0)| < \varepsilon$.

If I can do this no matter what $\varepsilon > 0$ you give me then $f$ is continuous at $x_0$ and I win the $(\varepsilon, \delta)$-game. If you succeed in finding an $\varepsilon > 0$ for which there is no such $\delta > 0$, then you win and $f$ is discontinuous at $x_0$. This becomes the following definition.

**Definition.** A function $f$: $\mathbb{R} \rightarrow \mathbb{R}$ (or $\mathbb{C} \rightarrow \mathbb{C}$) is *continuous* at $x_0$ if for every $\varepsilon > 0 \exists \delta > 0$ such that $|f(x) - f(x_0)| < \varepsilon \ \forall \ x$ such that $|x - x_0| < \delta$

**Note.** Another way to put this is that $\lim (x \rightarrow x_0) f(x) = f(x_0)$. (We do not have to exclude $x = x_0$ in the definition because we know $f$ is defined at $x_0$.)

If $f$ is continuous at every point in some set $S \subseteq \mathbb{R}$ or $\mathbb{C}$ then we say $f$ is *continuous on $S$.*

Let us sketch some curves and show why some are continuous and some are not...
**Examples.**

\[ x^2 \text{ is continuous everywhere.} \]

\[ \frac{1}{x} \text{ is continuous everywhere except } x = 0. \]

A simple step function

\[ f(x) = \begin{cases} 0, & \text{if } x \leq 1 \\ 1, & \text{if } x > 1 \end{cases} \]

will be continuous everywhere except at the step (in this case at \( x = 1 \)) where it is said to have a ‘jump discontinuity’.

\( \sin x \) is continuous everywhere - need to know theorems about when power series sum to continuous functions.

\( \sin \left( \frac{1}{x} \right) \) is discontinuous at \( x = 0 \).

Notice that at \( x = 0 \) the function does *not* have a step discontinuity and does *not* go to infinity.

Any polynomial is a continuous function for all real or complex values of the variable.
The definition of continuity just given has many of the expected properties. For example, if \( f \) and \( g \) are continuous at \( x_0 \) (or at all points in some set \( S \)) then the function \( h(x) = f(x) + g(x) \) is also continuous at \( x_0 \) (or on \( S \)), etc. This is an immediate consequence of Theorem 10.

We shall not dwell extensively on the notion of continuity, except to say that sometimes the consequences of definitions can be unexpected.

In this case one rather amazing fact which emerges is that it is possible to map the real interval \([0, 1]\) to the unit square in a 1-1, onto, continuous fashion. We may put up a Mathematica™ file which illustrates the iterative construction of such a function (called a Peano space filling curve).
Big O, Little o and \( \sim \)

There is some notation which is very useful and we choose to introduce at this point.

**Definition.** We say \( f(x) = O(g(x)) \) as \( x \to x_0 \) (or \( x \to \infty \)) if \( \exists \) some \( \delta > 0 \) and some absolute constant \( C > 0 \) such that

\[
|f(x)| \leq C|g(x)| \quad \forall x \text{ such that } 0 < |x - x_0| < \delta
\]

**Examples.**

\[
x^3 = O(x^2) \text{ as } x \to 0
\]

\[
x^2 = O(x^3) \text{ as } x \to \infty
\]

\[
sin(x) = O(1) \text{ as } x \to \infty \text{ (or as } x \to \text{ any } x_0 \in \mathbb{R})
\]

\[
n! = O\left(n^{n+\frac{1}{2}}\right) \text{ as } n \to \infty
\]

Thus a crude algorithm to sort a list into order might take a time \( O(n^2) \) as \( n \to \infty \) to run, where \( n \) is the number of elements being sorted. A faster sort may run in \( O(n \log n) \) time. The constants \( C \) here would depend on the hardware, compiler and the code.
Definition. We say \( f(x) = o(g(x)) \) as \( x \to x_0 \) (or \( x \to \infty \)) if
\[
\lim_{x \to x_0} \frac{f(x)}{g(x)} = 0
\]
i.e. if \( f(x) \) tends to zero faster than \( g(x) \).

E.g. \( x^3 = o(x^2) \) as \( x \to 0 \).

Definition. We say \( f(x) \sim g(x) \) as \( x \to x_0 \) (or \( x \to \infty \)) if
\[
\lim_{x \to x_0} \frac{f(x)}{g(x)} = 1
\]
i.e. the asymptotic behaviour of \( f(x) \) and \( g(x) \) are the same as \( x \to x_0 \).

E.g. \( x^2 + x^4 \sim x^2 \) as \( x \to 0 \).
Differentiation

The slope of the line between the point \((x, f(x))\) and the point \((x + h, f(x + h))\) is

\[
\frac{f(x + h) - f(x)}{x + h - x} = \frac{f(x + h) - f(x)}{h}
\]

**Definition.** The limit (if it exists)

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

is called the derivative of \(f\) at \(x\) and is often written as \(f'(x)\) or \(\frac{df}{dx}\).

It is the slope of the tangent to the curve at \(x\).
Example.

(i) The derivative of \( f(x) = x^2 \) is \( f'(x) = 2x \). For
\[
\lim_{h \to 0} \frac{(x + h)^2 - x^2}{h} = \lim_{h \to 0} (2x + h) = 2x = f'(x)
\]

(ii) The derivative of \( f(x) = 1/x \ (x \neq 0) \) is \( f'(x) = -1/x^2 \). For
\[
\lim_{h \to 0} \frac{1}{x + h} - \frac{1}{x} = \lim_{h \to 0} \frac{1}{x} - \frac{(x + h)}{x(x + h)} = \lim_{h \to 0} \frac{-1}{x(x + h)} = - \frac{1}{x} \cdot \frac{1}{x + h} = - \frac{1}{x^2} = f'(x)
\]

(iii) Exercise. In general we can use the binomial theorem to show that if \( f(x) = x^n \) then \( f'(x) = nx^{n-1} \) for any integer \( n \neq 0 \) and any \( x \) for which \( f(x) \) is defined.

Derivative of \( x^n \): Solution to (iii).

From the binomial theorem we have (for any fixed \( x \))
\[
\frac{(x + h)^n - x^n}{h} = \frac{x^n + \binom{n}{1}x^{n-1}h + O(h^2)}{h} - x^n
\]
\[
= nx^{n-1} + O(h) \to nx^{n-1} \quad \text{as} \quad n \to \infty
\]
which proves the result. |
Derivatives of sin and cos

We originally defined the functions sin and cos geometrically and later using power series. Later we remarked that a power series can be differentiated term by term inside its circle of convergence. (We used term-by-term differentiation to show that the derivative of the exponential functions is itself - the radius of convergence is infinite so the exponential function is everywhere differentiable.)

We can do exactly the same thing with the series for cos and sin, which again both have infinite radius of convergence. The power series definitions were

\[
\cos \theta = 1 - \frac{\theta^2}{2!} + \frac{\theta^4}{4!} - \ldots = \sum_{n=0}^{\infty} (-1)^n \frac{\theta^{2n}}{(2n)!}
\]

\[
\sin \theta = \theta - \frac{\theta^3}{3!} + \frac{\theta^5}{5!} - \ldots = \sum_{n=0}^{\infty} (-1)^n \frac{\theta^{2n+1}}{(2n+1)!}
\]

and if we differentiate term-by-term we obtain

\[
\frac{d}{d\theta} (\sin \theta) = 1 - \frac{3\theta^2}{3!} + \frac{5\theta^4}{4!} - \ldots = \cos \theta
\]

\[
\frac{d}{d\theta} (\cos \theta) = -2 \frac{\theta}{2!} + 4 \frac{\theta^3}{4!} - 6 \frac{\theta^5}{5!} + \ldots = -\sin \theta
\]

These are very useful results.
Properties of derivatives

It is not too hard to show that if \( f \) is differentiable (i.e. the derivative exists) at some point \( x \) then it is also continuous at \( x \). For if

\[
\left| \frac{f(x + h) - f(x)}{h} - l \right| < \varepsilon \quad \forall \ |h| < \delta
\]

then

\[
|f(x + h) - f(x)| \leq |h|(\|l\| + \varepsilon)
\]

and by making \( h \) as small as we please (but not zero of course) we establish that \( f \) is continuous at \( x \).

It is also easy to see (e.g. \( |x| \) at \( x = 0 \)) that the converse is not true (i.e. continuity does not imply differentiability).

**Theorem 19.** If the derivatives exist then

\[
\frac{d}{dx}(cf(x)) = cf'(x) \quad \text{for any constant } c
\]

\[
\frac{d}{dx}(f(x) + g(x)) = f'(x) + g'(x)
\]

**Proof.** The first is immediate from the definition and elementary properties of limits. The second is immediate from Theorem 10.
**Theorem 20.** (Chain rule) If \( f: X \to Y \) and \( g: Y \to Z \) are differential functions with \( X, Y, Z \subseteq \mathbb{R} \) then

\[
\frac{d}{dx} f(g(x)) = f'(g(x))g'(x)
\]

**Sketch proof.** From the definition of derivatives

\[
f(y + H) = f(y) + Hf'(y) + o(H)
\]

\[
g(x + h) = g(x) + hg'(x) + o(h)
\]

Put \( y = g(x) \) and \( H = hg'(x) + o(h) \) to obtain

\[
f(g(x + h)) = f(g(x)) + (hg'(x) + o(h))f'(y) + o(h)
\]

Hence

\[
f(g(x + h) - f(g(x) = h f'(g(x))g'(x) + o(h)
\]

and dividing by \( h \) the result follows.

**Example.**

\[
\frac{d}{dx} \exp(x^3) = \exp(x^3) \cdot 3x^2 = 3x^2 \exp(x^3)
\]
**Theorem 21** (Differentiation of a product). If the derivatives exist then
\[
\frac{d}{dx}(f(x)g(x)) = f'(x)g(x) + f(x)g'(x)
\]

**Example.**
\[
\frac{d}{dx}(\sin(x) \cdot \exp(x^2)) = \cos(x) \exp(x^2) + \sin(x) \exp(x^2) \cdot 2x
\]

**Theorem 22** (Differentiation of a ratio). If the derivatives exist and \(g(x) \neq 0\) then
\[
\frac{d}{dx}\left(\frac{f(x)}{g(x)}\right) = \frac{f'(x)g(x) - g'(x)f(x)}{g(x)^2}
\]

**Note.** There is no real need to remember this because it can readily be derived from the product rule by differentiating \(fg^{-1}\).

- You are expected to have practice in differentiating the standard functions. You should study section 8.3 *Techniques of differentiation* in the coursebook and work through Exercises 8.3.15.
The derivative of log

Let \( y = \log(x) \) then because \( \log \) is the inverse function of \( \text{Exp} \) we have \( \text{Exp}(y) = x \). Now differentiate both sides with respect to \( x \). Then by the Chain rule we have

\[
\frac{d}{dx}(\text{Exp}(y)) = \frac{d}{dx}(x) = 1
\]

\[
\text{Exp}(y) \frac{dy}{dx} = 1
\]

\[
\frac{dy}{dx} = \frac{1}{\text{Exp}(y)} = \frac{1}{x}
\]

Thus the derivative of \( \log(x) \) is \( 1/x \) \((x > 0)\).

If you look at the graph of \( \log(x) \) in Figure 1 you can see that as \( x \) increases the slope decrease (in fact like \( 1/x \) as we now know).
Maxima and minima (See section 9.2 of the coursebook).

A knowledge of the derivative of a function can be put to many uses. *Speed* is the derivative of distance travelled with respect to time and more generally equations involving derivatives (differential equations) are used to describe the evolution of complex dynamic systems through time.

Another use for the derivatives of a function is to find its local maxima and minima (turning points or extreme points).

If a differentiable function $f$ has a local maxima or minima at $x = a$ then $f'(a) = 0$ moreover if the second derivative exists then

- If $f''(a) > 0$ the extreme point is a minimum.
- If $f''(a) < 0$ the extreme point is a maximum.
- If $f''(a) = 0$ the extreme point may be either a maximum or a minimum or a point of inflection, and further investigation is required to determine which. (See the Mathematica file maxmim1.nb)
Taylor’s theorem

It is a remarkable fact that we can find a series expansion of a function, which must be differentiable as many times as we need, by calculating the derivatives of the function.

Notation. We shall often write $f^{(k)}(x)$ for the $k$th successive derivative of $f$ evaluated at $x$. For $k = 1$, $f^{(1)}(x) = f'(x)$. For $k = 2$, $f^{(2)}(x) = f''(x)$, i.e. the derivative of the derivative, etc.

Theorem 23 (Taylor). Suppose $f: \mathbb{R} \to \mathbb{R}$ and is $n + 1$ times differentiable then $\exists \theta, 0 < \theta < 1$, such that

$$f(x + h) = f(x) + \frac{h}{1!}f^{(1)}(x) + \frac{h^2}{2!}f^{(2)}(x) + ...$$

$$+ \frac{h^n}{n!}f^{(n)}(x) + \frac{h^{n+1}}{(n+1)!}f^{(n+1)}(x + \theta h)$$

Proof. We do not go into details of the proof, which can be found in most analysis textbooks.

However, this is a very important theorem and is the basis from which we can derive many useful power series expansions. So you should learn how to apply it to a given function.
Note. The last term can be viewed as the ‘error term’ when \( f(x + h) \) is approximated by the polynomial

\[
f(x) + \frac{h}{1!}f^{(1)}(x) + \frac{h^2}{2!}f^{(2)}(x) + \ldots + \frac{h^n}{n!}f^{(n)}(x)
\]

Finding power series expansions

If *all* derivatives exist (i.e. the function is ‘infinitely’ differentiable) and we can show that the `error term' in the Taylor expansion goes to zero as \( n \to \infty \), then we can use a knowledge of the derivatives of a function \( f \) to derive a power series expansion for \( f \).

**Example.** (Power series for Log(1+\( x \))). Let us try to find a power series expansion for Log(\( x \)). Since the function has a singularity at \( x = 0 \) we cannot expect to find a series expansion near \( x = 0 \). Instead we examine the function \( f(x) = \text{Log}(1 + x) \). Thus provided \( x > -1 \) we have

\[
\begin{align*}
  f^{(1)}(x) &= 1/(1 + x), \\
  f^{(2)}(x) &= -1/(1 + x)^2, \\
  f^{(3)}(x) &= 2/(1 + x)^3, \\
  f^{(4)}(x) &= -3.2/(1 + x)^4 \\
  &\ldots
\end{align*}
\]

In general

\[
f^{(n)}(x) = \frac{(-1)^{n-1}(n - 1)!}{(1 + x)^n}
\]
So that at $x = 0$

$$f^{(n)}(0) = (-1)^{n-1}(n - 1)!$$

Also if $-1 < x < 1$ we have

$$\left| \frac{x^{n+1}f^{(n+1)}(x)}{(n+1)!} \right| \leq \left| \frac{x^{n+1}n!}{(n+1)!} \right| \leq \frac{1}{n+1} \to 0 \text{ as } n \to \infty$$

so that the ‘error term’ goes to zero.

Thus we obtain

$$\Log(1 + x) = x - \frac{x^2}{2} + \frac{x^3}{3} + \ldots \quad (-1 < x < 1)$$

Notice that this series does in fact converge at $x = 1$ (why?) but diverges at $x = -1$ (why?).
Mathematica note on ‘Series’.

A useful Mathematica function for determining the first few terms of a power series is Series. Thus

\[
\text{Series}[\log(1 + x), \{x, 0, 4\}]
\]

which says give me the first 4 terms of the power series expansion of \(\log(1 + x)\) about \(x = 0\), yields

\[
x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + O(x^5)
\]

*Mathematica* knows about the Log function and Taylor’s theorem and can perform the symbolic calculation we just did ourselves!
Binomial series

**Theorem 24** (Binomial series) If $|x| < 1$ then for any $\alpha \in \mathbb{R}$

$$(1 + x)^\alpha = 1 + \alpha x + \frac{\alpha(\alpha - 1)}{2!} x^2 + \frac{\alpha(\alpha - 1)(\alpha - 2)}{3!} x^3 + \ldots$$

**Proof.** EER. Just apply Taylor’s theorem.]

Remark. This generalises the Binomial theorem (which gives a finite sum for positive integral $\alpha$) but notice the series only converges for $|x| < 1$.  

L’Hôpital’s rule

We often wish to determine limits of the form

\[ \lim_{x \to x_0} \frac{f(x)}{g(x)} \]

where

\[ \lim_{x \to x_0} f(x) = \lim_{x \to x_0} g(x) = 0 \]

Under these circumstances if \( f \) and \( g \) are differentiable at \( x_0 \) we have

\[ \lim_{x \to x_0} \frac{f(x)}{g(x)} = \lim_{x \to x_0} \frac{f'(x)}{g'(x)} \]

provided that the limit on the RHS exists.

**Example.** To determine \( \lim_{x \to 0} \frac{\sin x}{x} \) we cannot apply the quotient rule for limits but

\[ \lim_{x \to 0} \frac{\sin x}{x} = \lim_{x \to 0} \frac{\cos x}{1} = 1 \]

For some examples we may have to do this several times before the final limit is immediate.
Mathematica note on ‘Limit’.

The Mathematica function \texttt{Limit} can also deal with this type of calculation. For example

\[
\text{Limit}[(x^2 - 1)/(x - 1), \; x \to 1]
\]

yields the correct answer 2.

\textit{Note.}

\[
\frac{x^2 - 1}{x - 1} = \frac{(x - 1)(x + 1)}{x - 1} = x + 1 \quad \text{provided } x \neq 1
\]

Hence

\[
\lim_{x \to 1} \frac{x^2 - 1}{x - 1} = 2
\]

\textbf{Exercise.} For any integer \( n > 0 \) find

\[
\lim_{x \to 1} \frac{x^n - 1}{x - 1}
\]
Integration

The area under a curve $y = f(x)$ which is written as
\[ \int_{x = a}^{b} f(x) \, dx \]

is defined by a limiting process.

We divide up the interval $[a, b]$ into a finite number of sub-intervals whose maximum length is $\delta$ (say). Next we form the upper and lower rectangles as shown in Figure 12.
For a partition into $n$ sub-intervals with $x_0 = a$ and $x_n = b$ with the maximum length of the sub-intervals being $\delta$, we write

\[
L(\delta) = \sum_{i=0}^{n-1} \left( \min_{x \in [x_i, x_{i+1}]} f(x) \right) (x_{i+1} - x_i)
\]

\[
U(\delta) = \sum_{i=0}^{n-1} \left( \max_{x \in [x_i, x_{i+1}]} f(x) \right) (x_{i+1} - x_i)
\]

So that

\[
L(\delta) \leq U(\delta)
\]

where $L(\delta)$ is the sum of the areas of the lower rectangles and $U(\delta)$ is the sum of the areas of the upper rectangles. (There is no problem in defining the area of a rectangle in the usual way.)

Then as $\delta \to 0$ one can show (under reasonable conditions on $f$) that $L(\delta)$ is monotonic increasing and $U(\delta)$ is monotonic decreasing. These functions are both bounded so that

\[
\lim_{\delta \to 0} L(\delta) \quad \text{and} \quad \lim_{\delta \to 0} U(\delta)
\]

both exist.
**Definition.** If both limits exists and are equal we define their common value to be

\[
\int_{x = a}^{b} f(x)\,dx
\]

This is called the *integral* of \( f \) from \( a \) to \( b \).

**Example.** Suppose we want to compute from first principles the integral of \( f(x) = x \) from 0 to \( b \).

Then, because \( f \) is monotonic increasing, we have

\[
L(\delta) = \sum_{i = 0}^{n-1} x_i(x_{i+1} - x_i)
\]

\[
U(\delta) = \sum_{i = 0}^{n-1} x_{i+1}(x_{i+1} - x_i)
\]

Now suppose the interval partition is into equal sub-intervals each of length \( \delta = 1/n \). Then

\[
x_i = \frac{ib}{n}
\]

\[
x_{i+1} - x_i = \frac{b}{n}
\]
Thus
\[ L(\delta) = \sum_{i=0}^{n-1} \frac{ib}{n} \cdot \frac{b}{n} = \frac{b^2}{n^2} \sum_{i=0}^{n-1} i = \frac{b^2}{2} \left( 1 - \frac{1}{n} \right) \]

\[ U(\delta) = \sum_{i=0}^{n-1} \frac{(i+1)b}{n} \cdot \frac{b}{n} = \frac{b^2}{n^2} \sum_{i=0}^{n-1} (i + 1) = \frac{b^2}{2} \left( 1 + \frac{1}{n} \right) \]

Hence
\[ \lim_{\delta \to 0} L(\delta) = \frac{b^2}{2} = \lim_{\delta \to 0} U(\delta) \]

so that
\[ \int_{x=0}^{b} x \, dx = \frac{b^2}{2} \]

This may seem a rather complicated way to find the area of a triangle. Nevertheless, all numerical integration techniques are extensions of this simple idea.

In practice we shall see that by hand there are other ways to compute integrals.
Properties of integrals

Areas defined in this way have the properties that one might reasonably expect such as

**Theorem 25** If $f$ is integrable and $a < c < b$ then

$$\int_{x = a}^{b} f(x)\,dx = \int_{x = a}^{c} f(x)\,dx + \int_{x = c}^{b} f(x)\,dx$$

**Theorem 26** (Integral of a sum) If $f$ and $g$ are integrable then

$$\int_{x = a}^{b} (f(x) + g(x))\,dx = \int_{x = a}^{b} f(x)\,dx + \int_{x = a}^{b} g(x)\,dx$$

etc.
Fundamental theorem of calculus

We recall that both derivatives and integrals are defined by entirely different limiting processes. It is therefore most remarkable that there is a very simple connection between the two.

Theorem 27 (Fundamental theorem of calculus). If \( f(x) \) is integrable in \([a, b]\) and \( F(x) \) is such that \( F'(x) = f(x) \) then

\[
\int_{x = a}^{b} f(x) \, dx = F(b) - F(a)
\]

Thus a problem of integration can often be viewed as the inverse of differentiation.

Example.

\[
\int_{x = a}^{b} (x^3 + x^2 + 1) \, dx = \left[ \frac{x^4}{4} + \frac{x^3}{3} + x \right]_{a}^{b}
\]

\[
= \left( \frac{b^4}{4} + \frac{b^3}{3} + b \right) - \left( \frac{a^4}{4} + \frac{a^3}{3} + a \right)
\]
Techniques of integration.

If the range of integration is not given then the integral is only determined to within a constant. It is then called an *indefinite integral*. As soon as the limits of integration are specified we can then determine the constant or the value of the integral.

Using the Fundamental theorem gives us a range of techniques to evaluate integrals.

**Examples.** Using the Fundamental theorem we can evaluate the following indefinite integrals.

\[
\int x^6 \, dx = \frac{1}{7}x^7 + c
\]

\[
\int e^{3x} \, dx = \frac{1}{3}e^{3x} + c
\]

\[
\int \sin(5x) \, dx = -\frac{1}{5}\cos(5x) + c
\]

\[
\int (2x + 1)^3 \, dx = \frac{1}{8}(2x + 1)^4 + c
\]

\[
\int \frac{2}{x} \, dx = 2\log(x) + c
\]
Partial fractions.

We remind ourselves of the properties of partial fractions.

It is often useful, particularly in integration, to be able to express a rational function \( p(x)/q(x) \) as a sum of simpler functions consisting of

(a) a polynomial (if \( \deg(p) > \deg(q) \)).
(b) simpler functions whose denominators are either linear or quadratic.

Examples.

\[
\frac{2x + 1}{(x - 2)(x + 1)(x - 3)} = \frac{5/3}{x - 2} - \frac{1/12}{x + 1} + \frac{7/4}{x - 3}
\]

\[
\frac{x^2 + 1}{(x + 1)(x - 1)(x^2 + 2x + 2)} = \frac{1}{x + 1} - \frac{1/5}{x - 1} + \frac{4x + 7}{5(x^2 + 2x + 2)}
\]

There are a series of rules for how to determine the numerators of the partial fraction expansion of a given rational function (see page 114 of the coursebook (2nd Edition)). You should learn how to do this, at least for the simplest case of simple linear factors. (The ‘cover-up’ rule).
Mathematica note on ‘Apart’.

*Mathematica* can compute partial fractions and, once you have learnt the rules, you can use Mathematica to check whether you have done any particular partial fraction expansion correctly. The function required is called *Apart*.

Thus

\[
\text{Apart}\left[\frac{2x + 1}{(x - 2)(x + 1)(x - 3)}\right]
\]

yields

\[
\frac{7}{4(-3 + x)} - \frac{5}{3(-2 + x)} - \frac{1}{12(1 + x)}
\]
Integrating rational functions using partial fractions.

Using the technique of partial fractions it is possible to symbolically integrate any rational function.

Example. Find the indefinite integral

\[
\int \frac{2x + 1}{(x - 2)(x + 1)(x - 3)} \, dx
\]

Solution. Using the cover-up rule we find

\[
\frac{2x + 1}{(x - 2)(x + 1)(x - 3)} = -\frac{5/3}{x - 2} - \frac{1/12}{x + 1} + \frac{7/4}{x - 3}
\]

Hence integrating each term separately we have

\[
\int \frac{2x + 1}{(x - 2)(x + 1)(x - 3)} \, dx
\]

\[
= -\frac{5}{3} \log(x - 2) - \frac{1}{12} \log(x + 1) + \frac{7}{4} \log(x - 3) + c
\]

- You are expected to have practice in integrating the standard functions. You should study section 8.6 Techniques of integration in the coursebook and work through Exercises 8.6.2 and 8.6.4.
Mathematica note on ‘Integrate’

Since *Mathematica* can expand into partial fractions it is not too surprising that it can also integrate the above function. The function is `Integrate`

Thus

\[
\text{Integrate}\left[\frac{2x + 1}{(x - 2)(x + 1)(x - 3)}\right]
\]

yields

\[
\frac{-5\log(2 - x)}{3} + \frac{7\log(3 - x)}{4} - \frac{\log(1 + x)}{12}
\]

However, this doesn’t begin to touch upon the capabilities of `Integrate` which can perform all the indefinite integrations we do in this course and a very great deal more besides.
Integration by substitution.

Suppose we have to integrate $F(x)dx$ and we do not know of any direct way of doing this. We might try the method of substitution. We put $x = f(t)$ and hope to do this in such a way that it is possible to integrate $F(f(t))dt$. This method is based on the fact that if $F$ is continuous and $x = f(t)$ has continuous derivative then

$$\int_{x=a}^{x=b} F(x)dx = \int_{t=f(a)}^{t=f(b)} F(f(t))f'(t)dt$$

**Note:** Don’t forget to change the limits of integration when integrating by substitution.

**Example.** Find

$$\int \frac{x^3}{\sqrt{x} - 1} \, dx$$

Put $x - 1 = t^2$ so that (in an obvious shorthand) $dx = 2tdt$ and hence

$$\int \frac{x^3}{\sqrt{x} - 1} \, dx = \frac{(t^2 + 1)^3}{t} \, 2tdt$$
Consequently

\[ \int \frac{x^3}{\sqrt{x - 1}} \, dx = 2 \int (t^2 + 1)^3 \, dt \]

\[ = 2 \int (t^6 + 3t^4 + 3t^2 + 1) \, dt \]

\[ = 2t \left( \frac{t^7}{7} + \frac{3t^5}{5} + t^3 + t \right) \]

\[ = \frac{2}{7}(x - 1)^4 + \frac{6}{5}(x - 1)^3 + 2(x - 1)^2 + 2(x - 1) \]

Integration by substitution is a powerful technique but in order to make best use of it is necessary to learn a library of standard substitutions and the appropriate circumstances in which to use them.
Integration by parts.

This is another standard tool for integrating awkward functions. It is based on the formula

\[ \int_{x=a}^{x=b} u(x) \, dv(x) = [u(x)v(x)]_{x=a}^{x=b} - \int_{x=a}^{x=b} v(x) \, du(x) \]

which is obtained by integrating both sides of the rule for differentiating a product.

Example. Find

\[ \int x \log x \, dx \]

Take \( u(x) = \log x \) and \( v(x) = \frac{1}{2} x^2 \) so that \( dv = x \, dx \) and \( du = \frac{1}{x} \, dx \). Then

\[ \int x \log x \, dx = \frac{1}{2} x^2 \log x - \int \frac{1}{2} x^2 \frac{1}{x} \, dx \]

\[ = \frac{1}{2} x^2 \log x - \frac{1}{2} \int x \, dx \]

\[ = \frac{1}{2} x^2 \log x - \frac{x^2}{4} + C \]

(You should always check any integration by differentiating the result.)
Integration Exercises

Exercises 8.6.2 and 8.6.4. of the coursebook.
Numerical integration

Many functions whose integrals exist cannot be integrated symbolically at all. In such cases we can nevertheless evaluate a definite integral of such a function numerically.

Numerical integration is the study of algorithms to perform such computations. Most such algorithms are based on dividing up the region of integration into small areas over which the function is nearly constant and then estimating the upper and lower sums.

In Mathematica such an algorithm is given by the function NIntegrate.

Example. If we evaluate

\[ \text{NIntegrate[Exp[-x^2], \{x, -10, 10\}] \] 

we obtain 1.7245. It is no co-incidence that \( \sqrt{\pi} \approx 1.7245 \).

It is an odd situation that one can actually prove that

\[ \int_{-\infty}^{+\infty} e^{-x^2} \, dx = \sqrt{\pi} \]

but Exp[-x^2] cannot be symbolically integrated in terms of standard functions.
What can be computed?

It is difficult to disentangle the history of networks and automata theory from that of computing and artificial intelligence. The cast of characters is largely the same and the development of these subjects went hand in hand. The original pioneers of artificial intelligence were the theoreticians who created computing itself. We mention some of the principal contributors below.

**Alan Turing** who defined the Turing machine, the basic model of all computational processes, and who founded much of the basic theory of computation was passionately interested in the subject of artificial intelligence. He is also famous for proposing the ‘Turing Test’ in which a man conversing on a teletype tries to determine whether the other end of the conversation is conducted by a man or by a computer.

**John von Neumann**, who was largely responsible for building the first electronic computer, had a very farsighted vision of the future of computing. Apart from his work on reliable computing with unreliable components (which now looks very much up to date in the context of VLSI) von Neumann had a vision of a different model of computation (see above).

**Claude Shannon** who first made clear the relationship between electronics and Boolean algebra also made contributions to computer game playing and symbolic computation. In addition Shannon played a leading role in developing the information theory foreseen by von Neumann.

It is an interesting exercise to follow the history of automata theory as a separate concept from that of digital computing. Indeed, this history may be seen as starting in 1937, ten years before the birth of the digital computer as we know it now (Burks, Goldstine and von Neumann, 1947). Automata theory, it can be plausibly argued, really began in the mind of Alan Turing as a cross fertilization of two ideas: he had an interest in computing machinery and was well aware of the potential of electronics. Also as a pure mathematician he was interested in the subject of ‘effective computability’.

Informally, the value of a function, or the solution to some problem, is said to be effectively computable if there exists an unambiguous finite set of rules which, if applied relentlessly, will after a finite number of steps produce the required result. It is a fairly easy matter to show that by this definition there are many functions which are not effectively computable. If the rules can be carried out by some very simple machine then there is no doubt that the specification is complete and that we have an ‘effective procedure’. Turing’s thesis, which at first sight may seem extreme, is a sort of converse to this observation:

- Any procedure which can naturally be called effective can be realized by a (simple) machine.

Note that an effective procedure need not terminate. For example, the procedure may be designed to produce an infinite list; if any particular member of the list can be guaranteed to be produced in a finite number of steps then the procedure would be effective. It is not termination, but rather the finiteness of the description of the procedure which is important.
In his 1936 paper Turing defined the class of abstract machines which now bear his name. They are exceedingly simple. Yet Turing discovered that he could set these machines up to perform very complex calculations. His thesis that these machines could perform any effective procedure is closely related to the work of Alonzo Church, Emil Post and S.C. Kleene. An excellent defense of his position can be found in his brilliant article *Computing Machines and Intelligence*, 1950.

Both Church and Kleene, like Turing himself, were concerned with ideas relating to ‘effective computability’. The notion of an effective procedure can be implemented with a system that includes:

- A language for describing rules of behaviour.
- A machine that obeys statements in the language.

Let us turn, for a moment, to the language in which the rules can be expressed. The idea of Post (1943) is that expressions or statements in a logical system or language, whatever else they may seem to be, are in the final analysis nothing but strings of symbols written in some finite alphabet. Even the most powerful mathematical or logical system is ultimately, nothing but a set of rules that tell how some string of symbols may be transformed into another string of symbols. Just as Turing was able to show the equivalence of his broadest notion of a computing machine with the very sharply restricted idea of a Turing machine, so Post was able to reduce his broadest concept of a string transformation system to a family of astoundingly special and simple symbol manipulation operations. We can summarize Post’s view as: any system for manipulation of symbols which could naturally be called a formal or logical system can be realized as one of Post’s ‘canonical systems’.

From the work of Turing and his contemporaries the gestation period lasted some twenty-one years and saw many supporting developments take place. Shannon’s theories of *Switching Logic* (1938) and *Statistical Communication theory* (1948) being two examples. Indeed, the period saw the coming of Weiner’s classic book on Cybernetics where, although he mentions meeting Turing

“I spent a total of three weeks in England...I had an excellent chance... to talk over the fundamental ideas of Cybernetics with Mr. Turing.”,

there is no further mention of Turing’s ideas on abstract automata. The major events of this period, however, were still to come.

**Turing machines.**

If one considers that biological intelligence is a computational process we should be as precise as possible in our definition of computation. As we have said this question was addressed by many of the early workers in the field. Indeed, many of the principal conclusions had been worked out before the first computer had been built. The simple description that emerged as what is now called the Turing machine does not bear any immediately obvious relation to a modern day computer, but nevertheless all such computers can be considered in the final analysis as equivalent to Turing machines.

The basic model of a Turing machine has a finite state machine controller, an input tape which is divided into cells, and a tape head which scans one cell of the tape at a time. The tape can be regarded as infinite in both directions, but we make the restriction that when the tape is started the tape must be blank, except for some finite number of cells. The tape head has three functions, all of which are used in each operation cycle of the FSM controller. These functions are:

- Reading the cell of the tape being scanned.
- Writing on the scanned square.

and

- Moving left or right to an adjacent cell (which becomes the scanned cell in the next operation cycle).
Each cell of the tape may hold exactly one of a finite set of tape symbols. When a symbol is printed on the tape the previous symbol in the cell is erased.

When a Turing machine, with its input tape is started it may be a very, very long time before it completes its computation and comes to a halt. For many machine-tape pairs this will never happen, the computation may go on for ever. It would be useful to have a decision procedure which would enable us, given any Turing machine T and tape t, to determine whether the process will ever halt.

If there is any decision procedure then it, by definition, can be executed by a Turing machine A which takes as input a pair (d(T),t), where d(T) is some description of T and t is the tape, and which is guaranteed to halt printing ‘YES’ or ‘NO’, meaning T does eventually halt given t or T does not halt given t, respectively.

If A can solve the problem for all pairs (d(T),t), then it can certainly do so for the special pairs (d(T),d(T)), where the tape t is a description of T, rather than any old tape. We can therefore construct a new Turing machine B, which requires only one copy of the description d(T) but otherwise behaves like A applied to (d(T),d(T)). Thus B scans d(T) and halts printing

\[ \text{YES if } T \text{ halts given } d(T), \quad \text{NO if } T \text{ never halts given } d(T). \]

Now modify B by adding a loop to the YES exit. This new Turing machine C obeys the rules:

- If T halts given d(T) print YES and go into a loop.
- If T never halts given d(T) print NO and halt.

Now for the killer: what happens if C is applied to the tape d(C)? Plainly it halts if C applied to d(C) does not halt, and never halts if C applied to d(C) halts.

This is a contradiction and so we must conclude that a Turing machine such as C, and hence B, and hence A, cannot exist. The unsolvability of the halting problem generalises to any computing machine which can manipulate data and interpret it as instructions.

We shall see that there is a similar limitation to the Halting problem in logic. These are the theorems about decidability and completeness and due to Gödel. These theorems relate to the concept of proof. The unsolvability of the Halting problem and Gödel's theorems have implications for any information processing system which aspires to reason. These results imply that:

- There are sharp limits on what can be computed whatever the precise architecture.

**Computability of logic.**

The theoretical importance of the Turing machine to computing, transformational grammar and logic is difficult to over-estimate. To understand the reasons why this is so we need to look back to 1900. At the turn of the century the mathematician Hilbert posed a list of outstanding problems for the mathematicians of the twentieth century. Perhaps the most important of these was the decision problem for the first-order predicate calculus.

Logical systems specify a set of legal strings of symbols called well-formed formulas (wffs), which, when appropriately interpreted become statements which are either true or false. Some wffs are designated as axioms and serve with rules of inference as the basis for the production of theorems. Each theorem has a proof, which is a finite sequence of string transformations leading from the axioms to the wff which is the statement of the theorem.

- Thus in any logical system there is a decision problem: is the set of theorems a decidable subset of the set of wffs?

The generally accepted definition of a 'decidable' subset has emerged as roughly: if and only if there is a Turing
machine which given a member of the set can effectively determine whether it is a member of the subset. Thus if the decision problem for a particular logical system is soluble it means that it is possible to discover theorems by submitting wffs to a Turing machine and letting it pronounce ‘YES’ or ‘NO’ when it halts (i.e. we can in principle decide if a theorem is true or false).

In each logical system theorems are valid wffs; the completeness problem for the system asks whether every wff that is identically true is a theorem. That is, it essentially asks whether every identically true statement that can be expressed within the system can also be proved within the system. This is clearly a desirable property for a logical system to have. We shall now briefly detail the propositional and predicate calculii, and consider their decision and completeness problems.

The propositional calculus, the ‘logic’ referred to in computer hardware, only allows wffs that (apart from fixed symbols) have simple variable names, say p, q, and r. For example in the wff

\[(p \land q) \lor r\]

one might interpret p as ‘the voltage on line one is 9 volts’. The decision problem is soluble by truth tables or Karnaugh maps. The propositional calculus is complete.

The predicate calculus, as its name suggests, allows one to name not only individuals but properties or predicates of variables. Given a predicate P(x) one can ask is it true for all x or (at least) if there exists some x for which it is true. Consequently the predicate calculus allows ‘quantification’: the use of fixed symbols to stand for ‘for every’ and ‘there exists’.

The first-order predicate calculus restricts the application of quantifiers to simple variables so that

\[(\forall x)\ (M(x) \implies H(x))\]

(which might be interpreted as every Man is Human) is a legal wff, whereas the principle of mathematical induction

\[(\forall P)(P(0) \land (\forall n \in \mathbb{N})(P(n) \implies P(n+1)))\implies (\forall n)(P(n))\]

is not. This last expression is a wff in the second-order predicate calculus, which differs from the first only in that quantification over predicate names is allowed.

It turns out that the first-order predicate calculus (and hence the second order also) is undecidable, a result first proved by Church (1936) in his original paper on computability.

The first-order predicate calculus is nevertheless complete, a result first proved by Kurt Gödel in 1930. However, the second-order predicate calculus is incomplete, a result proved by Gödel in 1931 and one which had about the same effect on mathematics and logic as Einstein’s theory of Relativity had on physics. Gödel’s theorem says that any sufficiently rich formal system, for example the second-order predicate calculus and hence most of mathematics, is either inconsistent (in the sense that both a theorem and its negation may be proved) or is incomplete, i.e. contains valid wffs which cannot be proved within the system.

This result came as a bombshell to nineteenth century mathematicians and finally put paid to Hilbert’s program for axiomatically constructing the whole of mathematics and proving it consistent.
<table>
<thead>
<tr>
<th>Logic type</th>
<th>Decidable</th>
<th>Complete</th>
</tr>
</thead>
<tbody>
<tr>
<td>Propositional</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>First-order</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Second-order</td>
<td>No</td>
<td>No</td>
</tr>
</tbody>
</table>
APPENDIX II Partial derivatives and Extreme points in higher dimensions

Partial differentiation

Consider a real valued function of several variables. E.g.

\[ f(x, y) = \sqrt{x^2 + y^2} \]

is a function of two variables.

It is a perfectly reasonable question to ask about the slope at \((x, y)\) in (say) the \(x\) direction. This will be the derivative with respect to \(x\) where \(y\) is held constant. We write

\[ \frac{\partial f}{\partial x} = \frac{1}{2}(x^2 + y^2)^{-1/2} \cdot 2x = \frac{x}{\sqrt{x^2 + y^2}} \]

where \(\partial f/\partial x\) replaces \(df/dx\) to indicate that all other variables are being held constant.

We can think of a function of two variables as a surface, where the height above the point \((x, y)\) is \(f(x, y)\).

With functions of more variables it is harder to visualise but we still often refer to the ‘surface’ \(f(x, y, z)\) for example.

With functions of several variables we can often find extreme points (local maxima, minima and saddle points, etc.) by looking at the partial derivatives and finding those points where all partial derivatives vanish. In this case if we try to solve
we run into the problem that the partial derivatives are not well defined at (0, 0) but more often than not we can find
extreme points in this way.

**Exercise.** Sketch the surface \( f(x, y) = x^2 + y^2 \) and find all extreme points.

**Mixed partial derivatives**

Continuing with our example we can write

\[
\frac{\partial f}{\partial x} = \frac{x}{\sqrt{x^2 + y^2}}, \quad \frac{\partial f}{\partial y} = \frac{y}{\sqrt{x^2 + y^2}}
\]

\[
\frac{\partial}{\partial y} \frac{\partial f}{\partial x} = \frac{\partial^2 f}{\partial y \partial x} = -\frac{xy}{(x^2 + y^2)^{3/2}}
\]

\[
\frac{\partial}{\partial x} \frac{\partial f}{\partial y} = \frac{\partial^2 f}{\partial x \partial y} = -\frac{xy}{(x^2 + y^2)^{3/2}}
\]

So that in this case

\[
\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x} \quad (x, y) \neq (0, 0)
\]

i.e the second order **mixed** partial derivatives are equal. In fact this is true for most well behaved functions but it is
not always true.

**Example**

![Figure 14](image-url)

Figure 14 The graph of a differentiable function whose mixed partial derivatives at the origin do not have the same sign.
Consider the function

\[ f(x, y) = \begin{cases} 
xy(x^2 + y^2), & \text{if } (x, y) \neq 0, \\
0, & \text{if } (x, y) = 0 
\end{cases} \]

The graph is shown in Figure 14.

Geometrically, this says that if you walk away from the origin along the y-axis, your shoes will tilt to an ever steeper slope in the y direction \((f_y(0, 0) = 1)\), whereas if you walk along the y-axis your shoes tilt more and more steeply in the negative x direction. \((f_x(0, 0) = -1)\).

This happens because the second order mixed partial derivatives are not continuous at \((0, 0)\).

**Extreme points of functions of several variables**

As we have seen with functions of several variables we can often find extreme points (local maxima, minima and saddle points, etc.) by looking at the partial derivatives and finding those points where all partial derivatives vanish.

**Example.** Consider the function

\[ f(x, y) = 3x \, \text{Exp}(y) - x^3 - \text{Exp}(3y) \]

\[ \frac{\partial f}{\partial x} = 3\text{Exp}(y) - 3x^2 \]

\[ \frac{\partial f}{\partial y} = 3x \, \text{Exp}(y) - 3\text{Exp}(y) \]

Equating both partial derivatives to zero we obtain \((x, y) = (1, 0)\) as the only solution. A simple examination shows that this is a local maximum.

Is it the global maximum?

**Only-Extreme-Point-In-Town test**

With a smooth function of one variable if there is a local maximum and no other extreme point then the local maximum must be a global maximum because if the function was somewhere greater than the local maximum, there would have to be a local minimum between the local maximum and the higher point.

This leads to the only-extreme-point-in-town test:

If \(f(x)\) has a single extreme point and that point is a local maximum (minimum) then \(f\) has a global maximum (minimum) there.

*This test fails for functions of two or more variables.* In fact our last example is a counter example.
Figure 15 illustrates a graph of the function in which $x$ and $y$ have been replaced by $\tan x$ and $\tan y$ and then arctan of the resulting value is taken. This has the effect of compressing the graph over the entire $x$-$y$ plane to a graph contained in the cube of side length $\pi$, centred at the origin.

The bump is a local maximum that is not a global maximum, yet there are no saddle points or extreme points elsewhere on the surface. (It looks as though there is a saddle point on the left boundary in the image, but because of the compression of the entire $x$-$y$ plane to a square, this represents a saddle point of $f$ at infinity.)
APPENDIX III Problems and solutions

Problems

1. (i) Find the sum of the series
\[ \sum_{r=1}^{n} (r + 1)(r + 3) \]

(ii) State the Principle of Induction.

(iii) Define \( f(n) = 10^n + 3(4^{n+2}) + 5 \) for \( n = 1, 2, 3, \ldots \). Prove by induction that \( f(n) \) is divisible by 9 for all positive integers \( n \).

2. (i) Prove by induction De Moivre’s theorem: For positive integral \( n \geq 1 \).
\[ (\cos \theta + i \sin \theta)^n = \cos n\theta + i \sin n\theta \]

[Assume any trigonometric formulae you may require, but these should be clearly stated.]

(ii) Hence show that
\[ \cos 3\theta = 4\cos^3 \theta - 3\cos \theta \]
and
\[ \sin 3\theta = 3\sin \theta - 4\sin^3 \theta \]

(iii) Expanding \( e^{i\theta} \) into a power series show why \( \cos \theta + i \sin \theta = e^{i\theta} \).

(iv) Using (iii) or otherwise differentiate
\[ e^{2x} \cos 3x \quad \text{and} \quad e^{2x} \sin 3x \]

with respect to \( x \).
3. (i) Write
\[ S_n = \sum_{r=1}^{n} \frac{1}{r(r + 1)(r + 2)} = \frac{1}{2.3} + \frac{1}{3.4} + \ldots + \frac{1}{n(n + 1)(n + 2)} \]

By substituting
\[ \frac{1}{r(r + 1)(r + 2)} = \frac{1}{2} \left( \frac{1}{r(r + 1)} - \frac{1}{(r + 1)(r + 2)} \right) \]

find \( \lim S_n \) as \( n \to \infty \).

(ii) Does the series
\[ \sum_{r=1}^{n} \frac{4r + 1}{3r^2 - 1} \]

converge or diverge (give detailed reasons)?

(iii) Expand \( \log(2 - 5x) \) in powers of \( x \) as far as the term in \( x^3 \) and give the general term. For what range of \( x \) is the expansion valid?

(iv) Show (by substitution or otherwise) that
\[ \int \frac{dx}{1 + x^2} = \arctan x + C \]

where \( C \) is an arbitrary constant. Hence by expanding \( (1 + x^2)^{-1} \) as a power series in \( x \) and integrating term-by-term derive a power series for \( \arctan x \). Deduce that
\[ \frac{\pi}{4} = 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \frac{1}{9} - \ldots \]

4. (a) Find all solutions of
\[(i) \quad 6 \cos^2(x) - \cos(x) - 1 = 0 \]
\[(ii) \quad 2\cos^2(x) - 5 \cos(x) + 2 = 0 \]

for \( x \) in the range \( 0 \leq x \leq 2\pi \)

(b) Prove the trigonometric identities
\[ (i) \quad (\sin(0) + \cos(0))(1 - \sin(0)\cos(0)) = \sin^2(0) + \cos^2(0) \]
\[ (ii) \quad \cos^2(0) - \sin^2(0) + 1 = 2\cos^2(0) \]
\[ (iii) \quad \sin(A + B) \cdot \sin(A - B) = \sin^2(A) - \sin^2(B) \]
\[ (iv) \quad \frac{\sin(A - B)}{\cos(A) \cdot \cos(B)} + \frac{\sin(B - C)}{\cos(A) \cdot \cos(C)} + \frac{\sin(C - A)}{\cos(C) \cdot \cos(A)} = 0 \]

5. Show that the function
\[ S_n(x) = \frac{1}{2}a_0 + \sum_{k=1}^{n} \left( a_k \cos(kx) + b_k \sin(kx) \right) \]
can be written in the form
\[ \sum_{k=-n}^{n} A_k e^{ikx} \]

where
\[ a_k = A_k + A_{-k}, \quad b_k = i(A_k - A_{-k}) \quad (1 \leq k \leq n) \]
\[ A_0 = \frac{a_0}{2} \]

6. (a) For each of the following statements determine the smallest integer \( d > 0 \) which makes the statement valid

i) \( x(x + 1) = O(x^d) \) as \( x \to 0 \)

ii) \( x(x + 1) = O(x^d) \) as \( x \to \infty \)

iii) \( x\sqrt{x} = o(x^d) \) as \( x \to 0 \)

iv) \( x\sqrt{x} = o(x^d) \) as \( x \to \infty \)

v) \( n(n + 1)(n + 2) - n^d \) as \( n \to \infty \)

(b) Find the following limits where they exist

i) \( \lim_{x \to 1} x^2(\sin \pi x + \cos \pi x) \)

ii) \( \lim_{x \to 0} \sin \left( \frac{1}{x} \right) \)

iii) \( \lim_{x \to 0} x \sin \left( \frac{1}{x} \right) \)

iv) \( \lim_{x \to 1} \frac{x^4 - 1}{x - 1} \)

7. (a) Briefly explain by means of a diagram how the expression
\[ \int_{a}^{b} f(x)dx \]

is defined in terms of a limiting process involving upper and lower sums.

(b) Show how the construction given in (a) can be used to compute
\[ \int_{0}^{b} xdx \]

where \( b > 0 \) directly from first principles.

8. (a) (Modular arithmetic) Solve \( 7x + 3 \equiv 2 \) in \( \mathbb{F}_{11} \), the finite field of integers \( \text{mod} \ 11 \)
Antonia J. Jones: 15 December 2003

(b) (Modular arithmetic) Solve
\[
\begin{align*}
2x + 8y & \equiv 1 \pmod{13} \\
3x - 5y & \equiv 3 \pmod{13}
\end{align*}
\]
in \(F_{13}\), the finite field of integers (mod 13).

(c) (Complex numbers) Express \((-1 + 3/3 + i(-3 - \sqrt{3}))(1 + i3)\) in the form \(a + ib\) \((a, b \in \mathbb{R})\) and as \(re^{i\theta}\), where \(r > 0\) and \(\theta\) are real.

(d) (Complex numbers) Show that if \(z = r(\cos \theta + i \sin \theta)\) where \(r > 0\) then for \(z \neq 1\)
\[
\frac{1}{1 + z} = \frac{1}{1 + 2r \cos \theta + r^2} - \frac{ir \sin \theta}{1 + 2r \cos \theta + r^2}
\]

*9. Let \(a_1\) and \(b_1\) be any two positive numbers and let \(a_1 < b_1\). For \(n \geq 2\) let \(a_n\) and \(b_n\) be defined by the equations
\[
a_n = \sqrt{a_{n-1}b_{n-1}}, \quad b_n = \frac{a_{n-1} + b_{n-1}}{2}
\]
Prove that

(a) \(a_n < b_n\) for all \(n \geq 1\).

(b) The sequence \(a_1, a_2, ...\) converges.

(c) The sequence \(b_1, b_2, ...\) converges.

(d) That the two sequences have the same limit.

[Hint: Use the theorem that a bounded monotonic sequence converges.]

(*) Indicates a challenging question - harder than typical examination questions.

10. (a) (i) A particle starts at the origin with zero velocity and moves along the \(x\)-axis so that its acceleration
\[
d^2x/dt^2 = 1 + \cos \pi t
\]
Find the position of the particle at time \(t\).

(ii) Sketch the graph of the velocity against time.

(b) (i) For \(a > 0\) show that
\[
\int_0^1 \frac{x^{a-1}}{1 + x} \, dx = \frac{1}{a} - \frac{1}{a+1} + \frac{1}{a+2} - \frac{1}{a+3} + \ldots \quad (|x| < 1)
\]

(ii) Hence deduce the sum of the series
\[
1 + \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \ldots
\]

11. (i) By considering \(2^n S_n - S_n\) (or by guessing and proving your guess by induction) find the sum of the series
\[
S_n = \sum_{r=1}^n r 2^r = 1.2 + 2.2^2 + 3.2^3 + \ldots + n 2^n
\]

128
for $n \geq 1$. [N.B. the dot in the expression for $S_n$ indicates multiplication.]

(ii) Define $f(n) = 11^n + 7(11^{n-1}) + 12$ for $n = 1, 2, 3, \ldots$. Prove by induction that $f(n)$ is divisible by 30 for all positive integers $n$.

12. For a positive integer $n$ let $\omega$ be a complex primitive $n$th root of unity, i.e. $\omega^n = 1$ and $\omega^r \neq 1$ for $0 < r < n - 1$.

(i) By considering the sum of a GP or otherwise show that for all $z \in \mathbb{C}$

$$1 + z + z^2 + \ldots + z^{n-1} = \begin{cases} \frac{z^n - 1}{z - 1} & \text{if } z \neq 1, \\ n & \text{if } z = 1. \end{cases}$$

(ii) Putting $z = \omega$ in the above formula show that

$$1 + \omega + \omega^2 + \ldots + \omega^{n-1} = 0$$

(iii) Suppose $m$ is not a multiple of $n$. By putting $z = \omega^m$ in the formula of part (i) show that $\omega^m (0 \leq m \leq n-1)$ are all of the $n$th roots of unity and that

$$(\omega^0)^m + (\omega^1)^m + (\omega^2)^m + \ldots + (\omega^{n-1})^m = 0$$

[Hint: first show $\omega^m = 1$.]

(iv) Now suppose $m$ is a multiple of $n$, say $m = kn$ where $k \in \mathbb{N}$. Show that

$$(\omega^0)^m + (\omega^1)^m + (\omega^2)^m + \ldots + (\omega^{n-1})^m = n$$

(v) Hence, using (iii) and (iv) with $n = 3$, show that the sum of every third term of the series

$$\text{Exp}(x) = \sum_{r=0}^{\infty} \frac{x^r}{r!}$$

is

$$\sum_{k=0}^{\infty} \frac{x^{3k}}{(3k)!} = \frac{1}{3} \left( \text{Exp}(x) + \text{Exp}(\omega x) + \text{Exp}(\omega^2 x) \right)$$

where $\omega$ is a complex cube root of unity.
Solutions

See also the Mathematica file QuestionsAndSolutions.nb for ways in which to use Mathematica to solve, or help to solve, these questions.

1. (i) The general term is \( r^2 + 4r + 3 \). Therefore

\[
\sum_{r=1}^{n} (r^2 + 4r + 3) = \sum_{r=1}^{n} r^2 + \sum_{r=1}^{n} 4r + \sum_{r=1}^{n} 3
\]

\[
= \frac{1}{6}n(n + 1)(2n + 1) + 4 \frac{1}{2}n(n + 1) + 3n
\]

\[
= \frac{1}{6}n^2(2n^2 + 15n + 31)
\]

(ii) **Principle of induction.** If \( P \) is a proposition about natural numbers such that \( P(1) \) is true and \( P(k) \) implies \( P(k+1) \) then for all \( k \in \mathbb{N} \) then \( P(n) \) is true for all \( \forall n \in \mathbb{N} \).

(iii) Consider

\[ f(n + 1) - f(n) = (10^{n+1} + 3.4^{n+3} + 5) - (10^n + 3.4^{n+2} + 5) \]

\[ = 10^n(10 - 1) + 3.4^{n+2}(4 - 1) \]

\[ = 9(10^n + 4^{n+2}) \]

Thus if \( f(n) \) is divisible by 9 then so is \( f(n + 1) \). But \( f(1) = 10 + 3.64 + 5 = 207 = 9 \cdot 23 \) is divisible by 9, so the required result follows by induction.

2. (i) For \( n = 1 \) the result is certainly true. If the result holds for \( n = k \) we have

\[ (\cos\theta + i\sin\theta)^k = \cos k\theta + i\sin k\theta \]

On multiplying both sides by \( \cos\theta + i\sin\theta \) we obtain

\[ (\cos\theta + i\sin\theta)^{k+1} = (\cos\theta + i\sin\theta)(\cos k\theta + i\sin k\theta) \]

\[ = (\cos(\cos k\theta) - \sin \theta \sin k\theta) + i(\sin(\cos k\theta) + \cos \theta \sin k\theta) \]

If we now use the addition formula for cosine and sine, i.e.

\[ \cos(A + B) = \cos A \cos B - \sin A \sin B \]

\[ \sin(A + B) = \sin A \cos B + \cos A \sin B \]

with \( A = \theta \) and \( B = k\theta \) we obtain

\[ (\cos\theta + i\sin\theta)^{k+1} = \cos(k+1)\theta + i\sin(k+1)\theta \]

Hence if the result is true for \( n = k \) then the result is true for \( n = k + 1 \). However, we have seen the result is true for \( n = 1 \). Hence by the principle of induction the result is true for all \( n \geq 1 \).

(ii) Using De Moivre and expanding we have
\[ \cos 3\theta + i \sin 3\theta = (\cos \theta + i \sin \theta)^3 \]
\[ = \cos^3 \theta + 3i \cos^2 \sin \theta + 3i^2 \cos \sin^2 \theta + i^3 \sin^3 \theta \]
\[ = \cos^3 \theta - 3\cos \sin^2 \theta + i(3\cos^2 \sin \theta - \sin^3 \theta) \]

Equating real and imaginary parts and using \( \cos^2 \theta + \sin^2 \theta = 1 \) we obtain the results.

(iii) Using
\[ e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \ldots \quad (\text{for all } x \in \mathbb{C}) \]
and putting \( x = i\theta \) we have
\[ e^{i\theta} = 1 + i\theta + \frac{(i\theta)^2}{2!} + \frac{(i\theta)^3}{3!} + \frac{(i\theta)^4}{4!} + \ldots \]
\[ = \left( 1 - \frac{\theta^2}{2!} + \frac{\theta^4}{4!} - \ldots \right) + i \left( \theta - \frac{\theta^3}{3!} + \frac{\theta^5}{5!} - \ldots \right) \]
\[ = \cos \theta + i \sin \theta \]

(iv) Let \( C = e^{2\cos 3x} \) and \( S = e^{2\sin 3x} \). Then \( C + iS = e^{(2+3i)x} \) and so
\[ \frac{d}{dx}(C + iS) = (2 + 3i)e^{(2+3i)x} \]
\[ = (2 + 3i)(\cos 3x + i\sin 3x) \]

Equating real and imaginary parts gives
\[ \frac{d}{dx}(e^{2\cos 3x}) = e^{2x}(2\cos 3x - 3\sin 3x) \]
and
\[ \frac{d}{dx}(e^{2\sin 3x}) = e^{2x}(3\cos 3x + 2\sin 3x) \]

3. (i) Using the difference method with the given identity we find
\[ S_n = \sum_{r=1}^{n} \frac{1}{r(r + 1)(r + 2)} = \sum_{r=1}^{n} \frac{1}{2} \left( \frac{1}{r(r + 1)} - \frac{1}{(r + 1)(r + 2)} \right) \]
\[ = \frac{1}{2} \left( \frac{1}{1 \cdot 2} - \frac{1}{2 \cdot 3} \right) + \frac{1}{2} \left( \frac{1}{2 \cdot 3} - \frac{1}{3 \cdot 4} \right) + \ldots + \frac{1}{2} \left( \frac{1}{(n-1)n} - \frac{1}{n(n+1)} \right) \]
\[ = \frac{1}{4} - \frac{1}{2(n+1)(n+2)} \]

Hence \( \lim S_n = 1/4 \).

(ii) Now
\[ \frac{4r + 1}{3r^2 - 1} > \frac{4r}{3r^2} = \frac{4}{3} \cdot \frac{1}{r} \]
and \( \sum 1/r \) is divergent. It follows that the original series is divergent by the Comparison test.
(iii) We write
\[
\log(2 - 5x) = \log 2 - \frac{5x}{2} = \log 2 + \log \left(1 - \frac{5x}{2}\right)
\]

We then expand the second term using the series \(\log(1 + z) = z - \frac{z^2}{2} + \frac{z^3}{3} - \ldots\) to obtain
\[
\log(2 - 5x) = \log 2 + \left[\frac{5}{2} - \frac{1}{2}\left(\frac{5x}{2}\right)^2 + \frac{1}{3}\left(\frac{5x}{2}\right)^3 + \ldots + \frac{1}{r}\left(\frac{5x}{2}\right)^r + \ldots\right]
\]
\[
= \log 2 - \frac{5x}{2} - \frac{25x^2}{8} - \frac{125x^3}{24} - \ldots - \frac{5r}{2^{2r-1}}x^r - \ldots
\]

The expansion is valid when \(-1 < (5/2)x < 1\), i.e. \(-2/5 < x < 2/5\).

(iv) Put \(x = \tan \theta\) so that \(\theta = \arctan x\). Then
\[
\int \frac{dx}{1 + x^2} = \int \frac{\sec^2 \theta d\theta}{1 + \tan^2 \theta} = \int d\theta = \theta + C
\]

Now write

\[
\arctan x = \int \frac{dx}{1 + x^2}
\]
\[
= \int dx - \int x^2 dx + \int x^4 dx - \ldots
\]
\[
= x - \frac{x^3}{3} + \frac{x^5}{5} - \ldots
\]

where \(C = 0\) because \(\arctan 0 = 0\). Now put \(\theta = \pi/4\), i.e. \(x = 1\). Then we obtain
\[
\arctan 1 = \frac{\pi}{4} = 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \frac{1}{9} - \ldots
\]
as required.

4. (a) Factorising we obtain 6\(\cos^2 x\) - \(\cos x\) + 2 = (3\(\cos x\) + 1)(2\(\cos x\) - 1) = 0. Thus

\[\text{Either } 3\cos x + 1 = 0 \text{ i.e. } \cos x = -1/3 \text{ Or } 2\cos x - 1 = 0 \text{ i.e. } \cos x = 1/2.\]

Now solutions to \(\cos x = -1/3\) in the range \(0 \leq x < 360\) follow from arccos(-1/3) = 109.47° or 250.53°. Similarly, for \(\cos x = 1/2\) we obtain \(x = 60°\) or \(x = 300°\). Thus all solutions to the original equation in the range \(0 \leq x < 360\) are given by \(x = 60°, 109.47°, 250.53°, 300°\).

Now in fact there are infinitely many solutions given by any one of the above plus \(360k\) (\(k = 0, \pm1, \pm2, \ldots\)).

[Marked would be reduced if no comment about infinitely many solutions]

(b) Factorising we obtain 2\(\cos^2 x\) - 5\(\cos x\) + 2 = (2\(\cos x\) - 1)(\(\cos x\) - 2) = 0. Thus

\[\text{Either } 2\cos x - 1 = 0 \text{ i.e. } \cos x = 1/2 \text{ Or } \cos x - 2 = 0 \text{ i.e. } \cos x = 2.\]

Now solutions to \(\cos x = 1/2\) in the range \(0 \leq x < 360\) are \(x = 60°\) or \(300°\). But \(\cos x = 2\) has no real solutions (it does have complex solutions but the question meant real solutions). Hence all solutions to the original equation are \(x = \ldots\)
60° + 360k or 300° + 360k (k = 0, ±1, ±2, ...).

[Marks would be reduced if the fact that there are no solutions to \( \cos x = 2 \) is not mentioned]

(c) Prove the trigonometric identities.

(i) Multiplying out the LHS we have

\[
\begin{align*}
\sin \theta (1 - \sin^2 \theta) + \cos \theta (1 - \sin^2 \theta) &= \sin \theta - \sin^3 \theta + \cos \theta - \sin \theta \\
&= \sin \theta - \cos \theta + \cos \theta - \sin \theta + \sin^2 \theta \\
&= \sin^2 \theta + \cos^2 \theta
\end{align*}
\]

(ii) Again using \( \cos 2\theta + \sin 2\theta = 1 \) we obtain using the difference of two squares

\[
\begin{align*}
\cos^4 \theta - \sin^4 \theta + 1 &= \cos^2 \theta - \sin^2 \theta + 1 + 1 \\
&= 2\cos^2 \theta
\end{align*}
\]

(iii) Using the addition formula for \( \sin(A + B) \) and the difference of two squares we have

\[
\begin{align*}
\sin(A + B) \cdot \sin(A - B) &= (\sin A \cos B + \cos A \sin B)(\sin A \cos B - \cos A \sin B) \\
&= (\sin A \cos B)^2 - (\cos A \sin B)^2 = \sin^2 A(1 - \sin^2 B) - (1 - \sin^2 A) \sin^2 B \\
&= \sin^2 A - \sin^2 A \sin^2 B - \sin^2 B + \sin^2 A \sin^2 B = \sin^2 A - \sin^2 B
\end{align*}
\]

(iv) Using \( \sin(A - B) = \sin A \cos B - \cos A \sin B \) we have

\[
\begin{align*}
\tan A - \tan B
\end{align*}
\]

and similarly for the other terms. Hence the LHS of the given identity becomes just

\[
\tan A - \tan B + \tan B - \tan C + \tan C - \tan A
\]

which obviously sums to zero.

5. Using

\[
\begin{align*}
\cos \theta &= \frac{1}{2} \left( e^{i \theta} + e^{-i \theta} \right) \\
\sin \theta &= \frac{1}{2i} \left( e^{i \theta} - e^{-i \theta} \right)
\end{align*}
\]

we obtain
Antonia J. Jones: 15 December 2003

\[ S_n(x) = \frac{a_0}{2} + \frac{1}{2} \sum_{k=1}^{n} \left( a_k e^{ikx} + e^{-ikx} \right) - ib_k (e^{ikx} - e^{-ikx}) \]

\[ = \frac{a_0}{2} + \frac{1}{2} \sum_{k=1}^{n} \left( a_k - ib_k \right) e^{ikx} + \left( a_k + ib_k \right) e^{-ikx} \]

Now

\[ a_k - ib_k = 2A_k, \quad a_k + ib_k = 2A_k \]

so we obtain

\[ S_n(x) = \frac{a_0}{2} + \frac{1}{2} \sum_{k=1}^{n} \left( 2A_k e^{ikx} + 2A_k e^{-ikx} \right) \]

\[ = \frac{a_0}{2} + \frac{1}{2} \sum_{k=1}^{n} \left( A_k e^{ikx} + A_k e^{-ikx} \right) \]

\[ = \frac{a_0}{2} + \sum_{k=-n}^{n} A_k e^{ikx} - A_0 = \sum_{k=-n}^{n} A_k e^{ikx} \]

as required.

6. (a) (i) \( d = 1 \) (ii) \( d = 2 \) (iii) \( d = 1 \) (iv) \( d = 2 \) (v) \( d = 3 \).

(b) (i) \( \cos \pi + \sin \pi = -1 \)

(ii) Limit does not exist because \( \sin(1/x) \) takes values \( \pm 1 \) arbitrarily close to \( x = 0 \).

(iii) \( \sin(1/x) \) remains between -1 and +1, but \( x \) approaches zero. So \( x \sin(1/x) \to 0 \) as \( x \to 0 \).

(iv) For \( x \neq 1 \) we have

\[ \frac{x^3 - 1}{x - 1} = \frac{(x - 1)(x^2 + x + 1)}{(x - 1)} \]

\[ = x^2 + x + 1 \to 4 \quad \text{as} \quad x \to 1 \]

Remark. Easily generalises to

\[ \lim_{x \to 1} \frac{x^n - 1}{x - 1} = n \]

for positive integer \( n \).
7. a) The area under a curve \( y = f(x) \) which is written as

\[
\int_{a}^{b} f(x)dx
\]

is defined by a limiting process. We divide up the interval \([a, b]\) into a finite number of sub-intervals whose maximum length is \(\delta\) (say). Next we form the upper and lower rectangles as shown in Figure 16.

For a partition into \( n \) sub-intervals with \( x_0 = a \) and \( x_n = b \) with the maximum length of the sub-intervals being \(\delta\), we write

\[
L(\delta) = \sum_{i=0}^{n-1} \left( \min_{x \in [x_i, x_{i+1}]} f(x) \right) (x_{i+1} - x_i)
\]

\[
U(\delta) = \sum_{i=0}^{n-1} \left( \max_{x \in [x_i, x_{i+1}]} f(x) \right) (x_{i+1} - x_i)
\]

Then

\[
L(\delta) \leq U(\delta)
\]

where \(L(\delta)\) is the sum of the areas of the lower rectangles and \(U(\delta)\) is the sum of the areas of the upper rectangles. (There is no problem in defining the area of a rectangle in the usual way.)

As \(\delta \to 0\) one can show (under reasonable conditions on \(f\)) that \(L(\delta)\) is monotonic increasing and \(U(\delta)\) is monotonic decreasing. These functions are both bounded so that

\[
\lim_{\delta \to 0} L(\delta) \quad \text{and} \quad \lim_{\delta \to 0} U(\delta)
\]

both exist.

**Definition.** If both limits exists and are equal we define their common value to be

\[
\int_{a}^{b} f(x)dx
\]

This is called the *integral* of \(f\) from \(a\) to \(b\).
c) Because $f$ is monotonic increasing, we have

$$L(\delta) = \sum_{i=0}^{n-1} x_i(x_{i+1} - x_i)$$

$$U(\delta) = \sum_{i=0}^{n-1} x_{i+1}(x_{i+1} - x_i)$$

Now suppose the interval partition is into equal sub-intervals each of length $\delta = 1/n$. Then

$$x_i = \frac{ib}{n}$$

$$x_{i+1} - x_i = \frac{b}{n}$$

Thus

$$L(\delta) = \sum_{i=0}^{n-1} \frac{ib}{n} \frac{b}{n} = \frac{b^2}{n^2} \sum_{i=0}^{n-1} i = \frac{b^2}{2} \left( 1 - \frac{1}{n} \right)$$

$$U(\delta) = \sum_{i=0}^{n-1} \frac{(i+1)b}{n} \frac{b}{n} = \frac{b^2}{n^2} \sum_{i=0}^{n-1} (i + 1) = \frac{b^2}{2} \left( 1 + \frac{1}{n} \right)$$

Hence

$$\lim_{\delta \to 0} L(\delta) = \frac{b^2}{2} = \lim_{\delta \to 0} U(\delta)$$

so that

$$\int_{x=0}^{b} x \, dx = \frac{b^2}{2}$$

8. (a) $7x + 3 \equiv 2 \pmod{11} \iff 7x \equiv -1 \pmod{11} \equiv 7x \equiv 10 \equiv x \equiv 3 \pmod{11}$.

(b) We first try to eliminate $y$. To do this we solve $8z \equiv 5 \pmod{13}$ so that when we multiply the first equation by $z$ we can then add both equations and eliminate $y$. We notice that $8(-1) \equiv 5 \pmod{13}$ so that a solution is $z \equiv -1 \equiv 12 \pmod{13}$. Multiplying (mod 13) the first equation by 12 we obtain

$$11x + 5y = 12 \pmod{13}$$

$$3x - 5y \equiv 3 \pmod{13}$$

Adding (mod 13) the two equations we obtain $x \equiv 2 \pmod{13}$. Substitution back into the first equation gives $8y \equiv -3 \pmod{13} \Rightarrow y \equiv -2 \equiv 11 \pmod{13}$. So the solution is $x \equiv 2 \pmod{13}$ and $y \equiv 11 \pmod{13}$.

(c) Express $(-1 + 3\sqrt{3} + i(-3 - 3\sqrt{3}))(1 + i3)$ in the form $a + ib$ ($a, b \in \mathbb{R}$) and as $re^{i\theta}$, where $r > 0$ and $\theta$ are real.
Antonia J. Jones: 15 December 2003

\[ z = \frac{-1 + 3\sqrt{3} + i(-3 - \sqrt{3})}{1 + i3} = \frac{(-1 + 3\sqrt{3} + i(-3 - \sqrt{3}))}{(1 + i3)(1 - i3)} \]

\[ = \frac{-1 + 3\sqrt{3} + i(-3 - \sqrt{3}) + i3 - 9i\sqrt{3} + 3(-3 - \sqrt{3})}{1 + 9} \]

\[ = \frac{1}{10}(-10 - 10i\sqrt{3}) = -1 - i\sqrt{3} \]

Now \(|-1 - i\sqrt{3}|^2 = (-1)^2 + (\sqrt{3})^2 = 1 + 3 = 4\) so \(r = 2\) and \(\text{Arg}(1 - i\sqrt{3}) = -2\pi/3 = 0\). Thus \(z = 2(\cos(-2\pi/3) + i\sin(-2\pi/3))\).

(d) Observing that

\[ \frac{1}{1 + z} = \frac{1}{1 + r\cos(\theta) + ir\sin(\theta)} \]

\[ = \frac{(1 + r\cos(\theta) - ir\sin(\theta))}{(1 + r\cos(\theta) + ir\sin(\theta))(1 + r\cos(\theta) - ir\sin(\theta))} \]

\[ = \frac{1 + r\cos(\theta) - ir\sin(\theta)}{(1 + r\cos(\theta))^2 + r^2\sin(\theta)} = \frac{1 + r\cos(\theta) - ir\sin(\theta)}{1 + 2r\cos(\theta) + r^2} \]

we have the required result on separating into real and imaginary parts.

9. a) Since \(a_{n,1}^2 - 2a_{n,1}b_{n,1} + b_{n,1}^2 \geq 0\) with equality if and only if \(a_{n,1} = b_{n,1}\) we have

\[ a_n^2 = a_{n,1}b_{n,1} \leq \frac{(a_{n,1} + b_{n,1})^2}{4} = b_n^2 \]

Thus by induction \(a_n < b_n\) for all \(n \geq 1\), with strict inequality since \(a_1 < b_1\).

(b) and (c). We observe that

\[ a_2 = \sqrt{a_1b_1} > \sqrt{a_1a_1} = a_1 \]

\[ b_2 = \frac{a_1 + b_1}{2} < \frac{b_1 + b_1}{2} = b_1 \]

which is the base case of the induction

\[ a_n = \sqrt{a_{n-1}b_{n-1}} > \sqrt{a_{n-1}a_{n-1}} = a_{n-1} \]

\[ b_n = \frac{a_{n-1} + b_{n-1}}{2} < \frac{b_{n-1} + b_{n-1}}{2} = b_{n-1} \]

provided \(a_n < b_n\) for all \(n\). But we have just shown this in (a). Hence by induction \(a_n\) is strictly monotonic increasing and \(b_n\) is strictly monotonic decreasing. But \(a_n < b_n < b_1\) and \(a_n > a_1 > b_1\). Hence \(a_n\) is bounded above and \(b_n\) is bounded below. Hence by the theorem on bounded monotonic sequences \(a_n\) and \(b_n\) are each convergent.

[Note: The *Mathematica* file QuestionsandSolutions.nb explains how to calculate the common limit which is related to a method of Gauss for quickly calculating elliptic integrals. These are integrals which turn up in a variety of physical problems - e.g. the simple pendulum - and which cannot be expressed in terms of standard functions.]
Antonia J. Jones: 15 December 2003

(d) Suppose \( a_n \to r \) and \( b_n \to s \) as \( n \to \infty \) then from the second defining equation we have
\[ s = \frac{1}{2}(r + s) \]
from which it follows that \( r = s \).

10. (a) (i) Integrating with respect to \( t \) we obtain
\[
\frac{dx}{dt} = t + \frac{1}{\pi} \sin \pi t + C
\]
Since \( \frac{dx}{dt} = 0 \) at \( t = 0 \) we have \( C = 0 \) and integrating once again we obtain
\[
x = \frac{t^2}{2} - \frac{1}{\pi} \cos \pi t + C
\]
Since \( x = 0 \) at \( t = 0 \) we have \( C = 1/\pi^2 \) and so the position at time \( t \) is given by
\[
x = \frac{t^2}{2} - \frac{1}{\pi} \cos \pi t + \frac{1}{\pi^2}
\]

(ii) Sketch of velocity against time.

![Figure 17 Velocity against time.](image)

(b) (i) We have
\[
\int_{0}^{1} \frac{x^{a-1}}{1+x} \, dx = \int_{0}^{1} x^{a-1}(1 - x + x^2 - x^3 + x^4 - x^5 + ...) \, dx
\]
\[
= \int_{0}^{1} (x^{a-1} - x^a + x^{a+1} - x^{a+2} + ...) \, dx
\]
\[
= \left[ \frac{x^a}{a} - \frac{x^{a+1}}{a+1} + \frac{x^{a+2}}{a+2} - \frac{x^{a+3}}{a+3} + ... \right]_{x=0}^{x=1}
\]
\[
= \frac{1}{a} - \frac{1}{a+1} + \frac{1}{a+2} - \frac{1}{a+3} + ...
\]

(ii) Putting \( a = 1 \) gives
\[
\int_{0}^{1} \frac{dx}{1+x} = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + ...
\]
But
\[ \int_0^1 \frac{dx}{1 + x} = \left[ \log(1 + x) \right]_0^1 = \log[2] \]

which is the required sum.

11. (i) Write

\[ S_n = 1.2 + 2.2^2 + 3.2^3 + \ldots + n2^n \]

\[ 2S_n = 1.2^2 + 2.2^3 + \ldots + (n - 1)2^{n-1} + n2^{n+1} \]

Hence

\[ S_n = 2S_n - S_n = n2^{n+1} - 2^n - 2^{n-1} - \ldots - 2^2 - 2 \]

\[ = n2^{n+1} - 2(2^{n-1} + 2^{n-2} + \ldots + 1) \]

\[ = n2^{n+1} - 2(2^n - 1) \quad \text{(sum of a G.P.)} \]

\[ = (n - 1)2^{n-1} + 2 \quad (n \geq 1) \]

which is the required sum.

(ii) Consider

\[ f(n + 1) - f(n) = (11^{n+1} + 7(11^n) + 12) - (11^n + 7(11^{n-1}) + 12) \]

\[ = 11^n(11 - 1) + 7.11^{n-1}(11 - 1) \]

\[ = 10(11 + 7).11^{n-1} = 30.6.11^{n-1} \]

Thus if \( f(n) \) is divisible by 30 then so is \( f(n+1) \). But \( f(1) = 11 + 7 + 12 = 30 \) is divisible by 30, so the required result follows by induction.

12. For a positive integer \( n \) let \( \omega \) be a complex primitive \( n \)th root of unity, i.e. \( \omega^n = 1 \) and \( \omega^r \neq 1 \) for \( 0 < r < n-1 \).

(i) By considering the sum of a GP or otherwise show that for all \( z \in \mathbb{C} \)

\[ 1 + z + z^2 + \ldots + z^{n-1} = \begin{cases} 
\frac{z^n - 1}{z - 1} & \text{if } z \neq 1, \\
1 & \text{if } z = 1.
\end{cases} \]

The sum of the GP

\[ 1 + z + z^2 + \ldots + z^{n-1} \]

with \( n \) terms and common ratio \( z \) is given by

\[ 1 + z + z^2 + \ldots + z^{n-1} = \frac{z^n - 1}{z - 1} \quad \text{if } z \neq 1 \]

If \( z = 1 \) the sum is obviously \( n \) and the required result follows.

(ii) Putting \( z = \omega \) in (*) show that

\[ 1 + \omega + \omega^2 + \ldots + \omega^{n-1} = 0 \]

139
Antonia J. Jones: 15 December 2003

Now putting \( z = \omega \neq 1 \) in (*) we obtain

\[
1 + \omega + \omega^2 + \ldots + \omega^{n-1} = \frac{\omega^n - 1}{\omega - 1} = 0
\]

as required.

(iii) Suppose \( m \) is not a multiple of \( n \). By putting \( z = \omega^m \) in (*) show that \( \omega^m (0 \leq m \leq n-1) \) are all of the \( n \)th roots of unity and that

\[
(\omega^0)^m + (\omega^1)^m + (\omega^2)^m + \ldots + (\omega^{n-1})^m = 0
\]

Plainly \((\omega^m)^n = (\omega^n)^m = 1, \) so \( \omega^m \) is an \( n \)th root of unity. Since \( \omega \) is a primitive \( n \)th root of unity the numbers 1, \( \omega, \) \( \omega^2, \ldots, \omega^{n-1} \) are all distinct. But by Gauss’ theorem \( z^n = 1 \) has exactly \( n \) roots. Hence the complex numbers 1, \( \omega, \omega^2, \ldots, \omega^{n-1}, \) where \( \omega \) is any primitive \( n \)th root of unity, must be all of the roots of the equation.

If \( m \) is not a multiple of \( n \) then by the remainder theorem it must be of the form \( m = hn + r, \) where \( h \) is an integer and \( 0 < r \leq n-1. \) Then \( \omega^m = \omega^{hn+r} = \omega^r = 1 \) since \( \omega \) is a primitive \( n \)th root of unity and \( 0 < r \leq n-1. \) Hence putting \( z = \omega^m = 1 \) in (*) we obtain

\[
1^m + (\omega^1)^m + (\omega^2)^m + \ldots + (\omega^{n-1})^m = \frac{\omega^m - 1}{\omega^m - 1} = 0 \quad \text{since} \quad \omega^m = 1, \quad \text{and} \quad \omega^m = 1.
\]

(iv) Now suppose \( m \) is a multiple of \( n, \) say \( m = kn \) where \( k \in \mathbb{N}. \) Show that

\[
(\omega^0)^m + (\omega^1)^m + (\omega^2)^m + \ldots + (\omega^{n-1})^m = n
\]

If \( m = kn \) then \( \omega^m = \omega^m = (\omega^n)^k = 1 \) since \( \omega^n = 1. \) Thus by (*) with \( z = \omega^m = 1 \) we have

\[
(\omega^0)^m + (\omega^1)^m + \ldots + (\omega^{n-1})^m = (\omega^0)^m + (\omega^1)^m + (\omega^2)^m + \ldots + (\omega^{n-1})^m = n
\]

as required.

(v) Hence, using (iii) and (iv) with \( n = 3, \) show that the sum of every third term of the series

\[
\exp(x) = \sum_{r=0}^{\infty} \frac{x^r}{r!}
\]

is

\[
\sum_{k=0}^{\infty} \frac{x^{3k}}{(3k)!} = \frac{1}{3} \left( \exp(x) + \exp(\omega x) + \exp(\omega^2 x) \right)
\]

where \( \omega \) is a complex cube root of unity.

For \( n = 3 \) either of \( \omega = (-1 + i \sqrt{3})/2 \) or \( (-1 - i \sqrt{3})/2 \) is a primitive cube root of unity. Hence, using the power series expansion for the exponential function and absolute convergence over the whole complex plane, we have

\[
\exp(x) + \exp(\omega x) + \exp(\omega^2 x) = \sum_{r=0}^{\infty} \sum_{h=0}^{\infty} \frac{(\omega^h)^r}{h!} = \sum_{h=0}^{\infty} \frac{(1^h + \omega^h + \omega^{2h})x^h}{h!} = 3 \sum_{k=0}^{\infty} \frac{x^{3k}}{(3k)!}
\]
since \( 1^h + \omega^h + \omega^{2h} \) is 0, if \( h \) is not a multiple of 3 by (iii), and is 3, if \( h \) is a multiple of 3, say \( h = 3k \), by (iv). The result then follows.

Not asked. As a matter of interest we note that

\[
\exp(\alpha x) = \exp\left( -\frac{1 + i\sqrt{3}}{2} x \right) = e^{-\frac{x}{\sqrt{3}}} \left( \cos \frac{\sqrt{3}}{2} x + i \sin \frac{\sqrt{3}}{2} x \right)
\]

\[
\exp(\omega^2 x) = \exp\left( -\frac{1 - i\sqrt{3}}{2} x \right) = e^{-\frac{x}{\sqrt{3}}} \left( \cos \frac{\sqrt{3}}{2} x - i \sin \frac{\sqrt{3}}{2} x \right)
\]

Hence

\[
\frac{1}{3} (\exp(x) + \exp(\alpha x) + \exp(\omega^2 x)) = \frac{1}{3} \exp(x) + \frac{2}{3} e^{-\frac{x}{\sqrt{3}}} \cos \frac{\sqrt{3}}{2} x
\]

which is (of course) real if \( x \) is real.
PRACTICE EXAM QUESTIONS #1

1. (a) (i) (Modular arithmetic) Solve \(5x - 7 = 5\) in \(\mathbb{F}_{13}\), the finite field of integers \((\text{mod } 13)\) \([5]\)

(ii) (Modular arithmetic) Solve the simultaneous equations

\[
2x + 8y \equiv 1 \pmod{11}
\]
\[
3x - 5y \equiv 3 \pmod{11}
\]

in the field of integers \((\text{mod } 11)\). \([10]\)

(b) (i) State the Principle of Induction. \([5]\)

(ii) By considering \(3S_n - S_n\) (or by guessing and proving your guess by induction) find the sum of the series

\[
S_n = \sum_{r=1}^{n} r3^r = 1.3 + 2.3^2 + 3.3^3 + \ldots + n3^n
\]

for \(n \geq 1\). [N.B. the dot in the expression for \(S_n\) indicates multiplication.] \([10]\)

(iii) Define \(f(n) = 5^n + 6(5^{n-1}) + 1\) for \(n = 1, 2, 3, \ldots\). Prove by induction that \(f(n)\) is divisible by 4 for all positive integers \(n\) \([10]\)

(c) In how many ways can a committee of 5 lecturers and 3 professors be formed from 10 lecturers and 8 professors? \([10]\)

2. (a) Express the following complex numbers in the form \(a + ib\), where \(a\) and \(b\) are real numbers:

(i) \((5 + 3i)(2 - i) - (3 + i)\) \([2]\)

(ii) \((3 - 2i)^2\) \([2]\)

(iii) \(\frac{1}{2} - \frac{3 - 4i}{5 - 8i}\) \([2]\)

(iv) Express \((-1 + 3\sqrt{3} + i(-3 - \sqrt{3}))(1 + i3)\) in the form \(a + ib\) \((a, b \in \mathbb{R})\) and as \(re^{i\theta}\), where \(r > 0\) and \(\theta\) are real. \([4]\)

(v) Express the following complex number in the form \(a + ib\), where \(a\) and \(b\) are real numbers:

\[
\left(\frac{1}{\sqrt{2}} + \frac{i}{\sqrt{2}}\right)^{12}
\]

[Hint: De Moivre’s theorem.]

(b) Prove by induction De Moivre’s theorem: For positive integral \(n \geq 1\)

\[
(cos \theta + i sin \theta)^n = cos n\theta + i sin n\theta
\]

[Assume any trigonometric formulae you may require, but these should be clearly stated.] \([5]\)
(c) If the point \( P \) in the complex plane corresponds to the complex number \( z = x + iy \) show that if
\[
|z - 1| = 2|z - 2 - 3i|
\]

then the locus of \( P \) is a circle centre at -3 + 4i and find the radius of the circle. \[10\]

(d) Given that \( 2 + i \) is a root of the equation \( z^3 - 11z + 20 = 0 \) find the remaining roots. \[5\]
[Hint: Page 14]
PRACTICE EXAM QUESTIONS #2

1. (a) State the integral test for convergence/divergence of the series \( \sum a_n \), and give a sketch illustrating the idea. [7]

(b) Discuss the example

\[ a_n = \frac{1}{n^3} \]  

when the integral test is applied to the series \( \sum a_n \).

(c) Explain the notion of the ‘radius of convergence’ of a power series. [2]

(d) Determine the radius of convergence of the power series

(i) \( \sum_{n=1}^{\infty} n! x^n \)  

(ii) \( \sum_{n=1}^{\infty} \frac{x^n}{n!} \)  

(iii) \( \sum_{n=1}^{\infty} \frac{x^n}{n(n+1)} \)  

2. (a) Expand the function up to the first four terms as a power series in \( x \). What is its radius of convergence? [5]

(b) \( f \) is a real valued function \( f: \mathbb{R} \to \mathbb{R} \). What does it mean to say that \( f \) is differentiable at \( x = a \)? [2]

(c) Let \( u(x) \), \( v(x) \) be real valued differentiable functions. State the rule for differentiating the product \( u(x) \cdot v(x) \) with respect to \( x \). [3]

(d) Differentiate the following functions with respect to \( x \).

(i) \( x^2 \cos x \)  

(ii) \( (\sin 2x)/(x^2 + 2) \)  

(e) Locate the maxima and minima of the function \( y = 2x^3 - 9x^2 + 12x + 6 \) and roughly sketch the curve. [5]

3. (a) Integrate the following functions with respect to \( x \).

(i) \( x \cos x \)  [Hint: By parts] [4]

(ii) \( 1/(x^2 - 1) \)  [Hint: partial fractions] [4]
(b) Evaluate the integrals

\[
\int \frac{x \cos x}{\sin^2 x} \, dx \quad \int x^2 e^x \, dx \\
\int x^4 \sqrt{1 - x^2} \, dx \quad \int x^m \log x \, dx \quad (m \neq -1)
\]

[3 Marks each]
Triangle inequality for complex numbers

We need the following simple results

\[ \overline{a + b} = \overline{a} + \overline{b} \]
\[ \overline{ab} = \overline{a} \overline{b} \]
\[ \overline{\overline{a}} = a \]
\[ a + \overline{a} = 2\text{Re}(a) \quad (\text{Re} \text{ denotes the real part of } a) \]
\[ \text{Re}(a) \leq |a| \]
\[ |ab| = |a||b| \]
\[ |\overline{a}| = |a| \]

which are left as exercises for the reader (they all follow directly from the definitions of modulus and complex conjugate).

**Proposition.** For any complex numbers \( z, w \) we have

\[ |z + w| \leq |z| + |w| \quad \forall \ z, w \in \mathbb{C} \]

**Proof.** We have

\[
|z + w|^2 = (z + w)(\overline{z} + \overline{w}) = (z + w)(\overline{z} + w)
= z\overline{z} + z\overline{w} + w\overline{z} + w\overline{w} = |z|^2 + z\overline{w} + \overline{z}w + |w|^2
= |z|^2 + 2\text{Re}(zw) + |w|^2 \leq |z|^2 + 2|z||w| + |w|^2
\leq (|z| + |w|)^2
\]

Taking square roots the result follows since all terms are positive.