Foundations of Mathematical Physics: Vectors, Tensors and Fields

2007 – 2008

John Peacock

www.roe.ac.uk/japwww/teaching/vtf.html

Textbooks The standard recommended text for this course (and later years) is Riley, Hobson & Bence *Mathematical Methods for Physics and Engineering* (Cambridge). A slightly more sophisticated approach, which can often be clearer once you know what you are doing, is taken by Arfken & Weber *Mathematical Methods for Physicists* (Academic Press).

1 Vectors

1.1 The meaning of vectors

Because we inhabit a world with more than one spatial dimension, physical phenomena frequently require us to distinguish between

Scalar : a quantity specified by a single number;

Vector : a quantity specified by a number (magnitude) and a **direction**;

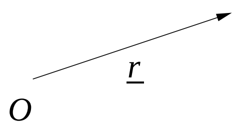
e.g. speed is a scalar, velocity is a vector. Vector algebra is an essential physics tool for describing vector quantities in a compact fashion. Modern notation is not that old: it was invented in the 1880s by Gibbs and by Heaviside. Earlier physicists from Newton to Maxwell had to work much harder to solve their problems.

Notation: Textbooks often denote vectors by boldface: $\boldsymbol{A}$, or occasionally the arrow notation: $\vec{A}$. But for writing vectors, the easiest notation is the underline: $\underline{A}$. Denote a vector by $\underline{A}$ and its magnitude by $|\underline{A}|$ or A. *Always* underline a vector to distinguish it from its magnitude. A unit vector is often denoted by a hat $\underline{\hat{A}} = \underline{A}\ /\ A$ and represents a direction.

The main intention of this course is to develop skill in using vector methods to solve problems in physics. As such, it deliberately repeats some material that has been seen before (in MM3 and AM3). The approach will be relatively informal; but this is no excuse for lack of rigour. It is important to be able to derive the key results in the subject.

1.1.1 Geometrical view: position vectors

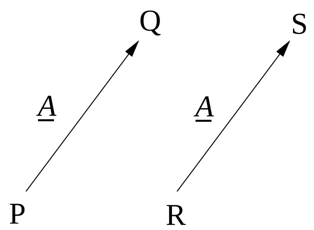
A vector is fundamentally a geometrical object, as can be seen by starting with the most basic example, the **position vector**. This is drawn as a line between an origin and a given point, with an arrow showing the direction. It is often convenient to picture this vector in a concrete way, as a thin rod carrying a physical arrowhead.



The position vector of a point relative to an origin O is normally written $\underline{r}$, which has length r (the radius of the point from the origin) and points along the unit vector $\underline{\hat{r}}$.

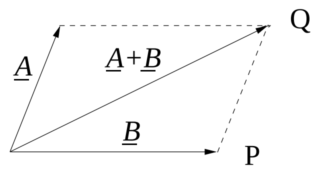
Formally speaking, this 'directed line segment' is merely a **representation** of the more abstract idea of a vector, and different kinds of vectors can be represented by a position vector: *e.g.* for a velocity vector we would draw a position vector pointing in the same direction as the velocity, and set the length proportional to the speed. This geometrical viewpoint suffices to demonstrate some of the basic properties of vectors:

Independence of origin



Vectors are unchanged by being **transported**: as drawn, both displacements from P to Q and from R to S represent the same vector. In effect, both are position vectors, but with P and R treated as the origin: the choice of origin is arbitrary.

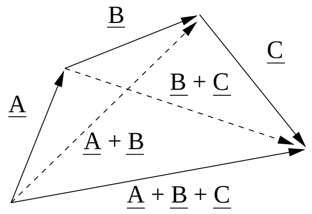
Addition of vectors: parallelogram law



$$\underline{A} + \underline{B} = \underline{B} + \underline{A} \text{ (\textbf{commutative})}.$$

From this, we see that the vector $\underline{A}$ that points from P to Q is just the position vector of the last point minus that of the first: we write this as $\underline{PQ} = \underline{OQ} - \underline{OP} = \underline{r}_Q - \underline{r}_P$. We prove this by treating $\underline{A}$ and $\underline{A} + \underline{B}$ as the two position vectors in the above diagram.

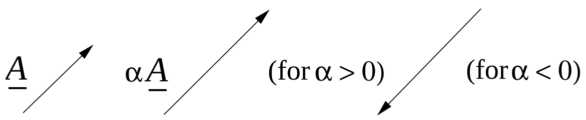
This generalises to any number of vectors: the **resultant** is obtained by adding the vectors nose to tail. This lets us prove that vector addition is **associative**:



A geometrical demonstration that $(\underline{A} + \underline{B}) + \underline{C} = \underline{A} + (\underline{B} + \underline{C})$.

Multiplication by scalars

A vector may be multiplied by a scalar to give a new vector *e.g.*



Also

$$\begin{aligned} |\alpha \underline{A}| &= |\alpha||\underline{A}| & \\ \alpha(\underline{A} + \underline{B}) &= \alpha \underline{A} + \alpha \underline{B} & \text{(distributive)} \\ (\alpha + \beta)\underline{A} &= \alpha \underline{A} + \beta \underline{A} & \text{(distributive)} \\ \alpha(\beta \underline{A}) &= (\alpha\beta)\underline{A} & \text{(associative).} \end{aligned}$$

In summary, as far as addition of vectors is concerned, or of multiplication by scalars, the power of vector notation is just that you treat vectors as if they were just a number (a 'directed number'). The important exception of multiplication of vectors will be dealt with shortly. In the meantime, there are already some common mistakes to avoid:

1. You can add vectors, but you can't add vectors and scalars.
2. Check that all quantities in a vector equation are of the same type: e.g. any equation *vector* = *scalar* is clearly wrong. (The only exception to this is if we lazily write *vector* = 0 when we mean $\underline{0}$.)
3. Never try to divide by a vector – there is no such operation.

1.1.2 Coordinate geometry

Although the geometrical view of vectors is fundamental, in practice it is often easier to convert vectors to a set of numbers: this is the approach to geometry pioneered by Descartes in 1637 (hence Cartesian coordinates). Now, a position vector is represented by either a row or column of numbers (**row vector** or **column vector**):

$$\underline{r} = (x, y, z) \quad \text{or} \quad \begin{pmatrix} x \\ y \\ z \end{pmatrix},$$

assuming three dimensions for now. These numbers are the **components** of the vector. When dealing with matrices, we will normally assume the column vector to be the primary form – but otherwise it is most convenient to use row vectors.

It should be clear that this xyz triplet is just a representation of the vector. But we will commonly talk as if (x, y, z) *is* the vector itself. The coordinate representation makes it easy to prove all the results considered above: to add two vectors, we just have to add the coordinates. For example, associativity of vector addition then follows just because addition of numbers is associative.

Basis vectors

A more explicit way of writing a Cartesian vector is to introduce basis vectors denoted by either $\underline{i}$, $\underline{j}$ and $\underline{k}$ or $\underline{e}_x$, $\underline{e}_y$ and $\underline{e}_z$ which point along the x, y and z-axes. These basis vectors are **orthonormal**: *i.e.* they are all unit vectors that are mutually perpendicular. The $\underline{e}_z$ vector is related to $\underline{e}_x$ and $\underline{e}_y$ by the r.h. screw rule.

The key idea of basis vectors is that any vector can be written as a **linear superposition** of different multiples of the basis vectors. If the components of a vector $\underline{A}$ are A_x, A_y, A_z, then we write

$$\underline{A} \;=\; A_x\,\underline{i} + A_y\,\underline{j} + A_z\,\underline{k} \quad \text{or} \quad \underline{A} = A_x\,\underline{e}_x + A_y\,\underline{e}_y + A_z\,\underline{e}_z\,.$$

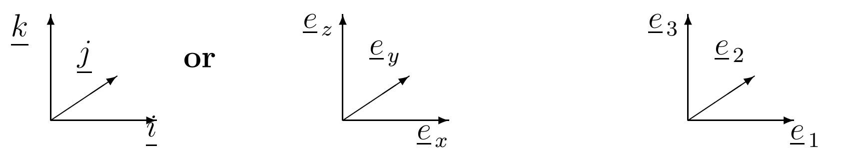
In row-vector notation, the basis vectors themselves are just

$$\underline{i} = \underline{e}_x = (1, 0, 0) \quad \underline{j} = \underline{e}_y = (0, 1, 0) \quad \underline{k} = \underline{e}_z = (0, 0, 1)$$

1.1.3 Suffix or Index notation

A more systematic labelling of basis vectors is by $\underline{e}_1$, $\underline{e}_2$ and $\underline{e}_3$. i.e. instead of $\underline{i}$ we write $\underline{e}_1$, instead of $\underline{j}$ we write $\underline{e}_2$, instead of $\underline{k}$ we write $\underline{e}_3$. This scheme is known as the **suffix** notation. Its great advantages are that it generalises easily to any number of dimensions and greatly simplifies manipulations and the verification of various identities (see later in the course).

Old Notation **New Notation**



$\underline{r} = x\underline{i} + y\underline{j} + z\underline{k}$ $\quad$ $\underline{r} = x\underline{e}_x + y\underline{e}_y + z\underline{e}_z$ $\quad$ $\underline{r} = x_1\underline{e}_1 + x_2\underline{e}_2 + x_3\underline{e}_3$

Thus any vector $\underline{A}$ in N dimensions can be written in this new notation as

$$\underline{A} = A_1\, \underline{e}_1 + A_2\, \underline{e}_2 + A_3\, \underline{e}_3 + \cdots = \sum_{i=1}^{N} A_i\, \underline{e}_i\ .$$

Free and dummy indices

We have written the components of $\underline{A}$ as A_i. This can be a common source of confusion: if we had written A_j instead, does that make any difference? In words, A_i means "the i^{th} component of $\underline{A}$". Thus, i is a **free index**: it means some integer that we have not yet specified, and indeed we might as well have called it j. The only important thing is to be consistent: if vectors $\underline{A}$ and $\underline{B}$ are equal, then all their individual components are equal, so we can write $A_i = B_i$ and get a relation that is implicitly true for any value i without having to write them all out. The relation $A_j = B_j$ expresses exactly the same result and is just a good a choice. But $A_i = B_j$ would be meaningless, because the indices don't balance. Where the value of an index is summed over, as in $\underline{A} = \sum_{i=1}^{N} A_i\, \underline{e}_i$, all possible values of i are used, and it is called a **dummy index**. Again, we can happily replace i by j or whatever, provided it is done consistently:

$$\boxed{\underline{A} = \sum_{i=1}^{3} A_i\, \underline{e}_i = \sum_{j=1}^{3} A_j\, \underline{e}_j.}$$

As a more complicated example, consider the vector equation

$$\underline{a} - (\underline{b} \cdot \underline{c})\, \underline{d} + 3\underline{n} = 0,$$

which must hold for each component separately:

$$a_i - (\underline{b} \cdot \underline{c})\, d_i + 3n_i = 0 \quad \text{for} \quad i = 1, 2, 3.$$

The free index i occurs **once and only once** in each term of the equation. In general every term in the equation must be of the same kind *i.e.* have the same free indices. In order to write the scalar product that appears in the second term in suffix notation, we should be careful to avoid using i as a summation index, since as we have already used it as a free index. It is better to write

$$a_i - \left(\sum_{k=1}^{3} b_k c_k \right) d_i + 3n_i = 0 \quad \text{for} \quad i = 1, 2, 3$$

rather than

$$a_i - \left(\sum_{i=1}^{3} b_i c_i \right) d_i + 3n_i = 0 \quad \text{for} \quad i = 1, 2, 3$$

which would be ambiguous if the brackets were removed – whereas in the first form the brackets are not essential.

Example: epicycles

To illustrate the idea of shift of origin, consider the position vectors of two planets 1 & 2 (Earth and Mars, say) on circular orbits: $\underline{r}_1 = r_1(\cos\omega_1 t, \sin\omega_1 t)$, $\underline{r}_2 = r_2(\cos\omega_2 t, \sin\omega_2 t)$. The position vector of Mars as seen from Earth is

$$\underline{r}_{21} = \underline{r}_2 - \underline{r}_1 = (r_2\cos\omega_2 t - r_1\cos\omega_1 t,\ r_2\sin\omega_2 t - r_1\sin\omega_1 t).$$

This describes an *epicycle*: Mars moves on a small circle of radius r_1 whose centre is carried round in a circle of radius r_2. Although epicycles have a bad press, notice that this is an exact description of Mars's motion, using the same number of free parameters as the heliocentric view. Fortunately, planetary orbits are not circles, otherwise the debate over whether the Sun or the Earth made the better origin might have continued much longer.

1.1.4 Vector physics: independence of basis

Although the coordinate approach is convenient and practical, expressing vectors as components with respect to a basis is against the spirit of the power of vectors as a tool for physics – which is that *physical laws relating vectors must be true independent of the coordinate system being used*. Consider the case of **vector dynamics**:

$$\underline{F} = m\underline{a} = m\,\frac{d}{dt}\,\underline{v} = m\,\frac{d^2}{dt^2}\,\underline{r}.$$

In one compact statement, this equation says that $F = ma$ is obeyed separately by all the components of F and a. The simplest case is where one of the basis vectors points in the direction of $\underline{F}$, in which case there is only one scalar equation to consider. But the vector equation is true whatever basis we use. We will return to this point later when we consider how the components of vectors alter when the basis used to describe them is changed.

Example: centripetal force

As an example of this point in action, consider again circular motion in 2D: $\underline{r} = r(\cos\omega t, \sin\omega t)$. What force is needed to produce this motion? We get the acceleration by differentiating twice w.r.t. t:

$$\underline{F} = m\underline{a} = m\,\frac{d^2}{dt^2}\underline{r} = m\,r(-\omega^2\sin\omega t, -\omega^2\cos\omega t) = -m\omega^2\,\underline{r}.$$

Although we have used an explicit set of components as an intermediate step, the final result just says that the required force is $m\omega^2 r$, directed radially inwards.

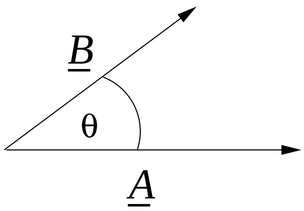
1.2 Multiplying vectors

So far, vector notation is completely pain-free: we just treat vectors as if they were numbers and the algebra of addition or subtraction is identical. What about multiplication? What could $\underline{A}\ \underline{B}$ mean? To see this, we have to think geometrically, and there are two aspects that resemble multiplication: the projection of one vector onto another, and the area of the parallelogram formed by two vectors.

1.2.1 Scalar or dot product

The scalar product (also known as the dot product) between two vectors is defined as

$$\boxed{(\underline{A}\cdot\underline{B}) \equiv AB\cos\theta, \text{ where } \theta \text{ is the angle between } \underline{A} \text{ and } \underline{B}}$$

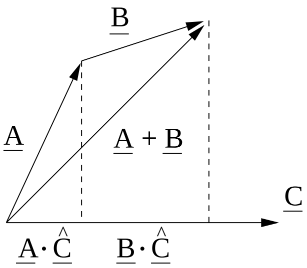


$(\underline{A}\cdot\underline{B})$ is a scalar — *i.e.* a single number. By definition, the scalar product is **commutative**: $(\underline{A}\cdot\underline{B}) = (\underline{B}\cdot\underline{A})$.

.

The geometrical significance of the scalar product is that it **projects** one vector onto another: $\underline{A}\cdot\hat{\underline{B}}$ is the component of A in the direction of the unit vector $\hat{\underline{B}}$, and its magnitude is $A\cos\theta$. This viewpoint makes it easy to prove that the scalar product is

Distributive over addition $\boxed{(\underline{A}+\underline{B})\cdot\underline{C} = \underline{A}\cdot\underline{C} + \underline{B}\cdot\underline{C}}$.



The components of $\underline{A}$ and $\underline{B}$ along $\underline{C}$ clearly add to make the components of $\underline{A}+\underline{B}$ along $\underline{C}$.

Scalar product in terms of components

You know very well that the scalar product is worked out in practice using the components of the vector:

$$\boxed{\underline{A}\cdot\underline{B} = \sum_{i=1}^{3} A_i\, B_i\ ;}$$

Let's prove that these two definitions are identical; this requires the distributive property of the scalar product. If $\underline{A} = \sum_i A_i \underline{e}_i$, then

$$\underline{A}\cdot\underline{e}_1 \;=\; (A_1\underline{e}_1 + A_2\underline{e}_2 + A_3\underline{e}_3)\cdot\underline{e}_1 = A_1,$$

so the orthonormality of the basis vectors picks out the projection of $\underline{A}$ in the direction of $\underline{e}_1$, and similarly for the components A_2 and A_3. In general we may write

$$\boxed{\underline{A}\cdot\underline{e}_i = \underline{e}_i\cdot\underline{A} \equiv A_i \quad \text{or sometimes} \quad (\underline{A})_i.}$$

If we now write $\underline{B} = \sum_i B_i\underline{e}_i$, then $\underline{A}\cdot\underline{B}$ amounts to picking out in turn A_i and multiplying it by B_i. So we recover the standard formula for the scalar product based on (i) distributivity; (ii) orthonormality of the basis.

Summary of properties of scalar product

(i) $\underline{A} \cdot \underline{B} = \underline{B} \cdot \underline{A}$; $\underline{A} \cdot (\underline{B} + \underline{C}) = \underline{A} \cdot \underline{B} + \underline{A} \cdot \underline{C}$

(ii) $\underline{\hat{n}} \cdot \underline{A}$ = the scalar projection of $\underline{A}$ onto $\underline{\hat{n}}$, where $\underline{\hat{n}}$ is a unit vector

(iii) $(\underline{\hat{n}} \cdot \underline{A})\, \underline{\hat{n}}$ = the vector projection of $\underline{A}$ onto $\underline{\hat{n}}$

(*iv*) $\boxed{\underline{A} \cdot \underline{A} = |\underline{A}|^2}$ which defines the magnitude of a vector. For a unit vector $\underline{\hat{A}} \cdot \underline{\hat{A}} = 1$

Example: parallel and perpendicular components

A vector may be resolved with respect to some direction $\underline{\hat{n}}$ into a parallel component $\underline{A}_{\parallel} = (\underline{\hat{n}} \cdot \underline{A})\underline{\hat{n}}$. There must therefore be a perpendicular component $\underline{A}_{\perp} = \underline{A} - \underline{A}_{\parallel}$. If this reasoning makes sense, we should find that $\underline{A}_{\parallel}$ and $\underline{A}_{\perp}$ are at right angles. To prove this, evaluate

$$\underline{A}_{\perp} \cdot \underline{\hat{n}} = (\underline{A} - (\underline{\hat{n}} \cdot \underline{A})\underline{\hat{n}}) \cdot \underline{\hat{n}} = \underline{A} \cdot \underline{\hat{n}} - \underline{\hat{n}} \cdot \underline{A} = 0$$

(because $\underline{\hat{n}} \cdot \underline{\hat{n}} = 1$).

Example: the cosine rule

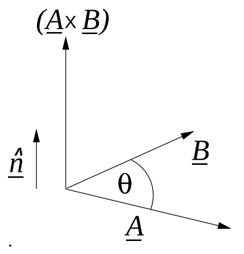
Consider two position vectors $\underline{A}$ and $\underline{B}$. They define a triangle, whose third side is the vector $\underline{C} = \underline{A} - \underline{B}$. $C^2 = (\underline{A} - \underline{B})^2 = (\underline{A} - \underline{B}) \cdot (\underline{A} - \underline{B}) = A^2 + B^2 - 2\underline{A} \cdot \underline{B}$. Hence we have a simple derivation of $C^2 = A^2 + B^2 - 2AB\cos\theta$, where θ is the angle between $\underline{A}$ and $\underline{B}$.

1.2.2 The vector or 'cross' product

The vector product represents the fact that two vectors define a parallelogram. This geometrical object has an area, but also an orientation in space – which can be represented by a vector.

$$\boxed{(\underline{A} \times \underline{B}) \;\equiv\; AB\sin\theta\, \underline{\hat{n}}\text{ , where } \underline{\hat{n}} \text{ in the 'right-hand screw direction'}}$$

i.e. $\underline{\hat{n}}$ is a unit vector normal to the plane of $\underline{A}$ and $\underline{B}$, in the direction of a right-handed screw for rotation of $\underline{A}$ to $\underline{B}$ (through $< \pi$ radians).

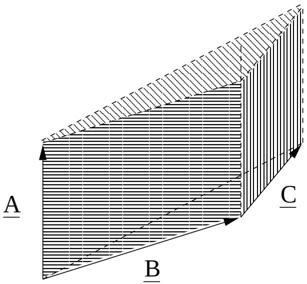


$(\underline{A} \times \underline{B})$ is a vector — i.e. it has a direction and a length.

It is important to note that the idea of the vector product is unique to three dimensions. In 2D, the area defined by two vectors is just a scalar: there is no choice about the orientation. In N dimensions, it turns out that $N(N-1)/2$ numbers are needed to specify the size and orientation of an element of area. So representing this by a vector is only possible for $N = 3$. The idea of such an 'oriented area' always exists, of course, and the general name for it is the **wedge product**, denoted by $\underline{A} \wedge \underline{B}$. You can feel free to use this notation instead of $\underline{A} \times \underline{B}$, since they are the same thing in 3D.

The vector product shares a critical property with the scalar product, but unfortunately one that is not as simple to prove; it is

Distributive over addition $\boxed{\underline{A} \times (\underline{B} + \underline{C}) = \underline{A} \times \underline{B} + \underline{A} \times \underline{C}}$.



It is easy enough to see that $\underline{A} \times (\underline{B} + \underline{C}) = \underline{A} \times \underline{B} + \underline{A} \times \underline{C}$ if all three vectors lie in the same plane. The various parallelograms of interest (shaded in this figure) differ by the triangles at top and bottom, which are clearly of identical shape.

When the three vectors are not in the same plane, the proof is more involved. What we shall do for now is to *assume* that the result is true in general, and see where it takes us. We can then work backwards at the end.

1.2.3 The vector product in terms of components

Because of the distributive property, when we write $\underline{A} \times \underline{B}$ in terms of components, the expression comes down to a sum of products of the basis vectors with each other:

$$\underline{A} \times \underline{B} = \left(\sum_{i=1}^{3} A_i \, \underline{e}_i \right) \times \left(\sum_{j=1}^{3} B_j \, \underline{e}_j \right) = \sum_{i=1}^{3} \sum_{j=1}^{3} A_i \, B_j \, (\underline{e}_i \times \underline{e}_j) \, .$$

Almost all the cross products of basis vectors vanish. The only one we need is the one that defines the z axis:

$$\boxed{\underline{e}_1 \times \underline{e}_2 = \underline{e}_3,}$$

and cyclic permutations of this. If the order is reversed, so is the sign: $\underline{e}_2 \times \underline{e}_1 = -\underline{e}_3$. In this way, we get

$$\underline{A} \times \underline{B} = \underline{e}_1(A_2 B_3 - A_3 B_2) + \underline{e}_2(A_3 B_1 - A_1 B_3) + \underline{e}_3(A_1 B_2 - A_2 B_1)$$

from which we deduce that

$$(\underline{A} \times \underline{B})_1 \; = \; (A_2 B_3 - A_3 B_2) \; , \text{ etc.}$$

So finally, we have recovered the familiar expression in which the vector product is written as a determinant:

$$\boxed{\underline{A} \times \underline{B} = \begin{vmatrix} \underline{e}_1 & \underline{e}_2 & \underline{e}_3 \\ A_1 & A_2 & A_3 \\ B_1 & B_2 & B_3 \end{vmatrix}}.$$

If we were to take this as the *definition* of the vector product, it is easy to see that the distributive property is obeyed. But now, how do we know that the geometrical properties of $\underline{A}\times\underline{B}$ are satisfied? One way of closing this loop is to derive the determinant expression in another way. If $\underline{A}\times\underline{B} = \underline{C}$, then $\underline{C}$ must be perpendicular to both vectors: $\underline{A}\cdot\underline{C} = \underline{B}\cdot\underline{C} = 0$. With some effort, these simultaneous equations can be solved to find the components of $\underline{C}$ (within some arbitrary scaling factor, since there are only two equations for three components). Or we can start by assuming the determinant and show that $\underline{A}\times\underline{B}$ is perpendicular to $\underline{A}$ and $\underline{B}$ and has magnitude $AB\sin\theta$. Again, this is quite a bit of algebra. For the present, we content ourselves with checking a simple test case, where we choose $\underline{A} = (A, 0, 0)$ along the x axis, and $\underline{B} = (B_1, B_2, 0)$ in the xy plane. This gives $\underline{A}\times\underline{B} = (0, 0, AB_2)$, which points in the z direction as required. From the scalar product, $\cos\theta = AB_1/AB = B_1/B$ (where $B = \sqrt{B_1^2 + B_2^2}$), so $\sin\theta = \sqrt{1-\cos^2\theta} = \sqrt{1 - B_1^2/B^2} = B_2/B$. Hence $AB_2 = AB\sin\theta$, as required.

Summary of properties of vector product

(i) $\underline{A}\times\underline{B} = -\underline{B}\times\underline{A}$

(ii) $\underline{A}\times\underline{B} = 0$ if $\underline{A}, \underline{B}$ are parallel

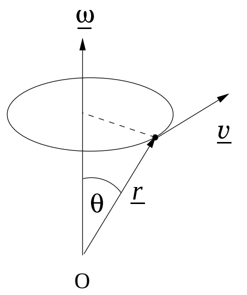
(iii) $\underline{A}\times(\underline{B}+\underline{C}) = \underline{A}\times\underline{B} + \underline{A}\times\underline{C}$

(iv) $\underline{A}\times(\alpha\underline{B}) = \alpha\underline{A}\times\underline{B}$

1.2.4 The vector product in Physics

(i) Angular velocity

Consider a point in a rigid body rotating with **angular velocity** $\underline{\omega}$: $|\underline{\omega}|$ is the angular speed of rotation measured in radians per second and $\underline{\hat{\omega}}$ lies along the axis of rotation. Let the position vector of the point with respect to an origin O on the axis of rotation be $\underline{r}$.



You should convince yourself that $\underline{v} = \underline{\omega}\times\underline{r}$ by checking that this gives the right direction for $\underline{v}$; that it is perpendicular to the plane of $\underline{\omega}$ and $\underline{r}$; that the magnitude $|\underline{v}| = \omega r\sin\theta = \omega\times$ radius of circle in which the point is travelling

(ii) Angular momentum

Now consider the **angular momentum** of the particle defined by $\underline{L} = \underline{r}\times(m\underline{v})$ where m is the mass of the particle.

Using the above expression for $\underline{v}$ we obtain

$$\underline{L} = m\underline{r}\times(\underline{\omega}\times\underline{r}) = m\left[\underline{\omega}r^2 - \underline{r}(\underline{r}\cdot\underline{\omega})\right]$$

where we have used the identity for the vector triple product. Note that $\underline{L} = 0$ if $\underline{\omega}$ and $\underline{r}$ are parallel. Note also that only if $\underline{r}$ is perpendicular to $\underline{\omega}$ do we obtain $\underline{L} = m\underline{\omega}r^2$, which means that only then are $\underline{L}$ and $\underline{\omega}$ in the same direction.

(iii) Torque

This last result sounds peculiar, but makes a good example of how physical laws are independent of coordinates. The **torque** or **moment** of a force about the origin $\underline{G} = \underline{r} \times \underline{F}$ where $\underline{r}$ is the position vector of the point where the force is acting and $\underline{F}$ is the force vector through that point. Torque causes angular momentum to change:

$$\frac{d}{dt}\underline{L} = \dot{\underline{r}} \times m\underline{v} + \underline{r} \times m\underline{a} = 0 + \underline{r} \times \underline{F} = \underline{G}.$$

If the origin is in the centre of the circle, then the centripetal force generates zero torque – but otherwise G is nonzero. This means that $\underline{L}$ has to change with time. Here, we have assumed that the vector product obeys the product rule under differentiation. This will be easy to prove a little later.

1.2.5 The scalar triple product

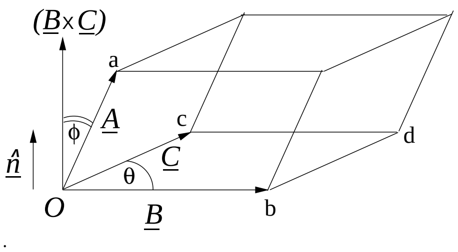
By introducing a third vector, we extend the geometrical idea of an area to the volume of the parallelipiped. The scalar triple product is defined as follows

$$(\underline{A}, \underline{B}, \underline{C}) \equiv \underline{A} \cdot (\underline{B} \times \underline{C})$$

Notes

(i) If $\underline{A}$, $\underline{B}$ and $\underline{C}$ are three concurrent edges of a parallelepiped, the volume is $(\underline{A}, \underline{B}, \underline{C})$.

To see this, note that:



$$\begin{aligned}
\text{area of the base} &= \text{area of parallelogram } Obdc \\
&= B\,C\,\sin\theta = |\underline{B} \times \underline{C}| \\
\text{height} &= A\cos\phi = \hat{\underline{n}} \cdot \underline{A} \\
\text{volume} &= \text{area of base} \times \text{height} \\
&= B\,C\,\sin\theta\, \hat{\underline{n}} \cdot \underline{A} \\
&= \underline{A} \cdot (\underline{B} \times \underline{C})
\end{aligned}$$

(ii) If we choose $\underline{C}, \underline{A}$ to define the base then a similar calculation gives volume $= \underline{B} \cdot (\underline{C} \times \underline{A})$
We deduce the following symmetry/antisymmetry properties:

$$(\underline{A}, \underline{B}, \underline{C}) = (\underline{B}, \underline{C}, \underline{A}) = (\underline{C}, \underline{A}, \underline{B}) = -(\underline{A}, \underline{C}, \underline{B}) = -(\underline{B}, \underline{A}, \underline{C}) = -(\underline{C}, \underline{B}, \underline{A})$$

(iii) If $\underline{A}, \underline{B}$ and $\underline{C}$ are **coplanar** (*i.e.* all three vectors lie in the same plane) then $V = (\underline{A}, \underline{B}, \underline{C}) = 0$, and vice-versa.

It is easy to write down an expression for the scalar triple product in terms of components:

$$\begin{aligned}
\underline{A}\cdot(\underline{B}\times\underline{C}) &= \sum_{i=1}^{3} A_i(\underline{B}\times\underline{C})_i \\
&= A_1(B_2C_3 - C_2B_3) - A_2(B_1C_3 - C_1B_3) + A_3(B_1C_2 - C_1B_2) \\
&= \begin{vmatrix} A_1 & A_2 & A_3 \\ B_1 & B_2 & B_3 \\ C_1 & C_2 & C_3 \end{vmatrix}.
\end{aligned}$$

The symmetry properties of the scalar triple product may be deduced from this by noting that interchanging two rows (or columns) changes the value by a factor -1.

1.2.6 The vector triple product

There are *several* ways of combining 3 vectors to form a new vector.
e.g. $\underline{A}\times(\underline{B}\times\underline{C})$; $(\underline{A}\times\underline{B})\times\underline{C}$, etc. Note carefully that *brackets are important*, since

$$\underline{A}\times(\underline{B}\times\underline{C}) \;\neq\; (\underline{A}\times\underline{B})\times\underline{C}\,.$$

Expressions involving two (or more) vector products can be simplified by using the identity

$$\boxed{\underline{A}\times(\underline{B}\times\underline{C}) \;=\; \underline{B}(\underline{A}\cdot\underline{C}) - \underline{C}(\underline{A}\cdot\underline{B})}\,.$$

This is a result you must memorise (say "back cab" and picture a 'black cab' taxi reversing). If you worry that you may have misremembered the bracket order, remember that $(\underline{A}\times\underline{B})\times\underline{C}$ would have to be orthogonal to $\underline{C}$ and hence made up from $\underline{A}$ and $\underline{B}$.

To prove this (or at least make it plausible), we can exploit the freedom to choose a basis, and take coordinates such that $\underline{C} = (0, 0, C)$ points along the z axis. In this case,

$$\underline{B}\times\underline{C} = \begin{vmatrix} \underline{e}_1 & \underline{e}_2 & \underline{e}_3 \\ B_1 & B_2 & B_3 \\ 0 & 0 & C \end{vmatrix} = (CB_2, -CB_1, 0).$$

$$\underline{A}\times(\underline{B}+\underline{C}) = \begin{vmatrix} \underline{e}_1 & \underline{e}_2 & \underline{e}_3 \\ A_1 & A_2 & A_3 \\ CB_2 & -CB_1 & 0 \end{vmatrix} = (CA_3B_1, CA_3B_2, -[CA_1B_1 + CA_2B_1]).$$

Finally, we write the result as a relation between vectors, in which case it becomes independent of coordinates, in the same way as we deduced the centripetal force earlier.

1.2.7 Summary of coordinate-geometry approach to vectors

A **vector** is *represented* (in some orthonormal basis $\underline{e}_1, \underline{e}_2, \underline{e}_3$) by an ordered set of 3 numbers with certain laws of addition.

$$\begin{aligned}
&\textit{e.g. } \underline{A} && \text{is represented by} && (A_1, A_2, A_3)\,; \\
&\underline{A}+\underline{B} && \text{is represented by} && (A_1+B_1, A_2+B_2, A_3+B_3)\,.
\end{aligned}$$

The various 'products' of vectors are defined as follows:

The Scalar Product is denoted by $\underline{A} \cdot \underline{B}$ and **defined** as

$$\underline{A} \cdot \underline{B} \;\equiv \sum_i A_i B_i \;.$$

$$\underline{A} \cdot \underline{A} = A^2 \text{ defines the magnitude } A \text{ of the vector.}$$

The Vector Product is denoted by $\underline{A} \times \underline{B}$, and is **defined** in a right-handed basis as

$$\underline{A} \times \underline{B} = \begin{vmatrix} \underline{e}_1 & \underline{e}_2 & \underline{e}_3 \\ A_1 & A_2 & A_3 \\ B_1 & B_2 & B_3 \end{vmatrix} .$$

The Scalar Triple Product

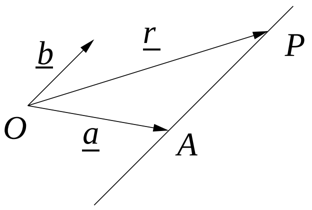
$$\begin{aligned} (\underline{A}, \underline{B}, \underline{C}) \;&\equiv\; \sum_i A_i (\underline{B} \times \underline{C})_i \\ &= \begin{vmatrix} A_1 & A_2 & A_3 \\ B_1 & B_2 & B_3 \\ C_1 & C_2 & C_3 \end{vmatrix} . \end{aligned}$$

In all the above formula the summations imply sums over each index taking values $1, 2, 3$.

1.3 Equations of lines and planes

1.3.1 The Equation of a Line

Suppose that P lies on a line which passes through a point A which has a position vector $\underline{a}$ with respect to an origin O. Let P have position vector $\underline{r}$ relative to O and let $\underline{b}$ be a vector through the origin in a direction parallel to the line.



We may write

$$\underline{r} = \underline{a} + \lambda \underline{b}$$

which is the **parametric equation of the line** *i.e.* as we vary the parameter λ from $-\infty$ to ∞, $\underline{r}$ describes all points on the line.

Rearranging and using $\underline{b} \times \underline{b} = 0$, we can also write this as

$$(\underline{r} - \underline{a}) \times \underline{b} \;=\; 0$$

or

$$\boxed{\underline{r} \times \underline{b} = \underline{c}}$$

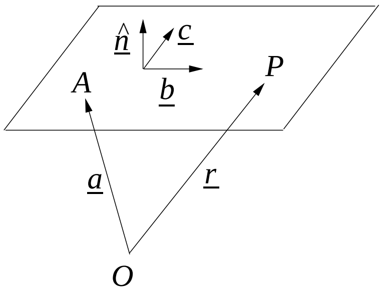
where $\underline{c} = \underline{a} \times \underline{b}$ is normal to the plane containing the line and origin.

Notes

(i) $\underline{r} \times \underline{b} = \underline{c}$ is an **implicit equation** for a line

(ii) $\underline{r} \times \underline{b} = 0$ is the equation of a line through the origin.

1.3.2 The Equation of a Plane



$\underline{r}$ is the position vector of an arbitrary point P on the plane
$\underline{a}$ is the position vector of a fixed point A in the plane
$\underline{b}$ and $\underline{c}$ are parallel to the plane but non-collinear: $\underline{b} \times \underline{c} \neq 0$.

We can express the vector $\underline{AP}$ in terms of $\underline{b}$ and $\underline{c}$, so that:

$$\underline{r} \;=\; \underline{a} + \underline{AP} \;=\; \underline{a} + \lambda \underline{b} + \mu \underline{c}$$

for some λ and μ. This is the **parametric equation of the plane**.

We define the unit normal to the plane

$$\hat{\underline{n}} \;=\; \frac{\underline{b} \times \underline{c}}{|\underline{b} \times \underline{c}|} \; .$$

Since $\underline{b} \cdot \hat{\underline{n}} = \underline{c} \cdot \hat{\underline{n}} = 0$, we have the implicit equation

$$(\underline{r} - \underline{a}) \cdot \hat{\underline{n}} \;=\; 0 \; .$$

Alternatively, we can write this as

$$\boxed{\underline{r} \cdot \hat{\underline{n}} \;=\; p} \; ,$$

where $p = \underline{a} \cdot \hat{\underline{n}}$ is the perpendicular distance of the plane from the origin. **Note**: $\underline{r} \cdot \underline{a} \;=\; 0$ is the equation for a plane through the origin (with unit normal $\underline{a}/|\underline{a}|$).

This is a very important equation which you must be able to recognise. In algebraic terms, it means that $ax + by + cz + d = 0$ is the equation of a plane.

1.3.3 Examples of Dealing with Vector Equations

Example 1: The intersection of two planes. Given two planes specified by normal vectors $\hat{\underline{n}}$ and $\hat{\underline{m}}$ and distances of closest approach to the origin of α and β respectively, what is the equation of the line where they intersect? This line has the equation $\underline{r} = \lambda \underline{p} + \underline{a}$, where $\underline{p}$ lies in both planes, and $\underline{a}$ is the position vector of a point somewhere on the line. Now, $\underline{p}$ has to be perpendicular to the two normals, so $\underline{p} = \hat{\underline{n}} \times \hat{\underline{m}}$ does the job nicely. How do we find a point on the line? It can be seen that one obvious point is the point of closest approach, which must be in the plane that holds the vectors to the closest points of the plane – i.e. $\hat{\underline{n}}$ and $\hat{\underline{m}}$. Therefore, $\underline{a}$ must be a mixture of these two vectors – even if they are not perpendicular, we can write

$$\underline{a} = x\,\hat{\underline{n}} + y\,\hat{\underline{m}}.$$

The numbers x & y can be found by requiring

$$\underline{a} \cdot \hat{\underline{n}} = \alpha; \quad \underline{a} \cdot \hat{\underline{m}} = \beta.$$

This gives a pair of simultaneous equations for x & y, whose solution is

$$x = \frac{\alpha - \beta(\underline{\hat{n}} \cdot \underline{\hat{m}})}{1 - (\underline{\hat{n}} \cdot \underline{\hat{m}})^2}; \quad y = \frac{\beta - \alpha(\underline{\hat{n}} \cdot \underline{\hat{m}})}{1 - (\underline{\hat{n}} \cdot \underline{\hat{m}})^2}.$$

Example 2: Is the following set of equations consistent?

$$\underline{r} \times \underline{b} = \underline{c}$$
$$\underline{r} = \underline{a} \times \underline{c}$$

Geometrical interpretation: the first equation is the (implicit) equation for a line whereas the second equation is the (explicit) equation for a point. Thus the question is whether the point is on the line. If we insert the 2nd into the into the l.h.s. of the first we find

$$\underline{r} \times \underline{b} = (\underline{a} \times \underline{c}) \times \underline{b} = -\underline{b} \times (\underline{a} \times \underline{c}) = -\underline{a}\,(\underline{b} \cdot \underline{c}) + \underline{c}\,(\underline{a} \cdot \underline{b})$$

Now $\underline{b} \cdot \underline{c} = \underline{b} \cdot (\underline{r} \times \underline{b}) = 0$ thus the previous equation becomes

$$\underline{r} \times \underline{b} = \underline{c}\,(\underline{a} \cdot \underline{b})$$

so that, on comparing with the first given equation, we require

$$\underline{a} \cdot \underline{b} = 1$$

for the equations to be consistent.

1.4 More on suffix notation

So far, we have been revising material that should have been relatively familiar. Now it is time to introduce some more powerful tools – whose idea is to make vector calculations quicker and easier to perform.

1.4.1 The Kronecker delta symbol δ_{ij}

We define the symbol δ_{ij} (pronounced "delta i j"), where i and j can take on the values 1 to 3, such that

$$\boxed{\begin{aligned} \delta_{ij} &= 1 \text{ if } i = j \\ &= 0 \text{ if } i \neq j \end{aligned}}$$

i.e. $\delta_{11} = \delta_{22} = \delta_{33} = 1$ and $\delta_{12} = \delta_{13} = \delta_{23} = \cdots = 0$.

The equations satisfied by the **orthonormal basis vectors** $\underline{e}_i$ can all now be written as

$$\boxed{\underline{e}_i \cdot \underline{e}_j = \delta_{ij}}\ .$$

e.g. $\quad \underline{e}_1 \cdot \underline{e}_2 = \delta_{12} = 0 \quad ; \quad \underline{e}_1 \cdot \underline{e}_1 = \delta_{11} = 1$

Notes

(i) Since there are two free indices i and j, $\underline{e}_i \cdot \underline{e}_j \;=\; \delta_{ij}$ is equivalent to 9 equations

(ii) $\delta_{ij} \;=\; \delta_{ji}$ [*i.e.* δ_{ij} is *symmetric* in its indices.]

(iii) $\sum_{i=1}^3 \delta_{ii} \;=\; 3 \quad (\,=\delta_{11}+\delta_{22}+\delta_{33})$

(iv) $\sum_{j=1}^3 A_j\delta_{jk} \;=A_1\delta_{1k}+A_2\delta_{2k}+A_3\delta_{3k}$

Remember that k is a free index. Thus if $k=1$ then only the first term on the rhs contributes and rhs $=A_1$, similarly if $k=2$ then rhs $=A_2$ and if $k=2$ then rhs $=A_3$. Thus we conclude that

$$\boxed{\sum_{j=1}^3 A_j\delta_{jk} \;=A_k}$$

In other words, the Kronecker delta picks out the kth term in the sum over j. This is in particular true for the multiplication of two Kronecker deltas:

$\sum_{j=1}^3 \delta_{ij}\delta_{jk} \;=\delta_{i1}\delta_{1k}+\delta_{i2}\delta_{2k}+\delta_{i3}\delta_{3k} \;=\delta_{ik}$

Generalising the reasoning in (iv) implies the so-called **sifting property**:

$$\boxed{\sum_{j=1}^3 (\text{anything })_j\delta_{jk} \;=(\text{anything })_k}$$

where $(\text{anything})_j$ denotes any expression that has a free index j. In effect, δ_{ij} is a kind of mathematical virus, whose main love in life it to replace the index you first thought of with one of its own. Once you get used to this behaviour, it's a very powerful trick.

Examples of the use of this symbol are:

1.
$$\begin{aligned}
\underline{A}\cdot\underline{e}_j &= \left(\sum_{i=1}^3 A_i\underline{e}_i\right)\cdot\underline{e}_j \;=\; \sum_{i=1}^3 A_i\,(\underline{e}_i\cdot\underline{e}_j)\\
&= \sum_{i=1}^3 A_i\delta_{ij} \;=\; A_j, \quad \text{since terms with } i\neq j \text{ vanish.}
\end{aligned}$$

2.
$$\begin{aligned}
\underline{A}\cdot\underline{B} &= \left(\sum_{i=1}^3 A_i\underline{e}_i\right)\cdot\left(\sum_{j=1}^3 B_j\underline{e}_j\right)\\
&= \sum_{i=1}^3\sum_{j=1}^3 A_iB_j\,(\underline{e}_i\cdot\underline{e}_j) \;=\; \sum_{i=1}^3\sum_{j=1}^3 A_iB_j\delta_{ij}\\
&= \sum_{i=1}^3 A_iB_i \quad\text{or}\quad \sum_{j=1}^3 A_jB_j.
\end{aligned}$$

Matrix representation of δ_{ij}

We may label the elements of a (3×3) matrix $\underline{\underline{M}}$ as M_{ij},

$$\underline{\underline{M}} \;=\; \begin{pmatrix} M_{11} & M_{12} & M_{13}\\ M_{21} & M_{22} & M_{23}\\ M_{31} & M_{32} & M_{33}\end{pmatrix}.$$

Note the 'double-underline' convention that we shall use to denote matrices and distinguish them from vectors and scalars. Textbooks normally use a different convention and denote a matrix in bold, **M**, but this is not practical for writing matrices by hand.

Thus we see that if we write δ_{ij} as a matrix we find that it is the identity matrix $\underline{\underline{I}}$.

$$\delta_{ij} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} .$$

1.4.2 Einstein summation convention

As you will have noticed, the novelty of writing out summations in full soon wears thin. A way to avoid this tedium is to adopt the Einstein summation convention; by adhering strictly to the following rules the summation signs are suppressed.

Rules

(i) Omit summation signs

(ii) If a suffix appears twice, a summation is implied *e.g.* $A_iB_i = A_1B_1 + A_2B_2 + A_3B_3$
Here i is a *dummy* index.

(iii) If a suffix appears only once it can take any value *e.g.* $A_i = B_i$ holds for $i = 1, 2, 3$
Here i is a *free* index. Note that there may be more than one free index. **Always** check that the free indices match on both sides of an equation *e.g.* $A_j = B_i$ is WRONG.

(iv) A given suffix **must not** appear more than twice in any term of an expression. Again, **always** check that there are no multiple indices *e.g.* $A_iB_iC_i$ is WRONG.

Examples

$\underline{A} = A_i\underline{e}_i$ *here i is a dummy index.*

$\underline{A} \cdot \underline{e}_j = A_i\underline{e}_i \cdot \underline{e}_j = A_i\delta_{ij} = A_j$ *here i is a dummy index but j is a free index.*

$\underline{A} \cdot \underline{B} = (A_i\underline{e}_i) \cdot (B_j\underline{e}_j) = A_iB_j\delta_{ij} = A_jB_j$ *here i,j are dummy indices.*

$(\underline{A} \cdot \underline{B})(\underline{A} \cdot \underline{C}) = A_iB_iA_jC_j$ *again i,j are dummy indices.*

Armed with the summation convention one can rewrite many of the equations of the previous lecture without summation signs *e.g.* the sifting property of δ_{ij} now becomes

$$\boxed{[\cdots]_j\delta_{jk} = [\cdots]_k}$$

so that, for example, $\delta_{ij}\delta_{jk} = \delta_{ik}$

From now on, except where indicated, the summation convention will be assumed. You should make sure that you are completely at ease with it.

1.4.3 Levi-Civita symbol ϵ_{ijk}

We saw in the last lecture how δ_{ij} could be used to greatly simplify the writing out of the orthonormality condition on basis vectors.

We seek to make a similar simplification for the vector products of basis vectors *i.e.* we seek a simple, uniform way of writing the equations

$$\underline{e}_1 \times \underline{e}_2 = \underline{e}_3 \quad \underline{e}_2 \times \underline{e}_1 = -\underline{e}_3 \quad \underline{e}_1 \times \underline{e}_1 = 0 \quad etc.$$

To do so we define the Levi-Civita symbol ϵ_{ijk} (pronounced 'epsilon i j k'), where i, j and k can take on the values 1 to 3, such that

$$\boxed{\begin{aligned} \epsilon_{ijk} &= +1 \text{ if } ijk \text{ is an } \textit{even} \text{ permutation of } 123\,; \\ &= -1 \text{ if } ijk \text{ is an } \textit{odd} \text{ permutation of } 123\,; \\ &= \;\; 0 \text{ otherwise } (\textit{i.e.}\ 2 \text{ or more indices are the same})\,. \end{aligned}}$$

An *even* permutation consists of an *even* number of transpositions.
An *odd* permutations consists of an *odd* number of transpositions.

For example,
$$\begin{aligned} \epsilon_{123} &= +1\,; \\ \epsilon_{213} &= -1 \;\{\text{ since } (123) \to (213) \text{ under } \textit{one} \text{ transposition } [1 \leftrightarrow 2]\}\,; \\ \epsilon_{312} &= +1 \;\{(123) \to (132) \to (312);\ 2 \text{ transpositions; } [2 \leftrightarrow 3][1 \leftrightarrow 3]\}\,; \\ \epsilon_{113} &= 0\,; \quad \epsilon_{111} \;=\; 0\,; \text{ etc.} \end{aligned}$$

$$\boxed{\epsilon_{123} = \epsilon_{231} = \epsilon_{312} = +1\ ; \quad \epsilon_{213} = \epsilon_{321} = \epsilon_{132} = -1\ ; \quad \text{all others} = 0\ .}$$

1.4.4 Vector product

The equations satisfied by the vector products of the (right-handed) orthonormal basis vectors $\underline{e}_i$ can now be written uniformly as

$$\boxed{\underline{e}_i \times \underline{e}_j \;=\; \epsilon_{ijk}\,\underline{e}_k} \qquad (i, j = 1{,}2{,}3)\ .$$

For example,

$$\underline{e}_1 \times \underline{e}_2 \;=\; \epsilon_{121}\,\underline{e}_1 + \epsilon_{122}\,\underline{e}_2 + \epsilon_{123}\,\underline{e}_3 \;=\; \underline{e}_3 \quad ; \quad \underline{e}_1 \times \underline{e}_1 \;=\; \epsilon_{111}\,\underline{e}_1 + \epsilon_{112}\,\underline{e}_2 + \epsilon_{113}\,\underline{e}_3 \;=\; 0$$

Also,

$$\begin{aligned}\underline{A} \times \underline{B} &= A_i\, B_j\, \underline{e}_i \times \underline{e}_j \\ &= \epsilon_{ijk}\, A_i B_j\, \underline{e}_k\end{aligned}$$

Thus we can see the very important and useful relation for the components of the vector product:

$$\boxed{(\underline{A} \times \underline{B})_k \;=\; \epsilon_{ijk}\, A_i B_j}$$

Example: differentiation of $\underline{A} \times \underline{B}$

Now we can prove the result we assumed earlier in differentiating $\underline{A} \times \underline{B}$:

$$\frac{d}{dt}\,\underline{A} \times \underline{B} = \frac{d}{dt}\,(\epsilon_{ijk}\, A_i B_j\, \underline{e}_k),$$

but ϵ_{ijk} and $\underline{e}_k$ are not functions of time, so we can use the ordinary product rule on the numbers A_i and B_j:

$$\frac{d}{dt}\,\underline{A} \times \underline{B} = (\epsilon_{ijk}\, \dot{A}_i B_j\, \underline{e}_k) + (\epsilon_{ijk}\, A_i \dot{B}_j\, \underline{e}_k) = \dot{\underline{A}} \times \underline{B} + \underline{A} \times \dot{\underline{B}}.$$

Cyclic symmetry of ϵ_{ijk}

The epsilon symbol obeys the useful relation

$$\epsilon_{ijk} \;=\; \epsilon_{kij} \;=\; \epsilon_{jki} \;=\; -\epsilon_{jik} \;=\; -\epsilon_{ikj} \;=\; -\epsilon_{kji}$$

This holds for *any* choice of i, j and k. To understand this note that converting e.g. ijk to kij requires two pairwise swaps involving k, so there is no sign change after a cyclic permutation. This means that we can find alternative forms for the components of the vector product:

$$\begin{aligned}(\underline{A} \times \underline{B})_k &= \epsilon_{ijk}\, A_i B_j = \epsilon_{kij}\, A_i B_j \\ \text{or relabelling indices} \quad & \quad k \to i \quad i \to j \quad j \to k \\ (\underline{A} \times \underline{B})_i &= \epsilon_{jki}\, A_j B_k = \epsilon_{ijk}\, A_j B_k\ .\end{aligned}$$

The scalar triple product can also be written using ϵ_{ijk}

$$(\underline{A}, \underline{B}, \underline{C}) = \underline{A} \cdot (\underline{B} \times \underline{C}) = A_i (\underline{B} \times \underline{C})_i$$

$$\boxed{(\underline{A}, \underline{B}, \underline{C}) = \epsilon_{ijk}\, A_i B_j C_k\ .}$$

Example: cyclic symmetry of the scalar triple product

$$\begin{aligned}(\underline{A}, \underline{B}, \underline{C}) &= \epsilon_{ijk}\, A_i B_j C_k \\ &= -\epsilon_{ikj}\, A_i B_j C_k \qquad \text{interchanging two indices of } \epsilon_{ikj} \\ &= +\epsilon_{kij}\, A_i B_j C_k \qquad \text{interchanging two indices again} \\ &= \epsilon_{ijk}\, A_j B_k C_i \qquad \text{relabelling indices} \quad k \to i \quad i \to j \quad j \to k \\ &= \epsilon_{ijk}\, C_i A_j B_k = (\underline{C}, \underline{A}, \underline{B})\end{aligned}$$

1.4.5 Product of two Levi-Civita symbols

The product of two Levi-Civita symbols obeys the following truly horrible-looking identity:

$$\epsilon_{ijk}\,\epsilon_{rsk} = \delta_{ir}\delta_{js} - \delta_{is}\delta_{jr}.$$

Proving that this is true is a painful exercise, which is certainly *not* examinable (nor is remembering the result). The reason we might want to think about such a combination is if we wanted a more rigorous proof of 'BAC-CAB' rule. Consider the i^{th} component of $\underline{A} \times (\underline{B} \times \underline{C})$:

$$\begin{aligned}
\left[\underline{A} \times (\underline{B} \times \underline{C})\right]_i &= \epsilon_{ijk}\,A_j(\underline{B} \times \underline{C})_k \\
&= \epsilon_{ijk}\,A_j\,\epsilon_{krs}\,B_rC_s \;=\; \epsilon_{ijk}\,\epsilon_{rsk}\,A_jB_rC_s \\
&= (\delta_{ir}\,\delta_{js} - \delta_{is}\,\delta_{jr})\,A_jB_rC_s \\
&= (A_jB_iC_j - A_jB_jC_i) \\
&= B_i(\underline{A} \cdot \underline{C}) - C_i(\underline{A} \cdot \underline{B}) \\
&= \left[\underline{B}(\underline{A} \cdot \underline{C}) - \underline{C}(\underline{A} \cdot \underline{B})\right]_i
\end{aligned}$$

This gives the same result that we obtained more simply by a special choice of z axis.

1.5 Coordinate transformations and change of basis

1.5.1 Linear transformation of basis

Suppose $\{\underline{e}_i\}$ and $\{\underline{e}_i{}'\}$ are two different orthonormal bases. The new basis vectors must have components in the old basis, so clearly $\underline{e}_1{}'$ can be written as a linear combination of the vectors $\underline{e}_1, \underline{e}_2, \underline{e}_3$:

$$\underline{e}_1{}' = \lambda_{11}\underline{e}_1 + \lambda_{12}\underline{e}_2 + \lambda_{13}\underline{e}_3$$

with similar expressions for $\underline{e}_2{}'$ and $\underline{e}_3{}'$. In summary,

$$\boxed{\underline{e}_i{}' \;=\lambda_{ij}\,\underline{e}_j}$$

(assuming summation convention) where λ_{ij} $(i = 1, 2, 3$ and $j = 1, 2, 3)$ are the 9 numbers relating the basis vectors $\underline{e}_1{}'$, $\underline{e}_2{}'$ and $\underline{e}_3{}'$ to the basis vectors $\underline{e}_1$, $\underline{e}_2$ and $\underline{e}_3$.

Notes

(*i*) λ_{ij} are nine numbers defining the change of basis or 'linear transformation'. They are sometimes known as 'direction cosines'.

(i) Since $\underline{e}_i{}'$ are orthonormal

$$\underline{e}_i{}' \cdot \underline{e}_j{}' \;=\; \delta_{ij}\;.$$

Now the l.h.s. of this equation may be written as

$$(\lambda_{ik}\,\underline{e}_k) \cdot (\lambda_{j\ell}\,\underline{e}_\ell) = \lambda_{ik}\lambda_{j\ell}\,(\underline{e}_k \cdot \underline{e}_\ell) = \lambda_{ik}\lambda_{j\ell}\delta_{k\ell} = \lambda_{ik}\,\lambda_{jk}$$

(in the final step we have used the sifting property of $\delta_{k\ell}$) and we deduce

$$\boxed{\lambda_{ik}\lambda_{jk} \;=\; \delta_{ij}}$$

(ii) In order to determine λ_{ij} from the two bases consider

$$\underline{e}_i{}' \cdot \underline{e}_j \;=\; (\lambda_{ik}\, \underline{e}_k) \cdot \underline{e}_j = \lambda_{ik}\, \delta_{kj} = \lambda_{ij}\,.$$

Thus

$$\boxed{\underline{e}_i{}' \cdot \underline{e}_j = \lambda_{ij}}$$

1.5.2 Inverse relations

Consider expressing the unprimed basis in terms of the primed basis and suppose that

$$\underline{e}_i = \mu_{ij}\, \underline{e}_j{}'.$$

Then $\quad \lambda_{ki} \;=\; \underline{e}_k{}' \cdot \underline{e}_i \;=\mu_{ij}\, (\underline{e}_k{}' \cdot \underline{e}_j{}') \;= \mu_{ij}\, \delta_{kj} = \mu_{ik} \qquad$ so that

$$\boxed{\mu_{ij} = \lambda_{ji}}$$

Note that $\underline{e}_i \cdot \underline{e}_j \;=\; \delta_{ij} \;= \lambda_{ki}\, (\underline{e}_k{}' \cdot \underline{e}_j) \;= \lambda_{ki}\lambda_{kj}$ and so we obtain a second relation

$$\boxed{\lambda_{ki}\lambda_{kj} \;=\; \delta_{ij}}\,.$$

1.5.3 The transformation matrix

We may label the elements of a 3×3 matrix $\underline{\underline{M}}$ as M_{ij}, where i labels the row and j labels the column in which M_{ij} appears:

$$\underline{\underline{M}} \;=\; \begin{pmatrix} M_{11} & M_{12} & M_{13} \\ M_{21} & M_{22} & M_{23} \\ M_{31} & M_{32} & M_{33} \end{pmatrix}.$$

The summation convention can be used to describe matrix multiplication. The ij component of a product of two 3×3 matrices M, N is given by

$$\boxed{(MN)_{ij} = M_{i1}\, N_{1j} + M_{i2}\, N_{2j} + M_{i3}\, N_{3j} = M_{ik}\, N_{kj}}$$

Likewise, recalling the definition of the transpose of a matrix $(M^T)_{ij} = M_{ji}$

$$(M^T N)_{ij} = (M^T)_{ik} N_{kj} = M_{ki} N_{kj}$$

We can thus arrange the numbers λ_{ij} as elements of a square matrix, denoted by λ and known as the **transformation matrix**:

$$\underline{\underline{\lambda}} \;=\; \begin{pmatrix} \lambda_{11} & \lambda_{12} & \lambda_{13} \\ \lambda_{21} & \lambda_{22} & \lambda_{23} \\ \lambda_{31} & \lambda_{32} & \lambda_{33} \end{pmatrix}$$

We denote the matrix transpose by $\underline{\underline{\lambda}}^T$ and define it by $(\lambda^T)_{ij} = \lambda_{ji}$ so we see that $\underline{\underline{\mu}} = \underline{\underline{\lambda}}^T$ is the transformation matrix for the inverse transformation.

We also note that δ_{ij} may be thought of as elements of a 3×3 unit matrix:

$$\begin{pmatrix} \delta_{11} & \delta_{12} & \delta_{13} \\ \delta_{21} & \delta_{22} & \delta_{33} \\ \delta_{31} & \delta_{32} & \delta_{33} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \underline{\underline{I}}.$$

i.e. the matrix representation of the Kronecker delta symbol is the **unit matrix** $\underline{\underline{I}}$.

The inverse relation $\lambda_{ik}\lambda_{jk} = \lambda_{ki}\lambda_{kj} = \delta_{ij}$ can be written in matrix notation as

$$\boxed{\underline{\underline{\lambda}}\,\underline{\underline{\lambda}}^T = \underline{\underline{\lambda}}^T\,\underline{\underline{\lambda}} = \underline{\underline{I}}}\ , \quad i.e.\ \boxed{\underline{\underline{\lambda}}^{-1} = \underline{\underline{\lambda}}^T}\ .$$

This is the definition of an **orthogonal matrix** and the transformation (from the $\underline{e}_i$ basis to the $\underline{e}_i{}'$ basis) is called an **orthogonal transformation**.

Now from the properties of determinants, $|\underline{\underline{\lambda}}\,\underline{\underline{\lambda}}^T| = |\underline{\underline{I}}| = 1 = |\lambda|\,|\lambda^T|$ and $|\lambda^T| = |\lambda|$, we have that $|\lambda|^2 = 1$ hence

$$|\lambda| = \pm 1\ .$$

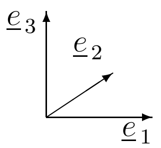
If $|\lambda| = +1$ the orthogonal transformation is said to be 'proper'
If $|\lambda| = -1$ the orthogonal transformation is said to be 'improper'

Handedness of basis

An improper transformation has an unusual effect. In the usual Cartesian basis that we have considered up to now, the basis vectors $\underline{e}_1$, $\underline{e}_2$, and $\underline{e}_3$ form a *right-handed* basis, that is, $\underline{e}_1 \times \underline{e}_2 = \underline{e}_3$, $\underline{e}_2 \times \underline{e}_3 = \underline{e}_1$ and $\underline{e}_3 \times \underline{e}_1 = \underline{e}_2$.

However, we could choose $\underline{e}_1 \times \underline{e}_2 = -\underline{e}_3$, and so on, in which case the basis is said to be *left-handed*.

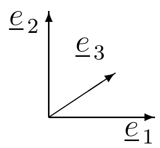
right handed



$$\begin{aligned} \underline{e}_3 &= \underline{e}_1 \times \underline{e}_2 \\ \underline{e}_1 &= \underline{e}_2 \times \underline{e}_3 \\ \underline{e}_2 &= \underline{e}_3 \times \underline{e}_1 \end{aligned}$$

$$\boxed{(\underline{e}_1, \underline{e}_2, \underline{e}_3) = 1}$$

left handed

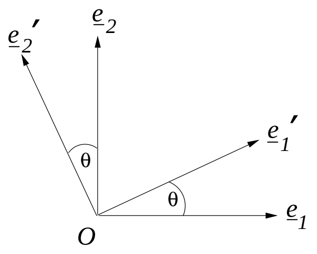


$$\begin{aligned} \underline{e}_3 &= \underline{e}_2 \times \underline{e}_1 \\ \underline{e}_1 &= \underline{e}_3 \times \underline{e}_2 \\ \underline{e}_2 &= \underline{e}_1 \times \underline{e}_3 \end{aligned}$$

$$\boxed{(\underline{e}_1, \underline{e}_2, \underline{e}_3) = -1}$$

1.5.4 Examples of orthogonal transformations

Rotation about the $\underline{e}_3$ axis. We have $\underline{e}_3{}' = \underline{e}_3$ and thus for a rotation through θ,



$$
\begin{aligned}
\underline{e}_3{}' \cdot \underline{e}_1 &= \underline{e}_1{}' \cdot \underline{e}_3 = \underline{e}_3{}' \cdot \underline{e}_2 = \underline{e}_2{}' \cdot \underline{e}_3 = 0\,, \quad \underline{e}_3{}' \cdot \underline{e}_3 = 1 \\
\underline{e}_1{}' \cdot \underline{e}_1 &= \cos\theta \\
\underline{e}_1{}' \cdot \underline{e}_2 &= \cos(\pi/2 - \theta) \;=\; \sin\theta \\
\underline{e}_2{}' \cdot \underline{e}_2 &= \cos\theta \\
\underline{e}_2{}' \cdot \underline{e}_1 &= \cos(\pi/2 + \theta) \;=\; -\sin\theta
\end{aligned}
$$

Thus

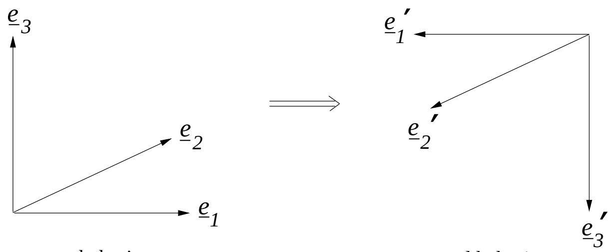
$$\underline{\underline{\lambda}} = \begin{pmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

It is easy to check that $\underline{\underline{\lambda}}\,\underline{\underline{\lambda}}^T = \underline{\underline{I}}$. Since $|\lambda| = \cos^2\theta + \sin^2\theta = 1$, this is a *proper* transformation. Note that rotations cannot change the handedness of the basis vectors.

Inversion or Parity transformation. This is defined such that $\underline{e}_i{}' = -\underline{e}_i$.

$$\text{i.e. } \lambda_{ij} \;=\; -\delta_{ij} \qquad \text{or} \qquad \lambda \;=\; \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \;=\; -\underline{\underline{I}}\,.$$

Clearly $\underline{\underline{\lambda}}\,\underline{\underline{\lambda}}^T = \underline{\underline{I}}$. Since $|\lambda| = -1$, this is an *improper* transformation. Note that the handedness of the basis is reversed: $\underline{e}_1{}' \times \underline{e}_2{}' = -\underline{e}_3{}'$



r.h. basis   l.h. basis

Reflection. Consider reflection of the axes in $\underline{e}_2$–$\underline{e}_3$ plane so that $\underline{e}_1{}' = -\underline{e}_1$, $\underline{e}_2{}' = \underline{e}_2$ and $\underline{e}_3{}' = \underline{e}_3$. The transformation matrix is

$$\underline{\underline{\lambda}} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Since $|\lambda| = -1$, this is an *improper* transformation. Again the handedness of the basis changes.

1.5.5 Products of transformations

Consider a transformation $\underline{\underline{\lambda}}$ to the basis $\{\underline{e}_i{}'\}$ followed by a transformation $\underline{\underline{\rho}}$ to another basis $\{\underline{e}_i{}''\}$

$$\underline{e}_i \overset{\lambda}{\Longrightarrow} \underline{e}_i{}' \overset{\rho}{\Longrightarrow} \underline{e}_i{}''$$

Clearly there must be an orthogonal transformation $\underline{e}_i \overset{\xi}{\Longrightarrow} \underline{e}_i{}''$

Now

$$\underline{e}_i{}'' = \rho_{ij}\underline{e}_j{}' = \rho_{ij}\lambda_{jk}\underline{e}_k = (\rho\lambda)_{ik}\underline{e}_k \quad \text{so} \quad \boxed{\underline{\underline{\xi}} = \underline{\underline{\rho}}\,\underline{\underline{\lambda}}}$$

Note the order of the product: the matrix corresponding to the first change of basis stands to the right of that for the second change of basis. In general, transformations do not commute so that $\underline{\underline{\rho}}\,\underline{\underline{\lambda}} \neq \underline{\underline{\lambda}}\,\underline{\underline{\rho}}$. Only the inversion and identity transformations commute with all transformations.

Example: Rotation of θ about $\underline{e}_3$ then reflection in yz plane

$$\begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} -\cos\theta & -\sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

whereas, if we reverse the order,

$$\begin{pmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} -\cos\theta & \sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

1.5.6 Improper transformations

We may write any improper transformation $\underline{\underline{\xi}}$ (for which $|\xi| = -1$) as $\underline{\underline{\xi}} = \left(-\underline{\underline{I}}\right) \underline{\underline{\lambda}}$ where $\underline{\underline{\lambda}} = -\underline{\underline{\xi}}$ and $|\lambda| = +1$ Thus an improper transformation can always be expressed as a proper transformation followed by an inversion.

e.g. consider $\underline{\underline{\xi}}$ for a reflection in the $1-3$ plane which may be written as

$$\underline{\underline{\xi}} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

Identifying $\underline{\underline{\lambda}}$ from $\underline{\underline{\xi}} = \left(-\underline{\underline{I}}\right) \underline{\underline{\lambda}}$ we see that λ is a rotation of π about $\underline{e}_2$.

1.6 Transformation properties of vectors and scalars

1.6.1 Transformation of vector components

Let $\underline{A}$ be any vector, with components A_i in the basis $\{\underline{e}_i\}$ and A'_i in the basis $\{\underline{e}_i{}'\}$ *i.e.*

$$\boxed{\underline{A} \;=A_i\,\underline{e}_i \;=A'_i\,\underline{e}_i{}'} \; .$$

The components are related as follows, taking care with dummy indices:

$$A'_i \;=\; \underline{A}\cdot\underline{e}_i{}' \;=\; (A_j\underline{e}_j)\cdot\underline{e}_i{}' = (\underline{e}_i{}'\cdot\underline{e}_j)A_j = \lambda_{ij}A_j$$

$$\boxed{A'_i = \lambda_{ij}A_j}$$

So the new components are related to the old ones by the same matrix transformation as applies to the basis vectors. The inverse transformation works in a similar way:

$$A_i \;=\; \underline{A} \cdot \underline{e}_i \;=\; (A'_k \underline{e}_k{}') \cdot \underline{e}_i = \lambda_{ki} A'_k \;= (\lambda^T)_{ik} A'_k .$$

Note carefully that we do *not* put a prime on the vector itself – there is only one vector, $\underline{A}$, in the above discussion.

However, the *components* of this vector are different in different bases, and so are denoted by A_i in the basis $\{\underline{e}_i\}$, A'_i in the basis $\{\underline{e}_i{}'\}$, etc.

In matrix form we can write these relations as

$$\boxed{\begin{pmatrix} A'_1 \\ A'_2 \\ A'_3 \end{pmatrix} = \begin{pmatrix} \lambda_{11} & \lambda_{12} & \lambda_{13} \\ \lambda_{21} & \lambda_{22} & \lambda_{23} \\ \lambda_{31} & \lambda_{32} & \lambda_{33} \end{pmatrix} \begin{pmatrix} A_1 \\ A_2 \\ A_3 \end{pmatrix} = \lambda \begin{pmatrix} A_1 \\ A_2 \\ A_3 \end{pmatrix}}$$

Example: Consider a rotation of the axes about $\underline{e}_3$

$$\begin{pmatrix} A'_1 \\ A'_2 \\ A'_3 \end{pmatrix} = \begin{pmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} A_1 \\ A_2 \\ A_3 \end{pmatrix} = \begin{pmatrix} \cos\theta\; A_1 + \sin\theta\; A_2 \\ \cos\theta\; A_2 - \sin\theta\; A_1 \\ A_3 \end{pmatrix}$$

A direct check of this using trigonometric considerations is significantly harder!

1.6.2 The Transformation of the scalar product

Let $\underline{A}$ and $\underline{B}$ be vectors with components A_i and B_i in the basis $\{\underline{e}_i\}$ and components A'_i and B'_i in the basis $\{\underline{e}_i{}'\}$. In the basis $\{\underline{e}_i\}$, the scalar product, denoted by $(\underline{A} \cdot \underline{B})$, is

$$(\underline{A} \cdot \underline{B}) \;= A_i B_i \,.$$

In the basis $\{\underline{e}_i{}'\}$, we denote the scalar product by $(\underline{A} \cdot \underline{B})\,'$, and we have

$$\begin{aligned} (\underline{A} \cdot \underline{B})\,' &= A'_i B'_i \;= \lambda_{ij} A_j \lambda_{ik} B_k \;= \delta_{jk} A_j B_k \\ &= A_j B_j = (\underline{A} \cdot \underline{B})\,. \end{aligned}$$

Thus the scalar product is the same evaluated in any basis. This is of course expected from the geometrical definition of scalar product which is independent of basis. We say that the scalar product is *invariant* under a change of basis.

Summary We have now obtained an algebraic definition of scalar and vector quantities. Under the orthogonal transformation from the basis $\{\underline{e}_i\}$ to the basis $\{\underline{e}_i{}'\}$, defined by the transformation matrix $\lambda :\; \underline{e}_i{}' = \lambda_{ij}\, \underline{e}_j$, we have that

- A **scalar** is a single number ϕ which is invariant:

$$\boxed{\phi' \;=\; \phi}\,.$$

Of course, not all scalar quantities in physics are expressible as the scalar product of two vectors *e.g.* mass, temperature.

- A **vector** is an 'ordered triple' of numbers A_i which transforms to A'_i :

$$\boxed{A'_i \;=\lambda_{ij}A_j} \; .$$

1.6.3 Transformation of the vector product

Improper transformations have an interesting effect on the vector product. Consider the case of inversion, so that $\underline{e}_i{}' = -\underline{e}_i$, and all the signs of vector components are flipped: $A'_i = -A_i$ etc.

If we now calculate the vector product $\underline{C} = \underline{A} \times \underline{B}$ in the new basis using the determinant formula, we obtain

$$\begin{vmatrix} \underline{e}_1{}' & \underline{e}_2{}' & \underline{e}_3{}' \\ A'_1 & A'_2 & A'_3 \\ B'_1 & B'_2 & B'_3 \end{vmatrix} = (-)^3 \begin{vmatrix} \underline{e}_1 & \underline{e}_2 & \underline{e}_3 \\ A_1 & A_2 & A_3 \\ B_1 & B_2 & B_3 \end{vmatrix} = - \begin{vmatrix} \underline{e}_1 & \underline{e}_2 & \underline{e}_3 \\ A_1 & A_2 & A_3 \\ B_1 & B_2 & B_3 \end{vmatrix}$$

which is $-\underline{C}$ as calculated in the original basis!

This shouldn't be surprising, since we defined the vector product direction in the first place assuming a right-hand screw rule, and the determinant expression was derived assuming that the basis was also right-handed. If we use this same formula in a left-handed basis, the direction of the vector product produced by the formula is reversed compared to our original definition.

The real solution to this is to stick to the geometrical definition, in which case the vector product is always a well-defined entity (the notion of a right-handed screw is not changed just because we use a left-handed basis).

But if we regard the simple determinant form as the *definition* of the vector product (which is frequently what is assumed), then we have to live with the possibility of $\underline{A} \times \underline{B}$ flipping sign. It is then an example of what is known as a 'pseudovector' or an 'axial vector' which has the transformation law

$$\boxed{C'_i = (\det \lambda)\lambda_{ij}C_j}$$

so that it behaves just like a vector under proper transformations, for which $\det \lambda = +1$, but picks up an extra minus sign under improper transformations, for which $\det \lambda = -1$.

1.6.4 Summary of story so far

We take the opportunity to summarise some key-points of what we have done so far. N.B. this is NOT a list of everything you need to know.

Key points from geometrical approach

You should recognise on sight that

$$\begin{aligned} \underline{r} \times \underline{b} = \underline{c} \quad & \text{is a line } (\underline{r} \text{ lies on a line}) \\ \underline{r} \cdot \underline{a} = d \quad & \text{is a plane } (\underline{r} \text{ lies in a plane}) \end{aligned}$$

Useful properties of scalar and vector products to remember

$$\begin{aligned} \underline{a} \cdot \underline{b} = 0 \quad & \Leftrightarrow \quad \text{vectors orthogonal} \\ \underline{a} \times \underline{b} = 0 \quad & \Leftrightarrow \quad \text{vectors collinear} \\ \underline{a} \cdot (\underline{b} \times \underline{c}) = 0 \quad & \Leftrightarrow \quad \text{vectors co-planar or linearly dependent} \\ \underline{a} \times (\underline{b} \times \underline{c}) \quad & = \quad \underline{b}(\underline{a} \cdot \underline{c}) - \underline{c}(\underline{a} \cdot \underline{b}) \end{aligned}$$

Key points of suffix notation

We label orthonormal basis vectors $\underline{e}_1, \underline{e}_2, \underline{e}_3$ and write the expansion of a vector $\underline{A}$ as

$$\underline{A} = \sum_{i=1}^{3} A_i \underline{e}_i$$

The Kronecker delta δ_{ij} can be used to express the orthonormality of the basis

$$\underline{e}_i \cdot \underline{e}_j = \delta_{ij}$$

The Kronecker delta has a very useful sifting property

$$\sum_j [\cdots]_j \delta_{jk} = [\cdots]_k$$

$$(\underline{e}_1, \underline{e}_2, \underline{e}_3) = \pm 1 \quad \text{determines whether the basis is right- or left-handed}$$

Key points of summation convention

Using the summation convention we have for example

$$\underline{A} = A_i \underline{e}_i$$

and the sifting property of δ_{ij} becomes

$$[\cdots]_j \delta_{jk} = [\cdots]_k$$

We introduce ϵ_{ijk} to enable us to write the vector products of basis vectors in a r.h. basis in a uniform way

$$\underline{e}_i \times \underline{e}_j = \epsilon_{ijk} \underline{e}_k$$

.

The vector products and scalar triple products in a r.h. basis are

$$\underline{A} \times \underline{B} = \begin{vmatrix} \underline{e}_1 & \underline{e}_2 & \underline{e}_3 \\ A_1 & A_2 & A_3 \\ B_1 & B_2 & B_3 \end{vmatrix} \quad \text{or equivalently} \quad (\underline{A} \times \underline{B})_i \;=\; \epsilon_{ijk} A_j B_k$$

$$\underline{A} \cdot (\underline{B} \times \underline{C}) = \begin{vmatrix} A_1 & A_2 & A_3 \\ B_1 & B_2 & B_3 \\ C_1 & C_2 & C_3 \end{vmatrix} \quad \text{or equivalently} \quad \underline{A} \cdot (\underline{B} \times \underline{C}) \;=\; \epsilon_{ijk} A_i B_j C_k$$

Key points of change of basis

The new basis is written in terms of the old through

$$\underline{e}_i{}' = \lambda_{ij}\underline{e}_j \quad \text{where } \lambda_{ij} \text{ are elements of a } 3 \times 3 \text{ transformation matrix } \underline{\underline{\lambda}}$$

$\underline{\underline{\lambda}}$ is an orthogonal matrix, the defining property of which is $\underline{\underline{\lambda}}^{-1} = \underline{\underline{\lambda}}^T$ and this can be written as

$$\underline{\underline{\lambda}}\,\underline{\underline{\lambda}}^T = \underline{\underline{I}} \quad \text{or} \quad \lambda_{ik}\lambda_{jk} = \delta_{ij}$$

$|\lambda| = \pm 1$ decides whether the transformation is proper or improper i.e. whether the handedness of the basis is changed

Key points of algebraic approach

A **scalar** is defined as a number that is invariant under an orthogonal transformation

A **vector** is defined as an object $\underline{A}$ represented in a basis by numbers A_i which transform to A_i' through

$$A_i' = \lambda_{ij}A_j.$$

or in matrix form

$$\begin{pmatrix} A_1' \\ A_2' \\ A_3' \end{pmatrix} = \underline{\underline{\lambda}} \begin{pmatrix} A_1 \\ A_2 \\ A_3 \end{pmatrix}$$