Mathematics

Mathematics 3

Advanced Higher

7534

Summer 2000

HIGHER STILL

Mathematics

Mathematics 3

Advanced Higher

Support Materials

MATHEMATICS 3 (ADVANCED HIGHER)

Introduction

These support materials for Mathematics were developed as part of the Higher Still Development Programme in response to needs identified at needs analysis meetings and national seminars.

Advice on learning and teaching may be found in *Achievement for All* (SOEID 1996), *Effective Learning and Teaching in Mathematics* (SOEID 1993), *Improving Mathematics* (SEED 1999) and in the Mathematics Subject Guide.

These notes are intended to support teachers/lecturers in the teaching of Mathematics 3 (AH). The resources referred to within the material are:

Understanding Pure Mathematics,
AJ Sadler and DWS Thorning, O.U.P., 1987 ISBN 0-19-914243-2

The Complete A Level Mathematics,
Orlando Gough, Heinemann Educational Books, 1987, ISBN 0-435-51345-1

Mathematics In Action 6S,
John Hunter, Nelson Blackie, ISBN 0-17-441027-1

Calculus,
John Hunter (Blackie/Chambers) 1972, ISBN 0-216-89481-6

*Numerical Mathematics**
Elizabeth West, Chapter 2 (University of Paisley 1990)

Numerical Analysis
C.Dixon, (Blackie Chambers 1974)

Core Maths for A-level
Bostock & Chandler (Stanley Thornes) ISBN 0-7487-0067-6)

* Note that Elizabeth West's book has been rewritten for the unit Numerical Analysis 1 (AH) and has been issued to all centres by HSDU.

VECTORS

CONTENT
know the meaning of the terms position vector, unit vector, scalar triple product, vector product, components, direction vector, direction ratios/cosines calculate scalar and vector products in three dimensions know that $\boldsymbol{a} \times \boldsymbol{b} = -\boldsymbol{b} \times \boldsymbol{a}$ find $\boldsymbol{a} \times \boldsymbol{b}$ and $\boldsymbol{a}.\boldsymbol{b} \times \boldsymbol{c}$ in component form

Comments
This content was included in CSYS Paper 2.

Teaching notes
Vectors were introduced in Mathematics 3(H). It is assumed, in particular, students
- know the meaning of the terms: vector, magnitude (length), direction, scalar multiple, position vector, unit vector, directed line segment, component, scalar product;
- know and can apply the basis vectors $\boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k}$
- can use the scalar product to find the angle between two directed line segments.

Notation
Vectors can be written in component form both vertically and horizontally, e.g.

$\boldsymbol{a} = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}$ or $\boldsymbol{a} = (a_1, a_2, a_3)$. Both forms are used in the solutions offered to worked examples, although teachers/lecturers may prefer to use the vertical form exclusively, leaving the horizontal form to indicate co-ordinates of a point.

$\boldsymbol{a}.\boldsymbol{b} \times \boldsymbol{c}$ is the **scalar triple product** and is often written as $[\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}]$.

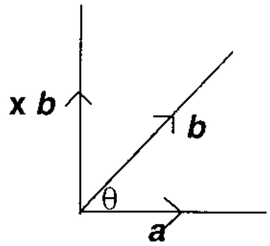
Order of presentation
The order of these notes follows the content order but teachers may wish students to study the work on lines and planes before proceeding to the vector product.

VECTOR PRODUCT OF TWO VECTORS

The **vector product** of two vectors $\boldsymbol{a}$ and $\boldsymbol{b}$ is a **vector** and is denoted by $\boldsymbol{a} \times \boldsymbol{b}$.

$\boldsymbol{a} \times \boldsymbol{b}$ has magnitude $|\boldsymbol{a}| \times |\boldsymbol{b}| \times \sin\theta$, where θ is the angle between $\boldsymbol{a}$ and $\boldsymbol{b}$.
has direction perpendicular to $\boldsymbol{a}$ and $\boldsymbol{b}$
has sense determined by the "right hand rule".

If the middle finger of the right hand is held perpendicular to the thumb and forefinger then if the thumb represents $\boldsymbol{a}$ *and the forefinger represents* $\boldsymbol{b}$ *then the middle finger represents* $\boldsymbol{a} \times \boldsymbol{b}$.



The basis vectors $\boldsymbol{i}, \boldsymbol{j}$ and $\boldsymbol{k}$ form a right handed system of mutually perpendicular unit vectors and an investigation of the vector products of these basis vectors reveals that:

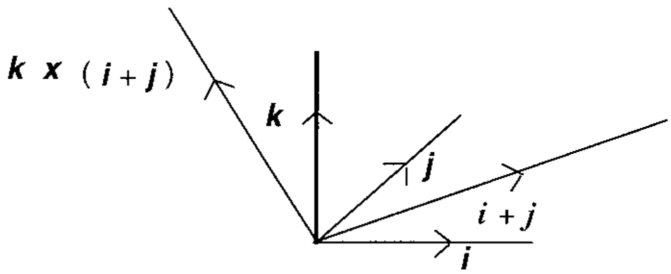
(i) $|\boldsymbol{i} \times \boldsymbol{i}| = |\boldsymbol{i}| \times |\boldsymbol{i}| \times \sin 0 = 1 \times 1 \times 0 = 0$ and so $\boldsymbol{i} \times \boldsymbol{i} = \boldsymbol{0}$
Similarly: $\boldsymbol{j} \times \boldsymbol{j} = \boldsymbol{0}$ and $\boldsymbol{k} \times \boldsymbol{k} = \boldsymbol{0}$

In general if $\boldsymbol{a}$ is any vector then $\boldsymbol{a} \times \boldsymbol{a} = \boldsymbol{0}$ (the zero vector)

(ii) $\boldsymbol{i} \times \boldsymbol{j} = \boldsymbol{k}$, $\quad \boldsymbol{j} \times \boldsymbol{k} = \boldsymbol{i}$ $\quad$ and $\boldsymbol{k} \times \boldsymbol{i} = \boldsymbol{j}$
$\boldsymbol{j} \times \boldsymbol{i} = -\boldsymbol{k}$, $\quad \boldsymbol{k} \times \boldsymbol{j} = -\boldsymbol{i}$ $\quad$ and $\boldsymbol{i} \times \boldsymbol{k} = -\boldsymbol{j}$

In general if $\boldsymbol{a}$ and $\boldsymbol{b}$ are any two vectors then $\boldsymbol{a} \times \boldsymbol{b} = -\boldsymbol{b} \times \boldsymbol{a}$

$\boldsymbol{k} \times (\boldsymbol{i} + \boldsymbol{j})$ is a vector - what is its relationship with $\boldsymbol{k} \times \boldsymbol{i}$ and $\boldsymbol{k} \times \boldsymbol{j}$?



Clearly, $\boldsymbol{i} + \boldsymbol{j}$ is in the plane defined by $\boldsymbol{i}$ and $\boldsymbol{j}$ and has magnitude $\sqrt{2}$ and makes an angle of $\frac{\pi}{4}$ with the direction of $\boldsymbol{i}$ as shown on the left.

$\boldsymbol{k} \times (\boldsymbol{i} + \boldsymbol{j})$ is $\perp$r to $\boldsymbol{k}$ so lies in the plane of $\boldsymbol{i}$ and $\boldsymbol{j}$

$\boldsymbol{k} \times (\boldsymbol{i} + \boldsymbol{j})$ is $\perp$r to $\boldsymbol{i} + \boldsymbol{j}$ so makes an angle of $\frac{3\pi}{4}$ with the +ve direction of $\boldsymbol{i}$

$|\boldsymbol{k} \times (\boldsymbol{i} + \boldsymbol{j})| = |\boldsymbol{k}| \times |\boldsymbol{i} + \boldsymbol{j}| \sin 90° = \sqrt{2}$.

Thus $\boldsymbol{k} \times (\boldsymbol{i} + \boldsymbol{j}) = \boldsymbol{j} - \boldsymbol{i}$

Now $\boldsymbol{k} \times \boldsymbol{i} = \boldsymbol{j}$ and $\boldsymbol{k} \times \boldsymbol{j} = -\boldsymbol{i}$ so $\boldsymbol{k} \times \boldsymbol{i} + \boldsymbol{k} \times \boldsymbol{j} = \boldsymbol{j} - \boldsymbol{i}$. Hence $\boldsymbol{k} \times (\boldsymbol{i} + \boldsymbol{j}) = \boldsymbol{k} \times \boldsymbol{i} + \boldsymbol{k} \times \boldsymbol{j}$

Can we generalise this statement for any three vectors $\boldsymbol{a}$, $\boldsymbol{b}$ and $\boldsymbol{c}$?

In fact we can: $\quad \boldsymbol{a} \times (\boldsymbol{b} + \boldsymbol{c}) = \boldsymbol{a} \times \boldsymbol{b} + \boldsymbol{a} \times \boldsymbol{c}$

Note: The above result is not stated in the content for this unit but needs to be assumed in order to express the vector product in component form. For completeness a justification of this result is included at the end of this section.

Vector product in component form

If $\boldsymbol{a} = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}$ and $\boldsymbol{b} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}$, then $\boldsymbol{a} \times \boldsymbol{b} = (a_1\boldsymbol{i} + a_2\boldsymbol{j} + a_3\boldsymbol{k}) \times (b_1\boldsymbol{i} + b_2\boldsymbol{j} + b_3\boldsymbol{k})$

$$\begin{aligned}\boldsymbol{a} \times \boldsymbol{b} &= a_1b_1\boldsymbol{i}\times\boldsymbol{i} + a_1b_2\boldsymbol{i}\times\boldsymbol{j} + a_1b_3\boldsymbol{i}\times\boldsymbol{k} \\ &\quad + a_2b_1\boldsymbol{j}\times\boldsymbol{i} + a_2b_2\boldsymbol{j}\times\boldsymbol{j} + a_2b_3\boldsymbol{j}\times\boldsymbol{k} \\ &\quad + a_3b_1\boldsymbol{k}\times\boldsymbol{i} + a_3b_2\boldsymbol{k}\times\boldsymbol{j} + a_3b_3\boldsymbol{k}\times\boldsymbol{k} \\ &= (a_2b_3 - a_3b_2)\boldsymbol{i} + (a_3b_1 - a_1b_3)\boldsymbol{j} + (a_1b_2 - a_2b_1)\boldsymbol{k}\end{aligned}$$

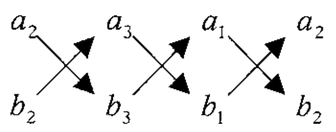
Hence $\boldsymbol{a} \times \boldsymbol{b} = \begin{pmatrix} a_2b_3 - a_3b_2 \\ a_3b_1 - a_1b_3 \\ a_1b_2 - a_2b_1 \end{pmatrix}$.

Determining the vector product

$$\boldsymbol{a} \times \boldsymbol{b} = (a_2b_3 - a_3b_2)\boldsymbol{i} + (a_3b_1 - a_1b_3)\boldsymbol{j} + (a_1b_2 - a_2b_1)\boldsymbol{k}$$

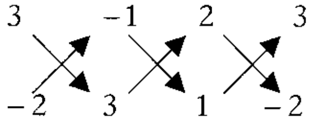
There are a number of techniques that are used to help in carrying out the process of calculating the components of $\boldsymbol{a} \times \boldsymbol{b}$. Two of these are mentioned here.

(i) Write down the components of the two vectors as shown – the first one on top and starting and finishing with the second component.



The components of $\boldsymbol{a} \times \boldsymbol{b}$ are the differences of the cross products.

For example find $\begin{pmatrix} 2 \\ 3 \\ -1 \end{pmatrix} \times \begin{pmatrix} 1 \\ -2 \\ 3 \end{pmatrix}$



Thus $\boldsymbol{a} \times \boldsymbol{b} = \begin{pmatrix} (9- & 2) \\ (-1-6) & \\ (-4- & 3) \end{pmatrix} = \begin{pmatrix} 7 \\ -7 \\ -7 \end{pmatrix}$

(ii) **using determinants**

$\boldsymbol{a} \times \boldsymbol{b} = (a_2b_3 - a_3b_2)\boldsymbol{i} + (a_3b_1 - a_1b_3)\boldsymbol{j} + (a_1b_2 - a_2b_1)\boldsymbol{k}$ can be written using 2 x 2 determinant notation giving:

$$\boldsymbol{a} \times \boldsymbol{b} = \begin{vmatrix} a_2 & a_3 \\ b_2 & b_3 \end{vmatrix} \boldsymbol{i} + \begin{vmatrix} a_3 & a_1 \\ b_3 & b_1 \end{vmatrix} \boldsymbol{j} + \begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix} \boldsymbol{k}$$

$$= \begin{vmatrix} a_2 & a_3 \\ b_2 & b_3 \end{vmatrix} \boldsymbol{i} - \begin{vmatrix} a_1 & a_3 \\ b_1 & b_3 \end{vmatrix} \boldsymbol{j} + \begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix} \boldsymbol{k}$$

Thus $\boldsymbol{a} \times \boldsymbol{b} = \begin{vmatrix} \boldsymbol{i} & \boldsymbol{j} & \boldsymbol{k} \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix}$

For example find $\begin{pmatrix} 2 \\ 3 \\ -1 \end{pmatrix} \times \begin{pmatrix} 1 \\ -2 \\ 3 \end{pmatrix}$

$\boldsymbol{a} \times \boldsymbol{b} = \begin{vmatrix} \boldsymbol{i} & \boldsymbol{j} & \boldsymbol{k} \\ 2 & 3 & -1 \\ 1 & -2 & 3 \end{vmatrix}$ = (9-2)$\boldsymbol{i}$ – (6-(-1))$\boldsymbol{j}$ +(-4 –3)$\boldsymbol{k}$ = 7$\boldsymbol{i}$ - 7$\boldsymbol{j}$ - 7$\boldsymbol{k}$

Note: Determinants appear later in the content of this unit. Teachers/lecturers may introduce them earlier if they favour the latter approach. If not then it should be included later as an application of the use of determinants

WORKED EXAMPLES

Example 1

P, Q and R are the points (0, 5, 1), (3, 2, 1) and (-2, 5, -3) respectively.

a) Find $\left|\overrightarrow{PQ}\right| \times \left|\overrightarrow{PR}\right|$

b) Hence calculate the area of triangle PQR.

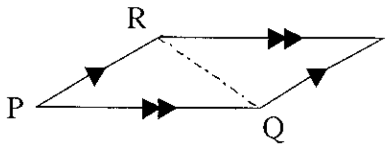
Solution

a) $\overrightarrow{PQ} = \begin{pmatrix} 3 \\ -3 \\ 0 \end{pmatrix}, \overrightarrow{PR} = \begin{pmatrix} -2 \\ 0 \\ -4 \end{pmatrix} \Rightarrow \overrightarrow{PQ} \times \overrightarrow{PR} = \begin{pmatrix} 3 \\ -3 \\ 0 \end{pmatrix} \times \begin{pmatrix} -2 \\ 0 \\ -4 \end{pmatrix} = \begin{pmatrix} 12 \\ 12 \\ -6 \end{pmatrix}$

b) Area of ΔPQR $= \frac{1}{2}$ area of parm.

$= \frac{1}{2}$ PQ.PR.sinP $= \frac{1}{2}\left|\overrightarrow{PQ} \times \overrightarrow{PR}\right|$

$= \frac{1}{2}\sqrt{144+144+36} = 9$sq units.



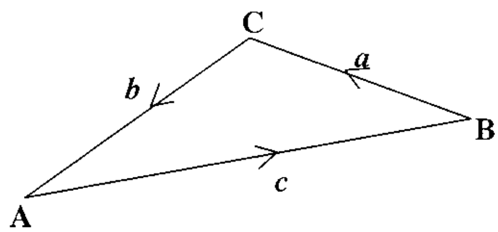
Example 2

The three sides of a triangle represent the vectors $\boldsymbol{a}$, $\boldsymbol{b}$ and $\boldsymbol{c}$ such that $\boldsymbol{a} + \boldsymbol{b} + \boldsymbol{c} = \boldsymbol{0}$

Prove

(i) $\boldsymbol{b} \times \boldsymbol{c} = \boldsymbol{c} \times \boldsymbol{a} = \boldsymbol{a} \times \boldsymbol{b}$

(ii) the Sine Rule for ΔABC



Solution

(i) $\boldsymbol{b} \times \boldsymbol{c} = -(\boldsymbol{a} + \boldsymbol{c}) \times \boldsymbol{c} = \boldsymbol{c} \times (\boldsymbol{a} + \boldsymbol{c}) = \boldsymbol{c} \times \boldsymbol{a} + \boldsymbol{c} \times \boldsymbol{c} = \boldsymbol{c} \times \boldsymbol{a}$

$\boldsymbol{c} \times \boldsymbol{a} = -(\boldsymbol{a} + \boldsymbol{b}) \times \boldsymbol{a} = \boldsymbol{a} \times (\boldsymbol{a} + \boldsymbol{b}) = \boldsymbol{a} \times \boldsymbol{a} + \boldsymbol{a} \times \boldsymbol{b} = \boldsymbol{a} \times \boldsymbol{b}$

(ii) Since $\boldsymbol{b} \times \boldsymbol{c} = \boldsymbol{c} \times \boldsymbol{a}$ $\quad |\boldsymbol{b}||c|\sin(180° - A) = |c||\boldsymbol{a}|\sin(180° - B)$

That is $bc\sin A = ca\sin B$ $\qquad$ Hence $b\sin A = a\sin B$ etc

SCALAR TRIPLE PRODUCT

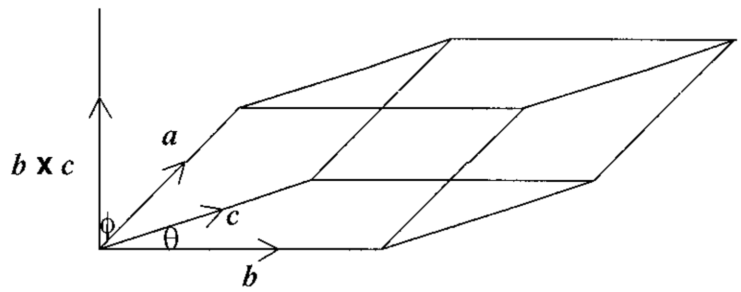
If $\boldsymbol{a}$, $\boldsymbol{b}$ and $\boldsymbol{c}$ are vectors then $\boldsymbol{a} \cdot (\boldsymbol{b} \times \boldsymbol{c})$ is a scalar and is an example of a **scalar triple product**

Now $\boldsymbol{b} \times \boldsymbol{c} = (b_2c_3 - b_3c_2)\boldsymbol{i} + (b_3c_1 - b_1c_3)\boldsymbol{j} + (b_1c_2 - b_2c_1)\boldsymbol{k}$
so $\boldsymbol{a} \cdot (\boldsymbol{b} \times \boldsymbol{c}) = a_1(b_2c_3 - b_3c_2) + a_2(b_3c_1 - b_1c_3) + a_3(b_1c_2 - b_2c_1)$

$$\text{Hence } \boldsymbol{a} \cdot (\boldsymbol{b} \times \boldsymbol{c}) = \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix}$$

Geometrical significance of scalar triple product

The diagram below shows three vectors $\boldsymbol{a}$, $\boldsymbol{b}$ and $\boldsymbol{c}$ and the vector product $\boldsymbol{b} \times \boldsymbol{c}$. The angle between $\boldsymbol{b}$ and $\boldsymbol{c}$ is θ and the angle between $\boldsymbol{b} \times \boldsymbol{c}$ and $\boldsymbol{a}$ is ϕ. The parallelepiped with $\boldsymbol{a}$, $\boldsymbol{b}$ and $\boldsymbol{c}$ as three edges has been drawn.



$\boldsymbol{b} \times \boldsymbol{c}$ has magnitude $|\boldsymbol{b}| \times |\boldsymbol{c}| \times \sin\theta$ and is perpendicular to the base of the parallelepiped.

Hence $\boldsymbol{a} \cdot (\boldsymbol{b} \times \boldsymbol{c})$ $= |\boldsymbol{a}| \times (|\boldsymbol{b}| \times |\boldsymbol{c}| \times \sin\theta)\cos\phi = |\boldsymbol{a}||\boldsymbol{b}||\boldsymbol{c}|\sin\theta \cos\phi$
$= (|\boldsymbol{b}||\boldsymbol{c}|\sin\theta)(|\boldsymbol{a}|\cos\phi)$ = area of base × height
= volume of parallelepiped

WORKED EXAMPLES

Example 1

If $\boldsymbol{a} = 2i + 3j - k$, $\quad \boldsymbol{b} = -i + j + 2k$, $\quad$ and $\quad \boldsymbol{c} = 3i - 2j + k$

Find (i) $\boldsymbol{a} \cdot (\boldsymbol{b} \times \boldsymbol{c})$ $\quad$ (ii) $\boldsymbol{b} \cdot (\boldsymbol{c} \times \boldsymbol{a})$ $\quad$ (iii) $\boldsymbol{c} \cdot (\boldsymbol{a} \times \boldsymbol{b})$

Solution

i) $\quad (\boldsymbol{b} \times \boldsymbol{c}) = \begin{pmatrix} -1 \\ 1 \\ 2 \end{pmatrix} \times \begin{pmatrix} 3 \\ -2 \\ 1 \end{pmatrix} = 5i + 7j - k \quad \therefore \boldsymbol{a} \cdot (\boldsymbol{b} \times \boldsymbol{c}) = 10 + 21 + 1 = 32$

ii) $(c \times a) = \begin{pmatrix} 3 \\ -2 \\ 1 \end{pmatrix} \times \begin{pmatrix} 2 \\ 3 \\ -1 \end{pmatrix} = -i + 5j + 13k \qquad \therefore \; b \,.\, (c \times a) = 1 + 5 + 26 = 32$

iii) $(a \times b) = \begin{pmatrix} 2 \\ 3 \\ -1 \end{pmatrix} \times \begin{pmatrix} -1 \\ 1 \\ 2 \end{pmatrix} = 7i - 3j + 5k \;\; \therefore \; c \,.\, (a \times b) = 21 + 6 + 5 = 32$

Example 2

In a tetrahedron OPQR, O is the origin and P, Q and R have position vectors relative to O given by $\boldsymbol{p} = 2i + 2k$, $\boldsymbol{q} = 2i + j$ and $\boldsymbol{r} = -i + 2j + k$ respectively.

i) Find $\boldsymbol{p} \times \boldsymbol{q}$ and hence prove that OR is perpendicular to the plane OPQ.

ii) Calculate the area of the triangle OPQ and the volume V, of the tetrahedron.

iii) Calculate $\boldsymbol{p}.(\boldsymbol{q} \times \boldsymbol{r})$ and verify that it is equal to 6V.

Solution

i) $\boldsymbol{p} = \begin{pmatrix} 2 \\ 0 \\ 2 \end{pmatrix}, \boldsymbol{q} = \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix}, \boldsymbol{r} = \begin{pmatrix} -1 \\ 2 \\ 1 \end{pmatrix} \Rightarrow \boldsymbol{p} \times \boldsymbol{q} = \begin{pmatrix} -2 \\ 4 \\ 2 \end{pmatrix} = 2\begin{pmatrix} -1 \\ 2 \\ 1 \end{pmatrix} = 2\boldsymbol{r}$

and so a normal vector for plane OPQ is parallel to $\boldsymbol{r}$, i.e. OR $\perp$r plane OPQ.

ii) Area of ΔOPQ $= \frac{1}{2}|\boldsymbol{p} \times \boldsymbol{q}| = \frac{1}{2}\sqrt{4+16+4} = \sqrt{6}$ units^2

$\text{V} = \frac{1}{3}\text{area of base} \times \text{height} = \frac{1}{3}\sqrt{6} \times |\boldsymbol{r}| = \frac{1}{3}\sqrt{6} \times \sqrt{6} = 2 \text{ units}^3$

iii) $\boldsymbol{q} \times \boldsymbol{r} = \begin{pmatrix} 1 \\ -2 \\ 5 \end{pmatrix} \qquad \boldsymbol{p}.\boldsymbol{q} \times \boldsymbol{r} = \begin{pmatrix} 2 \\ 0 \\ 2 \end{pmatrix} . \begin{pmatrix} 1 \\ -2 \\ 5 \end{pmatrix} = 2 + 0 + 10 = 12 = 6\text{V}$

Example 3

If two vectors $\boldsymbol{a}$ and $\boldsymbol{b}$ are perpendicular and equal in magnitude, show that vectors $\boldsymbol{u}$ and $\boldsymbol{v}$ defined by $\boldsymbol{u} = 2\boldsymbol{a} + \boldsymbol{b}$ and $\boldsymbol{v} = \boldsymbol{a} - 2\boldsymbol{b}$ are also perpendicular and equal in magnitude.

Show also that $\boldsymbol{u} \times \boldsymbol{v} = -5(\boldsymbol{a} \times \boldsymbol{b})$.

Solution

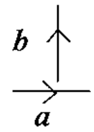
$a.b = 0, |a| = |b|$

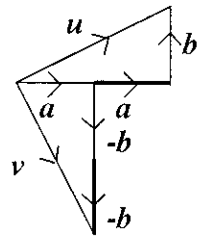
$u.v = (2a + b).(a - 2b)$

$= 2|a|^2 - 4a.b + b.a - 2|b|^2$

$= 2|a|^2 - 2|b|^2 = 0$

Hence u and v are perpendicular.





Now, $|u| = |v|$, (both $= \sqrt{5}|a|$) so $|u \times v| = |u||v|\sin 90° = 5|a|^2$.

But $|a \times b| = |a||b|\sin 90° = |a|^2$. Hence $|u \times v| = 5|a \times b|$

a, b, u and v, are coplaner. Using the right hand rule $u \times v$ is directed "into the page" and $a \times b$ is directed "out of the page". Hence, $u \times v = -5(a \times b)$.

Example 4

A plane passes through the points A, B and C with position vectors a, b and c respectively.
Show that $v = b \times c + c \times a + a \times b$ is normal to the plane unless $v = 0$.
If $v = 0$, how are the points A, B and C related?

Solution

$\overrightarrow{AB} = b - a$ and $\overrightarrow{BC} = c - b$.

The normal vector to the plane is parallel to $(b - a) \times (c - b)$

Now $(b - a) \times (c - b) = b \times (c - b) - a \times (c - b)$

$= b \times c - b \times b - a \times c + a \times b$

$= b \times c - 0 + c \times a + a \times b$ which is v

Hence v is normal to the plane.

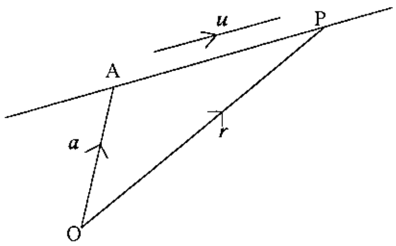
If $v = 0$, then $b - a$ and $c - b$ are parallel.

i.e. $\overrightarrow{AB}$ and $\overrightarrow{BC}$ are parallel, i.e. A, B and C are collinear.

CONTENT
know the equation of a line in vector form, parametric form and symmetric form
know the equation of a plane in vector form, parametric form, Cartesian form
find the equations of lines and planes given suitable defining information
find the angles between two lines, **two planes** [A/B] and between a line and a plane
find the intersection of two lines, a line and a plane, and two or three planes

Comments
This content was included in CSYS Paper 2.

THE EQUATION OF A LINE



The diagram shows a straight line, in the direction of the vector ***u***, passing through the point A, with position vector ***a***.

i) If P is any point on the line, with position vector ***r*** then:

$$\boldsymbol{r} = \boldsymbol{a} + \lambda \boldsymbol{u} \qquad (\text{ where } \lambda \text{ is a scalar})$$

This is the **vector equation** of the line.

ii) If A is the point (a_1, a_2, a_3) and $\boldsymbol{u} = \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix}$ and P is the point (x, y, z) then

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} + \lambda \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix}$$

Thus $x = a_1 + \lambda u_1, \quad y = a_2 + \lambda u_2, \quad z = a_3 + \lambda u_3$

These are the **parametric equations** of the line passing through (a_1, a_2, a_3) in the direction of $\begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix}$.

iii) If we eliminate λ from these equations we obtain

$$\frac{x-a_1}{u_1} = \frac{y-a_2}{u_2} = \frac{z-a_3}{u_3}$$

This is the equation of the line in **symmetric form** passing through (a_1, a_2, a_3) in the direction of $\begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix}$.

Example

Find the equation of the line passing through the points A(1, 2, 3) and B(2, 3, 5) in (a) vector form, (b) in parametric form and (c) in symmetric form.

Solution

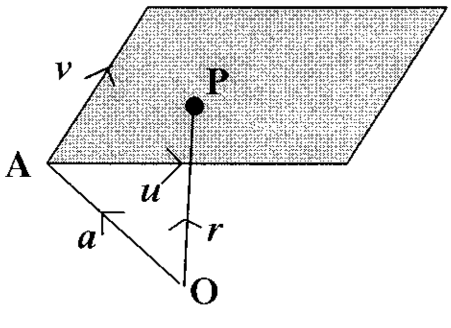
(a) $\boldsymbol{a} = \boldsymbol{i} + 2\boldsymbol{j} + 3\boldsymbol{k}$ and line is in the direction of $\overrightarrow{AB} = (1, 1, 2) = \boldsymbol{i} + \boldsymbol{j} + 2\boldsymbol{k}$

so the vector equation of the line is $\boldsymbol{r} = (\boldsymbol{i} + 2\boldsymbol{j} + 3\boldsymbol{k}) + \lambda(\boldsymbol{i} + \boldsymbol{j} + 2\boldsymbol{k})$

(b) the parametric equations are $x = 1 + \lambda$, $y = 2 + \lambda$, $z = 3 + 2\lambda$

(b) eliminating λ we obtain $\frac{x-1}{1} = \frac{y-2}{1} = \frac{z-3}{2}$

THE EQUATION OF A PLANE



i) Suppose that P (with position vector $\boldsymbol{r}$) lies in the plane which passes through the fixed point A (with position vector $\boldsymbol{a}$), and that $\boldsymbol{u}$ and $\boldsymbol{v}$ are two non-parallel vectors in the plane.

Then $\boxed{\boldsymbol{r} = \boldsymbol{a} + \lambda\boldsymbol{u} + \mu\boldsymbol{v}}$ (where λ and μ are scalars)

This is the **vector equation** of the plane.

ii) If $\boldsymbol{r} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$, $\boldsymbol{a} = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}$, $\boldsymbol{u} = \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix}$ and $\boldsymbol{v} = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix}$

then $\boxed{\begin{array}{l} x = a_1 + \lambda\, u_1 + \mu\, v_1 \\ y = a_2 + \lambda\, u_2 + \mu\, v_2 \\ z = a_3 + \lambda\, u_3 + \mu\, v_3 \end{array}}$ and these are **parametric equations** of the plane.

iii) Further, if vector $\boldsymbol{n} = \begin{pmatrix} a \\ b \\ c \end{pmatrix}$ is perpendicular to the plane, i.e normal to the plane

then $\boldsymbol{n} \,.\, \overrightarrow{AP} = 0$

$\boldsymbol{n} \,.\overrightarrow{AP} = 0 \Rightarrow \boldsymbol{n} \,.\, (\boldsymbol{r} - \boldsymbol{a}) = 0 \Rightarrow \boldsymbol{n} \,.\, \boldsymbol{r} = \boldsymbol{n} \,.\, \boldsymbol{a}$ *

$\boldsymbol{n} \,.\, \boldsymbol{r} = ax + by + cz$ and $\boldsymbol{n} \,.\, \boldsymbol{a} = aa_1 + ba_2 + ca_3$

so $ax + by + cz = aa_1 + ba_2 + ca_3$

Now, $aa_1 + ba_2 + ca_3$ is constant, for convenience let $aa_1 + ba_2 + ca_3 = -d$.

Hence $\boxed{ax + by + cz + d = 0}$

This is called the **Cartesian equation** of the plane.

It is important to note that the coefficients of x, y and z, in the Cartesian equation of the plane namely a, b and c are the components of a vector perpendicular to the plane.

A line at right angles to a plane is called a normal to the plane.

Hence the vector $\begin{pmatrix} a \\ b \\ c \end{pmatrix}$ is normal to the plane $ax + by + cz + d = 0$

* **Note** the equation $\boxed{\boldsymbol{n} \,.\, \boldsymbol{r} = \boldsymbol{n} \,.\, \boldsymbol{a}}$ is a vector equation and is sometimes referred to in textbooks as the vector equation of the plane (in dot product form).

Example

Find the equation of the plane passing through the points A(3, -2, 0), B(2, 0, 3) and C(1, -1, 1) in a) vector form, b) parametric form, and c) Cartesian form.

Solution

a) The position vector of A is $\begin{pmatrix} 3 \\ -2 \\ 0 \end{pmatrix}$ and $\overrightarrow{AB} = \begin{pmatrix} -1 \\ 2 \\ 3 \end{pmatrix}$ and $\overrightarrow{AC} = \begin{pmatrix} -2 \\ 1 \\ 1 \end{pmatrix}$

Thus the vector equation of the plane is $\boldsymbol{r} = \begin{pmatrix} 3 \\ -2 \\ 0 \end{pmatrix} + s\begin{pmatrix} -1 \\ 2 \\ 3 \end{pmatrix} + t\begin{pmatrix} -2 \\ 1 \\ 1 \end{pmatrix}$

b) From a) the parametric equations are

$$x = 3 - s - 2t$$
$$y = -2 + 2s + t$$
$$z = 3s + t$$

c) The vector $\overrightarrow{AB} \times \overrightarrow{AC} = \begin{pmatrix} -1 \\ 2 \\ 3 \end{pmatrix} \times \begin{pmatrix} -2 \\ 1 \\ 1 \end{pmatrix} = \begin{vmatrix} \boldsymbol{i} & \boldsymbol{j} & \boldsymbol{k} \\ -1 & 2 & 3 \\ -2 & 1 & 1 \end{vmatrix} = -\boldsymbol{i} - 5\boldsymbol{j} + 3\boldsymbol{k}$ is normal to the plane so the equation of the plane is $-x - 5y + 3z + d = 0$

But A(3, -2, 0) lies in the plane so $-3 + 10 + 0 + d = 0$ i.e $d = -7$

Hence equation of the plane is $x + 5y - 3z + 7 = 0$

THE ANGLE BETWEEN TWO PLANES [A/B]

The angle between two planes is defined to be the angle between their normals.

Example

Find the angle between the planes with equations

$$x + 3y - z - 5 = 0 \text{ and } 2x - y + z + 7 = 0$$

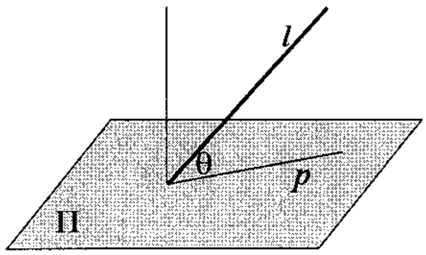
Solution

Normal vectors to the planes are $\boldsymbol{n}_1 = (1, 3, -1)$ and $\boldsymbol{n}_2 = (2, -1, 1)$ so angle between the planes is $\cos^{-1}\dfrac{\boldsymbol{n}_1.\boldsymbol{n}_2}{|\boldsymbol{n}_1| \times |\boldsymbol{n}_2|} = \cos^{-1}\dfrac{-2}{\sqrt{11} \times \sqrt{6}} = 104{\cdot}3°$.

Hence acute angle between the planes is 75.7°

THE ANGLE BETWEEN A LINE AND A PLANE

The angle between a line and a plane is defined to be the angle between the line and the projection of the line on the plane.



The diagram shows a line l and a plane Π.

The line p is the projection of l on Π.

The angle between the line l and the plane Π is θ.

In practice it is usually more convenient to find the angle between the line and a line perpendicular to the plane and then take the complement.

Example

Find the angle between the line with equation $\frac{x-2}{3} = \frac{y}{12} = \frac{z-5}{4}$ and the plane with equation $x - 2y - 2z - 5 = 0$

Solution

The line is in the direction of vector $\boldsymbol{v} = (3, 12, 4)$. A normal vector to the plane is $\boldsymbol{n} = (1, -2, -2)$.

Angle between line and normal to plane is $\cos^{-1} \frac{\boldsymbol{n}.\boldsymbol{v}}{|\boldsymbol{n}| \times |\boldsymbol{v}|} = \cos^{-1} \frac{-29}{3 \times 13}$.

Thus acute angle between the line and the normal to the plane is 42·0°.
Hence the acute angle between the line and the plane is 48·0°

INTERSECTIONS OF LINES AND PLANES

WORKED EXAMPLES

Example 1

Prove that the lines with equations

$$\frac{x-4}{3} = \frac{y+1}{-2} = \frac{z-2}{1} \quad (l_1) \qquad \text{and} \qquad \frac{x-5}{2} = \frac{y+1}{-1} = \frac{z-3}{1} \quad (l_2)$$

intersect, and find an equation for the plane containing them.

Solution

The lines will intersect if

(a) the direction vectors, say $\boldsymbol{u}$ and v, are such that $\boldsymbol{u} \neq k\boldsymbol{v}$ and

(b) they have a point in common.

The equations are given in symmetric form from which we can say that the direction vectors $\boldsymbol{u}$ and v are $\begin{pmatrix} 3 \\ -2 \\ 1 \end{pmatrix}$ and $\begin{pmatrix} 2 \\ -1 \\ 1 \end{pmatrix}$ respectively. Clearly $\boldsymbol{u} \neq k\boldsymbol{v}$ for any k.

To determine whether they have a point in common, we must write the equations in parametric form. Thus, for l_1: $\frac{x-4}{3} = \frac{y+1}{-2} = \frac{z-2}{1} = t$, say.

i.e. any point on l_1 is given by $(3t + 4, -2t - 1, t + 2)$

Similarly, any point on l_2 is given by $(2s + 5, -s - 1, s + 3)$, for some parameter s.

If the lines intersect, there will be a unique solution for the system of equations

$$3t + 4 = 2s + 5$$
$$-2t - 1 = -s - 1$$
$$t + 2 = s + 3$$

Solving the first two equations gives $t = -1$ and $s = -2$. Check this solution in the last equation. Hence $t = -1$ and $s = -2$ is a unique solution and the point of intersection is (1, 1, 1).

$$\boldsymbol{u} \times \boldsymbol{v} = \begin{pmatrix} 3 \\ -2 \\ 1 \end{pmatrix} \times \begin{pmatrix} 2 \\ -1 \\ 1 \end{pmatrix} = \begin{pmatrix} -1 \\ -1 \\ 1 \end{pmatrix}$$

$\begin{pmatrix} -1 \\ -1 \\ 1 \end{pmatrix}$ is a normal to the plane so, the equation of the plane is $-x - y + z = d$;

but (1, 1, 1) lies in the plane, so $-1 -1 + 1 = d$, i.e. $d = -1$

and the equation of the plane is $-x - y + z = -1$, which can be rewritten as $x + y - z = 1$.

Example 2

The plane Π_1 passes through A(3, 1, -1) and has a normal parallel to the vector (1, -2, 1). The plane Π_2 has equation $x - y = 1$.

Find:

(a) the size of the acute angle between Π_1 and Π_2,

(b) parametric equations for the line L of intersection of the planes.

Solution

(a) Angle between Π_1 and Π_2 = angle between the normals
(b)

i.e. angle between (1, -2, 1) and (1, -1, 0), say θ

$$\cos\theta = \frac{\begin{pmatrix}1\\-2\\1\end{pmatrix}.\begin{pmatrix}1\\-1\\0\end{pmatrix}}{\sqrt{6}\sqrt{2}} = \frac{1+2}{\sqrt{12}} = \frac{3}{2\sqrt{3}} = \frac{\sqrt{3}}{2} \quad \Rightarrow \theta = 30°$$

(c) L is perpendicular to both normals, i.e. parallel to their vector product

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

Equation of Π_1: $x-2y+z=d \Rightarrow 3-2-1=d$, since A lies on Π_1
i.e. $d = 0$ and equation is $x-2y+z=0$

Equation of Π_2 is $x-y=1$

Solving $x-2y+z=0$
$x-y=1$ Taking $z = 0$, we have $y = 1, x = 2$

Thus the point (2, 1, 0) lies in both planes and therefore on L, the line of intersection.

Equation of L : $\begin{pmatrix}x\\y\\z\end{pmatrix}-\begin{pmatrix}2\\1\\0\end{pmatrix}=t\begin{pmatrix}1\\1\\1\end{pmatrix}$,

i.e. $x = t + 2,\ y = t + 1,\ z = t$ are parametric equations of L.

Example 3

Show that the planes with equations

$$x-3y-2z=1$$
$$x-2y \qquad =5$$
$$2x-3y+2z=14$$

all intersect on a line and find parametric equations for this line.

Solution

$\Pi_1 : x - 3y - 2z = 1$, normal vector $\boldsymbol{a} = (1, -3, -2)$

$\Pi_2 : x - 2y = 5$, normal vector $\boldsymbol{b} = (1, -2, 0)$

$\Pi_3 : 2x - 3y + 2z = 14$, normal vector $\boldsymbol{c} = (2, -3, 2)$

For line of intersection of Π_1 and Π_2, direction vector is parallel to $\boldsymbol{a} \times \boldsymbol{b}$, i.e. (-4, -2, 1).

$(\boldsymbol{a} \times \boldsymbol{b}) \cdot \boldsymbol{c} = -4 \times 2 + (-2) \times (-3) + 1 \times 2 = 0$

so line of intersection of Π_1 and Π_2 must be parallel to Π_3.

To find a point on the line, put $x = 0$,

$$\left.\begin{array}{l} -3y - 2z = 1 \\ -2y = 5 \end{array}\right\} \Rightarrow y = -\frac{5}{2}, z = \frac{13}{4}, \text{ i.e. point } (0, -\frac{5}{2}, \frac{13}{4}) \text{ lies on line.}$$

Does this point lie in Π_3?

$2 \times 0 + (-3) \times (-\frac{5}{2}) + 2 \times \frac{13}{4} = 14$, i.e. point lies in Π_3.

Equation of line: $\begin{pmatrix} x \\ y \\ z \end{pmatrix} - \begin{pmatrix} 0 \\ -\frac{5}{2} \\ \frac{13}{4} \end{pmatrix} = t\begin{pmatrix} -4 \\ -2 \\ 1 \end{pmatrix} \Rightarrow x = -4t,\ y = -2t - \frac{5}{2},\ z = t + \frac{13}{4}$

Example 4

Find a vector equation for the line joining the points whose position vectors are $2\boldsymbol{i} - \boldsymbol{j} + \boldsymbol{k}$ and $-3\boldsymbol{i} + 2\boldsymbol{k}$, and hence find the point at which it meets the plane with equation $\boldsymbol{r} = p(\boldsymbol{i} + \boldsymbol{j}) + q(2\boldsymbol{i} - \boldsymbol{j} - \boldsymbol{k})$, where p and q are parameters.

Solution

Direction vector of line = $\begin{pmatrix} 2 \\ -1 \\ 1 \end{pmatrix} - \begin{pmatrix} -3 \\ 0 \\ 2 \end{pmatrix} = \begin{pmatrix} 5 \\ -1 \\ -1 \end{pmatrix}$

Equation is $\boldsymbol{r} - \begin{pmatrix} 2 \\ -1 \\ 1 \end{pmatrix} = t\begin{pmatrix} 5 \\ -1 \\ -1 \end{pmatrix}$ i.e $\boldsymbol{r} = (5t + 2)\boldsymbol{i} - (t + 1)\boldsymbol{j} + (1 - t)\boldsymbol{k}$

Where line meets plane $(5t + 2)\boldsymbol{i} - (t + 1)\boldsymbol{j} + (1 - t)\boldsymbol{k} = p(\boldsymbol{i} + \boldsymbol{j}) + q(2\boldsymbol{i} - \boldsymbol{j} - \boldsymbol{k})$

$= (p + 2q)\boldsymbol{i} + (p - q)\boldsymbol{j} - q\boldsymbol{k}$

It follows that $5t + 2 = p + 2q$
$t + 1 = -p + q \quad \Rightarrow \quad p = -2$
$1 - t = -q$

Also, $6t + 3 = 3q$ i.e. $2t + 1 = q$
$-t + 1 = -q \quad \Rightarrow t = -2$

and by substitution, $q = -3$

Hence $\boldsymbol{r} = -8\boldsymbol{i} + \boldsymbol{j} + 3\boldsymbol{k}$ i.e. the point is (-8, 1, 3)

TERMINOLOGY

The Direction of a line

The direction of one point from another is given by the ratios

x step: y step: z step.

For example if A is (1, 3, 2) and B is (3, 4, 1) then

x step: y step: z step = 2 : 1 : -1

and the direction of AB can be described in terms of the **direction ratios** 2 : 1 : -1

If $\overrightarrow{OP} = \begin{pmatrix} l \\ m \\ n \end{pmatrix}$ then its direction ratios are $l : m : n$

The direction of a line may also be determined by the angles the line makes with the positive direction of the axes.

Suppose $\overrightarrow{OP}$ above makes angles of θ, φ, and σ with the positive directions of $\boldsymbol{i}, \boldsymbol{j}$ and $\boldsymbol{k}$ respectively, then $\cos\theta = \dfrac{\boldsymbol{i}.|\overrightarrow{OP}|}{|\boldsymbol{i}|\times|\overrightarrow{OP}|} = \dfrac{l}{|\overrightarrow{OP}|}$

Similarly $\cos\phi = \dfrac{m}{|\overrightarrow{OP}|}$ and $\cos\sigma = \dfrac{n}{|\overrightarrow{OP}|}$

Hence if $r = |\overrightarrow{OP}| = \sqrt{l^2 + m^2 + n^2}$ the **direction cosines** of the line are $\dfrac{l}{r}$, $\dfrac{m}{r}$ and $\dfrac{n}{r}$

EXERCISE ON VECTORS

1. By using vector products, find the area of the triangle with vertices A(1, 3, -2), B(4, 3, 0) and C(2, 2, 1).

2. A line passes through a point A with position vector $\boldsymbol{a}$ and is parallel to a vector $\boldsymbol{b}$. If P is a general point on this line with position vector $\boldsymbol{r}$, state the vector equation of the line.
Show that the plane containing this line and parallel to another line with direction vector $\boldsymbol{c}$ is given by
$$\boldsymbol{r} \cdot (\boldsymbol{b} \times \boldsymbol{c}) = \boldsymbol{a} \cdot (\boldsymbol{b} \times \boldsymbol{c})$$

3. (a) Find parametric equations for the line l joining the points (-1, 1, -4) and (5, 4, -1).

   (b) Give the co-ordinates of the point of intersection of the line l and the plane Π, whose equation is $x + 2y - z = 11$, and find the angle at which l meets Π.

   (c) Find also an equation for the plane containing l which is perpendicular to Π.

4. Find parametric equations for the line l joining the points (3, 1, -2) and (-5, -1, 2).
Show that the plane which is parallel to l and contains the line $\dfrac{x-1}{1} = \dfrac{y}{-2} = \dfrac{z-4}{1}$ has equation $x + 2y + 3z = 13$.

5. Show that the lines with equations $\dfrac{x+2}{1} = \dfrac{y-5}{-1} = \dfrac{z+3}{2}$ and $\dfrac{x-3}{1} = \dfrac{y-8}{3} = \dfrac{z+1}{-2}$ intersect and find their point of intersection. Find an equation for the plane in which the lines lie.

6. (a) Find parametric equations for the line l of intersection of the planes
$$2x - 4y + z = 1 \quad \text{and} \quad x + y - z = 5$$

   (b) Given the co-ordinates of the point of intersection of the line l and the plane Π whose equation is $x + z = 1$, and find the size of the angle between l and Π.

7. Show that the three planes
$$3x - 5y + 2z + 1 = 0$$
$$2x + y - 3z + 5 = 0$$
$$5x - 7y + 2z + 3 = 0$$
meet in one line, L_1, and find parametric equations for L_1.

   Show that L_1 meets the line L_2 with equations $6x = 3y = 2z$, and find the co-ordinates of the point of intersection.

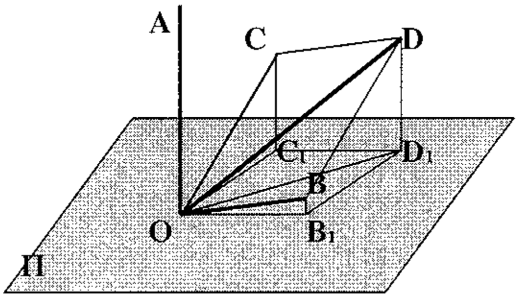
   Find also an equation for the plane containing L_1 and L_2.

ANSWERS

1. $\frac{1}{2}\sqrt{62}$
3. (a) $x = 2t - 1, y = t + 1, z = t - 4$ (b) (3, 3, -2); 30° (c) $x - y - z = 2$
4. $x = 4t + 3,\ y = t + 1,\ z = -2t - 2$
5. (1, 2, 3); $x - y - z + 4 = 0$
6. (a) $x = t,\ y = t - 2, z = 2t - 7$ (b) $(\frac{8}{3}, \frac{2}{3}, -\frac{5}{3})$; 60°
7. $x + 2 = y + 1 = z = t$; (1, 2, 3); $x - 2y + z = 0$

Justification of the result $\boldsymbol{a} \times (\boldsymbol{b} + \boldsymbol{c}) = \boldsymbol{a} \times \boldsymbol{b} + \boldsymbol{a} \times \boldsymbol{c}$

Let directed line segments $\overrightarrow{OA}, \overrightarrow{OB}, \overrightarrow{OC}, and\ \overrightarrow{OD}$ represent the vectors $\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}$ and $\mathbf{b} + \boldsymbol{c}$.



In the diagram on the left B_1, C_1 and D_1 are the projections of the points B, C and D onto the plane Π, the plane through O ⊥r to OA. Thus:

$OB_1 = OB\cos\angle BOB_1$,
$OD_1 = OD\cos\angle DOD_1$ and
$OC_1 = OC\cos\angle COC_1$

Since OBDC is a parallelogram then so is $OB_1D_1C_1$

Since the plane DOD_1 is perpendicular to Π it follows that $\overrightarrow{OA} \times \overrightarrow{OD}$ and $\overrightarrow{OA} \times \overrightarrow{OD_1}$ have the same direction and sense and lie in the plane Π .

$|\overrightarrow{OA} \times \overrightarrow{OD}| = OA.OD\sin\angle AOD = OA.OD\cos\angle\ DOD_1 = OA.OD_1 = |\overrightarrow{OA} \times \overrightarrow{OD_1}|$

$\therefore\ \overrightarrow{OA} \times \overrightarrow{OD} = \overrightarrow{OA} \times \overrightarrow{OD_1}$ ie $\boldsymbol{a} \times (\boldsymbol{b} + \boldsymbol{c}) = \overrightarrow{OA} \times \overrightarrow{OD_1}$

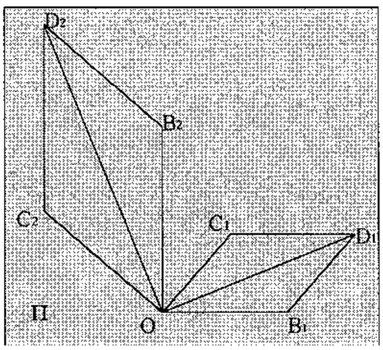
$\therefore\ |\boldsymbol{a} \times (\boldsymbol{b} + \boldsymbol{c})| = |\overrightarrow{OA}| \times |\overrightarrow{OD_1}|$ (since $OA \perp OD_1$) and $\boldsymbol{a} \times (\boldsymbol{b} + \boldsymbol{c})$ is ⊥r OD_1

Hence $\boldsymbol{a} \times (\boldsymbol{b} + \boldsymbol{c})$ can be represented by a directed line segment $\overrightarrow{OD_2}$ lying in the plane Π , perpendicular to OD_1 with length $OA.OD_1$

Similarly $\boldsymbol{a} \times \boldsymbol{b} = \overrightarrow{OA} \times \overrightarrow{OB_1}$ and hence $\boldsymbol{a} \times \boldsymbol{b}$ can be represented by a directed line segment $\overrightarrow{OB_2}$ in the plane Π , perpendicular to OB_1 with length $OA.OB_1$

Similarly $\boldsymbol{a} \times \boldsymbol{c} = \overrightarrow{OA} \times \overrightarrow{OC_1}$ and hence $\boldsymbol{a} \times \boldsymbol{c}$ can be represented by a directed line segment $\overrightarrow{OC_2}$ in the plane Π , perpendicular to OC_1 with length $OA.OC_1$

Drawing the directed line segments $\overrightarrow{OD_2}$, $\overrightarrow{OB_2}$, $\overrightarrow{OC_2}$ in the plane Π we obtain the following:



It follows that $OB_2C_2D_2$ is a parallelogram so $\overrightarrow{OD_2} = \overrightarrow{OB_2} + \overrightarrow{OC_2}$ and therefore

$$\boxed{\boldsymbol{a} \times (\boldsymbol{b} + \boldsymbol{c}) = \boldsymbol{a} \times \boldsymbol{b} + \boldsymbol{a} \times \boldsymbol{c}}$$

RESOURCES

Vectors & Coordinate Geometry

Understanding Pure Mathematics (Sadler & Thorning)
Section 2

Analytical Geometry and Vectors (Hunter)

The Complete A Level Mathematics Orlando Gough
Chapter 6
Section 6.1 – Vectors [**NB no ref to vector product or triple scalar product**]
Section 6.2 – Coordinate Geometry in two dimensions
Section 6.4 – Coordinate Geometry in three dimensions

MATRIX ALGEBRA

CONTENT
know the meaning of the terms matrix, element, row, column, order, identity matrix, inverse, determinant, singular, non-singular, transpose perform matrix operations: addition, subtraction, multiplication by a scalar, multiplication, establish equality of matrices know the properties: $A + B = B + A$ $AB \neq BA$ in general $(AB)C = A(BC)$ $A(B + C) = AB + AC$ $(A')' = A$ $(AB)' = B'A'$

Comments

This content was contained in CSYS Papers 1 and 2.

The list of properties is more extensive than the matrix content of CSYS Paper 1.

Teaching Notes

Matrices were introduced in Mathematics 1(AH) in the context of solving systems of equations and students should be familiar with the terms : matrix, element, row, column, order of a matrix, augmented matrix.

Operations on matrices were not covered in Mathematics 1(AH). Teachers may wish to consider some appropriate situations were matrices are used to record information in a rectangular array of numbers providing a context for introducing the addition and multiplication of matrices.

Most graphic calculators have a matrix facility and are especially useful when manipulating large matrices. Students could 'establish' some of the properties of matrices through the use of these calculators.

Matrix Notation
When we are considering a general matrix **A** it is common practice to denote the element in the **i th row** and the **j th** column by a_{ij} and write

$$A = [\, a_{ij} \,]$$

Some authors take this a stage further and write

$A = [\, a_{ij} \,]_{m\times n}$ to indicate the order of the matrix

The Arrangements do not specify any particular notation.

Using this notation we can write

If A and B are of the same order then:

$$\mathbf{A} + \mathbf{B} = [a_{ij} + b_{ij}]$$

If A and B are of orders $p \times q$ and $q \times r$, respectively then:

$$\mathbf{AB} = [\, c_{ij} \,] \text{ where } c_{ij} = \sum_{\lambda} a_{i\lambda} b_{\lambda j}$$

The transpose of A is A′ where

$$A' = [\, a'_{ij}] \text{ where } a'_{ij} = a_{ji}$$

MATRIX ALGEBRA

In the course of practising finding sums and products of matrices, if suitable examples are provided students should quickly conjecture that:

A + B = B + A but that AB ≠ BA in general
(AB)C = A(BC)
A(B + C) = AB + AC
(AB)′ = B′A′

They could also investigate the effect of multiplying appropriate matrices by matrices of the form $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ leading to the introduction of the identity matrix $I = [\, e_{ij} \,]$, where $e_{ij} = 1$ if $i = j$ and $e_{ij} = 0$ if $i \neq j$

Students do not need to be able to prove these results but establishing some of these results formally would illustrate direct proof and the usefulness of the notation.

Proof of: (AB)′ = B′A′

Suppose A has order $m \times n$ and is defined by $[\, a_{ij} \,]$ and B has order $n \times p$ and is defined by $[\, b_{ij} \,]$.

AB has order $m \times p$ and so the transpose matrix (AB)′ has order $p \times m$.

B′ has order $p \times n$ and A′ has order $n \times m$ so B′A′ also has order $p \times m$.

Since AB is defined by [c_{ij}] where $c_{ij} = \sum_{\lambda} a_{i\lambda} b_{\lambda j}$ then (AB)′ is defined by

[c'_{ij}] where $c'_{ij} = c_{ji} = \sum_{\mu} a_{j\mu} b_{\mu i}$,

but $b_{\mu i} = b'_{i\mu}$ and $a_{j\mu} = a'_{\mu j}$ so $\sum_{\mu} a_{j\mu} b_{\mu i} = \sum_{\mu} a'_{\mu j} b'_{i\mu} = \sum_{\mu} b'_{i\mu} a'_{\mu j}$

ie (AB)′ is defined by [c'_{ij}] where $c'_{ij} = \sum_{\mu} b'_{i\mu} a'_{\mu j}$

which is the matrix B′A′ of order $p \times m$,

i.e. (AB)′ = B′A′ .

CONTENT
know the properties $(AB)^{-1} = B^{-1}A^{-1}$ det (AB) = det A det B calculate the determinant of 2×2 and 3×3 matrices know the relationship of the determinant to invertability find the inverse of a 2×2 matrix

Comments

This content was contained in CSYS Papers 1 and 2.

Teaching Notes

In the algebra of numbers $ax = 1 \Rightarrow x = \frac{1}{a}$ and we write $x = a^{-1}$ i.e. $aa^{-1} = 1$.

In matrix algebra we investigate the possibility of finding for a matrix A, a matrix X such that AX = I

THE INVERSE OF A 2 × 2 MATRIX

Consider $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ then if AX = $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ it follows that X has order 2×2

Let X = $\begin{pmatrix} p & q \\ r & s \end{pmatrix}$ then AX = $\begin{pmatrix} ap+br & aq+bs \\ cp+dr & cq+ds \end{pmatrix}$

and AX = I $\quad \Rightarrow \begin{pmatrix} ap+br & aq+bs \\ cp+dr & cq+ds \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$

$\Rightarrow \quad ap + br = 1 \quad$ (i)

$$aq + bs = 0 \quad \text{(ii)}$$
$$cp + dr = 0 \quad \text{(iii)}$$
$$cq + ds = 1 \quad \text{(iv)}$$

$$d\text{(i)} - b\text{(iii)} \Rightarrow \quad (ad - bc)p = d$$
$$a\text{(iii)} - c\text{(i)} \Rightarrow \quad (ad - bc)r = -c$$
$$d\text{(ii)} - b\text{(iv)} \Rightarrow \quad (ad - bc)q = -b$$
$$a\text{(iv)} - c\text{(ii)} \Rightarrow \quad (ad - bc)s = a$$

Hence, provided $(ad - bc) \neq 0$ $\quad X = \dfrac{1}{ad - bc}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$

Thus A has an inverse, namely, $\dfrac{1}{ad - bc}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$ if $(ad - bc) \neq 0$

the value of the expression **$(ad - bc)$** determines whether or not A has an inverse and is called the **determinant of A** (cf the value of the expression $b^2 = 4ac$ in relation to the quadratic equation $ax^2 + bx + c = 0$)

The "determinant of A" is written as detA or $|A|$, and we denote $(ad - bc)$ by $\begin{vmatrix} a & b \\ c & d \end{vmatrix}$

The inverse of A is denoted by A^{-1} and $A^{-1} = \dfrac{1}{\det A}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$

Note

i) If $\det A = 0$, A has no inverse and is said to be **singular**.

ii) If $\det A \neq 0$, A has an inverse and is said to be **non-singular**.

iii) We have established that $AA^{-1} = I$.

Students should verify by multiplication that $A^{-1}A = I$

iv) Students should know that $\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \dfrac{1}{ad - bc}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$

$(AB)^{-1} = B^{-1}A^{-1}$

In the course of practising finding the inverses of suitable matrices through the use of appropriate examples students could be led to conjecture that $(AB)^{-1} = B^{-1}A^{-1}$

Proof of $(AB)^{-1} = B^{-1}A^{-1}$, where A and B are non-singular square matrices of the same order.

$$\begin{aligned}(B^{-1}A^{-1})(AB) &= B^{-1}(A^{-1}A)B \\ &= B^{-1}I\,B \\ &= B^{-1}B \\ &= I\end{aligned} \qquad \begin{aligned}(AB)(B^{-1}A^{-1}) &= A(B^{-1}B)A^{-1} \\ &= A\,I\,A^{-1} \\ &= AA^{-1} \\ &= I\end{aligned}$$

Hence $B^{-1}A^{-1}$ is the inverse of AB, i.e $(AB)^{-1} = B^{-1}A^{-1}$

THE DETERMINANT OF A 3 × 3 MATRIX

For the matrix $A = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix}$

$$\det A = \begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix} = a\begin{vmatrix} e & f \\ h & i \end{vmatrix} - b\begin{vmatrix} d & f \\ g & i \end{vmatrix} + c\begin{vmatrix} d & e \\ g & h \end{vmatrix} = a(ei - fh) - b(di - fg) + c(dh - eg)$$

Note: In the section of these notes dealing with the calculation of the components of a vector product two techniques were mentioned, one of which involved a 3 × 3 determinant. If the other method was used initially then the attention of students should now be drawn to the determinant method.

det (AB) = det A det B

Whilst practising calculating the determinants of 2 × 2 and 3 × 3 matrices students may be led to conjecture that det (AB) = det A det B.

Proof of det (AB) = det A det B

The proof offered here is for 2 × 2 matrices

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \text{ and } B = \begin{pmatrix} p & q \\ r & s \end{pmatrix} \Rightarrow AB = \begin{pmatrix} ap + br & aq + bs \\ cp + dr & cq + ds \end{pmatrix}$$

$$\begin{aligned}\det AB &= (ap + br)(cq + ds) - (cp + dr)(aq + bs) \\ &= apcq + brds + brcq + apds - cpaq - draq - cpbs - drbs \\ &= brcq + apds - draq - cpbs \\ &= ad(ps - qr) - bc(ps - qr) \\ &= (ad - bc)(ps - qr) \\ &= \det A \det B\end{aligned}$$

(The proof for 3 × 3 matrices is more difficult. The proof is in Hunter p.185.)

CONTENT
find the inverse, where it exists, of a 3×3 matrix by elementary row operations
know the role of the inverse matrix in solving linear systems

Comments
'know the role of the inverse matrix in solving linear systems' was in CSYS Paper 4

'find the inverse, where it exists, of a 3×3 matrix by elementary row operations' is new content.

Note: A^{-1} can also be found using the adjugate (adjoint) method. This method is illustrated at the end of this section.

Teaching Notes
Students will be familiar with elementary row operations through using them to solve systems of equations in Mathematics 1(AH)

FINDING THE INVERSE, WHERE IT EXISTS, OF A 3 X 3 MATRIX BY ELEMENTARY ROW OPERATIONS

The method is illustrated for the matrix

$$A = \begin{pmatrix} 2 & 1 & 4 \\ 1 & 0 & 2 \\ 2 & 3 & 1 \end{pmatrix}$$

We construct an augmented 3×6 matrix with the first three columns, the columns of A, and the other three columns, the columns of the 3×3 matrix I.

$$\left(\begin{array}{ccc|ccc} 2 & 1 & 4 & 1 & 0 & 0 \\ 1 & 0 & 2 & 0 & 1 & 0 \\ 2 & 3 & 1 & 0 & 0 & 1 \end{array}\right)$$

We then proceed to apply elementary row operations to the rows of the matrix with the purpose of obtaining a matrix whose first three columns are the columns of I.

$$\left(\begin{array}{ccc|ccc} 2 & 1 & 4 & 1 & 0 & 0 \\ 1 & 0 & 2 & 0 & 1 & 0 \\ 2 & 3 & 1 & 0 & 0 & 1 \end{array}\right) \rightarrow \left(\begin{array}{ccc|ccc} 1 & \frac{1}{2} & 2 & \frac{1}{2} & 0 & 0 \\ 1 & 0 & 2 & 0 & 1 & 0 \\ 2 & 3 & 1 & 0 & 0 & 1 \end{array}\right) \quad row_1 \rightarrow \frac{1}{2} row_1$$

$$\rightarrow \left(\begin{array}{ccc|ccc} 1 & \frac{1}{2} & 2 & \frac{1}{2} & 0 & 0 \\ 0 & -\frac{1}{2} & 0 & -\frac{1}{2} & 1 & 0 \\ 0 & 2 & -3 & -1 & 0 & 1 \end{array}\right) \quad \begin{array}{l} row_2 \rightarrow row_2 - row_1 \\ row_3 \rightarrow row_3 - 2row_1 \end{array}$$

$$\rightarrow \left(\begin{array}{ccc|ccc} 1 & \frac{1}{2} & 2 & \frac{1}{2} & 0 & 0 \\ 0 & 1 & 0 & 1 & -2 & 0 \\ 0 & 2 & -3 & -1 & 0 & 1 \end{array}\right) \quad row_2 \rightarrow -2row_2$$

$$\rightarrow \left(\begin{array}{ccc|ccc} 1 & \frac{1}{2} & 2 & \frac{1}{2} & 0 & 0 \\ 0 & 1 & 0 & 1 & -2 & 0 \\ 0 & 0 & -3 & -3 & 4 & 1 \end{array}\right) \quad row_3 \rightarrow row_3 - 2row_2$$

$$\rightarrow \left(\begin{array}{ccc|ccc} 1 & \frac{1}{2} & 2 & \frac{1}{2} & 0 & 0 \\ 0 & 1 & 0 & 1 & -2 & 0 \\ 0 & 0 & 1 & 1 & -\frac{4}{3} & -\frac{1}{3} \end{array}\right) \quad row_3 \rightarrow -\frac{1}{3}row_3$$

$$\rightarrow \left(\begin{array}{cccccc} 1 & \frac{1}{2} & 0 & -\frac{3}{2} & \frac{8}{3} & \frac{2}{3} \\ 0 & 1 & 0 & 1 & -2 & 0 \\ 0 & 0 & 1 & 1 & -\frac{4}{3} & -\frac{1}{3} \end{array}\right) \quad row_1 \rightarrow row_1 - 2row_3$$

$$\rightarrow \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & -2 & \frac{11}{3} & \frac{2}{3} \\ 0 & 1 & 0 & 1 & -2 & 0 \\ 0 & 0 & 1 & 1 & -\frac{4}{3} & -\frac{1}{3} \end{array}\right) \quad row_1 \rightarrow row_1 - \frac{1}{2}row_2$$

The first three columns of the resulting augmented matrix are the columns of I and the remaining columns produce $\begin{pmatrix} -2 & \frac{11}{3} & \frac{2}{3} \\ 1 & -2 & 0 \\ 1 & -\frac{4}{3} & -\frac{1}{3} \end{pmatrix}$

Now $\begin{pmatrix} -2 & \frac{11}{3} & \frac{2}{3} \\ 1 & -2 & 0 \\ 1 & -\frac{4}{3} & -\frac{1}{3} \end{pmatrix}\begin{pmatrix} 2 & 1 & 4 \\ 1 & 0 & 2 \\ 2 & 3 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$

and hence $A^{-1} = \begin{pmatrix} -2 & \frac{11}{3} & \frac{2}{3} \\ 1 & -2 & 0 \\ 1 & -\frac{4}{3} & -\frac{1}{3} \end{pmatrix}$

Note A justification of the above procedure for finding the inverse of a 3×3 matrix is included at the end of this section.

Knowing A^{-1} we could, for example, solve the system of equations

$$2x + y + 4z = 0$$
$$x \quad + 2z = -1$$
$$2x + 3y + z = 7$$

The system of equations can be represented by :

$$A\begin{pmatrix}x\\y\\z\end{pmatrix}=\begin{pmatrix}0\\-1\\7\end{pmatrix} \Rightarrow \quad A^{-1}A\begin{pmatrix}x\\y\\z\end{pmatrix}=A^{-1}\begin{pmatrix}0\\-1\\7\end{pmatrix} \Rightarrow \quad (A^{-1}A)\begin{pmatrix}x\\y\\z\end{pmatrix}=A^{-1}\begin{pmatrix}0\\-1\\7\end{pmatrix}$$

$$\Rightarrow \begin{pmatrix}x\\y\\z\end{pmatrix}=\begin{pmatrix}-2 & 1\tfrac{1}{3} & \tfrac{2}{3}\\1 & -2 & 0\\1 & -\tfrac{4}{3} & -\tfrac{1}{3}\end{pmatrix}\begin{pmatrix}0\\-1\\7\end{pmatrix}=\begin{pmatrix}1\\2\\-1\end{pmatrix} \qquad \Rightarrow x = 1, y = -2, z = -1$$

INTERSECTION OF THREE PLANES

There is an obvious link to the outcome on Vectors in this unit. Essentially, the equations of three planes can be represented as a system of three linear equations.

$a_1x + b_1y + c_1z = d_1$
$a_2x + b_2y + c_2z = d_2$ which can be written in matrix form:
$a_3x + b_3y + c_3z = d_3$

$$\begin{pmatrix}a_1 & b_1 & c_1\\a_2 & b_2 & c_2\\a_3 & b_3 & c_3\end{pmatrix}\begin{pmatrix}x\\y\\z\end{pmatrix}=\begin{pmatrix}d_1\\d_2\\d_3\end{pmatrix}$$

If $\begin{vmatrix}a_1 & b_1 & c_1\\a_2 & b_2 & c_2\\a_3 & b_3 & c_3\end{vmatrix} \neq 0$, a unique solution exists – the point of intersection of the 3 planes. However, if

$\begin{vmatrix}a_1 & b_1 & c_1\\a_2 & b_2 & c_2\\a_3 & b_3 & c_3\end{vmatrix} = 0$ then the equations do not have a unique solution.

Either (1) there are no solutions and the equations are inconsistent; or

(2) there are an infinite number of solutions, and

all the planes are the same,
or two of the planes are coincident
or the three planes have a common line.

Note Just as the determinant of a 3×3 matrix was expressed in terms of the determinants of three 2×2 matrices so the determinant of a 4×4 matrix can be expressed in terms of the determinants of four 3×3 matrices etc. It can be shown that, in general, a matrix A of order $n \times n$ has an inverse provided $\det A \neq 0$.

CONTENT
use 2×2 matrices to represent geometrical transformations in the (x, y) plane

Comments

This content was contained in CSYS Paper 2.

The transformations should include rotations, reflections, dilatations and the role of the transpose in orthogonal cases.

Teaching Notes

Under a given transformation, (x, y) is mapped to (x', y') where $x' = ax + by$, $y' = cx + dy$. This can be expressed in matrix form as

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x' \\ y' \end{pmatrix}$$

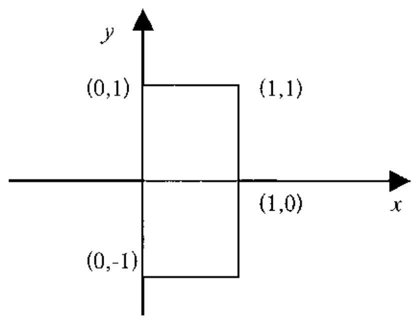
and the matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ is said to be associated with the transformation $(x, y) \to (x', y')$.

Note: $\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} a \\ c \end{pmatrix}$ and $\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} b \\ d \end{pmatrix}$

Thus if we know the images of (1, 0) and (0, 1) we can write down A. This is a useful technique to employ when finding the matrix of a matrix transformation.

Special cases

Reflection in the x-axis.



Consider the unit square.
Here (1, 0) $\to$ (1, 0)
and (0, 1) $\to$ (0, -1)

so $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ is the matrix associated with reflection in the x-axis.

check $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} a \\ -b \end{pmatrix}$ ie $(a, b) \to (a, -b)$, i.e. reflection in the x-axis.

Students should be given the task of finding the matrices associated with other simple transformations and compiling a table of results.

TRANSFORMATION	MATRIX
1. Identity	$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$
2. Reflection in x-axis	$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$
3. Reflection in y-axis	$\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$
4. Reflection in $y = x$	$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$
5. Reflection in $y = -x$	$\begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}$
6. Half turn about O	$\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$
7. Rotation about O of $\frac{\pi}{2}$ radians	$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$
8. Rotation about O of $-\frac{\pi}{2}$ radians	$\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$
9. Dilatation [O, k]	$\begin{pmatrix} k & 0 \\ 0 & k \end{pmatrix}$
10. Rotation of θ radians about O	$\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$

Investigating Further

Students could investigate the following

1. In transformations 1 to 6 and 9 the transpose of the matrix is the matrix itself
   i.e. $MM' = MM = M^2$
   In transformations 1 to 6, $MM' = I$ but this is not the case in transformation 9.
   What is the relationship between M^{-1} and M' in transformations 1 to 6?
   What is the inverse of the matrix in transformation 9?

2. The transpose of the matrix in transformation 7 is the matrix in transformation 8 and vice versa.
   What is MM' equal to in these two cases?
   What is the inverse of the matrix in transformation 7?
   What is the inverse of the matrix in transformation 8?

3. What is the transpose of the matrix in transformation 10?
   What is MM' equal to? What is M^{-1} equal to?

Link with complex numbers

In Mathematics 2(AH) students studied transformations in the complex plane defined in terms of operations on the complex numbers.

For example

a) in the transformation defined by $z \to \bar{z}$

$$x + iy \to x - iy$$

In co-ordinate terms $(x, y) \to (x, -y)$ ie reflection in the x – axis which we have noted is a matrix transformation with matrix $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$

b) in the transformation defined by $z \to iz$

$$x + iy \to -y + ix$$

In co-ordinate terms $(x, y) \to (-y, x)$ ie a rotation about O of $\frac{\pi}{2}$ radians which we have noted is a matrix transformation with matrix $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$.

WORKED EXAMPLES (miscellaneous)

Example 1

Matrix $A = \begin{pmatrix} p+3 & p \\ q & q+3 \end{pmatrix}$ and the matrix $X = (x \quad y)$.

If $3X = XA$, find x in terms of p, q and y.

If, further, $x + y = 1$, find x in terms of p and q.

Solution

$$3X = XA \Rightarrow (3x \quad 3y) = (x \quad y)\begin{pmatrix} p+3 & p \\ q & q+3 \end{pmatrix} = (px + 3x + qy \quad px + qy + 3y)$$

Equating corresponding entries

$$3x = px + 3x + qy \quad \Rightarrow \quad px = -qy \quad \Rightarrow \quad x = \frac{-qy}{p}$$

If $x + y = 1$, $y = 1 - x$, substitute in the above to get $x = \frac{-q(1-x)}{p}$

i.e. $px = -q + qx \quad \Rightarrow qx - px = q \quad \Rightarrow x(q - p) = q \quad \Rightarrow x = \frac{q}{q - p}$

Example 2

Evaluate $M\begin{pmatrix} 12 \\ 14 \end{pmatrix}$ where M is the inverse of the matrix $\begin{pmatrix} 3 & -2 \\ 1 & -4 \end{pmatrix}$.

Hence solve the system of equations

$$3x - 2y = 12$$
$$x - 4y = 14$$

Solution

$$\begin{vmatrix} 3 & -2 \\ 1 & -4 \end{vmatrix} = -10 \quad \Rightarrow \quad M = \frac{1}{-10}\begin{pmatrix} -4 & 2 \\ -1 & 3 \end{pmatrix}$$

So, $M\begin{pmatrix}12\\14\end{pmatrix} = \frac{1}{-10}\begin{pmatrix}-4 & 2\\-1 & 3\end{pmatrix}\begin{pmatrix}12\\14\end{pmatrix} = \frac{1}{-10}\begin{pmatrix}-20\\30\end{pmatrix} = \begin{pmatrix}2\\-3\end{pmatrix}$

System of equations can be represented as $\begin{pmatrix}3 & -2\\1 & -4\end{pmatrix}\begin{pmatrix}x\\y\end{pmatrix} = \begin{pmatrix}12\\14\end{pmatrix}$

with solution $\begin{pmatrix}3 & -2\\1 & -4\end{pmatrix}^{-1}\begin{pmatrix}3 & -2\\1 & -4\end{pmatrix}\begin{pmatrix}x\\y\end{pmatrix} = \begin{pmatrix}3 & -2\\1 & -4\end{pmatrix}^{-1}\begin{pmatrix}12\\14\end{pmatrix}$

i.e. $\begin{pmatrix}x\\y\end{pmatrix} = M\begin{pmatrix}12\\14\end{pmatrix}$ $\quad\Rightarrow x = 2$ and $y = -3$.

Example 3

P(2, -4) $\to$ P′(2, -6) under the transformation T with associated matrix $M = \begin{pmatrix}1 & a\\b & a\end{pmatrix}$

(i) Find the values of a and b.

(ii) Show that under this transformation all points of the plane are mapped on to a line, stating the equation of this line.

Solution

(i) $\begin{pmatrix}1 & a\\b & a\end{pmatrix}\begin{pmatrix}2\\-4\end{pmatrix} = \begin{pmatrix}2\\-6\end{pmatrix} \Rightarrow \begin{pmatrix}2-4a\\2b-4a\end{pmatrix} = \begin{pmatrix}2\\-6\end{pmatrix}$ $\quad\Rightarrow 2 - 4a = 2$ i.e. $a = 0$; $\Rightarrow 2b - 4a = -6$ i.e. $b = -3$

so $M = \begin{pmatrix}1 & 0\\-3 & 0\end{pmatrix}$

(ii) $\begin{pmatrix}1 & 0\\-3 & 0\end{pmatrix}\begin{pmatrix}x\\y\end{pmatrix} = \begin{pmatrix}x\\-3x\end{pmatrix}$ $\quad$ i.e. $(x, y) \to (x, -3x)$

and so all points of the plane are mapped on to the line with equation $y = -3x$.

Example 4

Matrix A represents the reflection of every point P(x, y) in the y-axis and matrix B represents the rotation of the plane through 90° anti-clockwise about the origin.

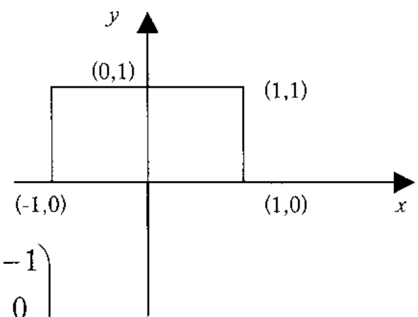
(i) Write down the 2×2 matrices A and B, and find the matrix products AB and BA.

(ii) Indicate on a diagram the positions Q and R of the images of P under the transformations given by AB and BA respectively.

(iii) Find matrix X such that $XAB = BA$ and state the transformation represented by X.

Solution

(i) $\begin{pmatrix}1\\0\end{pmatrix} \to \begin{pmatrix}-1\\0\end{pmatrix}, \begin{pmatrix}0\\1\end{pmatrix} \to \begin{pmatrix}0\\1\end{pmatrix}$ so $A = \begin{pmatrix}-1 & 0\\0 & 1\end{pmatrix}$



Similarly using the unit square for B

$\begin{pmatrix}1\\0\end{pmatrix} \to \begin{pmatrix}0\\1\end{pmatrix}$ and $\begin{pmatrix}0\\1\end{pmatrix} \to \begin{pmatrix}-1\\0\end{pmatrix}$ so $B = \begin{pmatrix}0 & -1\\1 & 0\end{pmatrix}$

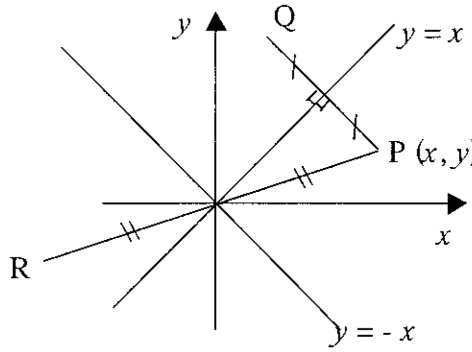
$$AB = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \text{and } BA = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}$$

(ii) AB: $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} y \\ x \end{pmatrix}$ i.e. $(x, y) \to (y, x)$

So that $P \to Q$ represents reflection in $y = x$

BA: $\begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} -y \\ -x \end{pmatrix}$ i.e. $(x, y) \to (-y, -x)$

So that $P \to R$ represents half turn about O



(iii) $XAB = BA \quad \Rightarrow \quad X\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}$

since $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}^{-1} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$,

the equation can be written $X = \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$

i.e. $X = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$ which represents reflection in $y = -x$.

EXERCISE ON MATRICES

1. Find the inverse of the matrix $\begin{pmatrix} 9 & 4 \\ 4 & 2 \end{pmatrix}$.

2. Evaluate the following determinants:

   (a) $\begin{vmatrix} 1 & 2 & 3 \\ 2 & 4 & 0 \\ -1 & 3 & 1 \end{vmatrix}$ (b) $\begin{vmatrix} 13 & 8 & 12 \\ 10 & 6 & 9 \\ 12 & 7 & 10 \end{vmatrix}$

3. Let $A = \begin{pmatrix} 1 & -1 & 0 \\ -1 & 0 & -1 \\ -1 & 1 & 0 \end{pmatrix}$.

   Write down the matrix $A - \lambda I$, where $\lambda \in \mathbf{R}$, and I is the 3 x 3 identity matrix. Find the values of λ for which the determinant of $A - \lambda I$ is zero.

4. Verify that $(AB)^{-1} = B^{-1}A^{-1}$ when $A = \begin{pmatrix} 6 & 2 \\ 10 & 4 \end{pmatrix}$, $B = \begin{pmatrix} 3 & 4 \\ 1 & 2 \end{pmatrix}$.

5. If $B = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$, prove that $B^2 = 4B - 3I$ where $I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ and then express B^3 in terms of B and I.

6. Using elementary row operations, find the inverse of the matrix $A = \begin{pmatrix} 1 & 3 & 3 \\ 4 & 13 & 12 \\ 2 & 7 & 5 \end{pmatrix}$

7. $A(\theta) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$ where θ is in radian measure. Prove that $[A(\theta)]^2 = A(2\theta)$ and then find the geometrical transformation corresponding to $\left[A\left(\frac{\pi}{2}\right)\right]^2$.

8. O is the origin and A, B and C are the points (1, 0), (1, 1) and (0, 1) respectively.
   (a) Calculate the co-ordinates of the images of O, A, B and C under the transformation T whose associated matrix is $\begin{pmatrix} 2 & 1 \\ -1 & 2 \end{pmatrix}$.
   (b) Prove that the image of $OABC$ under T is a square.
   (c) Write T as the composition of two of the following transformations:
   E : translation $\begin{pmatrix} u \\ v \end{pmatrix}$; F : reflection in the line $x = a$;
   G : clockwise rotation of $\theta°$ about O; H : dilatation [O, k]
   and find the unknowns in each one.

Answers

1. $\begin{pmatrix} 1 & -2 \\ -2 & 9/2 \end{pmatrix}$

2. 2. (a) 30 (b) 1

3. $\begin{pmatrix} 1-\lambda & -1 & 0 \\ -1 & -\lambda & -1 \\ -1 & 1 & -\lambda \end{pmatrix}$; $\det A = \lambda^2(1 - \lambda) \Rightarrow \lambda = 0, 1$

5. $B^3 = 13B - 12I$

6. $A^{-1} = \begin{pmatrix} 19 & -6 & 3 \\ -4 & 1 & 0 \\ -2 & 1 & -1 \end{pmatrix}$

7. Half turn about O

8. (a) (0, 0); (2, -1); (3, 1); (1, 2) (c) $G \circ H$ or $H \circ G$ where $\theta = 26.6°$, $k = \sqrt{5}$

1.

Justification of the procedure for finding the inverse of a 3 × 3 matrix

Consider the following products of matrices

a) $$\begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} = \begin{pmatrix} 2a & 2b & 2c \\ d & e & f \\ g & h & i \end{pmatrix}$$

b) $$\begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} 2a & 2b & 2c \\ d & e & f \\ g & h & i \end{pmatrix} = \begin{pmatrix} 2a+g & 2b+h & 2c+i \\ d & e & f \\ g & h & i \end{pmatrix}$$

c) $$\begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} 2a & 2b & 2c \\ d & e & f \\ g & h & i \end{pmatrix} = \begin{pmatrix} 2a+g & 2b+h & 2c+i \\ d & e & f \\ g & h & i \end{pmatrix}$$

These products illustrate that

(a) the row operation "double the first row" is equivalent to pre-multiplying a matrix by some matrix, R_a, say.

(b) the row operation "add the third row to the first row" is equivalent to pre-multiplying a matrix by some matrix, R_b, say.

(c) the result of "double the first row" followed by "add the third row to the first row" is equivalent to pre-multiplying a matrix by the matrix R_bR_a.

The above is true for all row operations and for any number of successive operations.

Hence,

i) carrying out a succession of row operations on the matrix **A** to obtain the matrix **I** is equivalent to pre-multiplying the matrix **A** by some matrix **X**.

ii) carrying out the same succession of row operations on the matrix **I** is equivalent to pre-multiplying the matrix **I** by the same matrix **X**.

In the above process the augmented matrix **(A | I)** was transformed into the augmented matrix **(I | B)** by a succession of row operations. In other words, the matrix **I** was transformed into a matrix **B** by the same succession of row operations that transformed matrix **A** into the matrix **I** .

In so doing we 'found' a matrix X such that

$$XA = I \qquad - (1)$$

and $$XI = B \qquad - (2)$$

Statement (1) indicates that $A^{-1} = X$

Statement (2) indicates that $X = B$

Finding A^{-1} using the adjugate (adjoint) method

$A^{-1} = \frac{1}{\det A} adjA$, where $adjA$ denotes the adjugate of A and equals $[A_{ij}]'$, the transpose of the matrix of cofactors.

$$[A_{ij}] = \begin{pmatrix} \begin{vmatrix} 0 & 2 \\ 3 & 1 \end{vmatrix} & -\begin{vmatrix} 1 & 2 \\ 2 & 1 \end{vmatrix} & \begin{vmatrix} 1 & 0 \\ 2 & 3 \end{vmatrix} \\ -\begin{vmatrix} 1 & 4 \\ 3 & 1 \end{vmatrix} & \begin{vmatrix} 2 & 4 \\ 2 & 1 \end{vmatrix} & -\begin{vmatrix} 2 & 1 \\ 2 & 3 \end{vmatrix} \\ \begin{vmatrix} 1 & 4 \\ 0 & 2 \end{vmatrix} & -\begin{vmatrix} 2 & 4 \\ 1 & 2 \end{vmatrix} & \begin{vmatrix} 2 & 1 \\ 1 & 0 \end{vmatrix} \end{pmatrix} = \begin{pmatrix} -6 & 3 & 3 \\ 11 & -6 & -4 \\ 2 & 0 & -1 \end{pmatrix}$$

$$\Rightarrow adjA = [A_{ij}]' = \begin{pmatrix} -6 & 11 & 2 \\ 3 & -6 & 0 \\ 3 & -4 & -1 \end{pmatrix} \text{ and } \det A = 2\begin{vmatrix} 0 & 2 \\ 3 & 1 \end{vmatrix} - \begin{vmatrix} 1 & 2 \\ 2 & 1 \end{vmatrix} + 4\begin{vmatrix} 1 & 0 \\ 2 & 3 \end{vmatrix} = 3$$

$$\Rightarrow A^{-1} = \frac{1}{\det A} adjA = \begin{pmatrix} -2 & 11/3 & 2/3 \\ 1 & -2 & 0 \\ 1 & -4/3 & -1/3 \end{pmatrix}$$

RESOURCES

Matrix Algebra

Understanding Pure Mathematics Sadler & Thorning
Section 6 (for basic concepts and transformations)
Section 17 (for determinants and inverses of 3 x 3 matrices)

Modern Mathematics for Schools Book 9 (2nd edition) (for transformations)

Algebra and Number Systems (Hunter)
Chapter 8

Mathematics in Action 6S
Chapter 10

FURTHER SEQUENCES AND SERIES

CONTENT
know the term power series understand and use the MacLaurin series : $f(x) = \sum_{r=0}^{\infty} \frac{x^r}{r!} f^{(r)}(0)$ find the MacLaurin series of simple functions: e^x, $\sin x$, $\cos x$, $\tan^{-1}x$, $(1+x)^{\alpha}$, $\ln(1+x)$, knowing their range of validity find the MacLaurin expansions for simple composites, such as e^{2x} use the MacLaurin series expansion to find power series for simple functions to a stated number of terms

Comments

This content was in CSYS Paper 2 (unrevised).

Teaching Notes

Graphics software provides a powerful illustration of the concept of convergence. The difficulties of differentiating or integrating power series term by term should be stressed, pointing out that these processes are only valid within the interval of convergence.

There is ample scope to link the work on power series with other parts of the course, e.g. binomial theorem, differentiation, integration and summation of infinite series.

Power series

In Mathematics 2(AH) we expanded $\frac{1}{1-r}$ as the geometric series

$1 + r + r^2 + r^3 + r^4 + \ldots\ldots$ provided $|r| < 1$

This was extended to expressing $\frac{1}{2x+1}$ *in the form*

$$1 - 2x + 4x^2 - 8x^3 + \ldots \quad |x| < \frac{1}{2}$$

$1 - 2x + 4x^2 - 8x^3 + \ldots$ is an example of a power series and the above is an example of expressing a function namely, $f(x) = \frac{1}{2x+1}$, as a power series provided, in this case, that $|x| < \frac{1}{2}$ which is known as the **range of validity** of the expansion.

In this section we are concerned with finding the power series of functions given that such expansions exist.

THE MACLAURIN SERIES

Suppose that $f(x)$ is a function which, throughout a certain interval including $x = 0$,

a) is differentiable any number of times
b) is the sum of a convergent power series

Let this series be $f(x) = a_0 + a_1x + a_2x^2 + a_3x^3 + a_4x^4 + ...$

Differentiating term by term successively we obtain

$$f(x) = a_0 + a_1x + a_2x^2 + a_3x^3 + a_4x^4 + ... \qquad \therefore\ f(0) = a_0$$
$$f'(x) = a_1 + 2a_2x^1 + 3a_3x^2 + 4a_4x^3 + ... \qquad \therefore\ f'(0) = a_1$$
$$f''(x) = 2a_2 + 6a_3x^1 + 12a_4x^2 + ... \qquad \therefore\ f''(0) = 2a_2$$
$$f'''(x) = 6a_3 + 24a_4x^1 + ... \qquad \therefore\ f'''(0) = 2.3a_3$$
$$f''''(x) = 24a_4 + ... \qquad \therefore\ f''''(0) = 2.3.4a_4$$

so that $a_0 = f(0)$, $a_1 = f'(0)$, $a_2 = \dfrac{f''(0)}{2!}$, $a_3 = \dfrac{f'''(0)}{3!}$, ... $a_r = \dfrac{f^{(r)}(0)}{r!}$ ....

Hence $f(x) = f(0) + f'(0)x + \dfrac{f''(0)}{2!}x^2 + \dfrac{f'''(0)}{3!}x^3 + ... + \dfrac{f^{(r)}}{r!}x^r$ ....

This is known as MacLaurin's series (after the Scots mathematician Colin MacLaurin, 1698-1746)

This result can be applied to some simple functions

The MacLaurin series for e^x

Let $f(x) = e^x$, then

$f(0) = 0$

$f'(x) = e^x \therefore f'(0) = 1$

$f''(x) = e^x \therefore f''(0) = 1$

$f'''(x) = e^x \therefore f'''(0) = 1$

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + ... + \frac{x^r}{r!} + ...$$

Putting $x = 1$ we obtain $e = 1 + 1 + \dfrac{1}{2!} + \dfrac{1}{3!} + \dfrac{1}{4!} + ...$

which allows us to calculate e as accurately as we wish.

Note This expansion was obtained in Mathematics 2(AH) as the limit of $\left(1 + \dfrac{1}{n}\right)^n$ as $n \to \infty$ as an extension of the binomial expansion.

Power series as an approximation to a function
A useful illustration of how the power series approximates to the function can be obtained by comparing the graph of $f(x) = e^x$ successively with the graphs of :

$$x \to 1 + x$$

$$x \to 1 + x + \frac{x^2}{2}$$

$$x \to 1 + x + \frac{x^2}{2} + \frac{x^3}{6}$$

$$x \to 1 + x + \frac{x^2}{2} + \frac{x^3}{6} + \frac{x^4}{24}$$

This approach can also be used to determine the likely range of validity.

In the case of e^x this is $\forall\, x \in \mathbf{R}$

Similarly

(a) $\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \ldots\ldots\ldots$ which can be shown to be valid $\forall\, x \in \mathbf{R}$

(b) $\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \ldots\ldots\ldots$ which can be shown to be valid $\forall\, x \in \mathbf{R}$

(c) $\tan^{-1} x = x - \frac{x^3}{3} + \frac{x^5}{5} - \frac{x^7}{7} + \ldots$ which can be shown to be valid for $-1 \leq x \leq 1$

(d) $(1+x)^\alpha = 1 + \alpha x + \frac{\alpha(\alpha-1)}{2!}x^2 + \frac{\alpha(\alpha-1)(\alpha-2)}{3!}x^3 + \ldots\ldots$ which can be shown to be valid for $-1 < x < 1$

(e) $\ln(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \ldots\ldots.$ which can be shown to be valid for $-1 < x \leq 1$

Range of validity of a MacLaurin expansion
Students are expected to know the range of validity of these power series. As indicated above in the case of $f(x) = e^x$, graphic calculators provide a powerful means of illustrating the likely range of validity of the McLaurin expansion of a function $f(x)$ by comparing the graph of $f(x)$ successively with the graphs of:

$$x \to f(0) + f'(0)x$$

$$x \to f(0) + f'(0)x + \frac{f''(0)}{2!}x^2$$

$$x \to f(0) + f'(0)x + \frac{f''(0)}{2!}x^2 + \frac{f'''(0)}{3!}x^3$$

etc.

Proving that these series do converge for a particular range of values of x is outwith the scope of this Unit. A more detailed justification is provided at the end of this section.

MACLAURIN EXPANSIONS FOR SIMPLE COMPOSITES

Given the MacLaurin expansion of a function we can find the expansions for simple composite functions as follows.

(a) Given that $e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + ...$ find an expansion for e^{2x}

It follows that $\quad e^{2x} = 1 + 2x + \frac{(2x)^2}{2!} + \frac{(2x)^3}{3!} + \frac{(2x)^4}{4!} + ... \forall\, x \in \mathbf{R}$

ie $\quad e^{2x} = 1 + 2x + 2x^2 + \frac{4x^3}{3} + \frac{2x^4}{3} + ... \forall\, x \in \mathbf{R}$

(b) Given that $(1+x)^\alpha = 1 + \alpha x + \frac{\alpha(\alpha-1)}{2!}x^2 + \frac{\alpha(\alpha-)(\alpha-2)}{3!}x^3 + ...$ for $-1 < x < 1$ find an expansion for $(2-x)^{-1}$ stating the range of validity of the series.

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

$$\Rightarrow \; (1-x)^{-1} = 1 + x + x^2 + x^3 + ... \qquad -1 < x < 1$$

Now $(2-x)^{-1} = \frac{1}{2-x} = \frac{1}{2(1-\frac{1}{2}x)} = \frac{1}{2}(1-\frac{1}{2}x)^{-1}$

Now apply the above result to get

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

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

with range of validity: $-1 < \frac{1}{2}x < 1$ i.e. $-2 < x < 2$

Note

i.) Lest students get the impression that we can find a power series expansion for all functions attention could be drawn to the functions $\ln x$ and $|x|$. $\ln x$ is not defined for $x = 0$ and although $|0|$ exists $|x|$ is not differentiable at $x = 0$.

ii.) Power series provide a means of evaluating particular values of functions – cf calculators

iii.) Within this Unit we are concerned solely with MacLaurin expansions of functions defined on **R**, the set of real numbers. Interesting results arise when we extend this to complex functions. One such result, is at the end of this section on MacLaurin Series.

WORKED EXAMPLES

Example 1

Given the MacLaurin expansions of e^x and $\sin x$ find expansions, as far as the term in x^4, of (a) $e^x \sin x$ and (b) $e^{\sin x}$

Solution

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \ldots\ldots \qquad \forall x \in \mathbf{R}$$

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \ldots. \qquad \forall x \in \mathbf{R}$$

(a)
$$e^x \sin x = (1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \ldots\ldots)(x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \ldots.)$$

$$= (x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \ldots.) + x\,(x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \ldots.)$$

$$+ \frac{x^2}{2!}\,(x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \ldots.) + \frac{x^3}{3!}\,(x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \ldots.)$$

$$= x + x^2 - \frac{x^3}{2!} + \frac{x^3}{3!} - \frac{x^4}{3!} + \frac{x^4}{3!} + \text{terms in } x^5$$

$$= x + x^2 + \frac{x^3}{3} + \ldots.\text{terms in } x^5$$

(b)
$$e^{\sin x} = 1 + \sin x + \frac{(\sin x)^2}{2!} + \frac{(\sin x)^3}{3!} + \frac{(\sin x)^4}{4!} + \ldots.$$

$$= 1 + x - \frac{x^3}{3!} + \frac{1}{2}(x - \frac{x^3}{3!})^2 + \frac{1}{6}(x^3 + \ldots.) + \frac{1}{24}(x^4 + \ldots) \ldots..$$

$$= 1 + x - \frac{x^3}{6} + \frac{1}{2}(x^2 - \frac{2x^4}{6}) + \frac{1}{6}x^3 + \frac{1}{24}x^4$$

$$= 1 + x + \frac{1}{2}x^2 - \frac{1}{6}x^4 + \frac{1}{24}x^4 + \ldots.$$

$$= 1 + x + \frac{1}{2}x^2 - \frac{1}{8}x^4 + \ldots.$$

Example 2

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

find the first three terms of the series for $(1 + 3x)^{1/3}$ and $(1 - 3x)^{-1/3}$. Hence show that, if x is so small that x^3 and higher powers of x can be neglected, the function $\left(\frac{1+3x}{1-3x}\right)^{1/3}$ can be expressed as $1 + 2x + 2x^2$.

Solution

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

$$= 1 + x - x^2 + \ldots..$$

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

Note: both series valid for $-\frac{1}{3} < x < \frac{1}{3}$.

$$\left(\frac{1+3x}{1-3x}\right)^{1/3} = (1+3x)^{1/3}(1-3x)^{-1/3}$$
$$= (1 + x - x^2 + \ldots)(1 + x + 2x^2 + \ldots)$$
$$= 1 + x + 2x^2 + x + x^2 + 2x^3 - x^2 - x^3 + \ldots$$
$$= 1 + 2x + 2x^2 \qquad \text{(ignoring } x^3 \text{ and higher powers)}$$

Example 3

Use MacLaurin's Theorem to find a power series for $\tan x$ as far as the term in x^3.

Solution

$f(x) = \tan x \qquad f(0) = 0$

$f'(x) = \sec^2 x \qquad f'(0) = 1$

$f''(x) = 2\sec x.\sec x.\tan x = 2\sec^2 x.\tan x \qquad f''(0) = 0$

$f'''(x) = 4\sec x.\sec x.\tan x.\tan x + 2\sec^2 x.\sec^2 x$

$\quad = 4\sec^2 x.\tan^2 x + 2\sec^4 x \qquad f'''(0) = 2$

Thus, since

$$f(x) = f(0) + f'(0).x + \frac{f''(0)}{2!}x^2 + \frac{f'''(0)}{3!}x^3 + ..$$
$$\tan x = 0 + 1.x + 0 + \frac{2}{3!}x^3 + \ldots = x + \frac{1}{3}x^3 + \ldots.$$

Note: This example links with Mathematics 1 (AH) – ' find the derivatives of $\tan x$ and $\sec x$'.

Example 4

(a) What is the domain of validity of the series $(1 + x^2)^{-1} = 1 - x^2 + x^4 - x^6 + x^8$ - ..

(b) Using this series find the power series for $\tan^{-1}x$

(c) Using these results find the sum of each of the following infinite series:

(i) $1 - \frac{1}{4} + \frac{1}{4^2} - \frac{1}{4^3} + \frac{1}{4^4} - \ldots$

(ii) $1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \frac{1}{9} - \ldots$

(iii) $1 - \frac{1}{3\times3} + \frac{1}{3^2\times5} - \frac{1}{3^3\times7} + \frac{1}{3^4\times9} - \ldots$

Solution

(a) $-1 < x^2 < 1$

(b) Since $\int (1 + x^2)^{-1}dx = \tan^{-1}x$

$$\therefore \tan^{-1}x = \int 1 - x^2 + x^4 - x^6 + x^8 - .. \, dx$$

$$= x - \frac{x^3}{3} + \frac{x^5}{5} - \frac{x^7}{7} + \ldots$$

(c)

(i), put $x = ½$ in $(1 + x^2)^{-1}$

So $1 - \frac{1}{4} + \frac{1}{4^2} - \frac{1}{4^3} + \frac{1}{4^4} - \ldots = (1 + \frac{1}{4})^{-1} = \frac{4}{5}$

(ii), put $x = 1$ in $\tan^{-1}x$

So $1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \frac{1}{9} - \ldots = \tan^{-1} 1 = \frac{\pi}{4}$

(iii), $1 - \frac{1}{3\times3} + \frac{1}{3^2\times5} - \frac{1}{3^3\times7} + \frac{1}{3^4\times9} - \ldots$

$$= 3^{½}\left(\frac{1}{3^{½}} - \frac{1}{3}.\frac{1}{(3^{½})^3} + \frac{1}{5}.\frac{1}{(3^{½})^5} - \frac{1}{7}.\frac{1}{(3^{½})^7} + \ldots\right)$$

$$= \sqrt{3}\tan^{-1}\frac{1}{\sqrt{3}} = \sqrt{3}\frac{\pi}{6}$$

Note

1. This question links with $\int \frac{1}{1+x^2}dx$ in Mathematics 2 (AH).
2. This question requires the application of knowledge in a more complex context

EXERCISE ON FURTHER SEQUENCES AND SERIES

1. Find the first four non-zero terms in the power series for $e^x\cos x$.

2. (a) Write down the binomial series for $(1 + x)^{-1/2}$ as far as the term in x^3.

   (b) Deduce the power series for $\dfrac{1}{\sqrt{(1-x^2)}}$ and by integrating find the first three non-zero terms of the power series for $\sin^{-1} x$.

   (c) Write down the series expansion for $\ln(1 + x)$ and hence that for $\ln(1 + \sin^{-1}x)$ as far as the term x^4.

3. Use the binomial theorem to evaluate $\sqrt{101}$ to five significant figures.

4. State the power series for $\ln(1 + x)$ and deduce that :

$$\ln\left(\frac{1+x}{1-x}\right) = 2\left(x + \frac{x^3}{3} + \frac{x^5}{5} + \frac{x^7}{7} + \ldots\right)$$

   By choosing a suitable value for x, evaluate ln 3 to 2 significant figures.

   Why could you not use the original series for $\ln(1 + x)$ to evaluate ln 3 ?

5. Using MacLaurin's Theorem to find the relevant power series, evaluate

$$\lim_{x\to 0}\frac{\sqrt{(1-x)} - \sqrt{(1-x^2)}}{x}$$

Answers

1. $1 + x - \frac{1}{3}x^3 - \frac{1}{5}x^4$

2. (a) $1 - \frac{1}{2}x + \frac{3}{8}x^2 - \frac{5}{16}x^3 + ..$ (b) $1 + \frac{1}{2}x^2 + \frac{3}{8}x^4$; $x + \frac{x^3}{6} + \frac{3}{40}x^5 + ..$

   (c) $x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \ldots$; $x - \frac{1}{2}x^2 + \frac{1}{2}x^3 - \frac{5}{12}x^4 + \ldots$

3. 10.050

4. $x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4}$; put $x = ½$ to give 1.1 (2 s.f.)

5. $-\frac{1}{2}$

Range of Validity of a MacLaurin Series – Further Justification

If $u_1, u_2, u_3, u_4, \ldots\ldots$ is a sequence of positive terms then the sequence of partial sums $S_1 = u_1$, $S_2 = u_1 + u_2$, $S_3 = u_1 + u_2 + u_3$, $S_4 = u_1 + u_2 + u_3 + u_4$, ...converges if the limit of $\frac{u_{n+1}}{u_n}$ as $n \to \infty$ is k and $k < 1$. It diverges if $k \geq 1$

Clearly, if $u_1, u_2, u_3, u_4, \ldots\ldots$ is a sequence of alternatively positive and negative terms then the sequence of partial sums $S_1 = u_1$, $S_2 = u_1 + u_2$, $S_3 = u_1 + u_2 + u_3$, $S_4 = u_1 + u_2 + u_3 + u_4$, ...converges if the limit of $\left|\frac{u_{n+1}}{u_n}\right|$ as $n \to \infty$ is k where $k < 1$. It diverges if $k > 1$. Because the terms are alternatively positive and negative it may converge if $k = 1$.

For example,

range of validity of the expansion for e^x

Here $\frac{u_{n+1}}{u_n} = \frac{x^n}{n!}\frac{(n-1)!}{x^{n-1}}$ and so

limit of $\left|\frac{u_{n+1}}{u_n}\right|$ as $n \to \infty$ is $\left|\frac{x}{n}\right|$

Now $\frac{x}{n} \to 0$ as $n \to \infty$ $\therefore$ expansion is valid for all real x

range of validity of the expansion for $\sin x$

Here $\frac{u_{n+1}}{u_n} = \frac{x^{2n+1}}{(2n+1)!}\frac{(2n-1)!}{x^{2n-1}} = \frac{x^2}{(2n+1)2n}$. Limit of $\left|\frac{u_{n+1}}{u_n}\right|$ as $n \to \infty$ is 0

$\therefore$ expansion is valid for all real x

range of validity of the expansion for $\log(1 + x)$

Here $\frac{u_{n+1}}{u_n} = -\frac{x^{n+1}}{n}\frac{(n-1)}{x^n} = -x(1-\frac{1}{n})$. Limit of $\left|\frac{u_{n+1}}{u_n}\right|$ as $n \to \infty$ is $|\,x\,|$

$\therefore$ expansion is certainly valid for $-1 < x < 1$

However in this case the actual range of validity is $-1 < x \leq 1$

Extending work on MacLaurin Series to complex numbers

Although this is not an explicit part of the content it links the work in complex numbers in Mathematics 2(AH) to work in this unit on further differential equations.

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \ldots + \frac{x^r}{r!} + \ldots\ldots \quad \forall x \in \mathbf{R}$$

Suppose we can extend this to complex numbers i.e.

$$e^z = 1 + z + \frac{z^2}{2!} + \frac{z^3}{3!} + \frac{z^4}{4!} + \ldots + \frac{z^r}{r!} + \ldots\ldots$$

Let us consider a particular complex number $i\theta$.

Then $e^{i\theta} = 1 + i\theta + \frac{(i\theta)^2}{2!} + \frac{(i\theta)^3}{3!} + \frac{(i\theta)^4}{4!} + \ldots + \frac{(i\theta)^r}{r!} + \ldots\ldots$

$\Rightarrow \quad e^{i\theta} = 1 + i\theta + \frac{(i\theta)^2}{2!} + \frac{(i\theta)^3}{3!} + \frac{(i\theta)^4}{4!} + \ldots + \frac{(i\theta)^r}{r!} + \ldots\ldots$

$\Rightarrow \quad e^{i\theta} = 1 + i\theta - \frac{\theta^2}{2!} - i\frac{\theta^3}{3!} + \frac{\theta^4}{4!} + i\frac{\theta^5}{5!} - \frac{\theta^6}{6!} \ldots\ldots\ldots$

$\Rightarrow \quad e^{i\theta} = (1 - \frac{\theta^2}{2!} + \frac{\theta^4}{4!} - \frac{\theta^6}{6!} + \ldots) + i(\theta - \frac{\theta^3}{3!} + \frac{\theta^5}{5!} - \ldots)$

$\Rightarrow \quad e^{i\theta} = \cos\theta + i\sin\theta$

It follows that $e^{-i\theta} = \cos(-\theta) + i\sin(-\theta) = \cos\theta - i\sin\theta$

and so $\cos\theta = \frac{1}{2}(e^{i\theta} + e^{-i\theta})$ and $\sin\theta = \frac{1}{2}(e^{i\theta} - e^{-i\theta})$

Note If we replace θ by π in $e^{i\theta} = \cos\theta + i\sin\theta$ we obtain

$$e^{i\pi} = \cos\pi + i\sin\pi = -1$$

Hence, $e^{i\pi} + 1 = 0$ which unites in a single relationship the five most important numbers in mathematics, namely, 0, 1, e, i, and π.

CONTENT
use iterative schemes of the form $x_{n+1} = g(x_n)$, $n = 0, 1, 2, \ldots$ to solve equations where $x = g(x)$ is a rearrangement of the original equation use graphical techniques to locate approximate solution x_0 know the condition for convergence of the sequence $\{x_n\}$ given by $x_{n+1} = g(x_n)$, $n = 0, 1, 2, \ldots$ and the meaning of the terms first and second order of convergence

Comments

Students should be made aware that the equation $f(x) = 0$ can be rearranged in a variety of ways and that some of these may yield suitable iterative schemes while others may not.

e.g. $f(x) = xe^x - 1 = 0 \Rightarrow xe^x = 1 \Rightarrow x = e^{-x}$ to give the scheme

$x_{n+1} = e^{-x_n}$ with $x_0 = 0.5$

Cobweb and staircase diagrams help to demonstrate the test for convergence for finding a fixed point α ($\alpha = g(\alpha)$) in a neighbourhood of α, namely $|g'(x)| < 1$.

e.g. $x_{n+1} = \frac{1}{2}\left\{x_n + \frac{c}{x_n}\right\}$ for estimating $\sqrt{c}$ is a second order method.

ITERATION

Iteration is the successive repetition of a mathematical process using the result of one stage as the input for the next.

When a non-linear equation $f(x) = 0$ can be rearranged in the form $x = g(x)$ it **may** be possible to solve the equation by the iterative formula $x_{n+1} = g(x_n)$, $n = 0, 1, 2, \ldots$

If the sequence defined by $x_{n+1} = g(x_n)$, does converge to a fixed point then $|g'(x)| < 1$ in the neighbourhood of this fixed point.

Proof – Suppose a is a root of the equation $x = g(x)$ so that $g(a) = a$.

As $x \to a$, the value of the expression $\frac{g(x) - g(a)}{x - a} \to g'(a)$ (c.f. first principles differentiation).

If x_n is a value of x close to a, it follows that $g(x_n) - g(a) \approx (x_n - a)\, g'(a)$

Since $g(x_n) = x_{n+1}$ and $g(a) = a$, this gives $x_{n+1} - a \approx (x_n - a)\, g'(a)$

Now, if $g'(a) > 0$, x_n and x_{n+1} will both be greater than a or both less than a (and if $g'(a) < 0$, a will lie between x_n and x_{n+1}).

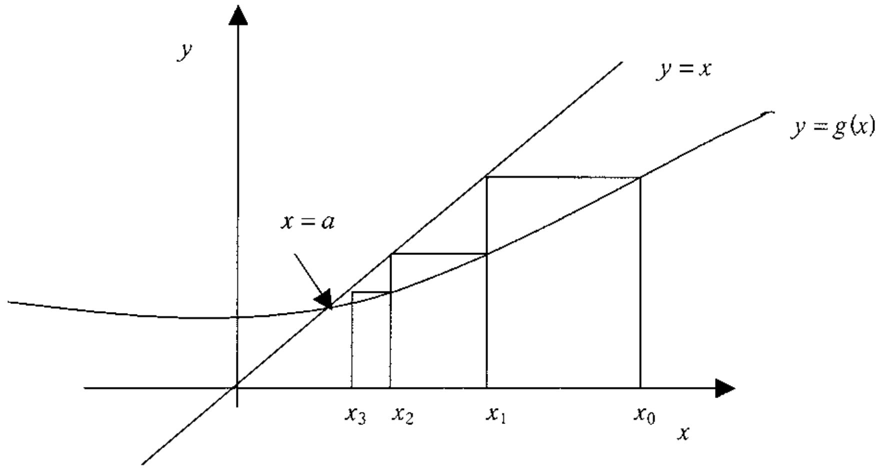
Also, since $|x_{n+1} - a| \approx |(x_n - a)|\, |g'(a)|$, $|x_{n+1} - a|$ will be less than $|(x_n - a)|$ if

$|g'(a)| < 1$, i.e. when $|g'(a)| < 1$, x_{n+1} will be closer to a than x_n and the sequence of approximations to the root will converge.

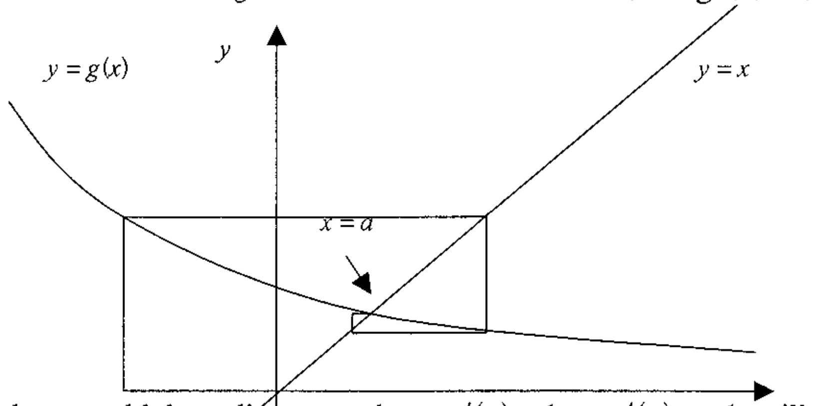
In practice, $|g'(a)|$ needs to be considerably less than 1 for convergence to be rapid.

This test for convergence can be demonstrated using staircase and cobweb diagrams.

(i) **Staircase** (convergence from one side of the root, $0 < g'(a) < 1$)



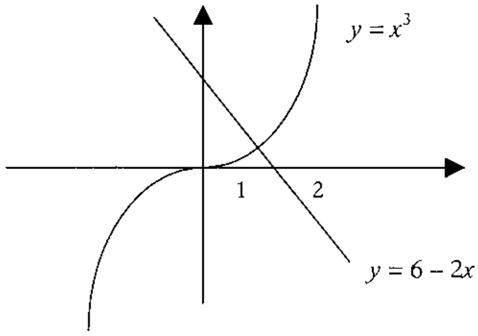
(ii) **Cobweb** (convergence from alternate sides of the root, $-1 < g'(a) < 0$)



Students could draw diagrams where $g'(a) > 1$ or $g'(a) < -1$ to illustrate divergence.

To illustrate the usefulness of the test for convergence, consider the equation:
$x^3 + 2x - 6 = 0$

A sketch of $y = x^3$ against $y = 6 - 2x$ reveals a positive root between 1 and 2.



The following rearrangements are possible:

(a) $x = (6 - 2x)^{1/3}$ (b) $x = \frac{1}{2}(x^3 - 6)$ (c) $x = \frac{6-2x}{x^2}$

Apply the test in the neighbourhood of the root, say $x = 1.5$

(a) $g(x) = (6 - 2x)^{1/3} \Rightarrow \quad g'(x) = -\frac{2}{3}(6-2x)^{-2/3}$

$g'(1.5) \approx -0.3 \Rightarrow |g'(1.5)| < 1$

(b) $g(x) = \frac{1}{2}(x^3 - 6) \Rightarrow \quad g'(x) = \frac{3}{2}x^2 \Rightarrow g'(1.5) \approx 3.4 \Rightarrow |g'(1.5)| > 1$

(c) $g(x) = \frac{6}{x^2} - \frac{2}{x} \Rightarrow g'(x) = -\frac{12}{x^3} + \frac{2}{x^2} \Rightarrow g'(1.5) \approx -2.7 \Rightarrow |g'(1.5)| > 1$

Only rearrangement (a) will give convergence to the root.

(a) Iterative formula is $x_{n+1} = (6 - 2x_n)^{1/3}$, $x_0 = 1.5$

$x_1 = 1.4422$	$x_6 = 1.4562$
$x_2 = 1.4605$	$x_7 = 1.4562$
$x_3 = 1.4548$	
$x_4 = 1.4566$	i.e. the root (or fixed point) is 1.456 (3 d.p.)
$x_5 = 1.4560$	

This example illustrates the various stages of using simple iteration viz.

1. Use a graphical method to locate an approximate solution (or initial value), x_0
2. Rearrange the original equation in the form $x = g(x)$
3. Apply the test for convergence.
4. Perform the iteration using $x_{n+1} = g(x_n)$

FIRST AND SECOND ORDER CONVERGENCE

For simple iteration, convergence occurs when $|g'(a)| < 1$ – see previous section – and the rate of convergence will increase as $|g'(a)|$ decreases.

It can be shown that successive errors in the sequence of approximations are connected by $e_{n+1} \approx e_n.g'(a)$ provided $g'(a) \neq 0$, (where $e_n = x_n - a$, the error in the nth iterate). In this case the process is said to be first order.

However, if $g'(a) = 0$ then $e_{n+1} \approx \frac{1}{2}e_n.g''(a)$ provided $g''(a) \neq 0$ and this process is called second order.

As a general rule, second order processes involve more complicated iterative formulae but converge more quickly.

The example given in the content section of the Arrangements document for this topic is $x_{n+1} = \frac{1}{2}\left\{x_n + \frac{c}{x_n}\right\}$ (which is the formula for finding $\sqrt{c}$ using Newton-Raphson method)

Here $g(x) = \frac{1}{2}(x + cx^{-1}) \quad \Rightarrow \quad g'(x) = \frac{1}{2}(1 - cx^{-2})$, therefore $g'(c) = \frac{1}{2}(1 - \frac{c}{c}) = 0$

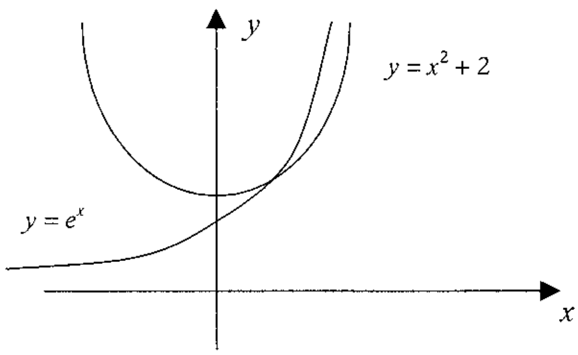
i.e. the process is second order.

WORKED EXAMPLES

Example 1
Solve the equation $e^x - x^2 - 2 = 0$, giving your answer to 2 d.p.

Solution

(i) Sketch to find approximate root.



Root lies between 1 and 2.
Take $x_0 = 1.5$

(ii) Rearrangement: $e^x = x^2 + 2 \Rightarrow x = \ln(x^2 + 2)$

(iii) Test: $g(x) = \ln(x^2 + 2)$

$$g'(x) = \frac{2x}{x^2+2} \qquad \Rightarrow \quad g'(1.5) \approx 0.7\text{, i.e. } |\, g'(1.5)\,| < 1$$

so convergence will take place (slowly because of high value of $|g'(x)|$)

(iv) Iteration: $x_{n+1} = \ln(x^2 + 2)$, $x_0 = 1.5$

$x_1 = 1.435$	$x_6 = 1.339$	$x_{11} = 1.323$
$x_2 = 1.401$	$x_7 = 1.333$	$x_{12} = 1.322$
$x_3 = 1.377$	$x_8 = 1.329$	$x_{13} = 1.321$
$x_4 = 1.360$	$x_9 = 1.326$	$x_{14} = 1.320$
$x_5 = 1.348$	$x_{10} = 1.324$	$x_{15} = 1.320$

i.e. root is 1.32 (2 d.p.)

Example 2

(a) Find the order of convergence of the following iterative processes:

(i) $x_{n+1} = \frac{1}{2} - x_n^3$, for a root of the equation $x^3 + x - \frac{1}{2} = 0$, taking $x_0 \in [0, ½]$

(ii) $x_{n+1} = \frac{1}{2}x_n(3 - 5x_n^2)$ for finding $\frac{1}{\sqrt{5}}$.

(b) In (ii), find $\frac{1}{\sqrt{5}}$ correct to 4 d.p.

Solution

(a) (i) $g(x) = \frac{1}{2} - x^3 \quad \Rightarrow \quad g'(x) = -3x^2$

Clearly $|g'(x)| < 1$ for a value of x in the interval $[0, ½]$ and so convergence takes place. Also, if λ represents the root $g'(\lambda) \neq 0$ since $0 < \lambda < ½$.
Hence the process is first order.

(ii) $g(x) = \frac{1}{2}x(3-5x^2) = \frac{3}{2}x - \frac{5}{2}x^3 \quad \Rightarrow \quad g'(x) = \frac{3}{2} - \frac{15}{2}x^2$

We know in this case that the root is $\frac{1}{\sqrt{5}}$. So $g'\left(\frac{1}{\sqrt{5}}\right) = 0$. Clearly, convergence takes place and since $g'\left(\frac{1}{\sqrt{5}}\right) = 0$, the process is second order.

(b) Take $x_0 = 0.5$ $\qquad x_1 = \frac{1}{2} \times 0.5(3 - 5 \times 0.5^2) = 0.4375$

$x_2 = 0.44690$ $\qquad x_3 = 0.44721$

$x_4 = 0.44721$ $\qquad$ and so the root is 0.4472 to 4 d.p.

The quick convergence is a feature of second order processes.

EXERCISE ON SIMPLE ITERATION

1. Prove that the equation $x^3 - 2x - 1 = 0$ has only one root in the interval $1 < x < 3$. Verify that the equation can be rewritten as $x = (2x+1)^{1/3}$.

   By using the iterative scheme $x_{n+1} = (2x_n + 1)^{1/3}$ with $x_0 = 2$, obtain an approximation to the root correct to two decimal places.

2. By using a suitable rearrangement of the equation $x^3 - 6x^2 + 11x - 8 = 0$, obtain an approximation to the root correct to three decimal places using an appropriate iterative scheme.

   You should draw a graph to help find a suitable starting value of the root and apply the test for convergence to ensure that your iterative process will converge.

3. Show that the iterative process $x_{n+1} = 1 - \sin x_n$ with $x_0 \in (0, \pi)$, for a root of the equation $\sin x + x - 1 = 0$, is first order.

4. By drawing an appropriate sketch, show that the equation $e^x - 4\sin x = 0$ has two roots between $x = 0$ and $x = 1.5$.

   Using a suitable rearrangement of the equation and a simple iterative method, find the smaller of these two roots correct to two decimal places.

Answers

1. 1.62
2. 3.521
4. 0.37

RESOURCES
Further Sequences and Series

MacLaurin Expansions

Calculus (Hunter)
Chapter 7, p.192 onwards gives a thorough treatment
(more than is required for this paper)

Core Maths for A-level (Bostock & Chandler)
Chapter 36

Iterative schemes

Numerical Mathematics (Elizabeth West)
Chapter 2

Numerical Analysis (C.Dixon)

Note that Elizabeth West's book has been rewritten and has forms the basis of the Support Materials for Numerical Analysis 1 (AH), issued to all centres by HSDU.

FURTHER ORDINARY DIFFERENTIAL EQUATIONS

CONTENT
solve first order linear differential equations using the integrating factor method find general solutions and solve initial value problems

Comments

This content was contained in CSYS Paper 2 (unrevised).

Teaching Notes

Linear differential equations are differential equations in which y and its derivatives occur only in linear combination even though their coefficients may be functions of x. The general form of a **first order linear differential equation** is

$$a(x)\frac{dy}{dx}+b(x)y=g(x)$$

First order linear differential equations occur in

a) motion subject to gravity and air-resistance, resulting in $\frac{dv}{dt}+cv=g$

b) a simple electrical circuit consisting of a resistance R and an inductance L, connected in series, with a constant electromotive force V which gives rise to the differential equation $L\frac{dI}{dt}+RI=V$

SOLVING FIRST ORDER LINEAR DIFFERENTIAL EQUATIONS

Consider the linear differential equation $x^3\frac{dy}{dx}+3x^2y=2x$

Now $\frac{d}{dx}\left(x^3y\right)=x^3\frac{dy}{dx}+3x^2y$ and so our equation can be written as $\frac{d}{dx}\left(x^3y\right)=2x$ which has solution $x^3y=x^2+c$

Thus the general solution of $x^3\frac{dy}{dx}+3x^2y=2$ is $y=\frac{1}{x}+\frac{c}{x^3}$ $\qquad (x\neq 0)$

The above is an example of an exact equation i.e. an equation whose left-hand side is an exact derivative.

If the original equation had been written in the equivalent form $x^2\frac{dy}{dx}+3xy=2$ it would not have been exact since the LHS is not an exact derivative. However multiplying throughout by x gives $x^3\frac{dy}{dx}+3x^2y=2x$ which as we have seen is exact.

The question now arises – can we always find an expression to multiply throughout by in order to make a first order linear differential equation exact?

The general linear equation $a(x)\frac{dy}{dx}+b(x)y=g(x)$ is usually first written in the form $\frac{dy}{dx}+P(x)y=f(x)$ which is known as the standard form. It is unlikely that this will be exact.

Suppose we multiply throughout by a function R of x to obtain

$$R\frac{dy}{dx}+RP(x)y=Rf(x).$$

We now try to choose R so that this equation is exact.

The existence of the term $R\frac{dy}{dx}$ on the LHS of the above suggests the LHS of the exact equation will be $\frac{d}{dx}(Ry)$

Thus we choose R so that $\frac{d}{dx}(Ry)=R\frac{dy}{dx}+RP(x)y$

$$\frac{d}{dx}(Ry)=R\frac{dy}{dx}+\frac{dR}{dx}y \quad \text{so} \quad R\frac{dy}{dx}+\frac{dR}{dx}y = R\frac{dy}{dx}+RP(x)y$$

Hence we choose R so that $\frac{dR}{dx}=P(x)R$. *

Now $\frac{dR}{dx}=P(x)R \Rightarrow \int\frac{1}{R}dR=\int P(x)dx \Rightarrow \ln R=\int P(x)dx \Rightarrow R=e^{\int P(x)dx}$

$R=e^{\int P(x)dx}$ and is called the **integrating factor**

* 'variables separable' differential equation encountered in Mathematics 2(AH).

In summary, to solve the equation $a(x)\frac{dy}{dx}+b(x)y=g(x)$

1. express it in standard form i.e. in the form $\frac{dy}{dx}+P(x)y=f(x)$
2. evaluate $e^{\int P(x)dx}$
3. multiply throughout by $e^{\int P(x)dx}$ to obtain
   $e^{\int P(x)dx}\frac{dy}{dx}+e^{\int P(x)dx}P(x)y=e^{\int P(x)dx}f(x)$
4. write this in the equivalent form $\frac{d}{dx}\left(e^{\int P(x)dx}y\right)=e^{\int P(x)dx}f(x)$
5. solve $e^{\int P(x)dx}y=\int e^{\int P(x)dx}f(x)\,dx$

WORKED EXAMPLES

Example 1

Solve the equation $x\frac{dy}{dx}=x+y$

Solution

Step 1: Expressing the equation in standard form we get $\frac{dy}{dx} - \frac{1}{x}y = 1$

Step 2: Integrating factor is $e^{\int \frac{-1}{x}dx} = e^{-\ln x} = \frac{1}{e^{\ln x}} = \frac{1}{x}$

Step 3: Multiplying throughout by "$\frac{1}{x}$" we obtain $\frac{1}{x}\frac{dy}{dx} - \frac{1}{x^2}y = \frac{1}{x}$

Step 4: $\frac{1}{x}\frac{dy}{dx} - \frac{1}{x^2}y = \frac{1}{x} \Leftrightarrow \frac{d}{dx}\left(\frac{y}{x}\right) = \frac{1}{x}$

Step 5: $\frac{d}{dx}\left(\frac{y}{x}\right) = \frac{1}{x} \quad \Leftrightarrow \quad \frac{y}{x} = \int \frac{1}{x}dx \quad \Leftrightarrow \quad \frac{y}{x} = \ln x + c$

Hence general solution is $y = x(\ln x + c)$

Example 2

Solve the differential equation $2x\frac{dy}{dx} - 3y = x^2 - x$, given that $y = 4$ when $x = 1$

Solution

Step 1: Expressing the equation in 'standard form' gives $\frac{dy}{dx} - \frac{3}{2x}y = \frac{x}{2} - \frac{1}{2}$

Step 2: Integrating factor = $e^{\int -\frac{3}{2x}dx} = e^{-\frac{3}{2}\ln x} = x^{-\frac{3}{2}}$

Step 3: Multiplying throughout by $x^{-\frac{3}{2}}$ gives $x^{-\frac{3}{2}}\frac{dy}{dx} - \frac{3}{2}x^{-\frac{5}{2}}y = x^{-\frac{3}{2}}(\frac{x}{2} - \frac{1}{2})$

Step 4: $x^{-\frac{3}{2}}\frac{dy}{dx} - \frac{3}{2}x^{-\frac{5}{2}}y = x^{-\frac{3}{2}}(\frac{x}{2} - \frac{1}{2})$

$\Leftrightarrow \quad \frac{d}{dx}(yx^{-3/2}) = x^{-3/2}\left(\frac{x}{2} - \frac{1}{2}\right) = \frac{1}{2}x^{-1/2} - \frac{1}{2}x^{-3/2}$

Step 5: $\frac{d}{dx}(yx^{-3/2}) = \frac{1}{2}x^{-1/2} - \frac{1}{2}x^{-3/2}$

$\Rightarrow \quad yx^{-3/2} = \frac{1}{2}\int (x^{-1/2} - x^{-3/2})dx = \frac{1}{2}(2x^{1/2} + 2x^{-1/2}) + C = x^{1/2} + x^{-1/2} + C$

$\Rightarrow \quad y = x^2 + x + Cx^{3/2}$ (general solution)

When $x = 1, y = 4 \Rightarrow 4 = 1 + 1 + C$ i.e. $C = 2$

Thus $y = x^2 + 2x^{3/2} + x$

Students should be encouraged to check solutions by substituting back into the original equation.

Example 3

Find the solution of the differential equation $\tan x \dfrac{dy}{dx} + 2y = \sin x \cos x$ for which $y = \dfrac{\pi}{16}$ when $x = \dfrac{\pi}{4}$.

Solution

Step 1: Expressing the equation in 'standard form' gives $\dfrac{dy}{dx} + \dfrac{2}{\tan x} y = \cos^2 x$

Step 2: Integrating factor = $e^{\int \frac{2\cos x}{\sin x} dx} = e^{\ln \sin^2 x} = \sin^2 x$

Step 3: Multiplying throughout by $\sin^2 x$ gives

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

Step 4:
$$\sin^2 x \frac{dy}{dx} + (2\sin x \cos x) y = \sin^2 x \cos^2 x$$

$$\Leftrightarrow \frac{d}{dx}(y \sin^2 x) = \sin^2 x \cos^2 x$$

Step 5:
$$\frac{d}{dx}(y \sin^2 x) = \sin^2 x \cos^2 x \Rightarrow y \sin^2 x = \int \sin^2 x \cos^2 x \, dx$$

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

$$\Rightarrow y = \frac{1}{8} x \operatorname{cosec}^2 x - \frac{1}{32} \sin 4x \operatorname{cosec}^2 x + C \operatorname{cosec}^2 x$$

If $y = \dfrac{\pi}{16}$ when $x = \dfrac{\pi}{4}$, $\dfrac{\pi}{16} = \dfrac{1}{8} . \dfrac{\pi}{4} .2 - 0 + 2C = \dfrac{\pi}{16} + 2C \Rightarrow C = 0$

Therefore, the particular solution is $y = \dfrac{1}{32} \operatorname{cosec}^2 x (4x - \sin 4x)$

Note The complexities of this question including the link with other units would place it at the A/B level.

Example 4

The differential equation $L\frac{dI}{dt}+RI=V$ occurs in electrical theory. It arises in a simple electrical circuit consisting of a fixed resistance R and a fixed inductance L, connected in series, with a constant electromotive force V.
Find I as a function of t, given that $I = 0$ when $t = 0$. Hence show that as t increases I approaches the value $\frac{V}{R}$.

Solution

Step 1: Expressing the equation in 'standard form' gives $\frac{dI}{dt}+\frac{R}{L}I=\frac{V}{L}$

Step 2: Integrating factor = $e^{\int\frac{R}{L}dt}=e^{\frac{Rt}{L}}$

Step 3: Multiplying throughout by $e^{\frac{Rt}{L}}$ gives $e^{\frac{Rt}{L}}\frac{dI}{dt}+e^{\frac{Rt}{L}}\frac{R}{L}I=e^{\frac{Rt}{L}}\frac{V}{L}$

Step 4: $e^{\frac{Rt}{L}}\frac{dI}{dt}+e^{\frac{Rt}{L}}\frac{R}{L}I=e^{\frac{Rt}{L}}\frac{V}{L}\Leftrightarrow\frac{d}{dt}\left(Ie^{\frac{Rt}{L}}\right)=\frac{V}{L}e^{\frac{Rt}{L}}$

Step 5: $\frac{d}{dt}\left(Ie^{\frac{Rt}{L}}\right)=\frac{V}{L}e^{\frac{Rt}{L}}\Rightarrow Ie^{\frac{Rt}{L}}=\int\frac{V}{L}e^{\frac{Rt}{L}}\Rightarrow Ie^{\frac{Rt}{L}}=\frac{V}{L}\frac{L}{R}e^{\frac{Rt}{L}}+c$

$\Rightarrow Ie^{\frac{Rt}{L}}=\frac{V}{R}e^{\frac{Rt}{L}}+c$

When $t = 0, I = 0$ so $0=\frac{V}{R}+c$ i.e. $c=-\frac{V}{R}$, hence $I=\frac{V}{R}(1-e^{-\frac{Rt}{L}})$

As t increases $e^{-\frac{Rt}{L}}\to 0$ and so $I\to\frac{V}{R}$

Example 5

In the previous example if a variable electromotive force $V\sin\omega t$ is applied then the corresponding differential equation is $L\frac{dI}{dt}+RI=V\sin\omega t$. If $I = 0$ when $t = 0$, find an expression for I and find what happens in this case as t increases.

[You may assume that $\int e^{kt}\sin\omega t\,dt=\frac{e^{kt}}{k^2+\omega^2}(k\sin\omega t-\omega\cos\omega t$]

Solution

Step 1: Expressing the equation in 'standard form' gives $\frac{dI}{dt}+\frac{R}{L}I=\frac{V}{L}\sin\omega t$

Step 2: Integrating factor $= e^{\int\frac{R}{L}dt}=e^{\frac{Rt}{L}}$

Step 3: Multiplying throughout by $e^{\frac{Rt}{L}}$ gives $e^{\frac{Rt}{L}}\frac{dI}{dt}+e^{\frac{Rt}{L}}\frac{R}{L}I=e^{\frac{Rt}{L}}\frac{V}{L}\sin\omega t$

Step 4: $e^{\frac{Rt}{L}}\frac{dI}{dt}+e^{\frac{Rt}{L}}\frac{R}{L}I=e^{\frac{Rt}{L}}\frac{V}{L}\sin\omega t \Leftrightarrow \frac{d}{dt}\left(Ie^{\frac{Rt}{L}}\right)=\frac{V}{L}e^{\frac{Rt}{L}}\sin\omega t$

Step 5: $\frac{d}{dt}\left(Ie^{\frac{Rt}{L}}\right)=\frac{V}{L}e^{\frac{Rt}{L}}\sin\omega t \Rightarrow Ie^{\frac{Rt}{L}}=\int\frac{V}{L}e^{\frac{Rt}{L}}\sin\omega t\,dt$

$$\Rightarrow Ie^{\frac{Rt}{L}}=\frac{V}{L}\frac{e^{\frac{R}{L}t}}{\frac{R^2}{L^2}+\omega^2}\left(\frac{R}{L}\sin\omega t-\omega\cos\omega t\right)+A$$

$$\Rightarrow Ie^{\frac{Rt}{L}}=\frac{Ve^{\frac{R}{L}t}}{R^2+L^2\omega^2}\left(R\sin\omega t-\omega L\cos\omega t\right)+A$$

$$\Rightarrow I=\frac{V}{R^2+L^2\omega^2}\left(R\sin\omega t-\omega L\cos\omega t\right)+Ae^{-\frac{R}{L}t}$$

when $t=0, I=0$ so $0=\frac{V}{R^2+L^2\omega^2}(-\omega L)+A$ hence $A=\frac{\omega LV}{R^2+L^2\omega^2}$

Thus $I=\frac{V}{R^2+L^2\omega^2}\left(R\sin\omega t-\omega L\cos\omega t+\omega Le^{-\frac{R}{L}t}\right)$

as $t\to\infty$ $\quad I\to\frac{V}{R^2+L^2\omega^2}\left(R\sin\omega t-\omega L\cos\omega t\right)$

Note **The example above has been included should teachers/lecturers wish to further emphasise the relevance of this topic and allow comparisons to be made with the previous example. It is clearly beyond level A/B with regard to assessment.**

CONTENT
know the meaning of the terms: second order linear differential equation with constant coefficients, homogeneous, non-homogeneous, auxiliary equation, complementary function and particular integral solve second order homogeneous ordinary differential equations with constant coefficients $a\frac{d^2y}{dx^2}+b\frac{dy}{dx}+cy=0$ find the general solution in the three cases where the roots of the auxiliary equation: (i) are real and distinct (ii) **coincide (are equal) A/B** (iii) **are complex conjugates A/B** **solve initial value problems**

Comments

Examples of contexts could include the motion of a spring, both with and without a damping term.
This content was contained in CSYS Paper 2 (unrevised).

Teaching Notes

A second order linear differential equation involves $\frac{d^2y}{dx^2}$, $\frac{dy}{dx}$ and y.

A simple example is $\frac{d^2y}{dx^2}=6x$

Now $\frac{d^2y}{dx^2}=6x \Rightarrow \frac{dy}{dx}=3x^2+c \Rightarrow y=x^3+cx+\text{d}$

This illustrates that the general solution of a second order differential equation will contain two arbitrary constants.

The differential equation $a\frac{d^2y}{dx^2}+b\frac{dy}{dx}+cy=f(x)$ is the general **second order differential equation with constant coefficients.**

Before proceeding to solve the general equation we attempt to solve the simpler equation

$$a\frac{d^2y}{dx^2}+b\frac{dy}{dx}+cy=0 \qquad \text{(A)}$$

which is the general **second order homogeneous ordinary differential equation with constant coefficients.**

It can easily be confirmed that if $y = u$ and $y = v$ are any two particular solutions of (A) then $y = \text{A}u + \text{B}v$, where A and B are arbitrary constants, is also a solution of (A).

Since A and B are arbitrary constants and provided that u and v are independent solutions (that is, one is not merely a multiple of the other), y = Au + Bv is the general solution of (A).

Thus, to solve equations of the form $a\frac{d^2y}{dx^2}+b\frac{dy}{dx}+cy=0$ we seek any two independent solutions.

Consider first the relatively simple example: Solve $\frac{d^2y}{dx^2}-y=0$

Now, $\frac{d^2y}{dx^2}-y=0 \Leftrightarrow \frac{d^2y}{dx^2}=y$

To solve $\frac{d^2y}{dx^2}=y$, we seek a function with second derivative is equal to itself.

This is a property of $x \rightarrow e^x$

i.e. if $y=e^x$ then $\frac{d^2y}{dx^2}=\frac{d}{dx}(\frac{dy}{dx})=\frac{d}{dx}e^x=e^x=y$

Thus $y=e^x$ is a solution of $\frac{d^2y}{dx^2}-y=0$.

Another solution is $y=e^{-x}$ since $\frac{d^2y}{dx^2}=\frac{d}{dx}(\frac{dy}{dx})=\frac{d}{dx}(-e^{-x})=e^{-x}=y$

Now $y=e^x$ and $y=e^{-x}$ are independent solutions thus the general solution is $y=Ae^x+Be^{-x}$, where A and B are arbitrary constants.

SOLVING THE EQUATION $a\frac{d^2y}{dx^2}+b\frac{dy}{dx}+cy=0$

The above example suggests that solutions of the form $y = e^{mx}$ should be sought for this equation.

If $y=e^{mx}$ then $\frac{dy}{dx}=me^{mx}$ and $\frac{d^2y}{dx^2}=m^2e^{mx}$

Hence $y = e^{mx}$ is a solution if and only if $am^2e^{mx} + bme^{mx} + ce^{mx} = 0$.

$am^2e^{mx} + bme^{mx} + ce^{mx} = 0 \Leftrightarrow \boldsymbol{am^2 + bm + c = 0}$ (since $e^{mx} \neq 0$)

Suppose that the roots of this equation are m_1 and m_2 then take $y=e^{m_1x}$

$$a\frac{d^2y}{dx^2}+b\frac{dy}{dx}+cy=am_1^2e^{m_1x}+bm_1e^{m_1x}+ce^{m_1x}=(am_1^2+bm_1+c)e^{m_1x}=0.$$

Similary, $y=e^{m_2x}$ is a solution and hence $y=Ae^{m_1x}+Be^{m_2x}$, where A and B are arbitrary constants, is the general solution.

The quadratic equation $\boldsymbol{am^2 + bm + c = 0}$ is called the **auxiliary equation** of the differential equation $a\frac{d^2y}{dx^2}+b\frac{dy}{dx}+cy=0$ and there are three possible situations which we need to explore.

(i) If $am^2 + bm + c = 0$ has real and distinct roots, say, α and β then, $e^{\alpha x}$ and $e^{\beta x}$ are both solutions of the original differential equation

Hence, $y = A e^{\alpha x} + B e^{\beta x}$, where A and B are arbitrary constants, is the general solution of the differential equation.

(ii) If $am^2 + bm + c = 0$ has equal roots, say α, then [A/B]

$e^{\alpha x}$ is a solution of the original differential equation.
It can be shown that $y = x e^{\alpha x}$ is also a solution.

Proof: $a\dfrac{d^2}{dx^2}(xe^{\alpha x}) + b\dfrac{d}{dx}(xe^{\alpha x}) + c(xe^{\alpha x})$

$= a(\alpha^2 xe^{\alpha x} + 2\alpha e^{\alpha x}) + b(\alpha xe^{\alpha x} + e^{\alpha x}) + c\alpha e^{\alpha x}$

$= (a\alpha^2 + b\alpha + c)xe^{\alpha x} + (2a\alpha + b)\, e^{\alpha x}$

$= 0 + (2a\alpha + b)\, e^{\alpha x}$

However, since α is a repeated root, $a(m - \alpha)(m - \alpha) \equiv am^2 + bm + c$

$\therefore\ 2a\alpha\, m = bm \qquad \Rightarrow 2\alpha = -\dfrac{b}{a} \qquad \Rightarrow (2a\alpha + b) = 0$

Therefore, $y = xe^{\alpha x}$ is a solution of the differential equation.

Hence, $y = A\, e^{\alpha x} + B\, xe^{\alpha x}$, where A and B are arbitrary constants, is the general solution of the differential equation.

(iii) When $am^2 + bm + c = 0$ has non-real roots, [A/B]
i.e. complex conjugates $\alpha \pm i\beta$, with α and β real and $\beta > 0$.

It can be shown that $e^{\alpha x}\cos \beta x$ and $e^{\alpha x}\sin \beta x$ are independent solutions of the differential equation.

Showing that $e^{\alpha x}\cos \beta x$ is a solution

$a\dfrac{d^2}{dx^2}(e^{\alpha x}\cos \beta x) + b\dfrac{d}{dx}(e^{\alpha x}\cos \beta x) + c(e^{\alpha x}\cos \beta x)$

$= a[(\alpha^2 - \beta^2)e^{\alpha x}\cos \beta x - 2\alpha\beta\, e^{\alpha x}\sin \beta x] + b[\alpha e^{\alpha x}\cos \beta x - \beta\, e^{\alpha x}\sin \beta x]$
$+ c\, e^{\alpha x}\cos \beta x$

$= [a(\alpha^2 - \beta^2) + b\alpha + c]\, e^{\alpha x}\cos \beta x - (2a\alpha + b)\beta\, e^{\alpha x}\sin \beta x$

Since $\alpha + i\beta$ is a root of $am^2 + bm + c = 0$, $a(\alpha + i\beta)^2 + b(\alpha + i\beta) + c = 0$

i.e. $a(\alpha^2 - \beta^2) + b\alpha + c + i(2a\alpha + b)\beta = 0$

i.e. $a(\alpha^2 - \beta^2) + b\alpha + c = 0$ and $2a\alpha + b = 0$

Therefore, $[a(\alpha^2 - \beta^2) + b\alpha + c]\, e^{\alpha x}\cos \beta x - (2a\alpha + b)\beta\, e^{\alpha x}\sin \beta x = 0$

i.e. $e^{\alpha x}\cos \beta x$ is a solution of the differential equation.

Similarly, $e^{\alpha x}\sin \beta x$ is also a solution.

Hence $y = A\,e^{\alpha x}\cos\beta x + B\,e^{\alpha x}\sin\beta x$, where A and B are arbitrary constants, is the general solution of the differential equation.

Curious readers will find a further explanation of this at the end of this section.

WORKED EXAMPLES

Example 1

Find the general solutions of the following differential equations:

(a) $\dfrac{d^2y}{dx^2}+5\dfrac{dy}{dx}-6y=0$ **(b)** $\dfrac{d^2y}{dx^2}-4\dfrac{dy}{dx}+4y=0$ **(c)** $\dfrac{d^2y}{dx^2}+2\dfrac{dy}{dx}+5y=0$

Solution

(a) Auxiliary equation is $m^2 + 5m - 6 = 0$

$m^2 + 5m - 6 = 0 \Rightarrow (m+6)(m-1) = 0$, i.e. $m = -6$ or 1

General solution is $y = Ae^{-6x} + Be^{x}$

(b) Auxiliary equation is $m^2 - 4m + 4 = 0$

$m^2 - 4m + 4 = 0 \Rightarrow (m-2)^2 = 0$, **i.e.** $m = 2$

General solution is $y = Ae^{2x} + Bxe^{2x}$

(c) Auxiliary equation is $m^2 + 2m + 5 = 0 \Rightarrow m = -1 \pm 2i$

$m^2 + 2m + 5 = 0 \Rightarrow m = -1 \pm 2i$

General solution is $y = e^{-x}(A\cos 2x + B\sin 2x)$

Example 2

When a body, suspended from a light spring hanging vertically, is displaced vertically from its equilibrium position and released its displacement from its equilibrium position after time t satisfies the differential equation $\dfrac{d^2x}{dt^2}+21\dfrac{dx}{dt}+98x=0$.

Hence find a formula for x in terms of t.

Solution

Auxiliary equation is $m^2 + 21m + 98 = 0$

$m^2 + 21m + 98 = 0 \Rightarrow (m+7)(m+14) = 0 \Rightarrow$ m = –7 or –14

General solution is $x = Ae^{-7t} + Be^{-14t}$

Example 3

When a body, suspended from a light spring hanging vertically, is displaced vertically from its equilibrium position and released its displacement from its equilibrium position after time t satisfies the differential equation $\dfrac{d^2x}{dt^2}+98x=0$.

Hence find a formula for x in terms of t.

Solution

Auxiliary equation is $m^2 + 98 = 0$

$m^2 + 98 = 0 \Rightarrow m = \pm\ \sqrt{98}i\ = \pm\ 7\sqrt{2}i$

General solution is $x = A\cos 7\sqrt{2}t + B\sin 7\sqrt{2}t$

Note For information.
Comparing the solutions to examples 2 and 3 above we see that, in the case of example 2, $x \to 0$ as $t \to \infty$ whereas in example 3, x is a periodic function of t with constant amplitude. Example 2 illustrates motion of a spring with a damping term and example 3 illustrates motion of a spring without a damping term.

CONTENT
solve second order non-homogeneous ordinary differential equations with constant coefficients: $a\frac{d^2y}{dx^2}+b\frac{dy}{dx}+cy=f(x)$ **using the auxiliary equation and particular integral method [A/B]**

Comments

$\frac{d^2y}{dx^2}-4\frac{dy}{dx}+13y=0$ with $y = -1$ and $\frac{dy}{dx}=2$ when $x = 0$ is an example of an initial value problem.

This content was contained in CSYS Paper 2 (unrevised).

Teaching notes

Second-order linear differential equations occur in

c) undamped motion of a weight attached to a spring, $m\frac{d^2y}{dt^2}=-ky$

d) forced resonant systems modelled by the equation

$$\frac{d^2y}{dt^2}+\omega^2y=B\omega^2\sin\omega t$$

$$a\frac{d^2y}{dx^2}+b\frac{dy}{dx}+cy=f(x)$$

It is reasonable to suppose that the general solution of the equation

$$a\frac{d^2y}{dx^2}+b\frac{dy}{dx}+cy=f(x)$$

is linked to the general solution of the equation

$$a\frac{d^2y}{dx^2}+b\frac{dy}{dx}+cy=0$$

in some way.

Suppose that $y_1 = Ae^{\alpha x}+Be^{\beta x}$ is the general solution of $a\frac{d^2y}{dx^2}+b\frac{dy}{dx}+cy=0$

Consider $y_2 = y_1 + u$, where u is a function of x then

$$\begin{aligned}a\frac{d^2y_2}{dx^2}+b\frac{dy_2}{dx}+cy_2 &= a\frac{d^2(y_1+u)}{dx^2}+b\frac{d(y_1+u)}{dx}+c(y_1+u)\\ &= a\frac{d^2y_1}{dx^2}+a\frac{d^2u}{dx^2}+b\frac{dy_1}{dx}+b\frac{du}{dx}+cy_1+cu\\ &= a\frac{d^2y_1}{dx^2}+b\frac{dy_1}{dx}+cy_1+a\frac{d^2u}{dx^2}+b\frac{du}{dx}+cu\\ &= a\frac{d^2u}{dx^2}+b\frac{du}{dx}+cu \quad \left[\text{since } a\frac{d^2y_1}{dx^2}+b\frac{dy_1}{dx}+cy_1=0\right]\end{aligned}$$

Hence $y_2 = y_1 + u$ is a general solution provided $a\frac{d^2u}{dx^2}+b\frac{du}{dx}+cu=f(x)$

i.e. provided u is a solution of $a\frac{d^2y}{dx^2}+b\frac{dy}{dx}+cy=f(x)$

TO FIND THE GENERAL SOLUTION OF $a\frac{d^2y}{dx^2}+b\frac{dy}{dx}+cy=f(x)$

The above suggests the following approach:

1. Find the general solution of the corresponding homogeneous differential equation $a\frac{d^2y}{dx^2}+b\frac{dy}{dx}+cy=0$. {called the **complementary function (CF)**}

2. Find a solution of $a\frac{d^2y}{dx^2}+b\frac{dy}{dx}+cy=f(x)$ {called a **particular integral (PI)**}

3. The **general solution** of the original equation is the sum of **the complementary function** and **the particular integral.**

Finding a Particular Integral of $a\frac{d^2y}{dx^2}+b\frac{dy}{dx}+cy=f(x)$

The form of $f(x)$ needs to be considered before the PI can be found.

To find the PI the following table which lists the most important forms for $f(x)$ and the corresponding expressions to be tried as PIs is very useful.

$f(x)$	Form for PI
$a_nx^n + a_{n-1}x^{n-1} + ..a_0$	$y = b_nx^n + b_{n-1}x^{n-1} + ..b_0$
$e^{\alpha x}$	$y = Ae^{\alpha x}$, if α is not a root of $am^2 + bm + c = 0$ $y = Axe^{\alpha x}$, if α is a non-repeated root of $am^2 + bm + c = 0$ $y = Ax^2e^{\alpha x}$, if α is a repeated root of $am^2 + bm + c = 0$
$x\,e^{\alpha x}$	$y = (Ax + B)e^{\alpha x}$, if α is not a root of $am^2 + bm + c = 0$ $y = (Ax^2 + Bx)e^{\alpha x}$, if α is a non-repeated root of $am^2 + bm + c = 0$ $y = (Ax^3 + Bx^2)e^{\alpha x}$, if α is a repeated root of $am^2 + bm + c = 0$
$C\cos\alpha x + D\sin\alpha x$	$y = A\cos\alpha x + B\sin ax$, $i\alpha$ is not a root of $am^2 + bm + c = 0$ $y = x(A\cos\alpha x + B\sin ax)$, $i\alpha$ is a root of $am^2 + bm + c = 0$

The PI is then determined by ensuring that it satisfies $a\frac{d^2y}{dx^2}+b\frac{dy}{dx}+cy=f(x)$.

WORKED EXAMPLES

Example 1

Find the general solutions of the following differential equations:

(a) $\dfrac{d^2y}{dx^2} + 5\dfrac{dy}{dx} - 6y = 7e^x$ (b) $\dfrac{d^2y}{dx^2} - 4\dfrac{dy}{dx} + 4y = 25\sin x$

(c) $\dfrac{d^2y}{dx^2} + 2\dfrac{dy}{dx} + 5y = e^{-x}$

Solution

(a) **Step1:** Solve $\dfrac{d^2y}{dx^2} + 5\dfrac{dy}{dx} - 6y = 0$

Auxiliary equation is $m^2 + 5m - 6 = 0 \Leftrightarrow (m + 6)(m - 1) = 0$

Thus the **complementary function (CF)** is $y = Ae^{-6x} + Be^x$.

Step 2: For the Particular Integral (PI), try $y = Cxe^x$ (1 is a non-repeated root).

$y = Cxe^x \quad \Rightarrow \quad \dfrac{dy}{dx} = Ce^x + Cxe^x$

$\Rightarrow \quad \dfrac{d^2y}{dx^2} = Ce^x + Ce^x + Cxe^x = 2Ce^x + Cxe^x$

Substitute these values in the differential equation to give:

$2Ce^x + Cxe^x + 5Ce^x + 5Cxe^x - 6Cxe^x = 7e^x$

$\Rightarrow \; 7\,Ce^x = 7e^x$

$\Rightarrow \; C = 1$

So, PI is $y = xe^x$

Step 3: General solution is (CF + PI) i.e $y = Ae^{-6x} + Be^x + xe^x$

(b) **Step 1:** Solve the equation $\dfrac{d^2y}{dx^2} - 4\dfrac{dy}{dx} + 4y = 0$

Auxiliary equation is $m^2 - 4m + 4 = 0 \Leftrightarrow (m - 2)(m - 2) = 0$

Thus the CF is $y = Ae^{2x} + Bxe^{2x}$.

Step 2: For PI, try $y = k\cos x + l\sin x$

$y = k\cos x + l\sin x \quad \Rightarrow \quad \dfrac{dy}{dx} = -k\sin x + l\cos x$

$\Rightarrow \quad \dfrac{d^2y}{dx^2} = -k\cos x - l\sin x$

Substituting in the differential equation gives:

$-k\cos x - l\sin x + 4k\sin x - 4l\cos x + 4k\cos x + 4l\sin x = 25\sin x$

$\Rightarrow \; (3k - 4l)\cos x + (4k + 3l)\sin x = 25\sin x$

equating coefficients $\quad 3k - 4l = 0$

$4k + 3l = 25 \quad \Rightarrow k = 4$ and $l = 3$

Thus, PI = $4\cos x + 3\sin x$

Step 3: General solution is CF + PI i.e. $y = Ae^{2x} + Bxe^{2x} + 4\cos x + 3\sin x$

(c) **Step 1**: Solve the equation $\frac{d^2y}{dx^2}+2\frac{dy}{dx}+5y=0$

Auxiliary equation is $m^2 + 2m + 5 = 0 \Rightarrow m = -1 \pm 2i$

Thus CF is $y = e^{-x}(A\cos 2x + B\sin 2x)$.

Step 2: For PI, try $y = ce^{-x}$

$$y = ce^{-x} \Rightarrow \frac{dy}{dx} = -ce^{-x} \Rightarrow \frac{d^2y}{dx^2} = ce^{-x}$$

Substitute in the differential equation to give:

$$ce^{-x} - 2ce^{-x} + 5ce^{-x} = e^{-x} \quad \Rightarrow 4ce^{-x} = e^{-x} \Rightarrow c = \frac{1}{4}$$

Thus, PI is $y = \frac{1}{4}e^{-x}$

Step 3: The general solution is CF + PI i.e. $y = e^{-x}(A\cos 2x + B\sin 2x) + \frac{1}{4}e^{-x}$

Example 2

Find the general solution of the differential equation:

$$\frac{d^2y}{dx^2} - y = \frac{1}{2}x^2$$

and the particular solution which satisfies $y = \frac{dy}{dx} = 0$ at $x = 0$.

Solution

Step 1: Solve the equation $\frac{d^2y}{dx^2} - y = 0$

Auxiliary equation is $m^2 - 1 = 0 \Rightarrow (m-1)(m+1) = 0, \Rightarrow m = \pm 1$

Thus CF is $y = Ae^{x} + Be^{-x}$

Step 2: For PI try $y = ax^2 + bx + c \Rightarrow \frac{dy}{dx} = 2ax + b \Rightarrow \frac{d^2y}{dx^2} = 2a$

Substitute in differential equation to give:

$$2a - ax^2 - bx - c = \frac{1}{2}x^2$$

equate coefficients:

$$-a = \frac{1}{2}, \ -b = 0, \ 2a - c = 0$$

i.e. $a = -\frac{1}{2}$, $b = 0$ and $c = -1$.

Thus PI is $y = -\frac{1}{2}x^2 - 1$

Step 3: General solution is $y = Ae^{x} + Be^{-x} - \frac{1}{2}x^2 - 1$

Step 4: $\frac{dy}{dx} = Ae^{x} + Be^{-x} - x$ At $x = 0, y = 0 \Rightarrow A + B - 1 = 0$

$$\frac{dy}{dx} = 0 \Rightarrow A + B = 0$$

i.e. $A = \frac{1}{2}$ and $B = \frac{1}{2}$

Particular solution is $y = \frac{1}{2}e^{x} + \frac{1}{2}e^{-x} - \frac{1}{2}x^{2} - 1$

Example 3

Earlier we solved the differential equation $L\frac{dI}{dt} + RI = V$ using the integrating factor method. If I = 0 when $t = 0$, try solving this equation using the method of this section.

Solution

Step1: Solve the equation $L\frac{dI}{dt} + RI = 0$

'Auxiliary equation' is $Lm + R = 0 \Rightarrow m = \frac{R}{L}$

Thus CF is $I = Ae^{-\frac{R}{L}t}$

Step 2: For PI try $I = k$

$I = k \Rightarrow \frac{dI}{dt} = 0$ and so substituting in the equation gives $Rk = V$

i.e. $k = \frac{V}{R}$

Step 3: General solution is $I = Ae^{-\frac{R}{L}t} + \frac{V}{R}$

When $t = 0$ $I = 0$ so $0 = A + \frac{V}{R}$ thus $A = -\frac{V}{R}$ and $I = \frac{V}{R}(1 - e^{-\frac{R}{L}t})$

Further advice on finding PIs

If $f(x)$ is a sum of terms, e.g. $f(x) = p(x) + q(x)$ and

$$y = u(x) \text{ is a PI of } a\frac{d^2y}{dx^2} + b\frac{dy}{dx} + cy = p(x)$$

and $$y = v(x) \text{ is a PI of } a\frac{d^2y}{dx^2} + b\frac{dy}{dx} + cy = q(x)$$

then it can easily be shown that

$$y = u(x) + v(x) \text{ is a PI of } a\frac{d^2y}{dx^2} + b\frac{dy}{dx} + cy = p(x) + q(x)$$

Example 4

Solve the equation $\frac{d^2y}{dx^2} - 4\frac{dy}{dx} + 3y = 9x + 65\cos 2x$

Step 1: Solve $\frac{d^2y}{dx^2} - 4\frac{dy}{dx} + 3y = 0$

Auxiliary equation is $m^2 - 4m + 3 = 0 \Rightarrow (m - 1)(m - 3) = 0$

Hence complementary function is $y = Ae^x + Be^{3x}$

Step 2: For PI try $y = ax + b + (c\sin 2x + d\cos 2x)$

$\frac{dy}{dx} = a + 2c\cos 2x - 2d\sin 2x$ and $\frac{d^2y}{dx^2} = -4c\sin 2x - 4d\cos 2x$

Substituting into the equation gives: -

$-4(c\sin 2x + d\cos 2x) - 4(a + 2c\cos 2x - 2d\sin 2x) + 3(ax + b + c\sin 2x + d\cos 2x) = 9x + 65\cos 2x$

i.e. $(8d - c)\sin 2x - (d + 8c)\cos 2x + 3ax + (3b - 4a) = 9x + 65\cos 2x$

Equating coefficients $(8d - c) = 0$; $(d + 8c) = -65$; $3a = 9$; $(3b - 4a) = 0$

Hence $a = 3$; $b = 4$; $d = -1$; $c = -8$ and so the PI is

$y = 3x + 4 - (8\sin 2x + \cos 2x)$

Step 3: General solution is $y = Ae^x + Be^{3x} + 3x + 4 - (8\sin 2x + \cos 2x)$

Example 5

If a potential V (volt) is applied to a simple circuit with inductance L (henry), resistance R (ohm) and capacitance C (farad), then the charge q (coulomb) on the capacitor after time t seconds is given by $L\frac{d^2q}{dt^2} + R\frac{dq}{dt} + \frac{q}{C} = V$.

Find a formula for q where, L = 5×10^{-3}, R = 6, C = 10^{-3} and V = 100.

Solution

Equation is $0\cdot005\frac{d^2q}{dt^2} + 6\frac{dq}{dt} + 1000q = 100$ i.e. $\frac{d^2q}{dt^2} + 1200\frac{dq}{dt} + 200000q = 20000$

Step 1: Solve the equation $\frac{d^2q}{dt^2} + 1200\frac{dq}{dt} + 200000q = 0$

Auxiliary equation is $m^2 + 1200m + 200000 = 0$

$m^2 + 1200m + 200000 = 0 \Rightarrow (m + 200)(m + 1000) = 0$

CF is $q = Ae^{-200t} + Be^{-1000t}$

Step 2: For PI try $q = k$. Substituting in equation $0 + 0 + 200000k = 20000$

Hence PI is $q = \frac{1}{10}$

Step 3: General solution is $q = Ae^{-200t} + Be^{-1000t} + \frac{1}{10}$

EXERCISE ON DIFFERENTIAL EQUATIONS

1. Find the solution of the differential equation $\frac{dy}{dx} + ay = b$, where a and b are non-zero constants, for which $y = 0$ when $x = 0$.
2. Find the general solution of the differential equation $\frac{dy}{dx} = x + y$.
3. Find the general solution of the differential equation $x\frac{dy}{dx} - 2y + \sqrt{x} = 0$.

4. Find the general solution of the differential equation $\cos x \frac{dy}{dx} + y \sin x = 1$.

   Find the solution satisfying the condition $y = \sqrt{2}$ when $x = \frac{\pi}{4}$, and the minimum value of y for this solution.

5. Find the solution of the differential equation $\frac{d^2y}{dx^2} + 6\frac{dy}{dx} + 5y = 4e^{-x}$ for which $y = \frac{dy}{dx} = 0$ at $x = 0$.

6. Find the general solution of the differential equation $\frac{d^2y}{dx^2} + 4y = 12\cos 4x$.

7. Find the general solution of the differential equation $2\frac{d^2y}{dx^2} + \frac{dy}{dx} - y = x$ and the particular solution for which $y(0) = 0$ and $y'(0) = 1$.

8. Find the general solution of the differential equation

$$\frac{d^2y}{dx^2} - \frac{dy}{dx} - 4y = 5 + 2x - 12x^2.$$

   If, in addition, $y = 12$ and $\frac{dy}{dx} = 10$ when $x = 0$, find the particular solution.

Answers

1. $y = \frac{b}{a}(1 - e^{-ax})$
2. $y = ce^x - x - 1$
3. $y = \frac{2}{3}x^{½} + cx^2$
4. $y = \sin x + k\cos x; y = \sin x + \cos x$; $-\sqrt{2}$
5. $y = \frac{1}{4}e^{-5x} - \frac{1}{4}e^{-x} + xe^{-x}$
6. $y = A\cos 2x + B\sin 2x - \cos 4x$
7. $y = Ae^{½x} + Be^{-x} - x - 1$; $A = 2, B = -1$
8. $y = Ae^{3x} + Be^{-2x} + 2x^2 - x$; $A = 7, B = 5$

Earlier we verified that $Ae^{\alpha x}\cos\beta x + Be^{\alpha x}\sin\beta x$, where A and B are arbitrary constants, is the general solution of the differential equation $a\dfrac{d^2y}{dx^2}+b\dfrac{dy}{dx}+cy=0$, where $\alpha \pm i\beta$, with α and β real and $\beta > 0$ are roots of the auxiliary equation $am^2 + bm + c = 0$. But what leads us to consider $Ae^{\alpha x}\cos\beta x + Be^{\alpha x}\sin\beta x$ in the first place?

When the auxiliary equation has real roots m_1 and m_2 the general solution is $Ae^{-m_1x} + Be^{-m_2x}$ and so, assuming that we can differentiate functions of a complex variable we would expect the general solution to be $y = Ce^{(\alpha+i\beta)x} + De^{(\alpha-i\beta)x}$, where C and D are arbitrary constants.

Now $Ce^{(\alpha+i\beta)x} + De^{(\alpha-i\beta)x} = Ce^{\alpha x}e^{i\beta x} + De^{\alpha x}e^{-i\beta x} = e^{\alpha x}(Ce^{i\beta x} + De^{-i\beta x})$

At the end of the section on power series it was shown that

$$e^{i\theta} = \frac{1}{2}(\cos\theta + i\sin\theta) \text{ and } e^{-i\theta} = \frac{1}{2}(\cos\theta - i\sin\theta)$$

$$\begin{aligned}
\text{Thus } Ce^{i\beta x} + De^{-i\beta x} &= \frac{1}{2}C(\cos\beta x + i\sin\beta x) + \frac{1}{2}D(\cos\beta x - i\sin\beta x)\\
&= \frac{1}{2}(C+D)\cos\beta x + \frac{1}{2}(C-D)i\sin\beta x\\
&= A\cos\beta x + B\sin\beta x\\
\text{Hence } Ce^{(\alpha+i\beta)x} + De^{(\alpha-i\beta)x} &= e^{\alpha x}(A\cos\beta x + B\sin\beta x)
\end{aligned}$$

RESOURCES
Differential equations

Understanding Pure Mathematics (Sadler & Thorning)
Chapter 20 (first order differential equations)
pp517 – 519 (second order differential equations (homogeneous))

Calculus (Hunter)
pp.200 - 218

FURTHER NUMBER THEORY AND FURTHER METHODS OF PROOF

CONTENT
know the terms necessary condition, sufficient condition, if and only if, converse, negation, contrapositive use further methods of mathematical proof: some simple examples involving the natural numbers direct methods of proof: sums of certain series and other straightforward results *e.g.* $x > 1 \Rightarrow x^2 > 1$, *the triangle inequality, the sum to n terms of an arithmetic or geometric series* further proof by contradiction *e.g. if* $x, y \in \mathbf{R}$ *such that* $x + y$ *is irrational then at least one of* x, y *is irrational* proof using the contrapositive *e.g. if* m, n *are integers and* $mn = 100$, *then either* $m \leq 10$ *or* $n \leq 10$ *e.g. is 7 is not a factor of* n^2 *then 7 is not a factor of* n

Teaching Notes

This is an extension of the work on proofs undertaken in Mathematics 2(AH) which involved methods of direct and indirect proof, including proof by induction and disproving a conjecture by use of a counter example.

In mathematics we encounter results like:

(a) if $x > 3$ then $x^2 > 9$, where $x \in \mathbf{Z}$

(b) if $f(x) = x^n$ then $f'(x) = nx^{n-1}$

(c) if $z = \cos\frac{\pi}{3} + i\sin\frac{\pi}{3}$ then $z^3 = -1$

These results are stated in the form "if p then q", where p and q stand for statements.

Many of other results can be expressed in this form e.g. the Theorem of Pythagoras can be expressed as - If, in triangle ABC, $\angle B = 90°$ then $AC^2 = AB^2 + BC^2$

Statements that can be proved to be true on the basis of clearly stated assumptions are usually called theorems. However statements of the form "if p then q" are not always true and may be an (informed) guess made on the basis of an observation made during an investigation which has then to be proved to be true or proved to be false.

For example,

(i) Are there an infinite number of twin primes, n and $n + 2$? (unproved)

(ii) Is $2^{2^n} + 1$ prime for all $n \in \mathbf{N}$? (erroneous, holds for $n = 1, 2, 3, 4$ but not for all higher values)

BACKGROUND

One approach to illustrating this topic is through the consideration of solution sets.

In example (a) above, viz "If $x > 3$ then $x^2 > 9$, where $x \in \mathbf{Z}$", p is the statement: $x > 3, x \in \mathbf{Z}$ and q is the statement: $x^2 > 9, x \in \mathbf{Z}$.

Let P be the set of those integers x which are greater than 3, written symbolically as $P = \{x: x > 3, x \in Z\}$. Then **if $x > 3, x \in$ Z** means the same as **if x is a member of P** which is written symbolically as **if $x \in$ P.**

Let $Q = \{x: x^2 > 9, x \in Z\}$. Then **then $x^2 > 9, x \in$ Z** means the same as **then $x \in$ Q.**

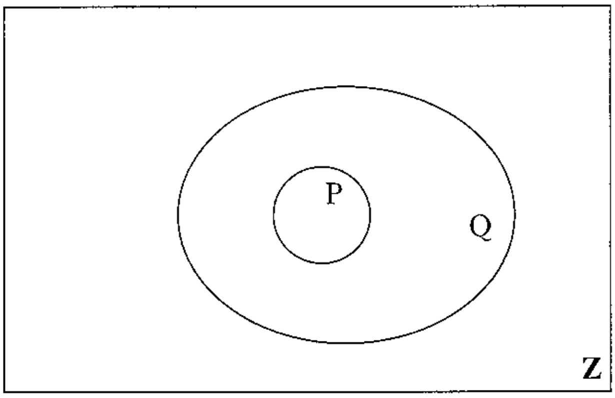
Hence the compound statement **if p then q** means the same as **if $x \in$ P then $x \in$ Q** and **if p then q** is true if every x in P is also in Q, ie if P is a subset of Q. $[P \subseteq Q]$

So proving that **if p then q** is true is the same as proving that **$P \subseteq Q$.**

Now $P = \{4,5,6,7,\ldots\ldots\ldots\ldots\}$ and $Q = \{\ldots.-5,-4,4,5,6,7,\ldots\ldots.\}$
Clearly $P \subseteq Q$ and so "if p then q" is true.

NECESSARY CONDITIONS and SUFFICIENT CONDITIONS

The relationship between the sets P and Q is illustrated in the Venn Diagram below



In 'mathematical English' the words necessary and sufficient have precise meanings.

From the diagram above it is seen that in the situation where $P \subseteq Q$

For $x \in$ Q it is **sufficient** but **not necessary** that $x \in$ P.
For $x \in$ P it is **necessary** but **not sufficient** that $x \in$ Q.

Similarly in the corresponding situation of proving "if p then q" is true;

It is a **sufficient condition** but **not a necessary condition** that p is true.
It is a **necessary condition** but **not a sufficient condition** that q is true.

CONVERSE

The **converse** of "if p then q" is " if q then p" and this is true if $Q \subseteq P$

Clearly $Q \not\subseteq P$ in the illustratory example and "if q then p" is false
i.e.in this case the statement is true but the converse is false.

(Had the initial statement been "if $x^2 > 9$ then $x > 3$" then it would have been false but the converse would have been true.)

NECESSARY AND SUFFICIENT CONDITIONS

In some situations the statement is true and its converse is also true.
For example "In ΔACB, if $\angle C = 90°$ then $c^2 = a^2 + b^2$" can be proved true as can its converse "In ΔACB, if $c^2 = a^2 + b^2$ then $\angle C = 90°$"

In a situation where "if p then q" and "if q then p" are both true we say "p is true **if and only** if q is true"

In this case $P \subseteq Q$ and $Q \subseteq P$ and so $P = Q$ and for $x \in Q$ it is sufficient and necessary that $x \in P$ and we say "A **necessary and sufficient condition** for q to be true is that p is true"

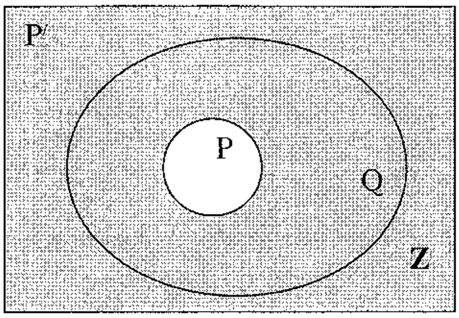
THE NEGATION OF A STATEMENT

The **negation** of a statement p is a statement saying the opposite to p.

For example if p is "$x > 3$" then its negation is "$x \not> 3$" or "$x \leq 3$".
We denote the negation of p by $\sim p$ i.e. if p is "$x > 3$" then $\sim p$ is "$x \leq 3$"

"if $\sim p$" is the same as "if $x \in \{x: x \leq 3\}$".
$\{x: x \leq 3, x \in Z\}$ is called the complement of $\{x: x > 3, x \in Z\}$

The complement of P is sometimes denoted by P′ as illustrated on the right



CONTRAPOSITIVE

Clearly $P \subseteq Q$ is equivalent to $Q' \subseteq P'$
i.e. "if p then q" is true $\Leftrightarrow$ "if $\sim q$ then $\sim p$" is true.

"if $\sim q$ then $\sim p$" is called the **contrapositive** of "if p then q"

METHODS OF PROOF

Direct Proof

In the Teachers' Notes for Mathematics 2(AH) direct proof was described in the following terms "One form of proof is to work from the truth of one or more already known statements to the statement we wish to prove true. Linking statements in this way provides a direct proof".

In the above context we work from the truth of the statement p *through linking true statements until we prove the truth of* q.

A number of examples of direct proof have appeared in the Teachers' Notes for Mathematics 1(AH), Mathematics 2(AH) and earlier in this unit.

Indirect Proof

On occasions it is difficult (impossible) to prove a statement by direct means and consequently we employ indirect methods

Proof using the contrapositive

As indicated above "if ~ q then ~ p" is called the contrapositive of "if p then q" and if one is true then so is the other. Sometimes it is easier to prove the contrapositive of a statement true than prove the statement true.

Example 1
Prove that if m, n are integers and $mn = 100$, then either $m \leq 10$ or $n \leq 10$

Solution
Statement: If m, n are integers and $mn = 100$, then either $m \leq 10$ **or** $n \leq 10$.

Contrapositive: If m, n are integers and $m > 10$ ***and*** $n > 10$ then $mn \neq 100$.

Proof of contrapositive
$m > 10$ *and* $n > 10$ $\Rightarrow mn > 100 \Rightarrow mn \neq 100$

Contrapositive is true. Hence the statement is true.

Example 2
Prove that if 7 is not a factor of n^2 then 7 is not a factor of n

Solution
Statement: If 7 is not a factor of n^2 then 7 is not a factor of n

Contrapositive: If 7 is a factor of n then 7 is a factor of n^2

Proof of contrapositive

7 is a factor of $n \Rightarrow n = 7m$ for some integer m

$\Rightarrow n^2 = (7m)^2 = 49m^2$

$\Rightarrow$ 7 is a factor of m^2

Contrapositive is true. Hence the statement is true.

Example 3
Prove that if x and y are integers and xy is odd then both x and y are odd.

Solution
Statement: If x and y are integers and xy is odd then both x and y are odd

Contrapositive: If x and y are integers and at least one of x or y is even the xy is even

Proof of contrapositive

at least one of x or y is even

$\Rightarrow x = 2m$ or $y = 2n$ for some integer m or n

$\Rightarrow xy = 2my$ or $xy = 2nx$ for some integer m or n

$\Rightarrow xy$ is even

Contrapositive is true. Hence statement is true.

Proof by Contradiction

The basis of this form of proof is that a statement and its negation cannot both be true. We make an initial assumption in the form of a statement r (usually denying the truth of the original statement) and then proceed to prove that $\sim r$ is also true. As this is an impossibility (a contradiction) so our initial assumption is false and hence the statement is true.

The Teaching Notes for Mathematics 2(AH) contain a proof by contradiction for the statement that $\sqrt{2}$ is irrational.

In the proof we assumed that $\sqrt{2}$ is rational i.e. that $\sqrt{2} = \frac{m}{n}$ where m and n are integers. We further assumed, without any loss of generality, that

r: "m and n have no common factor" is true

and then proceeded to prove that as a direct consequence of this assumption that

$\sim r$: "m and n have a common factor" is true

Since r and $\sim r$ cannot both be true we have a contradiction and so our initial assumption that $\sqrt{2}$ is rational is wrong and consequently $\sqrt{2}$ is irrational.

Example 1

Prove that if $x, y \in \mathbf{R}$ such that $x + y$ is irrational then at least one of x, y is irrational

Solution

Assume that "$x, y \in \mathbf{R}$ such that $x + y$ is irrational and x, y are both rational"

$x + y$ is irrational and x, y are both rational

$\Rightarrow x = \frac{m}{n}$ and $y = \frac{p}{q}$ for some integers m, n, p and q.

$\Rightarrow x + y = \frac{mq + np}{nq}$ for some integers m, n, p and q.

$\Rightarrow x + y$ is rational

This is a contradiction and so the initial assumption is wrong.
Hence if $x, y \in R$ such that $x + y$ is irrational then at least one of x, y is irrational is true.

Example 2

Let x and y be integers. If xy is an odd integer, prove that x and y are both odd.

Solution

Suppose xy is an odd integer and at least one of x and y, x say is even.

xy is an odd integer and at least one of x and y, x say is even

$\Rightarrow x = 2z$, where z is an integer.

$\Rightarrow xy = 2zy$

$\Rightarrow xy$ is even.

This is a contradiction and so x and y are both odd.

CONTENT
further proof by mathematical induction prove the following result : $\sum_{r=1}^{n} r^2 = \frac{1}{6}n(n+1)(2n+1)$; $n \in \mathbf{N}$

Teaching Notes

Proof by induction was introduced in Mathematics 2(AH) and the teaching notes provided an explanation of the process.

Prove that $\sum_{r=1}^{n} r^2 = \frac{1}{6}n(n+1)(2n+1)$; $n \in \mathbf{N}$

When $n = 1$, $\sum_{r=1}^{n} r^2 = 1^2 = 1$ **and** $\frac{1}{6}n(n+1)(2n+1) = \frac{1}{6}1.2.3 = 1$ so result true for $n = 1$

Assume that the result is true for a particular positive integer k,

i.e that $\sum_{r=1}^{k} r^2 = \frac{1}{6}k(k+1)(2k+1)$

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

$$= \frac{1}{6}(k+1)[(2k^2+k) + 6(k+1)]$$

$$= \frac{1}{6}(k+1)[2k^2+7k+6]$$

$$= \frac{1}{6}(k+1)[(k+2)(2k+3)]$$

$$= \frac{1}{6}[k+1]([k+1]+1)(2[k+1]+1)$$

Thus if result is true for $n = k$ it is true for $n = k + 1$.

Since it is true for n = 1 by induction it is true for all positive integers

WORKED EXAMPLES

Example 1

Prove by induction that, for all positive integers n,

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

Solution

For $n = 1$, LHS = 1; RHS = $\frac{1}{3}.1.1.3 = 1$, hence true for $n = 1$

Suppose true for a particular positive integer k.

Then $1^2 + 3^2 + 5^2 + \ldots.. + (2k-1)^2 + (2k+1)^2$

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

$$= \frac{1}{3}(2k+1)[k(2k-1) + 3(2k+1)]$$

$$= \frac{1}{3}(2k+1)(2k^2+5k+3)$$

$$= \frac{1}{3}(2k+1)(k+1)(2k+3)$$

$$= \frac{1}{3}(k+1)(2k+1)(2k+3)$$

$$= \frac{1}{3}[k+1](2[k+1]-1)(2[k+1]+1)$$

Thus if result is true for $n = k$ it is true for $n = k + 1$.

Result is true for $n = 1$, hence, by induction, result is true for all positive integers n.

Example 2

Prove, that for all positive integers n, $u_n = 9^n + 7$ is divisible by 8.

Solution

$u_1 = 9^1 + 7 = 16$ and 16 is divisible by 8 so result is true for $n = 1$.

Assume that result is true for a particular positive integer k
i.e. assume $u_k = 9^k + 7$ is divisible by 8.
i.e. assume $u_k = 8q$ for some positive integer q.

$u_{k+1} = 9^{k+1} + 7 = 9.9^k + 7 = (8+1)9^k + 7 = 8.\ 9^k + 9^k + 7 = 8.\ 9^k + u_k$

Hence, since u_k is divisible by 8, u_{k+1} is divisible by 8
Thus if result is true for $n = k$ it is true for $n = k + 1$.
But result true for $n = 1$ so by induction result is true for all positive integers n.

Example 3

Prove that for all positive integers n, that $n^3 - n$ is divisible by 6.

Solution

When $n = 1$, $n^3 - n = 0$. Now 0 is divisible by 6 so result is true for $n = 1$

Assume the result is true a particular positive integer k, i.e. assume $k^3 - k$ is divisible by 6

Consider $(k+1)^3 - (k+1)$

$(k+1)^3 - (k+1) = k^3 + 3k^2 + 3k + 1 - (k+1) = k^3 + 3k^2 + 2k = k^3 - k + 3k^2 + 3k$

Hence, $(k+1)^3 - (k+1) = (k^3 - k) + 3k(k+1)$

Either k or $k + 1$ is even so $k(k+1)$ is divisible by 2 and so $3\,k(k+1)$ is divisible by 6

Hence, since $k^3 - k$ is divisible by 6, $(k+1)^3 - (k+1)$ is divisible by 6
so if result true for $n = k$ it is true for $n = k + 1$

But result is true for $n = 1$ so by induction the result is true for all positive integers n.

CONTENT
know the result : $\sum_{r=1}^{n} r^3 = \frac{n^2(n+1)^2}{4}$ *e.g. proof of the result* $\sum_{r=1}^{n} r^3 = \left(\sum_{r=1}^{n} r\right)^2 = \frac{n^2(n+1)^2}{4}$ *is a useful extension* apply the above results and the one for $\sum_{r=1}^{n} r$ to prove by direct methods results concerning other sums *e.g.* $\sum_{r=1}^{n} r(r+1) = \frac{1}{3}n(n+1)(n+2)$ *e.g.* $\sum_{r=1}^{n} r(r+1)(r+2) = \frac{1}{4}n(n+1)(n+2)(n+3)$

The result : $\sum_{r=1}^{n} r^3 = \frac{n^2(n+1)^2}{4}$

Students could investigate the sums $4\sum_{r=1}^{n} r^3$ for n = 1, 2, 3, 4,....They should notice that the sums are each the square of a number which is the product of two consecutive integers, e.g $36 = 6^2 = (2\times3)^2$; $144 = 12^2 = (3\times4)^2$ e.t.c. and conjecture that $4\sum_{r=1}^{n} r^3 = n^2(n+1)^2$. The result can then be proved true by induction.

Example 1

Solution

Show that $\sum_{r=1}^{n} r(r+1) = \frac{1}{3}n(n+1)(n+2)$

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

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

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

$$= \frac{1}{3}n(n+1)(n+2)$$

Example 2

Show that $\sum_{r=1}^{n} r(r+1)(r+2) = \frac{1}{4}n(n+1)(n+2)(n+3)$

Solution

$$\sum_{r=1}^{n} r(r+1)(r+2) = \sum_{r=1}^{n} (r^3 + 3r^2 + 2r) = \sum_{r=1}^{n} r^3 + 3\sum_{r=1}^{n} r^2 + 2\sum_{r=1}^{n} r$$

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

$$= \frac{1}{4}\, n(n+1)[\, n(n+1) + 2(2n+1) + 4]$$

$$= \frac{1}{4}\, n(n+1)(n^2 + 5n + 6)$$

$$= \frac{1}{4}\, n(n+1)(n+2)(n+3)$$

CONTENT
know the division algorithm and proof
use Euclid's algorithm to find the greatest common divisor (g.c.d.) of two positive integer *i.e. use the division algorithm repeatedly*
know how to express the g.c.d. as a linear combination of the two integers [A/B]
use the division algorithm to write integers in terms of bases other than 10 [A/B]

Comments

The special cases of finding the g.c.d. of two Fibonacci numbers are possible extensions.

Teaching Notes

DIVISION ALGORITHM

For given integers a and b, with $b > 0$, there exist integers q and r such that

$$a = bq + r, \ 0 \quad 0 \leq r < b$$

Proof

On the real number line, the integers $a, a - b, a - 2b, \ldots$ form a decreasing sequence with only finitely many of the terms ≥ 0.

Thus there is a unique integer q for which

$$a - (q + 1)b < 0 \leq a - qb$$

and so $\quad 0 \leq a - qb < b$

Writing $r = a - qb$ then $a = qb + r, 0 \leq r < b$

To show that q and r are unique, suppose that $a = q_1b + r_1, 0 \leq r_1 < b$

Then $r_1 = a - q_1b$ and $0 \leq a - q_1b < b$

So that $\quad a - (q_1 + 1)b < 0 \leq a - q_1b$

i.e. q_1 is the integer defined above as q.

EUCLID'S ALGORITHM [A/B]

Euclid's algorithm is concerned with finding the greatest common divisor (g.c.d) of two positive integers and is based on repeated application of the division algorithm.

Thus, for integers a and b,

$$a = q_1 b + r_1, \qquad 0 \le r_1 < b$$
$$b = q_2 r_1 + r_2, \qquad 0 \le r_2 < r_1$$
$$r_1 = q_3 r_2 + r_3, \qquad 0 \le r_3 < r_2$$

until $\quad r_{n-2} = q_n r_{n-1} + r_n$

$$r_{n-1} = q_{n+1} r_n + 0$$

Starting with the last statement and working towards the first one we see that:
$r_n \mid r_{n-1}$ and so $r_n \mid r_{n-2}$ and so $r_n \mid r_{n-3}$ etc. until $r_n \mid r_1$ and $r_n \mid b$ and $r_n \mid a$

Hence $r_n \mid d$, where d is the g.c.d. of a and b (since $r_n \mid a$ and $r_n \mid b$)

Starting from the first statement and working to the last one we see that:
$r_1 = a - q_1 b \Rightarrow d \mid r_1$; $r_2 = b - q_2 r_1 \Rightarrow d \mid r_2 \ldots\ldots\ldots\ldots\ldots\Rightarrow d \mid r_n$

We have shown that $r_n \mid d$ and $d \mid r_n$ thus $r_n = d$.

A common notation used for the g.c.d. of a and b is (a, b)

USING EUCLID'S ALGORITHM TO FIND THE G.C.D. OF TWO POSITIVE INTEGERS

Example
Find (3657, 2703)

Solution

$$3657 = 1 \times 2703 + 954$$
$$2703 = 2 \times 954 + 795$$
$$954 = 1 \times 795 + 159$$
$$795 = 5 \times 159 + 0$$

$\therefore (3657, 2703) = 159$

EXPRESSING THE G. C. D. OF TWO POSITIVE INTEGERS AS A LINEAR COMBINATION OF THE TWO INTEGERS. (A/B)

Example

(i) Use Euclid's algorithm to find the g.c.d. of 2873 and 6643.
(ii)Hence find integers s and t to write this g.c.d. in the form $s.2873 + t.6643$

Solution

(i)
$$6643 = 2 \times 2873 + 897$$
$$2873 = 3 \times 897 + 182$$
$$897 = 4 \times 182 + 169$$
$$182 = 1 \times 169 + 13$$
$$169 = 13 \times 13 + 0$$

so, $(2873, 6643) = 13$

(ii) from 2^{nd} bottom line above,
$$13 = 182 - 1 \times 169$$
$$= 182 - [897 - 4 \times 182] = 5 \times 182 - 897$$
$$= 5[2873 - 3 \times 897] - 897 = 5 \times 2873 - 16 \times 897$$

$= 5 \times 2873 - 16[6643 - 2 \times 2873]$

$= 37 \times 2873 - 16 \times 6643$

i.e. $s = 37$ and $t = -16$

REPRESENTATION OF INTEGERS IN BASES OTHER THAN TEN [A/B]

Another application of the division algorithm is to express numbers in bases other than ten

For example, to write 797_{ten} in base seven we apply the division algorithm by dividing repeatedly by seven as follows:

$$\begin{aligned} 797 &= 7 \times 113 + 6 \\ 113 &= 7 \times 16 + 1 \\ 16 &= 7 \times 2 + 2 \\ 2 &= 7 \times 0 + 2 \end{aligned} \uparrow$$

Therefore $797_{ten} = 2216_{seven}$

Justification

$$\begin{aligned} 797 &= 7 \times 113 + 6 \\ &= 7\,(7 \times 16 + 1) + 6 \\ &= 7^2 \times 16 + 7 \times 1 + 6 \\ &= 7^2(7 \times 2 + 2) + 7 \times 1 + 6 \\ &= 7^3 \times 2 + 7^2 \times 2 + 7 \times 1 + 6 \\ &= 2.7^3 + 2.7^2 + 1.7 + 6 \end{aligned}$$

In general, if g is a positive integer greater than 1, then each positive integer a can be expressed uniquely in the form:

$$a = c_n g^n + c_{n-1} g^{n-1} + \dots + c_1 g + c_0, \quad \text{where } 0 \le c_i \le n-1, \text{ and } 0 < c_n \le g-1$$

EXERCISE ON FURTHER NUMBER THEORY

1. Give an example of two unequal positive integers m, n such that $\sqrt{m}$, $\sqrt{n}$ are both irrational while $\sqrt{(mn)}$ is rational.

2. Determine whether each of the following assertions about integers is true or false. If true, prove the result, and if false give a counter example.

   (i) If $b \mid (a^2 + 1)$ then $b \mid (a^4 + 1)$
   (ii) If $b \mid (a^2 + 1)$ then $b \mid (a^4 - 1)$
   (iii) If p is prime and $p \mid a$ and $p \mid (a^2 + b^2)$ then $p \mid b$
   (iv) If p is prime and $p \mid a$ and $p \mid (a^2 + 6b^2)$ then $p \mid b$

3. Prove by induction that, for all integers $n \geq 2$,

$$\binom{2}{2} + \binom{3}{2} + \dots + \binom{n}{2} = \binom{n+1}{3}$$

4. Prove that $\sqrt{6}$ is irrational.

5. Prove that, if n^2 is odd, where n is an integer, then n is odd.

6. Prove that $\sum_{r=1}^{n} r^2(r+1) = \frac{1}{12}n(n+1)(n+2)(3n+1)$

7. Use Euclid's algorithm to find the g.c.d. of 7293 and 798 and express this g.c.d as $7293x + 798y$ where x and y are integers.

8. Write 728 in the scale of 5.

9. Use Euclid's algorithm to find integers a and b such that $93a + 25b = 1$.

Answers

1. $m = 2, n = 8$ say

2.

(i) F e.g. $a = 2$ and $b = 5$
(ii) T
(iii) T
(iv) F e.g. $p = 2, a = 4, b = 3$

7. $(7293, 798) = 3$; $3 = (-115) \times 7293 + 1051 \times 798$, i.e. $x = -115$ and $y = 1051$

8. $728 = 10403_5$

9. $a = 7, b = -26$

RESOURCES

Further Number Theory and Further methods of proof

Algebra and Number Systems (Hunter)
Chapters 1, 5 and 6.

Mathematics in Action 6S
Chapter 2