CAMBRIDGE UNIVERSITY ENGINEERING DEPARTMENT

PART IA (First Year) 2009-2010

Paper 4 : Mathematical Methods

Lecture course	**:**	**Fast Maths Course, Lectures 1–8**
Lecturer	**:**	**Prof. J. B. Young**
Schedule	**:**	**Weeks 1–4, Michaelmas 2009**
Recommended book	**:**	**James G. Modern Engineering Mathematics. 3rd Edition. Addison–Wesley. 2001.**

VECTOR ALGEBRA

Reference :

James G., Modern Engineering Mathematics, Chapters 3 & 4 to 4.4.

1 VECTOR FUNDAMENTALS

1.1 Definition and basic rules of manipulation

In mechanics, Newton's Second Law is written $F = ma$, implying that when a body of mass m is acted on by a force F the body experiences an acceleration a in the same direction as the force. This notation is fine so long as we are dealing with a problem in one spatial dimension. When faced with a real-world 3D problem, however, we require three scalar equations relating the components of the force in the three cartesian directions (F_x, F_y, F_z) to the corresponding components of the acceleration (a_x, a_y, a_z) :

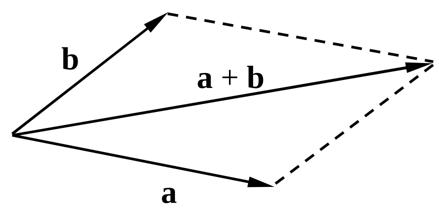
$$F_x = ma_x , \qquad F_y = ma_y , \qquad F_z = ma_z .$$

Vector methods have been developed as a shorthand so that the three scalar equations can be replaced by a single vector equation relating the vector force **F** to the vector acceleration **a** :

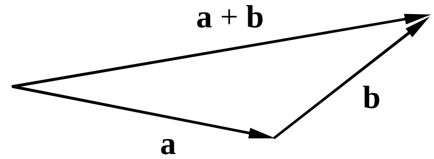
$$\mathbf{F} = m\mathbf{a}$$

Of course, the applications of vector algebra are not restricted to mechanics.

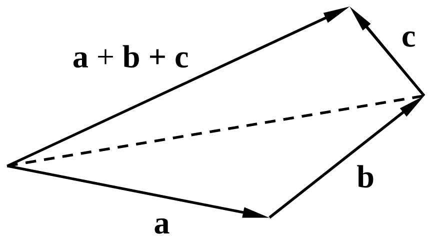
For our purposes, we take the definition of a vector as follows :

> Vectors are quantities possessing both magnitude and direction which obey the parallelogram rule of addition.
>
> 
>

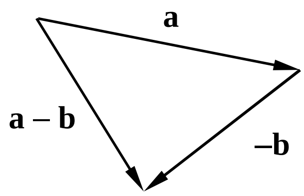
The parallelogram rule is equivalent to adding the vectors 'nose to tail' so in practice it is easier to work in terms of a *triangle* rule of addition :



Several vectors can be added geometrically by successive applications of the triangle rule :



Vectors are subtracted by adding the negative of the relevant vector : $\mathbf{a} - \mathbf{b} = \mathbf{a} + (-\mathbf{b})$.
Thus, if **a** and **b** are the same vectors as in the diagrams above :



Vectors, like scalars, obey the following algebraic rules :

$$\mathbf{a} + \mathbf{b} = \mathbf{b} + \mathbf{a}$$

$$(\mathbf{a} + \mathbf{b}) + \mathbf{c} = \mathbf{a} + (\mathbf{b} + \mathbf{c})$$

$$k(\mathbf{a} + \mathbf{b}) = k\mathbf{a} + k\mathbf{b} \qquad \text{(where } k \text{ is any scalar constant)}$$

In these notes :

- Vectors will be represented by **upright bold type** : **a**
- The magnitude of a vector will be represented by a pair of vertical lines : $|\mathbf{a}|$

A vector of unit length will be represented by a caret ^ : $\hat{\mathbf{a}}$ is a unit vector parallel to **a**.

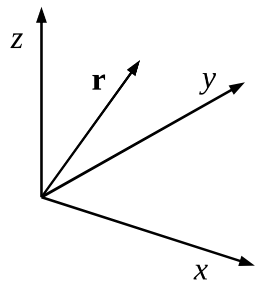
Two or more vectors which are perpendicular to each other are said to be *orthogonal*.

Orthogonal vectors of unit length are said to be *orthonormal*.

The vectors **i**, **j**, **k** are understood to be unit vectors in the directions of the *x*, *y*, *z* axes of a right-handed cartesian coordinate system.

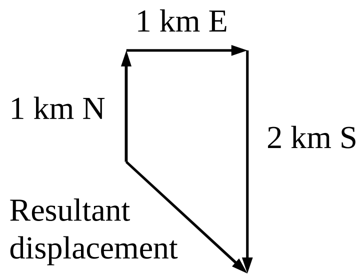
1.2 Examples of vectors

(i) *Position relative to a specified origin*.



The position vector **r**, illustrated with respect to a cartesian coordinate system.

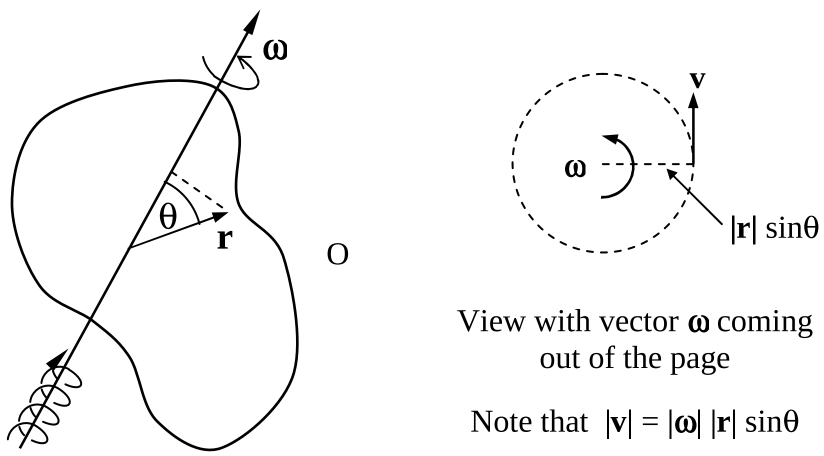
(ii) *Linear displacement*. Displacements can be added using the triangle law :



(iii) *Linear velocity* and *linear acceleration*.

(iv) *Angular velocity* and *angular acceleration*. The angular velocity vector **ω** of a rotating body is the vector whose :

- magnitude equals the body's rate of rotation (*e.g.*, in radians per second), and
- direction is that in which a right-hand screw with the same spin as the body would move through a stationary cork.



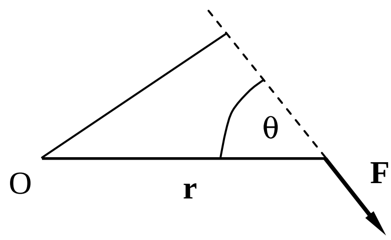
View with vector **ω** coming out of the page

Note that $|\mathbf{v}| = |\boldsymbol{\omega}|\,|\mathbf{r}|\sin\theta$

(v) *Force*. Forces can be added using the triangle law.

(vi) *The moment of a force.* The vector moment **M** of a vector force **F** applied at a point **r** about an origin O is the vector whose :

- magnitude equals the moment about O, and
- direction is that in which a right-hand screw would move if twisted by the moment.



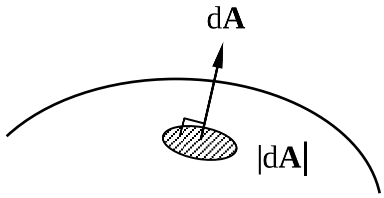
Vector moment **M** has magnitude

$$|\mathbf{M}| = |\mathbf{F}|\,|\mathbf{r}| \sin\theta$$

and direction into the page.

(vii) *The differential area on a surface*. The differential vector area d**A** is a vector whose :

- magnitude equals the area |d**A**| of a differential element on the surface, and
- direction is normal to, and outward from, the side of the surface defined as positive.



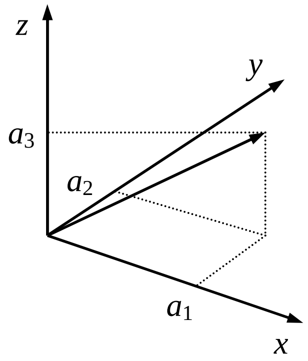
Quantities that have magnitude and direction but do not obey the parallelogram law are not vectors. For example, the <u>finite</u> angular rotation of a body is not a vector. This is because the application of a finite rotation θ_1 about the x axis followed by a finite rotation θ_2 about the z axis, say, gives a different result from a rotation θ_2 about the z axis followed by a rotation θ_1 about the x axis. However, the resultant of two <u>infinitesimal</u> rotations $d\theta_1$ and $d\theta_2$ is independent of the order of application so that angular velocities (of magnitude $d\theta_1/dt$ and $d\theta_2/dt$) do add vectorially (and so angular velocity is a vector).

It is important to appreciate that the magnitude and direction of a vector is independent of the coordinate system used to describe it. This is the great advantage of working with vectors. For example, the statement **c** = **a** + **b** is completely independent of any coordinate system. This means that we can manipulate vectors using the rules of vector algebra without specifying a coordinate system. It is only when we come to the stage of inserting numerical values that is it necessary to choose a coordinate system.

1.3 Representing vectors in component form

We can express the magnitude and direction of a vector numerically with respect to any chosen co-ordinate system. For a cartesian system, we set up mutually perpendicular axes *x*, *y* and *z*, and define **i**, **j** and **k** to be vectors of unit length along the axes. A vector **a** can then be expressed as a linear combination of the vectors **i**, **j** and **k** :

$$\mathbf{a} = a_1\,\mathbf{i} + a_2\,\mathbf{j} + a_3\,\mathbf{k}$$



a_1, a_2 and a_3 are called the components of **a**. These alone are sufficient to specify the vector once the coordinate axes have been defined, so **a** is often represented by,

$$\begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} \qquad \text{or} \qquad a_i \quad (i = 1, 2, 3)\ .$$

The convention is to use a *right-handed* set of axes, obtained as follows. Draw axes 1 and 2 (*x* and *y*) at right angles in a plane. Draw axis 3 (*z*) normal to the plane in the direction a right-handed screw would move were axis 1 rotated to lie on top of axis 2. It will now be found that axis 1 is in the direction a right-handed screw would move if axis 2 were rotated to lie on top of axis 3. Similarly, axis 2 is in the direction a right-handed screw would move if axis 3 were rotated to lie on top of axis 1. Note the cyclic order : 1 rotated to 2 gives 3, 2 rotated to 3 gives 1, 3 rotated to 1 gives 2.

Addition of vectors is accomplished by adding components. Thus,

$$\mathbf{c} = \mathbf{a} + \mathbf{b} = (a_1 + b_1)\mathbf{i} + (a_2 + b_2)\,\mathbf{j} + (a_3 + b_3)\mathbf{k}$$

or, alternatively,

$$\begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} + \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} = \begin{bmatrix} a_1 + b_1 \\ a_2 + b_2 \\ a_3 + b_3 \end{bmatrix}$$

or, more compactly,

$$c_i = a_i + b_i \quad (i = 1, 2, 3)$$

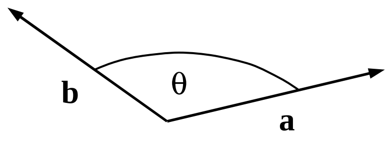
2 VECTOR MULTIPLICATION

2.1 The scalar product

The scalar product of two vectors **a** and **b** is written $\mathbf{a} \bullet \mathbf{b}$ and is defined by,

$$\mathbf{a} \bullet \mathbf{b} = |\mathbf{a}|\,|\mathbf{b}| \cos\theta$$

where θ is the included angle.



The scalar product can be thought of as :

- The magnitude of **a** multiplied by the component of **b** in the direction of **a**, or
- The magnitude of **b** multiplied by the component of **a** in the direction of **b**.

In forming the scalar product, we take two vectors and end up with a scalar (a single number). It can be seen from the definition that the order of the vectors does not matter, $\mathbf{b} \bullet \mathbf{a} = \mathbf{a} \bullet \mathbf{b}$. Also, the scalar product obeys the so-called *distributive* law,

$$\mathbf{a} \bullet (\mathbf{b} + \mathbf{c}) = \mathbf{a} \bullet \mathbf{b} + \mathbf{a} \bullet \mathbf{c}$$

If **a** and **b** are orthogonal, cos θ = 0 so $\mathbf{a} \bullet \mathbf{b} = 0$.

If **a** and **b** are parallel, cos θ = 1 so $\mathbf{a} \bullet \mathbf{b} = |\mathbf{a}||\mathbf{b}|$.

If **a** and **b** are anti-parallel, cos θ = −1 so $\mathbf{a} \bullet \mathbf{b} = -|\mathbf{a}||\mathbf{b}|$.

The unit vectors in cartesian coordinates therefore obey the following relationships :

$$\mathbf{i} \bullet \mathbf{i} = \mathbf{j} \bullet \mathbf{j} = \mathbf{k} \bullet \mathbf{k} = 1$$

$$\mathbf{i} \bullet \mathbf{j} = \mathbf{j} \bullet \mathbf{k} = \mathbf{k} \bullet \mathbf{i} = 0$$

If we express the vectors **a** and **b** in cartesian form and multiply out, we obtain,

$$\mathbf{a} \bullet \mathbf{b} \;=\; (a_x\mathbf{i} + a_y\mathbf{j} + a_z\mathbf{k}) \bullet (b_x\mathbf{i} + b_y\mathbf{j} + b_z\mathbf{k}) \;=\; a_x b_x + a_y b_y + a_z b_z$$

In particular,

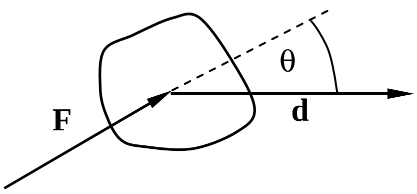
$$\mathbf{a} \bullet \mathbf{a} \;=\; (a_x\mathbf{i} + a_y\mathbf{j} + a_z\mathbf{k}) \bullet (a_x\mathbf{i} + a_y\mathbf{j} + a_z\mathbf{k}) \;=\; a_x^2 + a_y^2 + a_z^2 \;=\; |\mathbf{a}|^2$$

It is very easy to form the scalar product numerically as shown by this example :

$$\begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} \bullet \begin{bmatrix} 4 \\ 2 \\ 3 \end{bmatrix} \;=\; (1\times4) + (2\times2) + (1\times3) \;=\; 11$$

The scalar product is an *invariant*. This means it takes the same value whatever coordinate system is used. If, in the above example, we had decided to use 3D polar rather than cartesian coordinates, the components of **a** and **b** would be different but the scalar product would still be 11.

As a physical example of the use of the scalar product suppose a vector force **F** moves a body a vector distance **d**, not necessarily in the same direction as the force :



The component of **F** parallel to the direction of motion does work on the body but the normal component makes no contribution. Hence, the work done by the force is,

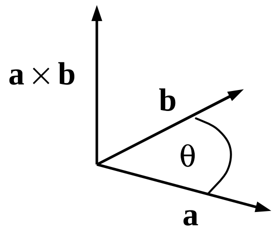
$$W \;=\; |\mathbf{F}| \cos\theta \, |\mathbf{d}| \;=\; \mathbf{F} \bullet \mathbf{d}$$

2.2 The vector product

The *vector product* of two vectors **a** and **b** (written $\mathbf{a}\times\mathbf{b}$ or $\mathbf{a}\wedge\mathbf{b}$ and pronounced 'a cross b') is defined to be a vector whose magnitude is given by

$$|\mathbf{a}\times\mathbf{b}| = |\mathbf{a}|\,|\mathbf{b}|\sin\theta$$

where θ is the included angle between the vectors (such that θ < 180° when the tails of the vectors are made to coincide). The direction of $\mathbf{a}\times\mathbf{b}$ is perpendicular to both **a** and **b**, such that if we rotate **a** to lie on top of **b**, the direction of motion of a right-handed screw is in the direction of $\mathbf{a}\times\mathbf{b}$.



In forming the vector product we take two <u>vectors</u> and obtain another <u>vector</u>. It can be seen from the definition that the order of the vectors is important and that,

$$\mathbf{b}\times\mathbf{a} = -\mathbf{a}\times\mathbf{b}$$

Like the scalar product, the vector product obeys the distributive law,

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

If **a** and **b** are orthogonal, sin θ = 1 so $|\mathbf{a}\times\mathbf{b}| = |\mathbf{a}|\,|\mathbf{b}|$.

If **a** and **b** are parallel or anti-parallel, sin θ = 0 so $|\mathbf{a}\times\mathbf{b}| = 0$.

The unit vectors in cartesian coordinates obey the following relationships :

$$\mathbf{i}\times\mathbf{j} = \mathbf{k}, \qquad \mathbf{j}\times\mathbf{k} = \mathbf{i}, \qquad \mathbf{k}\times\mathbf{i} = \mathbf{j}$$

$$\mathbf{j}\times\mathbf{i} = -\mathbf{k}, \qquad \mathbf{k}\times\mathbf{j} = -\mathbf{i}, \qquad \mathbf{i}\times\mathbf{k} = -\mathbf{j}$$

$$\mathbf{i}\times\mathbf{i} = 0, \qquad \mathbf{j}\times\mathbf{j} = 0, \qquad \mathbf{k}\times\mathbf{k} = 0$$

Expressing the vectors **a** and **b** in cartesian components and multiplying out, we obtain,

$$\mathbf{a}\times\mathbf{b} = (a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k}) \times (b_1\mathbf{i} + b_2\mathbf{j} + b_3\mathbf{k})$$

$$= (a_2b_3 - a_3b_2)\mathbf{i} + (a_3b_1 - a_1b_3)\mathbf{j} + (a_1b_2 - a_2b_1)\mathbf{k}$$

Like the scalar product, the vector product is invariant; its value does not depend on the coordinate system used.

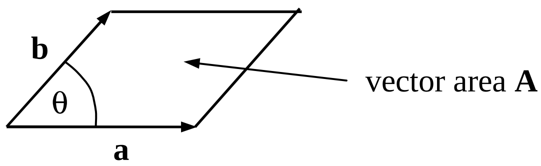
In the above form, the cartesian expansion for **a**×**b** is quite difficult to remember. The result can be found in the mathematics data book but the easiest way of getting the cartesian components correct is to remember the representation of **a**×**b** as a determinant (see later if you don't know what a determinant is) :

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

$$= (a_2b_3 - a_3b_2)\mathbf{i} + (a_3b_1 - a_1b_3)\mathbf{j} + (a_1b_2 - a_2b_1)\mathbf{k}$$

The following are examples of the use of the vector product :

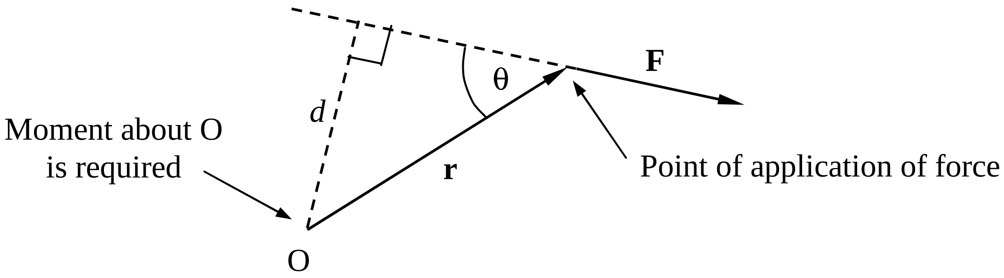
(i) The vector area of a parallelogram :



The magnitude of the area of the parallelogram is $|\mathbf{A}| = |\mathbf{a}|\,|\mathbf{b}| \sin\theta$. The vector area (direction out of the page) is,

$$\mathbf{A} = \mathbf{a}\times\mathbf{b} \qquad \text{(The order is important)}$$

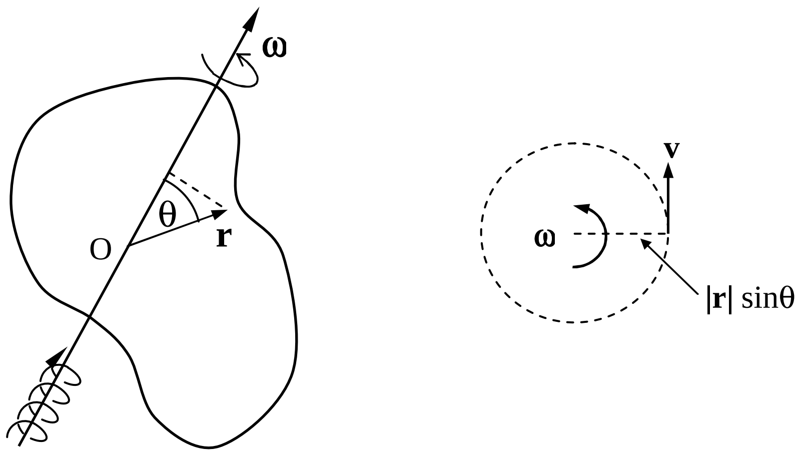
(ii) The vector moment of a force :



The magnitude of the moment about O of the force **F**, applied at a point on the body defined by the position vector **r**, is $M = |\mathbf{F}|\,|\mathbf{r}|\sin\theta$. The vector moment (direction into the page) is,

$$\mathbf{M} = \mathbf{r} \times \mathbf{F} \qquad \text{(The order is important)}$$

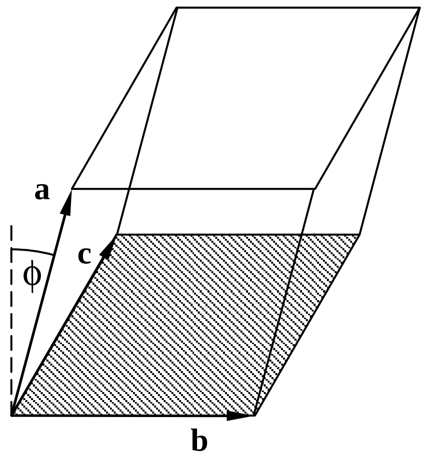
(iii) The velocity at a point in a rotating body :



When a body rotates with angular velocity **ω** as shown, the magnitude of the velocity at the point specified by the position vector **r** is $|\mathbf{v}| = |\boldsymbol{\omega}|\,|\mathbf{r}|\sin\theta$. The vector velocity (direction as shown) is,

$$\mathbf{v} = \boldsymbol{\omega} \times \mathbf{r} \qquad \text{(The order is important)}$$

2.3 The scalar triple product



The parallelepiped shown above is defined by the three vectors **a**, **b** and **c**. The area of the base is **b**×**c** (with direction 'upwards'). The volume V of the parallelepiped (a positive scalar quantity) is therefore,

$$V = |\mathbf{a}| \cos\phi \, |\mathbf{b}\times\mathbf{c}| = \mathbf{a}\bullet(\mathbf{b}\times\mathbf{c})$$

$\mathbf{a}\bullet(\mathbf{b}\times\mathbf{c})$ is known as the *scalar triple product*. In fact, the brackets are unnecessary as the expression $\mathbf{a}\bullet\mathbf{b}\times\mathbf{c}$ could not be evaluated in any other way.

The volume can be evaluated using any two of the vectors to form the base. Hence, we deduce that changing the vectors around in cyclic order does not change the value of the scalar triple product because,

$$\mathbf{a}\bullet(\mathbf{b}\times\mathbf{c}) = \mathbf{c}\bullet(\mathbf{a}\times\mathbf{b}) = \mathbf{b}\bullet(\mathbf{c}\times\mathbf{a}) = V$$

Changing the cyclic order, however, changes the sign of the scalar triple product,

$$\mathbf{a}\bullet(\mathbf{c}\times\mathbf{b}) = \mathbf{b}\bullet(\mathbf{a}\times\mathbf{c}) = \mathbf{c}\bullet(\mathbf{b}\times\mathbf{a}) = -V$$

From the above relationships we obtain,

$$\mathbf{a}\bullet(\mathbf{b}\times\mathbf{c}) = \mathbf{c}\bullet(\mathbf{a}\times\mathbf{b}) = (\mathbf{a}\times\mathbf{b})\bullet\mathbf{c}$$

which shows that we can interchange the $\bullet$ and the $\times$ in a scalar triple product without affecting the value. Because of this, the scalar triple product is sometimes written [**a**,**b**,**c**] as there can be no ambiguity.

Note that $[\mathbf{a},\mathbf{b},\mathbf{c}] = 0$ if the three vectors are coplanar (the parallelepiped volume is zero).

2.4 The vector triple product

The other possible combination of three vectors is the *vector triple product*, $\mathbf{a}\times(\mathbf{b}\times\mathbf{c})$. The result of this operation gives a vector which is orthogonal to the vectors $\mathbf{a}$ and $\mathbf{b}\times\mathbf{c}$. We can derive an important vector identity for simplifying calculations where the vector triple product appears. We do this by considering the x (or 1) component in a cartesian expansion :

$$\begin{aligned}
[\mathbf{a}\times(\mathbf{b}\times\mathbf{c})]_1 &= a_2\,(\mathbf{b}\times\mathbf{c})_3 - a_3\,(\mathbf{b}\times\mathbf{c})_2 \\
&= a_2\,(b_1c_2 - b_2c_1) - a_3\,(b_3c_1 - b_1c_3) \\
&= (a_2c_2 + a_3c_3)\,b_1 - (a_2b_2 + a_3b_3)\,c_1 \\
&= (a_1c_1 + a_2c_2 + a_3c_3)\,b_1 - (a_1b_1 + a_2b_2 + a_3b_3)\,c_1 \\
&= [(\mathbf{a}\bullet\mathbf{c})\mathbf{b} - (\mathbf{a}\bullet\mathbf{b})\mathbf{c}]_1
\end{aligned}$$

Note the clever trick in line 4 where $a_1b_1c_1$ was both added and subtracted. The same result is obtained for the 2 and 3 components (by symmetry), and hence we obtain the identity :

$$\mathbf{a}\times(\mathbf{b}\times\mathbf{c}) = (\mathbf{a}\bullet\mathbf{c})\mathbf{b} - (\mathbf{a}\bullet\mathbf{b})\mathbf{c}$$

If, instead, the first two vectors are bracketed, we easily find that,

$$(\mathbf{a}\times\mathbf{b})\times\mathbf{c} = -\mathbf{c}\times(\mathbf{a}\times\mathbf{b}) = (\mathbf{a}\bullet\mathbf{c})\mathbf{b} - (\mathbf{b}\bullet\mathbf{c})\mathbf{a}$$

As an example of vector manipulation, we simplify the expression $(\mathbf{a}\times\mathbf{b})\bullet(\mathbf{a}\times\mathbf{b})$ **:**

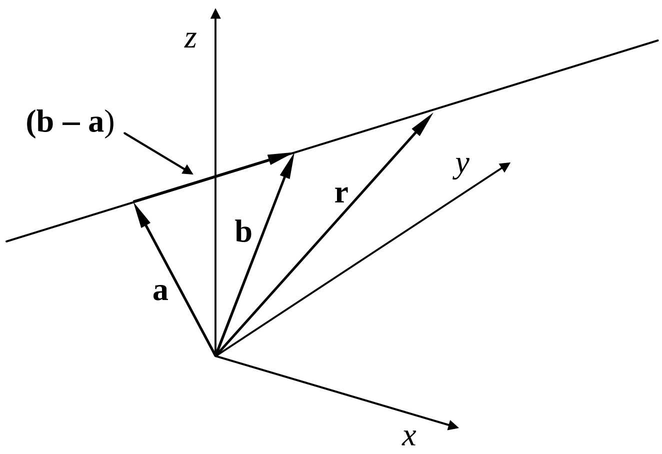
$$\begin{aligned}
(\mathbf{a}\times\mathbf{b})\bullet(\mathbf{a}\times\mathbf{b}) &= (\mathbf{a}\times\mathbf{b})\bullet\mathbf{c} && \text{(defining } \mathbf{c} = \mathbf{a}\times\mathbf{b}\text{)} \\
&= \mathbf{a}\bullet(\mathbf{b}\times\mathbf{c}) && \text{(interchanging the } \bullet \text{ and the } \times\text{)} \\
&= \mathbf{a}\bullet[\mathbf{b}\times(\mathbf{a}\times\mathbf{b})] \\
&= \mathbf{a}\bullet[(\mathbf{b}\bullet\mathbf{b})\mathbf{a} - (\mathbf{a}\bullet\mathbf{b})\mathbf{b}] && \text{(using the vector triple product identity)} \\
&= (\mathbf{a}\bullet\mathbf{a})\,(\mathbf{b}\bullet\mathbf{b}) - (\mathbf{a}\bullet\mathbf{b})\,(\mathbf{a}\bullet\mathbf{b}) \\
&= |\mathbf{a}|^2\,|\mathbf{b}|^2 - |\mathbf{a}\bullet\mathbf{b}|^2
\end{aligned}$$

Actually, in practice, it is rather easier to write,

$$(\mathbf{a}\times\mathbf{b})\bullet(\mathbf{a}\times\mathbf{b}) = |\mathbf{a}|^2\,|\mathbf{b}|^2\sin^2\theta = |\mathbf{a}|^2\,|\mathbf{b}|^2\,(1-\cos^2\theta) = |\mathbf{a}|^2\,|\mathbf{b}|^2 - |\mathbf{a}\bullet\mathbf{b}|^2 \quad (!)$$

3 VECTOR REPRESENTATION OF LINES AND PLANES

3.1 Straight lines in 3D



Two separate points define a straight line in 3D space. Let **a** and **b** be the position vectors of the two known points and let **r** be the position vector of a general point on the line. The vector equation of the line is then,

$$\mathbf{r} = \mathbf{a} + \lambda(\mathbf{b} - \mathbf{a})$$

where λ is a scalar parameter. When $\lambda = 0$, $\mathbf{r} = \mathbf{a}$, and when $\lambda = 1$, $\mathbf{r} = \mathbf{b}$. By choosing values of λ between $-\infty$ and $+\infty$, every point on the line is obtained. If **t** (for tangential) is a unit vector in the direction of the line, then the above equation can be written,

$$\mathbf{r} = \mathbf{a} + \mu\,\mathbf{t}$$

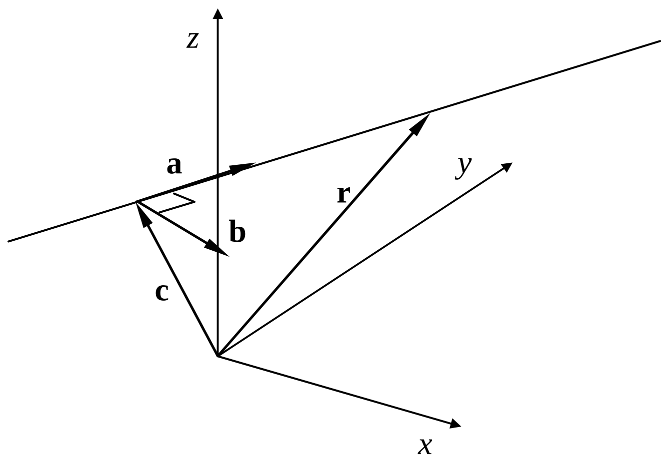
where μ is a different scalar parameter (unless $\mathbf{b}-\mathbf{a}$ happens to equal **t**). Note that lines can always be represented in terms of just one scalar parameter.

The vector equation for a line is really three scalar equations rolled into one. To find the scalar cartesian form, we decompose the vector equation by writing,

$$x = a_1 + \lambda(b_1 - a_1)$$
$$y = a_2 + \lambda(b_2 - a_2)$$
$$z = a_3 + \lambda(b_3 - a_1)$$

A more compact way of presenting these equations is,

$$\frac{x - a_1}{b_1 - a_1} = \frac{y - a_2}{b_2 - a_2} = \frac{z - a_3}{b_3 - a_3} \quad (= \lambda)$$



Another way of writing the vector equation of a line is,

$$\mathbf{a} \times \mathbf{r} = \mathbf{b}$$

where, as shown in the diagram, **a** is a vector tangential to the line and **b** is a particular vector normal to the line (which then defines the line). To convince yourself that the above equation really does represent a line, note that the position vector **c** of a point on the line must satisfy,

$$\mathbf{a} \times \mathbf{c} = \mathbf{b}$$

Equating the two expressions for **b** and using the distributive property of the vector product in reverse, we find,

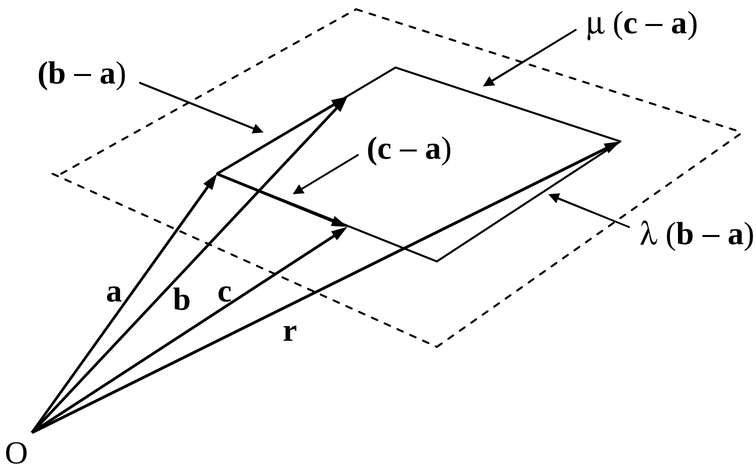
$$\mathbf{a} \times (\mathbf{r} - \mathbf{c}) = \mathbf{0}$$

Hence, (**r** – **c**) is a vector parallel to **a** and so we can write

$$\mathbf{r} - \mathbf{c} = \lambda\, \mathbf{a}$$

which is the equation of a line, in the direction of **a**, passing through the point defined by the position vector **c**.

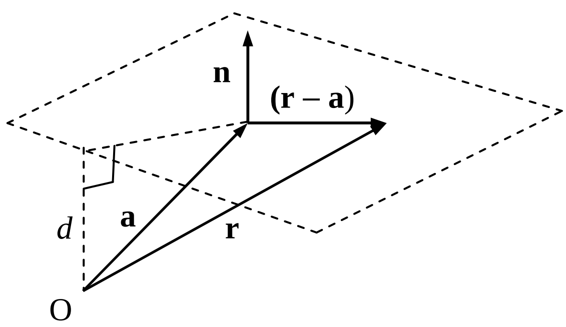
3.2 Planes



Three points define a plane so long as they do not lie on a single straight line. If **a**, **b** and **c** are the position vectors of the three known points, then it can be seen from the diagram that the position vector **r** of a general point on the plane is given by,

$$\mathbf{r} = \mathbf{a} + \lambda(\mathbf{b} - \mathbf{a}) + \mu(\mathbf{c} - \mathbf{a})$$

where λ and μ are scalar parameters that can take any values between $-\infty$ and $+\infty$. Note that planes can always be specified in terms of two scalar parameters.



An alternative form of the vector equation for a plane can be obtained in terms of a vector **n** normal to the plane. As can be seen from the diagram the equation is,

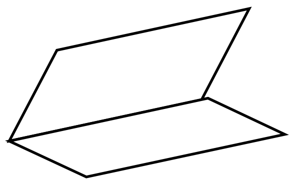
$$(\mathbf{r} - \mathbf{a}) \bullet \mathbf{n} = 0$$

Expanding in scalar form, we obtain the well-known cartesian equation for a plane,

$$n_1 x + n_2 y + n_3 z = \mathbf{a} \bullet \mathbf{n}$$

If **n** is a unit normal vector ($n_1^2 + n_2^2 + n_3^2 = 1$), it can be seen from the diagram that the scalar product $\mathbf{a} \bullet \mathbf{n} = d$ represents the perpendicular distance d of the origin O from the plane.

Two planes define a line unless they are parallel :



Suppose we have two planes defined by the equations $\mathbf{a} \bullet \mathbf{r} = p$ and $\mathbf{b} \bullet \mathbf{r} = q$. By comparison with the general equation, we see that **a** and **b** are vectors normal to the planes (not unit vectors, necessarily). The line of intersection will therefore be parallel to the vector $\mathbf{a}\times\mathbf{b}$ and can be written,

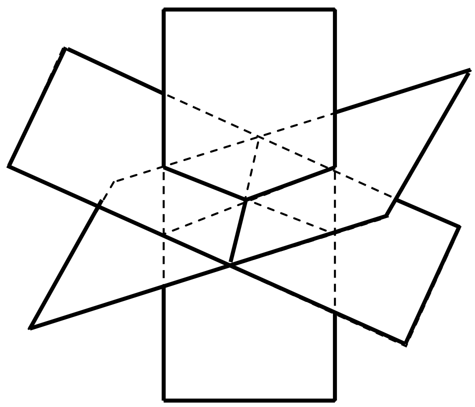
$$\mathbf{r} = \mathbf{c} + \lambda\,(\mathbf{a}\times\mathbf{b})$$

where **c** is the position vector of any point lying on the line and λ is a scalar parameter. Note that if the vector product $\mathbf{a}\times\mathbf{b} = \mathbf{0}$ the two normals are in the same direction so the planes are parallel and do not intersect. Furthermore, if,

$$\frac{\mathbf{a}}{p} = \frac{\mathbf{b}}{q}$$

the planes are one and the same; instead of a line, we have a plane of intersection.

Three planes define a point unless their normals are coplanar :

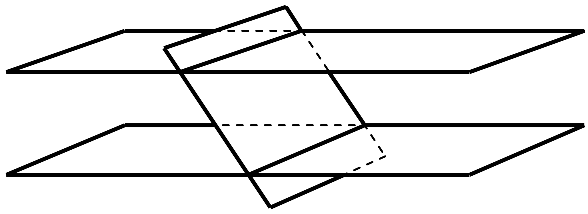


The point of intersection of three planes can be found by solving the three simultaneous equations, $\mathbf{a} \bullet \mathbf{r} = p$, $\mathbf{b} \bullet \mathbf{r} = q$ and $\mathbf{c} \bullet \mathbf{r} = s$. In cartesian component form, these are :

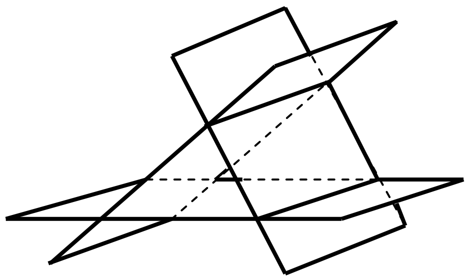
$$\begin{aligned} a_1 x + a_2 y + a_3 z &= p \\ b_1 x + b_2 y + b_3 z &= q \\ c_1 x + c_2 y + c_3 z &= s \end{aligned}$$

Geometrically, we expect a unique solution (a single point of intersection) unless :

(i) The planes are all parallel but separate, so there is no intersection, or
(ii) The planes are all the same, so there is a plane of intersection, or
(iii) The three planes intersect along a single line, so there is a line of intersection, or
(iv) Two planes are parallel, so there is no common intersection, see below, or



(v) We get a Toblerone situation, see below, so there is no common intersection.



In all these special cases, the normals to the three planes can be moved so that they are coplanar. Therefore, the volume of the parallelepiped they form is zero. Hence, we can recognise the special cases because $\mathbf{a} \bullet \mathbf{b} \times \mathbf{c} = 0$.

4 VECTORS AND MATRICES

4.1 Matrix-vector multiplication

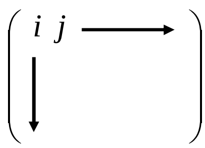
The most general way in which the components of a vector **a** can be transformed by linear operations into the components of a vector **b** is as follows :

$$\begin{aligned} b_1 &= A_{11}a_1 + A_{12}a_2 + A_{13}a_3 \\ b_2 &= A_{21}a_1 + A_{22}a_2 + A_{23}a_3 \\ b_3 &= A_{31}a_1 + A_{32}a_2 + A_{33}a_3 \end{aligned}$$

where A_{11}, A_{12} etc are constant coefficients. In setting up a shorthand notation to express these relationships, we define the *matrix* **A** as the array of coefficients A_{ij} :

$$\mathbf{A} = \begin{pmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{pmatrix}$$

A can also be written A_{ij} where i (= 1, 2, 3) enumerates the rows and j (= 1, 2, 3) the columns. An easy way of remembering how the subscripts change is provided by the mnemonic :



Writing the vectors **a** and **b** in column format (*i.e.*, as 3×1 matrices), we define matrix-vector multiplication so that,

$$\begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} = \begin{pmatrix} A_{11}a_1 + A_{12}a_2 + A_{13}a_3 \\ A_{21}a_1 + A_{22}a_2 + A_{23}a_3 \\ A_{31}a_1 + A_{32}a_2 + A_{33}a_3 \end{pmatrix} = \begin{pmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}$$

In order to satisfy these equalities, the i^{th} component of the column vector **b** must be obtained by multiplying the i^{th} row of the matrix **A** by the column vector **a** in the sense that,

$$b_i = \sum_{j=1}^{3} A_{ij}\, a_j = A_{i1}a_1 + A_{i2}a_2 + A_{i3}a_3$$

The matrix **A** is said to *operate* on the vector **a** to give the vector **b**. With the understanding that matrix-vector multiplication is implied, we can now write,

$$\mathbf{b} = \mathbf{Aa}$$

4.2 The transpose of a matrix

The *transpose* $\mathbf{A}^t$ of a matrix **A** is defined to be a matrix in which rows and columns are interchanged (row 1 of **A** becomes column 1 of $\mathbf{A}^t$ and so on). Thus,

$$\mathbf{A}^t = \begin{pmatrix} A_{11} & A_{21} & A_{31} \\ A_{12} & A_{22} & A_{32} \\ A_{13} & A_{23} & A_{33} \end{pmatrix} \qquad \text{or} \qquad A^t_{ij} = A_{ji}$$

A scalar can be thought of as a 1×1 matrix, so it transposes into itself, $\lambda^t = \lambda$. A vector **a** is usually thought of as a 3×1 matrix, although this is only a convention. If we need the components set out as a row, we simply write the vector as $\mathbf{a}^t$,

$$\mathbf{a}^t = (a_1 \quad a_2 \quad a_3)$$

This is the same vector but expressed in a different way.

4.3 The determinant of a matrix

Every <u>square</u> matrix **A** has a number associated with it called a determinant, written |**A**| or sometimes det **A**. Once again, this is done as a convenient shorthand notation.

The determinant of a 2×2 matrix $\mathbf{A} = \begin{pmatrix} a_1 & a_2 \\ b_1 & b_2 \end{pmatrix}$ is defined as,

$$|\mathbf{A}| = \begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix} = a_1 b_2 - a_2 b_1$$

The determinant of a 3×3 matrix $\mathbf{A} = \begin{pmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{pmatrix}$ is then given by,

$$|\mathbf{A}| = a_1 \begin{vmatrix} b_2 & b_3 \\ c_2 & c_3 \end{vmatrix} - a_2 \begin{vmatrix} b_1 & b_3 \\ c_1 & c_3 \end{vmatrix} + a_3 \begin{vmatrix} b_1 & b_2 \\ c_1 & c_2 \end{vmatrix}$$

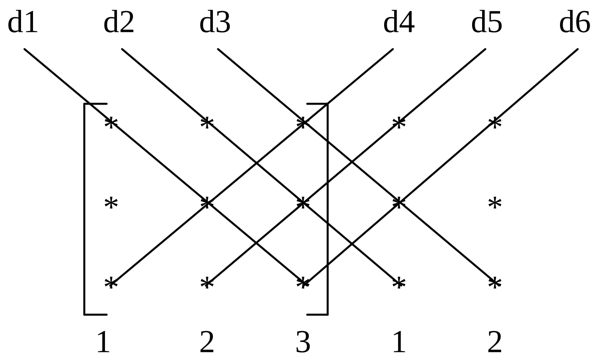
The 2×2 determinants are obtained by crossing out the first row and each column in turn of the original 3×3 matrix. They are called the *minors* of the 3×3 determinant. With the relevant sign in front (+, –, +) they are called the *cofactors*. The representation above is called 'expansion about the first row'. In fact, a determinant can be expanded about any row or column so long as proper attention is given to the signs.

When the 2×2 determinants are multiplied out, we find,

$$|\mathbf{A}| = a_1(b_2c_3 - b_3c_2) + a_2(b_3c_1 - b_1c_3) + a_3(b_1c_2 - b_2c_1)$$

$$= (a_1b_2c_3 + a_2b_3c_1 + a_3b_1c_2) - (a_1b_3c_2 + a_2b_1c_3 + a_3b_2c_1)$$

Note the right-handed cyclic order in the first bracket and the left-handed cyclic order in the second bracket. An easy way of obtaining this expression is to write out the matrix with the first and second columns repeated :



The value of the determinant is then the sum of the three products of the three numbers down the diagonals, d1, d2, d3, minus the sum of the three products of the three numbers down the diagonals d4, d5, d6.

4.4 Scalar and vector products as matrix operations

In order to express the scalar product as the operation of a matrix on a vector, we transpose the first vector into a 1×3 matrix. Thus,

$$\mathbf{a}^t\mathbf{b} = \begin{pmatrix} a_1 & a_2 & a_3 \end{pmatrix}\begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} = a_1b_1 + a_2b_2 + a_3b_3 = \mathbf{a} \bullet \mathbf{b}$$

Note that,

$$\mathbf{a}^t\mathbf{b} = \mathbf{a} \bullet \mathbf{b} = \mathbf{b} \bullet \mathbf{a} = \mathbf{b}^t\mathbf{a}$$

Expressing the vector product as the operation of a matrix on a vector is rather more complicated. Thus,

$$\begin{pmatrix} 0 & -a_3 & a_2 \\ a_3 & 0 & -a_1 \\ -a_2 & a_1 & 0 \end{pmatrix}\begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} = \begin{pmatrix} a_2b_3 - a_3b_2 \\ a_3b_1 - a_1b_3 \\ a_1b_2 - a_2b_1 \end{pmatrix} = \mathbf{a} \times \mathbf{b}$$

Note that the determinant of this matrix is zero so the matrix has no inverse. Hence, given the vector **a** and the vector product **a**×**b** we cannot solve for **b**. One consequence of this is that if **a**×**b** = **a**×**c**, we cannot assume that **b** = **c**. A similar situation exists for scalar products. If **a** • **b** and **a** • **c** are equal, we cannot assume that **b** = **c**.

As mentioned earlier, the vector product **a**×**b** can also be expressed as a determinant :

$$\mathbf{a}\times\mathbf{b} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix} = \mathbf{i}\begin{vmatrix} a_2 & a_3 \\ b_2 & b_3 \end{vmatrix} - \mathbf{j}\begin{vmatrix} a_1 & a_3 \\ b_1 & b_3 \end{vmatrix} + \mathbf{k}\begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix} = \begin{pmatrix} a_2b_3 - a_3b_2 \\ a_3b_1 - a_1b_3 \\ a_1b_2 - a_2b_1 \end{pmatrix}$$

This leads us on to prove that the scalar triple product can also be expressed as a determinant :

$$\mathbf{a}\bullet\mathbf{b}\times\mathbf{c} = (a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k})\bullet\begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix}$$

$$= (a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k})\bullet\left(\mathbf{i}\begin{vmatrix} b_2 & b_3 \\ c_2 & c_3 \end{vmatrix} - \mathbf{j}\begin{vmatrix} b_1 & b_3 \\ c_1 & c_3 \end{vmatrix} + \mathbf{k}\begin{vmatrix} b_1 & b_2 \\ c_1 & c_2 \end{vmatrix}\right)$$

$$= \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix}$$

It has already been shown that **a** • **b** × **c** is the volume of the parallelepiped formed by the vectors **a**, **b** and **c**, and that the sign is changed if we swap **b** and **c**. From this we deduce the rule for determinants that if two rows (or two columns) are swapped, the sign of the determinant changes. On the other hand, a cyclic permutation of rows and columns leaves the value of a determinant unchanged.

4.5 Simultaneous equations and the inverse of a matrix

So far, the matrix has been presented as a way of representing three linear equations relating each component of a vector **b** to the components of another vector **a**. More generally, however, any set of three linear equations can be written in matrix notation. For example, the equations describing the intersection of three planes (**a** • **r** = p, **b** • **r** = q and **c** • **r** = s) are,

$$\begin{aligned} a_1x + a_2y + a_3z &= p \\ b_1x + b_2y + b_3z &= q \\ c_1x + c_2y + c_3z &= s \end{aligned}$$

In matrix form these are written,

$$\begin{pmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} p \\ q \\ s \end{pmatrix}$$

Formally, we can obtain the position vector of the point of intersection $(x \ \ y \ \ z)^t$ by finding the inverse of the 3×3 matrix and using it to operate on the vector $(p \ \ q \ \ s)^t$. Thus,

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{pmatrix}^{-1} \begin{pmatrix} p \\ q \\ s \end{pmatrix}$$

This procedure assumes that the matrix really does have an inverse. However, if the normals to the three planes are coplanar, the volume of the parallelepiped formed is zero, $\mathbf{a} \bullet \mathbf{b} \times \mathbf{c} = 0$. But $\mathbf{a} \bullet \mathbf{b} \times \mathbf{c}$ also equals the determinant of the matrix and so we deduce that if the determinant of a matrix is zero the inverse does not exist.

It is important to note that the equation involving the inverse represents the formal solution to the original matrix equation. In practice, one never actually solves a set of simultaneous equations by calculating the inverse of the matrix, particularly on a computer. Rather one uses a much more efficient numerical technique such as *Gaussian elimination*.

As an example, we check to see if the following equations have a unique solution :

$$\begin{aligned} x_1 + 3x_2 + 19x_3 &= 1 \\ x_1 \qquad\quad - x_3 &= 3 \\ 4x_1 + 3x_2 + 16x_3 &= 2 \end{aligned}$$

The equations can be written,

$$\begin{pmatrix} 1 & 3 & 19 \\ 1 & 0 & -1 \\ 4 & 3 & 16 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 3 \\ 2 \end{pmatrix}$$

Using the diagonal multiplication trick for calculating the determinant of the matrix, we find,

$$\begin{vmatrix} 1 & 3 & 19 \\ 1 & 0 & -1 \\ 4 & 3 & 16 \end{vmatrix} = [1\times0\times16 + 3\times(-1)\times4 + 19\times1\times3] - [19\times0\times4 + 1\times(-1)\times3 + 3\times1\times16] = 0$$

So the inverse of the matrix does not exist and there is no unique solution to the equations.