Mathematical Techniques: Vector Algebra and Vector Fields

Lecture notes by W.B. von Schlippe

Part 1. Vector Algebra.

1. Definition of vectors; addition, subtraction, multiplication by scalar factor; associative law of addition; associative, commutative and distributive laws of multiplication by scalar factors.

2. Geometrical representation of vectors.

3. Scalar product; distributive and commutative law of scalar products; kinetic energy.

4. Modulus of a vector; unit vectors; base vectors; linear independence.

5. Vector product of two vectors; distributive law of vector product; non-commutative law of vector product;
   Examples: moment of force (torque); angular momentum.

6. Applications: Vector equations of lines and planes.

7. Triple products; triple scalar product; triple vector product.

8. Coordinate transformations; transformation properties of vectors under rotations; invariance of the dot-product under rotations; reflections; polar and axial vectors.

9. Exercises and problems.

Part 2. Vector Fields.

1. Definition of scalar and vector fields; field lines.

2. Gradient of a scalar field; potential

3. Flux of a vector field; divergence.

4. The nabla operator.

5. Theorems of Gauss and Green.

6. Curl of a vector field; Stokes' theorem.

7. Application: Maxwell equations of Electromagnetism.

Part 1. Vector Algebra.

1. Definition of vectors; addition, multiplication by scalar factor, subtraction; associative law of addition, commutative and distributive laws of multiplication by scalar factors.

A vector is a triplet of numbers. It is represented in the following notation:

$$\vec{r} = (x, y, z) \qquad \text{or} \qquad \vec{a} = (a_x, a_y, a_z) \tag{1}$$

Here x, y and z are the *components* of $\vec{r}$, and similarly a_x, a_y and a_z are the components of $\vec{a}$.

Vectors are defined by the following properties:

1. Given two vectors, $\vec{a} = (a_x, a_y, a_z)$ and $\vec{b} = (b_x, b_y, b_z)$, the *sum* of $\vec{a}$ and $\vec{b}$ is the vector $\vec{c}$ defined by

$$\vec{c} = \vec{a} + \vec{b} = (a_x + b_x, a_y + b_y, a_z + b_z) \tag{2}$$

2. The product of vector $\vec{a}$ with a scalar factor λ is the vector $\vec{c}$ defined by

$$\vec{c} = \lambda\vec{a} = (\lambda a_x, \lambda a_y, \lambda a_z) \tag{3}$$

3. Based on definitions 1 and 2 we can define the *subtraction* of vectors $\vec{a}$ and $\vec{b}$ as the addition of $\vec{a}$ and $(-\vec{b})$: thus, the vector difference of vectors $\vec{a}$ and $\vec{b}$ is the vector $\vec{c}$ defined by

$$\vec{c} = \vec{a} - \vec{b} = \vec{a} + (-\vec{b}) = (a_x - b_x, a_y - b_y, a_z - b_z) \tag{4}$$

From the above definitions one gets the following rules:

(i) The addition of vectors is associative:

$$\vec{d} = (\vec{a} + \vec{b}) + \vec{c} = \vec{a} + (\vec{b} + \vec{c}) \tag{5}$$

We can therefore drop the brackets in the sum of three vectors and write

$$\vec{d} = \vec{a} + \vec{b} + \vec{c}$$

(ii) The addition of vectors is commutative:

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

(iii) The multiplication by a scalar factor is distributive:

$$\lambda(\vec{a} + \vec{b}) = \lambda\vec{a} + \lambda\vec{b}$$

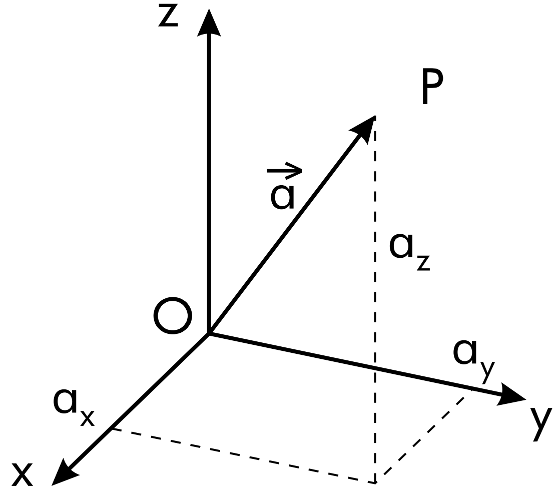

Figure 1: Geometrical representation of vectors.

(iv) The multiplication by a scalar factor is commutative:

$$\lambda\vec{a} = \vec{a}\lambda$$

Rules (i) - (iv) follow directly from the corresponding rules for the vector components, which are ordinary numbers.

A special vector is the *null vector*: this is the vector whose three components are all equal to nought. There is no universally agreed notation for the null vector; since there is usually no ambiguity, we shall denote it by zero, *i.e.* write $(0, 0, 0) = 0$.

2. Geometrical representation of vectors.

The geometrical representation of a vector is that of a *directed* quantity: thus the vector $\vec{a} = (a_x, a_y a_z)$ is represented in a cartesian coordinate system by the arrow from the origin to the point P whose coordinates are a_x, a_y and a_z, respectively (Fig. 1). *Fig. 1*

The law of addition of vectors leads to the representation of the sum of vectors $\vec{a}$ and $\vec{b}$ as the diagonal of the parallelogram constructed from $\vec{a}$ and $\vec{b}$ as shown in Fig. 2. *Fig. 2*

To construct the difference vector $\vec{c} = \vec{a} - \vec{b}$ we can either use the definition and construct first the vector $-\vec{b}$ and then use the parallelogram rule to add $\vec{a}$ and $(-\vec{b})$ (Fig. 3a). *Fig. 3a* Alternatively we can note that if $\vec{c} = \vec{a} - \vec{b}$, then $\vec{a} = \vec{b} + \vec{c}$. Therefore we get $\vec{c}$ by drawing the vector from $\vec{b}$ to $\vec{a}$ (Fig. 3b). *Fig. 3b*

The vector $\lambda\vec{a}$ points in the direction of $\vec{a}$ and has a length λ times the length of $\vec{a}$. Two vectors pointing in the same direction are called *collinear*. Vectors which point in opposite directions are called *anticollinear*. Thus $\vec{a}$ and $(-\vec{a})$ are anticollinear, but more generally also $\vec{a}$ and $(-\lambda\vec{a})$ if $\lambda > 0$.

3. Scalar product.

Distributive and commutative law of scalar products; work and kinetic energy.

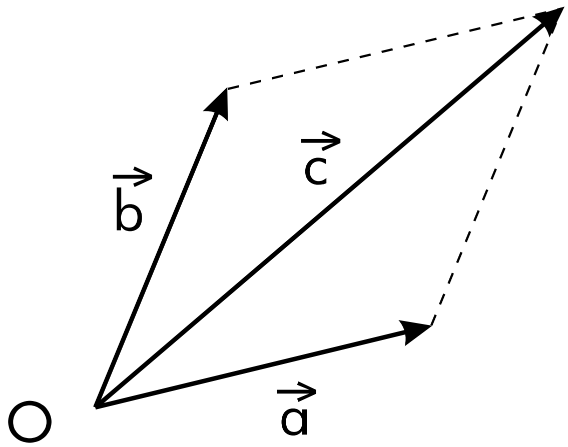

Figure 2: Addition of vectors $\vec{a}$ and $\vec{b}$.

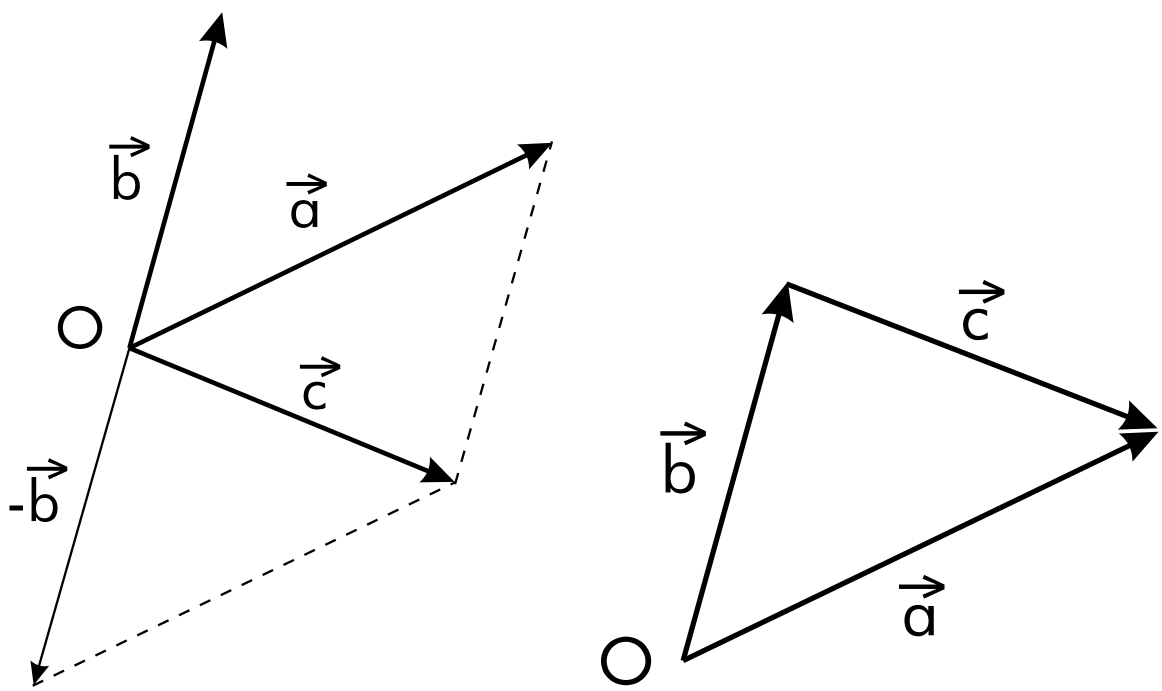

Figure 3: Subtraction of vectors $\vec{a}$ and $\vec{b}$.

The multiplication of vectors $\vec{a}$ and $\vec{b}$ is not uniquely defined. One can multiply the components of $\vec{a}$ with the components of $\vec{b}$ in nine different ways: $a_x b_x$, $a_x b_y$, $\dots a_z b_z$. These products are called *dyads*. One can also add and subtract dyads in various ways. Two of such combinations are particularly useful. They are the *scalar product* and the *vector product* of $\vec{a}$ and $\vec{b}$.

The scalar product is defined by

$$\vec{a} \cdot \vec{b} = a_x b_x + a_y b_y + a_z b_z \tag{6}$$

Because of the way the scalar product is written it is also called the *dot product*. In section 8 we will see that the scalar product does not change under a rotation of the coordinate axes. This is the property of scalars, and the invariance under rotations is the reason why the scalar product has its name.

The scalar product is distributive:

$$\vec{a} \cdot (\vec{b} + \vec{c}) = \vec{a} \cdot \vec{b} + \vec{a} \cdot \vec{c} \tag{7}$$

This follows directly from the the distributive law for the components. Thus for the x components we have

$$a_x(b_x + c_x) = a_x b_x + a_x c_x$$

and similarly for the y and z components.

The scalar product is commutative:

$$\vec{a} \cdot \vec{b} = \vec{b} \cdot \vec{a} \tag{8}$$

This also follows directly from the commutative law for the components.

The associative law has no meaning in relation to the scalar product. For instance, if we take the scalar product $\vec{a} \cdot \vec{b}$, then this is a scalar, and it is meaningless to form its dot product with a third vector.

The scalar product of $\vec{a}$ with itself is called the *modulus squared* of $\vec{a}$, and one writes

$$a^2 = \vec{a} \cdot \vec{a} = a_x^2 + a_y^2 + a_z^2 \tag{9}$$

The square root of a^2 is called the *modulus* of $\vec{a}$ and is denoted by $|\vec{a}|$. Thus

$$|\vec{a}| = \sqrt{a_x^2 + a_y^2 + a_z^2} \tag{10}$$

Consider the right-angled triangle OAB with sides $OA = a_x$, $AB = a_y$ and hypotenuse $OB = c$ (Fig. 4): by Pythagoras' theorem we have $c^2 = a_x^2 + a_y^2$. Then consider the right-angled triangle OBP: OP is the length of $\vec{a}$, and from the triangle we have $OP^2 = c^2 + a_z^2 = a_x^2 + a_y^2 + a_z^2$. Comparing this with the expression for the modulus of $\vec{a}$ we conclude that *the modulus of a vector is equal to its length.* *Fig. 4*

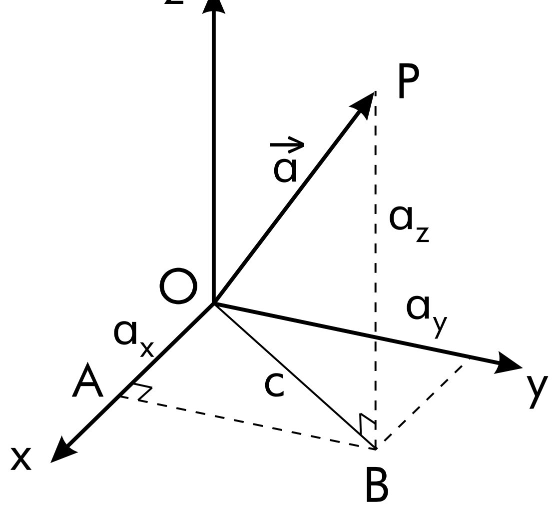

Figure 4: Modulus of vector $\vec{a}$.

There are three vectors which have a special significance, namely the vectors

$$\hat{\imath} = (1, 0, 0), \qquad \hat{\jmath} = (0, 1, 0) \qquad \text{and} \qquad \hat{k}(0, 0, 1) \tag{11}$$

which are of unit length and point in the x, y and z direction, respectively. Generally any vector of unit length is called a *unit vector*. To distinguish unit vectors from general vectors we shall always put a carat on unit vectors rather than an arrow.

A vector, divided by its modulus, is a unit vector. We can therefore write

$$\hat{a} = \frac{\vec{a}}{|\vec{a}|}$$

Consider the dot products of $\vec{a}$ with $\hat{\imath}$, $\hat{\jmath}$ and $\hat{k}$:

$$\vec{a} \cdot \hat{\imath} = a_x, \qquad \vec{a} \cdot \hat{\jmath} = a_y, \qquad \text{and} \qquad \vec{a} \cdot \hat{k} = a_z$$

Now, if $\vec{a}$ makes the angle α with the x axis, we have also $a_x = |\vec{a}| \cos\alpha$, and similarly $a_y = |\vec{a}| \cos\beta$ and $a_z = |\vec{a}| \cos\gamma$, if β and γ are the angles $\vec{a}$ makes with the y and z axis, respectively. We get therefore the following representation of $\vec{a}$:

$$\vec{a} = |\vec{a}|(\cos\alpha, \cos\beta, \cos\gamma) \tag{12}$$

Thus the cosines of the angles α, β and γ are seen to define the direction of $\vec{a}$. They are therefore called the *directional cosines* of $\vec{a}$. An important property of the directional cosines is found if we recall that $\hat{a}$ is a unit vector. Therefore we have

$$\left(\frac{\vec{a}}{|\vec{a}|}\right)^2 = \hat{a}^2 = 1 = \cos^2\alpha + \cos^2\beta + \cos^2\gamma \tag{13}$$

So there is one relation between the three directional cosines and one could conclude that two directional cosines are sufficient to define the direction of a vector. But that would be

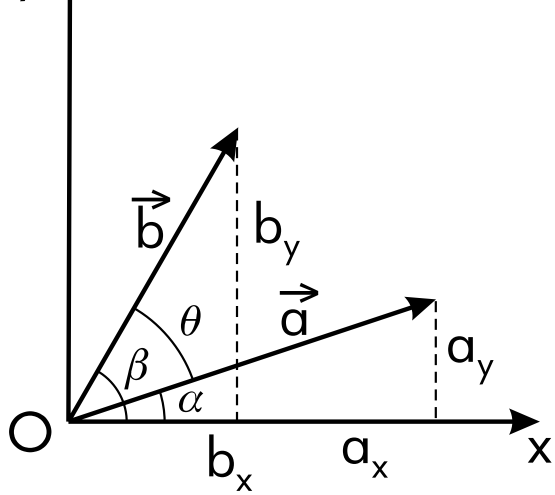

Figure 5: To derive $\vec{a} \cdot \vec{b} = ab\,\cos\theta$.

wrong because the relation between the directional cosines is quadratic rather than linear. Therefore two directional cosines define the third one only up to its sign.

Consider the dot products of $\hat{\imath}$, $\hat{\jmath}$ and $\hat{k}$ with each other: it follows from their definitions that

$$\hat{\imath} \cdot \hat{\jmath} = 0, \qquad \hat{\jmath} \cdot \hat{k} = 0, \qquad \text{and} \qquad \hat{k} \cdot \hat{\imath} = 0 \tag{14}$$

Now, $\hat{\imath}$ is perpendicular to $\hat{\jmath}$ and they are both perpendicular to $\hat{k}$. Another word for perpendicular is *othogonal*. We can therefore say that the three vectors $\hat{\imath}$, $\hat{\jmath}$ and $\hat{k}$ are *mutually orthogonal*. We have also seen that $\vec{a} \cdot \hat{\imath} = |\vec{a}| \cos\alpha$. Therefore if $\vec{a}$ and $\hat{\imath}$ are orthogonal, then their dot product is equal to zero, because then $\cos\alpha = \cos(\pi/2) = 0$. We shall see that generally the dot product of any two vectors $\vec{a}$ and $\vec{b}$ can be represented as

$$\vec{a} \cdot \vec{b} = ab\cos\theta \tag{15}$$

where we have put $a = |\vec{a}|$, $b = |\vec{b}|$ and we denote the angle between $\vec{a}$ and $\vec{b}$ by θ.[1]

For two vectors which lie in the xy plane the statement is easily shown. If $\vec{a}$ and $\vec{b}$ make angles α and β with the x axis (Fig. 5), then *Fig. 5*

$$\vec{a} = a(\cos\alpha, \sin\alpha, 0), \qquad \vec{b} = b(\cos\beta, \sin\beta, 0)$$

and hence

$$\vec{a} \cdot \vec{b} = ab(\cos\alpha\cos\beta + \sin\alpha\sin\beta) = ab\cos(\alpha - \beta) = ab\cos\theta$$

For general vectors, which do not lie in a coordinate plane, we can intuitively see in the following way that this statement remains true: it is possible to rotate the coordinate frame in such a way that after the rotation the vectors lie in a coordinate plane. This rotation affects neither the lengths of the vectors nor the angle between them. Therefore the expression

[1] Frequently this relation is taken as the definition of the dot product.

$ab\cos\theta$ for the dot product must be true in any coordinate frame. Later in this course, in connection with matrix algebra, we shall present this proof rigorously.

An important example of a scalar product in mechanics is the work done by a force $\vec{F}$ in moving an object through a distance $\delta\vec{r}$, which is defined by

$$\delta W = \vec{F} \cdot \delta\vec{r}$$

Another example is the kinetic energy T of a mass m moving with momentum $\vec{p}$:

$$T = \frac{\vec{p}^{\,2}}{2m}$$

where $\vec{p}^{\,2}$ is the dot product $\vec{p} \cdot \vec{p}$.

4. Base vectors; linear independence.

In Section 3 we have mentioned the unit vectors $\hat{\imath}$, $\hat{\jmath}$ and $\hat{k}$, and we have shown that the components of an arbitrary vector $\vec{a}$ could be written as

$$a_x = \vec{a} \cdot \hat{\imath}, \qquad a_y = \vec{a} \cdot \hat{\jmath}, \qquad \text{and} \qquad a_z = \vec{a} \cdot \hat{k}$$

We can therefore conclude that $\vec{a}$ can be represented as the *linear superposition* of the unit vectors $\hat{\imath}$, $\hat{\jmath}$ and $\hat{k}$:

$$\vec{a} = a_x\hat{\imath} + a_y\hat{\jmath} + a_z\hat{k} \tag{16}$$

where the components of $\vec{a}$ play the role of the *superposition coefficients*. Indeed, if we take the dot product of the latter equation with $\hat{\imath}$ and take into account the orthogonality of $\hat{\imath}$, $\hat{\jmath}$ and $\hat{k}$, then we recover the statement

$$\vec{a} \cdot \hat{\imath} = a_x$$

and similarly for the other two components.

An important property of the unit vectors $\hat{\imath}$, $\hat{\jmath}$ and $\hat{k}$ is their *linaer independence*. This is the statement that none of the three vectors can be represented as a linear superposition of the other two. For $\hat{\imath}$, $\hat{\jmath}$ and $\hat{k}$ their linear independence can be proved like this: assume that they are *not* linearly independent, i.e. assume that there is a linear relationship between them, e.g. $\hat{\imath} = a\hat{\jmath} + b\hat{k}$ or, more symmetrically

$$c_1\hat{\imath} + c_2\hat{\jmath} + c_3\hat{k} = 0$$

then, taking the dot product with $\hat{\imath}$, we get

$$c_1\hat{\imath} \cdot \hat{\imath} + c_2\hat{\jmath} \cdot \hat{\imath} + c_3\hat{k} \cdot \hat{\imath} = 0$$

but $\hat{\imath} \cdot \hat{\imath} = 1$ and $\hat{\jmath} \cdot \hat{\imath} = \hat{k} \cdot \hat{\imath} = 0$, hence $c_1 = 0$. Similarly, taking the dot products with $\hat{\jmath}$ and then with $\hat{k}$, we can show that $c_2 = 0$ and $c_3 = 0$, i.e. there is indeed no linear relationship between $\hat{\imath}$, $\hat{\jmath}$ and $\hat{k}$.

The other important property of the unit vectors $\hat{\imath}$, $\hat{\jmath}$ and $\hat{k}$ is their *completeness*. This is the property expressed by Eq. (16) which is true for *any* vector $\vec{a}$.

A set of orthogonal unit vectors like $\hat{\imath}$, $\hat{\jmath}$ and $\hat{k}$, which are linearly independent and complete, are called *base vectors*.

An interesting example of a set of base vectors are the vectors appropriate for polar coordinates:

$$\begin{aligned}
\hat{r} &= (\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta), \\
\hat{\theta} &= (\cos\theta\cos\phi, \cos\theta\sin\phi, -\sin\theta), \\
\hat{\phi} &= (\sin\phi, \cos\phi, 0)
\end{aligned}$$

It is left as an exercise to check that these three vectors have all the properties required of a set of base vectors.

The question of linear independence has a striking geometrical interpretation. For instance, if three vectors lie in one plane, then they are linearly dependent. Vectors which lie in one plane are said to be *coplanar*. Three non-coplanar vectors are linearly independent. More than three vectors in a three-dimensional space are always linearly dependent. It is left as an exercise to check all statements made in this paragraph are consistent with the algebraic definition of linear (in)dependence.

5. Vector product of two vectors;

Definition of vector product; distributive law of vector product; non-commutative law of vector product. Examples: moment of force (torque); angular momentum.

The vector product of $\vec{a}$ and $\vec{b}$ is defined as the vector $\vec{c}$, expressed in terms of the components of $\vec{a}$ and $\vec{b}$ by

$$\vec{c} = (a_y b_z - a_z b_y, a_z b_x - a_x b_z, a_x b_y - a_y b_z) \tag{17}$$

The vector product has its name because it is a vector, *i.e.* because its components transform under rotations of the coordinate axes like the components of a vector. We shall give the proof of this statment later on in this course in the chapter on matrices. Another name for the vector product is *cross product* because it is usually written in the form of

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

The cross product is distributive:

$$\vec{a} \times (\vec{b} + \vec{c}) = \vec{a} \times \vec{b} + \vec{a} \times \vec{b} \tag{18}$$

The proof is straightforward. It can be done for each component separately. Thus, for instance, for the x component we have

$$\begin{aligned}
[\vec{c} \times (\vec{a} + \vec{b})]_x &= c_y(a_z + b_z) - c_z(a_y + b_y) = c_y a_z + c_y b_z - c_z a_y - c_z b_y \\
&= (c_y a_z - c_z a_y) + (c_y b_z - c_z b_y) \\
&= (\vec{c} \times \vec{a})_x + (\vec{c} \times \vec{b})_x
\end{aligned}$$

and similarly for the y and z components.

The cross product is *not* associative:

$$\vec{a} \times (\vec{b} \times \vec{c}) \neq (\vec{a} \times \vec{b}) \times \vec{c} \tag{19}$$

The proof is straight forward. Consider, for instance, the x components of the two vectors:

$$(\vec{a} \times (\vec{b} \times \vec{c}))_x = a_y(\vec{b} \times \vec{c})_z - a_z(\vec{b} \times \vec{c})_y$$

and

$$((\vec{a} \times \vec{b}) \times \vec{c})_x = (\vec{a} \times \vec{b})_y c_z - (\vec{a} \times \vec{b})_z c_y = (a_z b_x - a_x b_z) c_z - \ldots$$

where we did not need to write any more terms because we can see that the former expression does not contain the x component of $\vec{a}$ whereas the latter expression does contain a_x.

The cross product is *not* commutative. More precisely, the cross product *anticommutes*:

$$\vec{a} \times \vec{b} = -\vec{b} \times \vec{a} \tag{20}$$

This can be shown separately for each component; for instance for the x component we have:

$$(\vec{a} \times \vec{b})_x = a_y b_z - a_z b_y$$

whereas

$$(\vec{b} \times \vec{a})_x = b_y a_z - b_z a_y = -(\vec{a} \times \vec{b})_x$$

and similarly for the other two components.

Consider the cross products of the base vectors $\hat{\imath} = (1, 0, 0)$, $\hat{\jmath} = (0, 1, 0)$, and $\hat{k} = (0, 0, 1)$:

$$\hat{\imath} \times \hat{\jmath} = (i_x, i_y, i_z) \times (j_x, j_y, j_z) = (i_y j_z - i_z j_y, i_z j_x - i_x j_z, i_x j_y - i_y j_x) = (0, 0, 1) = \hat{k}$$

and similarly $\hat{\jmath} \times \hat{k} = \hat{\imath}$ and $\hat{k} \times \hat{\imath} = \hat{\jmath}$. We see in these three examples that the cross product of $\hat{\imath}$ and $\hat{\jmath}$ is a vector, whose direction is orthogonal to the plane spanned by the vectors $\hat{\imath}$ and $\hat{\jmath}$, and similarly in the other two cases.

Next consider the cross product of $\hat{\imath}$ with itself. By direct calculation we find that

$$\hat{\imath} \times \hat{\imath} = 0$$

and similarly $\hat{\jmath} \times \hat{\jmath} = \hat{k} \times \hat{k} = 0$.

We can show more generally that

the cross product of any two vectors $\vec{a}$ and $\vec{b}$ is orthogonal to the plane spanned by these two vectors

and

the cross product of any two collinear vectors is the null vector.

The former one of these statements can be shown fairly simply for two vectors that lie in a coordinate plane, for instance for the vectors $\vec{a} = (a_x, a_y, 0)$ and $\vec{b} = (b_x, b_y, 0)$. We immediately see that the x and y components of their cross product are equal to nought because they have factors of a_z and b_z in them, which are both equal to nought. Thus the only nonzero component is the z component, which is $(\vec{a} \times \vec{b})_z = a_x b_y - a_y b_x$.

To see that the statement is generally true we appeal again to an intuitive argument: if the coordinate system is rotated such that the vectors $\vec{a}$ and $\vec{b}$ do not lie in a coordinate plane after the rotation, then the orientation of their cross product in relation to $\vec{a}$ and $\vec{b}$ remains unchanged, i.e. the cross product of $\vec{a}$ and $\vec{b}$ is always orthogonal to $\vec{a}$ and $\vec{b}$. This will be proved rigorously in the chapter on matrix algebra.

The discussion of the direction of the cross product is incomplete without spelling out the rule concerning the *sense* in which it points along the normal to the plane spanned by $\vec{a}$ and $\vec{b}$. This rule is as follows:

> The cross product of $\vec{a} \times \vec{b}$ is orthogonal to the plane spanned by $\vec{a}$ and $\vec{b}$ such that, looking down from the product vector into their plane, $\vec{a}$ can be rotated into the direction of $\vec{b}$ in *anticlockwise* direction through an angle of less than 180°.

To prove the statement that the cross product of collinear vectors is the null vector we note that collinear vectors are related by $\vec{b} = \lambda \vec{a}$, where λ is a scalar factor. Thus we have

$$\vec{a} \times \vec{b} = \lambda(\vec{a} \times \vec{a})$$

and it is easy to see by direct calculation that all components of the cross product of $\vec{a}$ with itself are equal to nought.

An interesting formula can be found for the modulus of the cross product:

$$|\vec{a} \times \vec{b}| = ab \sin\theta \tag{21}$$

where θ is the angle between $\vec{a}$ and $\vec{b}$. This statement is easy to show for vectors $\vec{a}$ and $\vec{b}$ that lie in a coordinate plane, for instance for $\vec{a} = (a_x, a_y, 0)$ and $\vec{b} = (b_x, b_y, 0)$. Assuming that $\vec{a}$ makes an angle α with the x axis we can also write $\vec{a} = a(\cos\alpha, \sin\alpha, 0)$, and similarly, if β is the angle that $\vec{b}$ makes with the x axis, we have $\vec{b} = b(\cos\beta, \sin\beta, 0)$, and hence

$$\vec{a} \times \vec{b} = ab(0, 0, \cos\alpha \sin\beta - \sin\alpha \cos\beta) = ab(0, 0, \sin(\alpha - \beta))$$

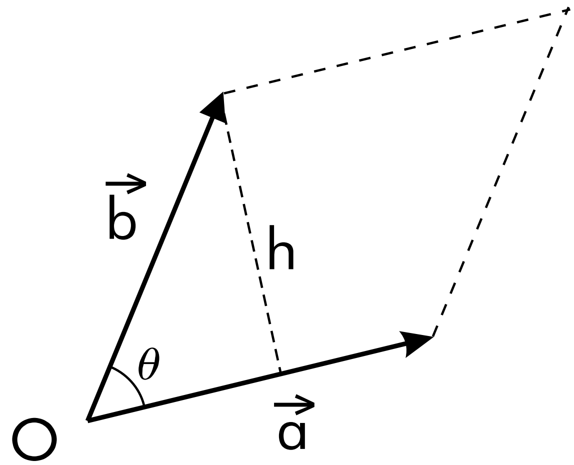

Figure 6: To derive $|\vec{a} \times \vec{b}| = ab\,\sin\theta$.

and since $(\alpha - \beta) = \theta$ is the angle between $\vec{a}$ and $\vec{b}$ we have proved the statement.

If $\vec{a}$ and $\vec{b}$ do not lie in a coordinate plane we invoke the by now familar argument that they can be brought into a coordinate plane by a rotation without changing their magnitudes and the angle between them, so the statement is true generally.

The result that the modulus of the cross product of two vectors is equal to the product of their moduli times the sine of the angle between them gives rise to an interesting geometrical interpretation. Indeed, consider the vectors $\vec{a}$ and $\vec{b}$ and draw the parallelogram as shown in Fig. 6. *Fig. 6*

The height h of the parallelogram is equal to $b\sin\theta$, and the base is a, and therefore the area of the parallelogram is $ah = ab\sin\theta$. We can therefore conclude that

> The modulus of the cross product of $\vec{a}$ and $\vec{b}$ is equal to the area of the parallelogram spanned by $\vec{a}$ and $\vec{b}$.

Let us return to the breakdown of the associative law of the vector product. We can see it now by noting that the two vectors in Eq. (19) lie in different planes: vector $\vec{a} \times (\vec{b} \times \vec{c})$ lies in the plane spanned by $\vec{b}$ and $\vec{c}$, whereas $(\vec{a} \times \vec{b}) \times \vec{c}$ lies in the plane spanned by $\vec{a}$ and $\vec{b}$.

6. Applications:

In this section we consider some applications of the concept of vectors in geometry.

6.1. As a first example consider the vector equation of a line:

$$\vec{r} = \lambda\hat{n} + \vec{a} \tag{22}$$

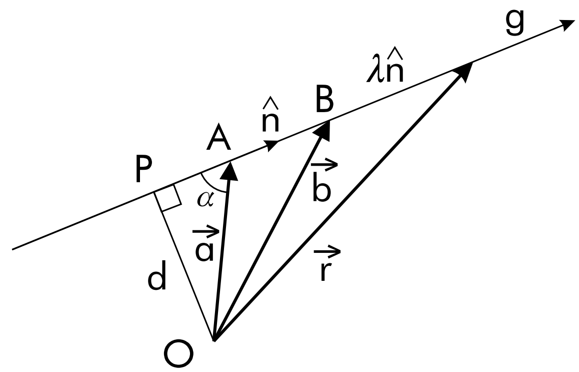

Figure 7: To derive Eq. (22)

where $\vec{a}$ is a constant vector, $\hat{n}$ is a fixed unit vector and λ is a real parameter that takes on any value from $-\infty$ to ∞. To show that this is indeed the equation of a line we proceed like this. Let A and B be two points with vectors $\vec{a}$ and $\vec{b}$, respectively (see Fig. 7). The straight line through these points can be constructed by first going to A and then in the direction of $(\vec{a}-\vec{b})$ by a variable distance. Setting $\hat{n}=(\vec{a}-\vec{b})/|\vec{a}-\vec{b}|$ and measuring the variable distance by λ we get the required equation. *Fig. 7*

Let us denote the line by g. The distance of g from the origin is equal to the length of the perpendicular, dropped from the origin onto g. From the triangle OAP we see that this length is $d=a\sin\alpha$, where α is the angle between $\vec{a}$ and $\hat{n}$. But $a\sin\alpha=|\vec{a}\times\hat{n}|$, hence

$$d=|\vec{a}\times\hat{n}|$$

<u>Exercise:</u> Denote the vector that is perpendicular to g by $\vec{r}_0$. This vector is also perpendicular to $\hat{n}$. To lie on g it must satisfy the equation of g with a parameter λ_0 to be determined. We get λ_0 by imposing the orthogonality condition:

$$\vec{r}_0\cdot\hat{n}=(\vec{a}+\lambda_0\hat{n})\cdot\hat{n}=0$$

hence $\lambda_0=-\vec{a}\cdot\hat{n}$. The distance of g from the origin is the length of $\vec{r}_0$, i.e.

$$|\vec{r}_0|=|\vec{a}+\lambda_0\hat{n}|=|\vec{a}-(\vec{a}\cdot\hat{n})\hat{n}|$$

Show that this expression for the distance is consistent with the one derived previously.

Let us also find the distance of a point M with vector $\vec{b}$ from g (see Fig. 8). Denote the foot of the perpendicular dropped from M onto g by Q. From the triangle AMQ we see that $MQ=AM\sin\alpha$ or, in vector notation, *Fig. 8*

$$d=|\vec{r}_1-\vec{b}|=|\vec{a}-\vec{b}|\sin\alpha=|(\vec{a}-\vec{b})\times\hat{n}|$$

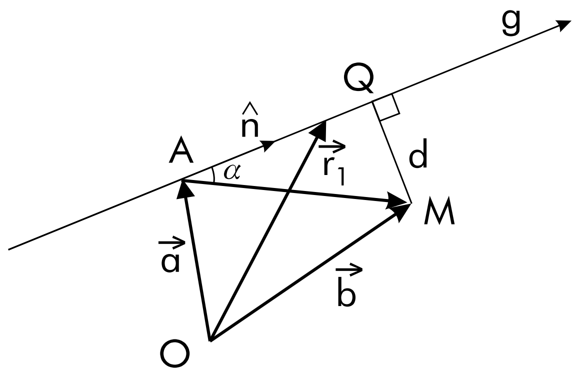

Figure 8: Distance of point M from line g.

6.2. Vector equation of a plane:

The equation of a plane through the points given by the vectors $\vec{a}$, $\vec{b}$ and $\vec{c}$ is

$$\vec{r} = \lambda\vec{a} + \mu\vec{b} + \nu\vec{c}, \qquad \text{where} \quad \lambda + \mu + \nu = 1 \tag{23}$$

This can be seen in the following way. First note that the vectors $\vec{a}-\vec{b}$, $\vec{b}-\vec{c}$ and $\vec{c}-\vec{a}$ are collinear. This is obvious from their geometrical representation, but can be easily checked by forming their triple scalar product.

Then consider the vector $\vec{r} = \vec{a}+\alpha(\vec{a}-\vec{b})+\beta(\vec{b}-\vec{c})$ where α and β are arbitrary constants. This vector points to a point in the plane which is spanned by $\vec{a}-\vec{b}$ and $\vec{b}-\vec{c}$ and passes through the point given by $\vec{a}$. Rearranging we get $\vec{r} = (1+\alpha)\vec{a} + (\beta-\alpha)\vec{b} - \beta\vec{c}$, hence, if we put $\lambda = 1-\alpha$, $\mu = \beta - \alpha$ and $\nu = -\beta$, we get Eq. (23).

7. Triple products;

Triple scalar product; triple vector product.

7.1. The scalar product of a vector $\vec{a}$ with the cross product $\vec{b}\times\vec{c}$ is a scalar. It is called the *triple scalar product* of the vectors $\vec{a}$, $\vec{b}$ and $\vec{c}$. If the angle between $\vec{b}$ and $\vec{c}$ is γ we have

$$\vec{b}\times\vec{c} = \hat{n}bc\sin\gamma$$

where $\hat{n}$ is the unit vector normal to the plane spanned by $\vec{b}$ and $\vec{c}$, and we know (see section 6) that $|\vec{b}\times\vec{c}|$ is the area of the parallelogram based on $\vec{b}$ and $\vec{c}$. Taking the dot product of $\vec{a}$ with $\vec{b}\times\vec{c}$ we get

$$\vec{a}\cdot(\vec{b}\times\vec{c}) = (\vec{a}\cdot\hat{n})bc\sin\gamma$$

but $h = \vec{a}\cdot\hat{n} = a\cos\delta$ is the height of the parallelepiped based on the vectors $\vec{a}$, $\vec{b}$ and $\vec{c}$ (see Fig. 9), and therefore we have the result that the triple scalar product $\vec{a}\cdot(\vec{b}\times\vec{c})$ is the *Fig. 9*

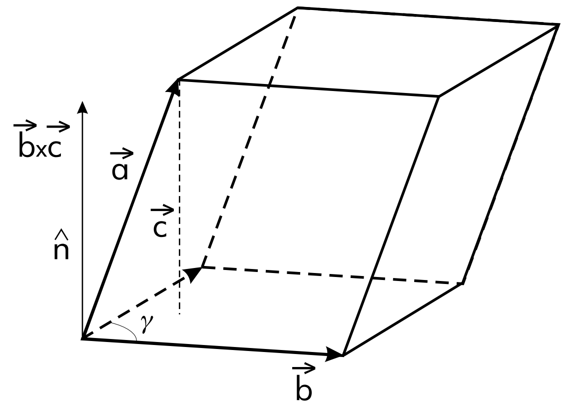

Figure 9: Triple scalar product.

volume of this parallelepiped.

Our discussion is deficient in so far as we can get a negative value for the volume. This is the case when δ is greater than $\pi/2$, i.e. when the vectors $\vec{a}$, $\vec{b}$ and $\vec{c}$, in that order, form a left-handed system. Thus, to get the volume as a positive quantity irrespective of the handedness of the three vectors, we say that the volume of the parallelepiped is equal to the *modulus* of the triple product:

$$V_{\text{parallelepiped}} = |\vec{a} \cdot (\vec{b} \times \vec{c})| \tag{24}$$

An interesting symmetry property of the triple scalar product is expressed by the equations

$$\vec{a} \cdot (\vec{b} \times \vec{c}) = \vec{b} \cdot (\vec{c} \times \vec{a}) = \vec{c} \cdot (\vec{a} \times \vec{b}) \tag{25}$$

which can be verified by direct evaluation of the three triple products. We can express this property in words by saying that *the triple product does not change under a cyclic permutation of the vectors.* Because of this symmetry one frequently uses the simplified, and more symmetric, notation $(\vec{a}, \vec{b}, \vec{c})$, i.e.

$$(\vec{a}, \vec{b}, \vec{c}) = \vec{a} \cdot (\vec{b} \times \vec{c}) = \vec{b} \cdot (\vec{c} \times \vec{a}) = \vec{c} \cdot (\vec{a} \times \vec{b}) \tag{26}$$

Note however that the sign of the triple product changes under an odd permutation of the vectors, e.g.

$$(\vec{a}, \vec{b}, \vec{c}) = -(\vec{a}, \vec{c}, \vec{b}) \tag{27}$$

From the latter equation we get the result that the triple scalar product vanishes if it contains one vector twice, for instance

$$(\vec{a}, \vec{b}, \vec{b}) = 0 \tag{28}$$

Indeed, if we set $\vec{c} = \vec{b}$ in Eq. (27), we get $(\vec{a}, \vec{b}, \vec{b}) = -(\vec{a}, \vec{b}, \vec{b})$, and since the only number that is equal to minus itself is zero, we get the result of Eq. (28).

More generally we can show that the triple scalar product vanishes if the three vectors are coplanar. Indeed, if they are coplanar, then we can write, for instance,

$$\vec{c} = \lambda\vec{a} + \mu\vec{b}$$

and hence

$$(\vec{a}, \vec{b}, \vec{c}) = (\vec{a}, \vec{b}, \lambda\vec{a} + \mu\vec{b}) = \lambda(\vec{a}, \vec{b}, \vec{a}) + \mu(\vec{a}, \vec{b}, \vec{b}) = 0 \tag{29}$$

because both scalar triple products on the right-hand side are equal to nought by Eq. (28).

The converse statement, that the vectors $\vec{a}$, $\vec{b}$ and $\vec{c}$ are coplanar if their triple scalar product is equal to nought, can also be proved. Thus the vanishing of the triple scalar product is the necessary and sufficient condition for the vectors to be coplanar.

Exercise: Prove that the vectors $\vec{a}$, $\vec{b}$ and $\vec{c}$ are coplanar if their triple scalar product is equal to nought.

An interesting extension of the scalar triple product arises if we take the dot product of two cross products. Then one can prove the following identity

$$(\vec{a} \times \vec{b}) \cdot (\vec{c} \times \vec{d}) = (\vec{a} \cdot \vec{c})(\vec{b} \cdot \vec{d}) - (\vec{a} \cdot \vec{d})(\vec{b} \cdot \vec{c}) \tag{30}$$

This identity, known as *Lagrange's identity*, can be proved by explicit evaluation. This is straight forward but tedious. More interesting is the following proof.
We first note that the left-hand side is a scalar, made of the four vectors $\vec{a}$ to $\vec{d}$, and that it depends linearly on each of the four vectors. Thus, if we want to rewrite it in terms of dot products, we must consider the following combinations:

$$(\vec{a} \cdot \vec{b})(\vec{c} \cdot \vec{d}), \quad (\vec{a} \cdot \vec{c})(\vec{b} \cdot \vec{d}) \quad \text{and} \quad (\vec{a} \cdot \vec{d})(\vec{b} \cdot \vec{c}) \tag{31}$$

Next we note that $(\vec{a} \times \vec{b}) \cdot (\vec{c} \times \vec{d})$ changes sign under the replacement $\vec{a} \leftrightarrow \vec{b}$ and under $\vec{c} \leftrightarrow \vec{d}$. Therefore we must discard the first one of the three expressions in Eq. (31), which changes into itself under either one of these transformations. The second and third ones of these expressions change into each other. We therefore get a change of sign under either one of the replacements if we take the difference of these two expressions. There remains an overall sign to fix. This can be done by choosing a particular case, for instance by setting $\vec{a} = \vec{c} = \hat{\imath}$ and $\vec{b} = \vec{d} = \hat{\jmath}$.

7.2. The triple vector product $(\vec{a} \times \vec{b}) \times \vec{c}$ has the interesting property that

$$(\vec{a} \times \vec{b}) \times \vec{c} = (\vec{a} \cdot \vec{c})\vec{b} - (\vec{b} \cdot \vec{c})\vec{a} \tag{32}$$

Proof: we note that $(\vec{a} \times \vec{b})$ is orthogonal to the plane spanned by $\vec{a}$ and $\vec{b}$, and therefore its cross product with $\vec{c}$ lies in that plane. Therefore the triple cross product is a linear combination of $\vec{a}$ and $\vec{b}$, i.e.

$$(\vec{a} \times \vec{b}) \times \vec{c} = \lambda\vec{a} + \mu\vec{b}$$

where the scalar factors λ and μ must depend linearly on $\vec{b}$ and $\vec{c}$ and on $\vec{a}$ and $\vec{c}$, respectively. They must therefore be of the form of $\alpha(\vec{b}\cdot\vec{c})$ and $\beta(\vec{a}\cdot\vec{c})$, where α and β are numerical factors. Now, the triple cross product changes sign under the replacement $\vec{a}\leftrightarrow\vec{b}$. Therefore we must have $\beta=-\alpha$, i.e.

$$(\vec{a}\times\vec{b})\times\vec{c}=\alpha[(\vec{a}\cdot\vec{c})\vec{b}-(\vec{b}\cdot\vec{c})\vec{a}]$$

The remaining numerical factor can be found by considering a particular case, e.g. $\vec{a}=\vec{c}=\hat{\imath}$ and $\vec{b}=\hat{\jmath}$.

Exercise: Use the identity (32) to prove that

$$\vec{a}\times(\vec{b}\times\vec{c})+\vec{b}\times(\vec{c}\times\vec{a})+\vec{c}\times(\vec{a}\times\vec{b})=0$$

An interesting extension of the triple vector product is the expression $(\vec{a}\times\vec{b})\times(\vec{c}\times\vec{d})$, for which we can deduce the following identities:

$$\begin{aligned}\vec{f}\equiv(\vec{a}\times\vec{b})\times(\vec{c}\times\vec{d}) &= \vec{b}(\vec{a},\vec{c},\vec{d})-\vec{a}(\vec{b},\vec{c},\vec{d})\\ &= \vec{c}(\vec{a},\vec{b},\vec{d})-\vec{d}(\vec{a},\vec{b},\vec{c})\end{aligned}\tag{33}$$

which shows that the vector $\vec{f}$ lies both in the plane spanned by $\vec{a}$ and $\vec{b}$ and in the plane spanned by $\vec{c}$ and $\vec{d}$, i.e. it lies along the intersection of these two planes.

8. Coordinate transformations.

Transformation properties of vectors under rotations; invariance of the dot product under rotations; reflections; polar and axial vectors.

In this section we discuss coordinate transformations in two dimensions. The generalization to the three dimensional case will be done in the chapter on matrix algebra.

Thus consider two Cartesian coordinate frames (xy) and (x',y') with common origin O (see Fig. 10). The x' axis (y' axis) makes the angle α with the x (y) axis. The base *Fig. 10* vectors in the xy frame are denoted $\hat{\imath}$ and $\hat{\jmath}$, those in the primed frame are $\hat{\imath}'$ and $\hat{\jmath}'$. Thus we have the following set of dot products:

$$\begin{aligned}\hat{\imath}^2 &= \hat{\jmath}^2=\hat{\imath}'^2=\hat{\jmath}'^2=1, \qquad \hat{\imath}\cdot\hat{\jmath}=\hat{\imath}'\cdot\hat{\jmath}'=0,\\ \hat{\imath}\cdot\hat{\imath}' &= \hat{\jmath}\cdot\hat{\jmath}'=\cos\alpha,\\ \hat{\imath}\cdot\hat{\jmath}' &= \cos\left(\frac{\pi}{2}+\alpha\right)=-\sin\alpha\\ \hat{\imath}'\cdot\hat{\jmath} &= \cos\left(\frac{\pi}{2}-\alpha\right)=\sin\alpha\end{aligned}\tag{34}$$

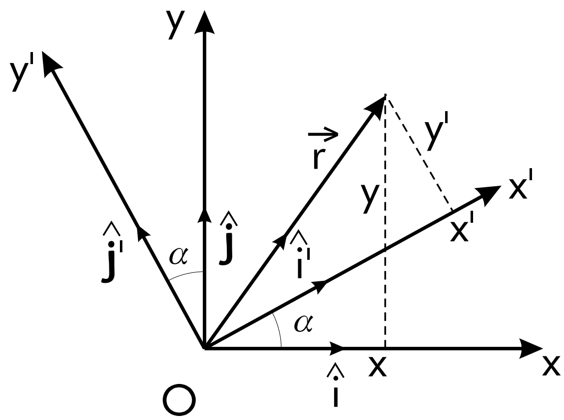

Figure 10: Rotation of coordinate frame.

Now consider the point P with radius vector $\vec{r}$. In the xy frame it is represented by $\vec{r} = x\hat{\imath} + y\hat{\jmath}$, and in the $x'y'$ frame by $\vec{r}\,' = x'\,\hat{\imath}' + y'\,\hat{\jmath}'$, and since these are different representations of the same vector we have the identity

$$\vec{r} = x\hat{\imath} + y\hat{\jmath} = x'\,\hat{\imath}' + y'\,\hat{\jmath}'$$

Taking the dot products first with $\hat{\imath}$ and then with $\hat{\jmath}$ we get

$$\begin{aligned} x &= x'\,\hat{\imath}'\cdot\hat{\imath} + y'\,\hat{\jmath}'\cdot\hat{\imath} \\ y &= x'\,\hat{\imath}'\cdot\hat{\jmath} + y'\,\hat{\jmath}'\cdot\hat{\jmath} \end{aligned}$$

and hence, with Eq. (34), we get

$$\begin{aligned} x &= x'\,\cos\alpha - y'\,\sin\alpha \\ y &= x'\,\sin\alpha + y'\,\cos\alpha \end{aligned} \qquad (35)$$

and similarly we get the formulas for the inverse transformation:

$$\begin{aligned} x' &= x\cos\alpha + y\sin\alpha \\ y' &= -x\sin\alpha + y\cos\alpha \end{aligned} \qquad (36)$$

Another important transformation is the *reflection at the origin*. This is defined as the transformation

$$x \to x' = -x, \qquad y \to y' = -y, \qquad \text{and} \qquad z \to z' = -z \qquad (37)$$

Under this transformation the vector $\vec{r}$ changes its sign:

$$\vec{r} \to \vec{r}\,' = -\vec{r} \qquad (38)$$

where $\vec{r}$ and $\vec{r}\,'$ refer to the same physical vector.

We can now restate our original definition of a vector more precisely:

a vector is a triplet of numbers, the components of the vector, which transform under rotations of the coordinate axes and under reflections like the components of $\vec{r}$

Let us examine various quantities which we have considered in this chapter as to their transformation properties.[2]

Consider how the modulus of $\vec{a}$ transforms under rotations:

$$\begin{aligned} |\vec{a}|^2 &= a_x^2 + a_y^2 \rightarrow a_x'^2 + a_y'^2 \\ &= (a_x \cos\alpha + a_y \sin\alpha)^2 + (-a_x \sin\alpha + a_y \cos\alpha)^2 \\ &= a_x^2 + a_y^2 \end{aligned}$$

and one says that the modulus of a vector is *invariant* under rotations. Similarly one can show that the modulus is invariant under reflections. These invariance properties are characteristic of *scalars*. One *defines* a scalar as a quantity that is invariant under rotations and reflections.

It is left as an exercise to verify that the dot product of two vectors is invariant under rotations and reflections. Thus the dot product of two vectors is a scalar.

Next consider the cross product of two vectors which lie in the xy plane. This vector has only a z component which we expect to be unaffected by a rotation in the xy plane. This can be checked using the transformation formulas (35). The question as to the full three-dimensional rotations must be postponed, but anticipating the result of that calculation we can say that the cross product of two vectors transforms under rotations like a vector. This is the justification for calling it a vector.

Let us also consider the behaviour of the cross product under reflections. We have

$$\vec{a} \times \vec{b} \rightarrow \vec{a}' \times \vec{b}' = (-\vec{a}) \times (-\vec{b}) = \vec{a} \times \vec{b}$$

i.e. we get the important result that under reflections the cross product of two vectors does *not* behave like a vector, i.e. it does not change its sign! This result gives rise to the distinction between two classes of vectors: those that do and those that do not change sign under reflections. The former are called *polar* vectors, whereas the latter are called *axial* vectors (or sometimes *pseudo vectors*). These names have their origin in applications of vector calculus in mechanics.

Of fundamental importance in mechanics are the polar vectors $\vec{r}$, the momentum $\vec{p}$ and the force $\vec{F}$. Examples of axial vectors in mechanics are the moment of force (or *torque*) $\vec{M} = \vec{r} \times \vec{F}$ and the angular momentum $\vec{J} = \vec{r} \times \vec{p}$.

Finally let us sharpen up the statement that the cross product of two vectors is an axial vector. In this statement it was tacitly assumed that the two vectors were polar vectors.

[2]for the time being we consider rotations only in the xy plane; reflections are done in three dimensions.

One can easily check that the cross product of two axial vectors is an axial vector and that the cross product of an axial vector and a polar vector is a polar vector. Let us denote for the sake of this discussion a general polar vector by $\vec{P}$ and a general axial vector by $\vec{A}$. Then we can formally express this result by the following scheme:

$$\vec{P} \times \vec{P} = \vec{A} \times \vec{A} = \vec{A}, \qquad \vec{P} \times \vec{A} = \vec{P}$$

The distinction between polar and axial vectors has also a bearing on the notion of scalars. The statement that the dot product of two vectors is invariant under reflections also needed the tacit assumption that the two vectors were polar vectors. If, on the other hand, we consider the dot product of a polar vector with an axial vector, then we get a quantity which changes sign under reflections (but is invariant under rotations). We therefore also distinguish between two classes of scalars, those which do not and those which do change sign under reflections. The former are called scalars, whereas the latter are called *pseudo scalars*.

9.) Exercises and problems.

1) Evaluate the moduli of the vectors $\vec{a}$ and $\vec{b}$, the scalar products $\vec{a}\cdot\vec{b}$ and find the angles between the $\vec{a}$ and $\vec{b}$ if

1.1 $\vec{a} = (2,\ -3,\ 1),\ \vec{b} = (4,\ 7,\ 5);$

1.2 $\vec{a} = \frac{1}{5}(3,\ -4,\ 0),\ \vec{b} = \frac{1}{\sqrt{14}}(2,\ 1,\ 3)$

2) Evaluate and compare the vector expressions $\vec{a}\cdot(\vec{b}+\vec{c})$ and $\vec{a}\cdot\vec{b}+\vec{a}\cdot\vec{c}$ if

2.1 $\vec{a} = (3,\ -2,\ -1),\ \vec{b} = (2,\ 3,\ 7),\ \vec{c} = (-4,\ 1,\ 3);$

2.2 $\vec{a} = (2.3,\ 1.2,\ 1.7),\ \vec{b} = (0.2,\ -3.5,\ 0.7),\ \vec{c} = (2.4,\ 3.1,\ -1.3).$

3) Evaluate the vector products $\vec{a}\times\vec{b}$ for the vectors of Q1. Using the expression $|\vec{a}\times\vec{b}| = ab\sin\theta$ find the angles theta between the vectors $\vec{a}$ and $\vec{b}$ and compare the results with the angles found in Q1.

4) Using the results of Questions 1 and 3 show that

$$(\vec{a}\times\vec{b})^2 + (\vec{a}\cdot\vec{b})^2 = a^2b^2$$

5) Show that the kinetic energy $T = \dfrac{p^2}{2m}$ of a particle of mass m travelling with momentum $\vec{p}$ can be represented as

$$T = \frac{p_r^2}{2m} + \frac{J^2}{2mr^2}$$

where $p_r = \dfrac{\vec{r}\cdot\vec{p}}{|\vec{r}|}$ is the radial momentum and $\vec{J} = \vec{r}\times\vec{p}$ is the angular momentum of the particle.

6) Use the triple scalar product to find the equation of a circle that lies on the intersection of a unit sphere with a plane through the centre of the sphere and passes through two points given by the unit vectors $\vec{a} = (\sin\theta_a\,\cos\phi_a, \sin\theta_a\,\sin\phi_a, \cos\theta_a)$ and $\vec{b} = (\sin\theta_b\,\cos\phi_b, \sin\theta_b\,\sin\phi_b, \cos\theta_b)$. The equation should be in the form of a relation between the polar angle θ and azimuth ϕ of a general point on the circle.
Find the minimum and maximum values of θ if $\theta_a = \pi/3$, $\phi_a = 0$, $\theta_b = 2\pi/3$ and $\phi_b = \pi/2$.
Answer: $\theta \in [50.77^o, 129.23^o]$ or $|\cot\theta| \le \sqrt{2/3}$.

Part 2. Vector Fields

---

1. Definition of scalar and vector fields.

In physics one calls a function of x, y, and z a field. We shall study in this section *scalar* and *vector* fields.

Examples of scalar fields are the gravitational and the electrostatic potentials. Examples of vector fields are the gravitational field, the electrostatic field and the velocity field of a fluid.

We shall usually denote scalar fields by $\Phi(x, y, z)$ or $\Phi(\vec{r})$, and vector fields by $\vec{A}(\vec{r})$ or $\vec{E}(\vec{r})$ etc.

A scalar field is characterised by a single value at each point in space.

A vector field is characterised by its magnitude and direction at each point in space.

To get an intuitive picture of scalar and vector fields one draws *equipotential lines* for scalar fields and *field lines* for vector fields.

In physics one finds empirically relations between different fields, for instance between magnetic field and current. Here we study the mathematical form of such relationships.

2. Gradient of a scalar field; potential.

Consider a scalar field, such as the gravitational potential or the electrostatic potential. Denote the scalar field by $\Phi(\vec{r}) = \Phi(x, y, z)$. Keep z fixed at some value z_0, i.e. consider the scalar field in a plane parallel to the (x, y) plane, and consider the relation

$$\Phi(x, y, z_0) = c$$

where c is a constant. This is an implicit equation of y as a function of x. If we represent this function in the (x, y) plane, we get a line at fixed potential or *equipotential line*. We get a set of equipotential lines if we let c assume different values (Fig. 1). Such a set of equipotential lines are also called "*contour lines*". If now z is allowed to vary continuously we get a set of *equipotential surfaces*.

Examples:

1. Let $\Phi(\vec{r}) = r^2$, then the equation of the equipotential surfaces takes on the form

$$\Phi(\vec{r}) = r^2 = x^2 + y^2 + z^2 = c$$

   i.e. the equation of a family of spherical surfaces.

2. Let

$$\Phi(\vec{r}) = \frac{1}{|\vec{r} - \vec{r}_0|} - \frac{1}{|\vec{r} + \vec{r}_0|}$$

   i.e. the potential of two unlike point charges at $\vec{r}_0$ and $-\vec{r}_0$. The contour lines for $z = 0$ are as shown in Fig. 2.

If we move along a contour line the potential does not change. If we move at an angle to an equipotential line we observe a change of the potential. The fastest variation of the potential is observed when we move at right angles to the contour lines. This is an intuitively obvious statement. We are now going to put it on a rigorous mathematical footing. To do this we introduce the concept of the *gradient* of a scalar field.

Consider the differential of a scalar field $\Phi(\vec{r})$. By the rules of calculus of functions of several variables we have

$$d\Phi(\vec{r}) = \frac{\partial \Phi}{\partial x}dx + \frac{\partial \Phi}{\partial y}dy + \frac{\partial \Phi}{\partial z}dz \tag{39}$$

and we see that this expression has the form of a scalar product of two vectors,

$$\left(\frac{\partial \Phi}{\partial x}, \frac{\partial \Phi}{\partial y}, \frac{\partial \Phi}{\partial z}\right) \tag{40}$$

and

$$(dx,\, dy,\, dz) \equiv d\vec{r}$$

The vector (40) is called the *gradient* of $\Phi(\vec{r})$ and is denoted $\mathrm{grad}\,\Phi(\vec{r})$. Thus

$$\mathrm{grad}\,\Phi(\vec{r}) = \left(\frac{\partial \Phi}{\partial x}, \frac{\partial \Phi}{\partial y}, \frac{\partial \Phi}{\partial z}\right) \tag{41}$$

and we can rewrite Eq. (39) as

$$d\Phi(\vec{r}) = \text{grad}\Phi(\vec{r}) \cdot d\vec{r} \tag{42}$$

Examples:

3.
$$\text{grad}\, r = \left(\frac{\partial r}{\partial x}, \frac{\partial r}{\partial y}, \frac{\partial r}{\partial z}\right)$$

but $r = \sqrt{x^2+y^2+z^2}$, hence $\dfrac{\partial r}{\partial x} = \dfrac{x}{\sqrt{x^2+y^2+z^2}} = \dfrac{x}{r}$, and similarly $\dfrac{\partial r}{\partial y} = \dfrac{y}{r}$ and $\dfrac{\partial r}{\partial z} = \dfrac{z}{r}$, hence

$$\text{grad}\, r = \left(\frac{x}{r}, \frac{y}{r}, \frac{z}{r}\right) = \frac{1}{r}\,(x,y,z) = \frac{\vec{r}}{r} = \hat{r}$$

4.
$$\text{grad}\frac{1}{r} = \left(\frac{\partial \frac{1}{r}}{\partial x}, \frac{\partial \frac{1}{r}}{\partial y}, \frac{\partial \frac{1}{r}}{\partial z}\right)$$

but

$$\frac{\partial \frac{1}{r}}{\partial x} = -\frac{1}{r^2}\frac{\partial r}{\partial x} = -\frac{1}{r^2}\frac{x}{r}, \qquad \frac{\partial \frac{1}{r}}{\partial y} = -\frac{1}{r^2}\frac{\partial r}{\partial y} = -\frac{1}{r^2}\frac{y}{r}, \qquad \frac{\partial \frac{1}{r}}{\partial z} = -\frac{1}{r^2}\frac{\partial r}{\partial z} = -\frac{1}{r^2}\frac{z}{r},$$

hence

$$\text{grad}\frac{1}{r} = -\frac{\vec{r}}{r^3} = -\frac{\hat{r}}{r^2}$$

Properties of the gradient.

If $u(\vec{r})$ and $v(\vec{r})$ are two scalar functions and a is a constant, then

(i)
$$\text{grad}(u(\vec{r}) + v(\vec{r})) = \text{grad}\, u(\vec{r}) + \text{grad}\, v(\vec{r}) \tag{43}$$

(ii)
$$\text{grad}(a u(\vec{r})) = a\, \text{grad}\, u(\vec{r}) \tag{44}$$

(iii)
$$\text{grad}(uv) = u\, \text{grad}\, v + v\, \text{grad}\, u \tag{45}$$

(iv)
$$\text{grad} f(u(\vec{r})) = \frac{f(u)}{du}\, \text{grad}\, u \tag{46}$$

(v)

$$\text{grad}\Phi(\vec{r}) \text{ is perpendicular to the equipotential surface } \Phi(\vec{r}) = c. \tag{47}$$

The proofs of properties (i) to (iv) are left to the reader as an exercise: they involve no more than the rules of partial differentiation of functions of several variables. The proof of property (v) proceeds as follows:
consider the equation of the equipotential surface $\Phi(\vec{r}) = c$. Take two points on this surface, separated by an infinitesimal displacement $\delta\vec{r}$, hence

$$d\Phi(\vec{r}) = \text{grad}\Phi(\vec{r}) \cdot \delta\vec{r} = 0$$

i.e. gradΦ is perpendicular to $\delta\vec{r}$. But in the limit when the two points coalesce, $\delta\vec{r} \to d\vec{r}$ lies in the tangential plane of the equipotential surface, and hence gradΦ is normal to the surface $\Phi = c$. Note that this is what we said about the gradient intuitively at the beginning of this section.

The directional derivative of $\Phi(\vec{r})$.

Related to the gradient is the *directional derivative* of a scalar field.

Consider the scalar field $\Phi(\vec{r})$, and let $\hat{s}$ be an arbitrary unit vector. The expression

$$\frac{\partial\Phi(\vec{r})}{\partial s} = \lim_{\varepsilon\to 0} \frac{\Phi(\vec{r}+\varepsilon\hat{s}) - \Phi(\vec{r})}{\varepsilon}$$

is called the (directional) derivative of $\Phi(\vec{r})$ in the direction of $\hat{s}$. To find the relationship between the directional derivative and the gradient, represent the unit vector $\hat{s}$ in terms of its directional cosines: $\hat{s} = (\cos\alpha, \cos\beta, \cos\gamma)$, hence

$$\Phi(\vec{r}+\varepsilon\hat{s}) = \Phi(x+\varepsilon\cos\alpha, y+\varepsilon\cos\beta, z+\varepsilon\cos\gamma)$$

and expanding this about $\varepsilon = 0$ we get

$$\Phi(\vec{r}+\varepsilon\hat{s}) = \Phi(\vec{r}) + \frac{\partial\Phi(\vec{r})}{\partial x}\varepsilon\cos\alpha + \frac{\partial\Phi(\vec{r})}{\partial y}\varepsilon\cos\beta + \frac{\partial\Phi(\vec{r})}{\partial z}\varepsilon\cos\gamma + \mathcal{O}(\varepsilon^2)$$

and hence

$$\frac{\Phi(\vec{r}+\varepsilon\hat{s}) - \Phi(\vec{r})}{\varepsilon} = \frac{\partial\Phi(\vec{r})}{\partial x}\cos\alpha + \frac{\partial\Phi(\vec{r})}{\partial y}\cos\beta + \frac{\partial\Phi(\vec{r})}{\partial z}\cos\gamma + \mathcal{O}(\varepsilon)$$

and in the limit of $\varepsilon \to 0$ the left-hand side becomes the directional derivative, and on the right-hand side we get the scalar product of $\Phi(\vec{r})$ with $\hat{s}$, thus

$$\frac{\partial\Phi(\vec{r})}{\partial s} = \hat{s}\cdot\text{grad}\,\Phi(\vec{r}) \tag{48}$$

Now, if the normal unit vector to the equipotential surface is denoted by $\hat{n}$, we can write $\text{grad}\,\Phi = \hat{n}|\text{grad}\,\Phi|$, and hence

$$\frac{\partial \Phi(\vec{r}\,)}{\partial s} = |\text{grad}\,\Phi(\vec{r}\,)|\cos\theta$$

where $\cos\theta = \hat{s}\cdot\hat{n}$. In the particular case when $\hat{s} = \hat{n}$, i.e. when the directional derivative is taken in the direction normal to the equipotential surface, we have

$$\frac{\partial \Phi(\vec{r}\,)}{\partial n} = |\text{grad}\,\Phi(\vec{r}\,)|$$

and because of $\cos\theta < 1$ we can also conclude that

$$|\text{grad}\,\Phi(\vec{r}\,)| = \max\frac{\partial \Phi(\vec{r}\,)}{\partial s}$$

i.e. the potential changes most rapidly in the direction of the normal to the equipotential surface ("direction of steepest descent").

Finally, let us remark that in physics one calls a vector field, which is the gradient of a scalar field, a *conservative* vector field.

3. Flux of a vector field; divergence.

Let S be a surface $z = f(x, y)$ and $\vec{A}(\vec{r}\,)$ a vector field. Consider a surface element dS whose normal is $\hat{n}$. The integral

$$\int_S \vec{A}(\vec{r}\,)\cdot\hat{n}\,dS$$

is called the flux of the vector field $\vec{A}(\vec{r}\,)$ through S. frequently one uses alternative notations. Observing that $A_n(\vec{r}\,) = \vec{A}(\vec{r}\,)\cdot\hat{n}$ is the normal component of $\vec{A}(\vec{r}\,)$ one can write for the flux

$$\int_S A_n(\vec{r}\,)dS$$

and denoting $\hat{n}\,dS$ by $d\vec{S}$ one can also write

$$\int_S \vec{A}(\vec{r}\,)\cdot d\vec{S}$$

In the particular case when the surface S is closed, the integral is usually denoted by the symbol $\oint_S$. The surface integral

$$\oint_S \vec{A}(\vec{r}\,)\cdot d\vec{S} = \oint_S \vec{A}(\vec{r}\,)\cdot\hat{n}dS$$

is the *net flux* of $\vec{A}$ *out of* the closed surface S, if $\hat{n}$ is the *outward* normal vector to dS. If all field lines of $\vec{A}$ begin and end outside of the volume V enclosed by S, then we expect the net flux to be zero. In this case one says that there are no sources of

$\vec{A}$ in V. On the other hand, if there is a source of $\vec{A}$ in V, such as for instance an electric charge that gives rise to a field $\vec{E}$, then the net flux has a positive value whose magnitude depends on the strength of the source. If the net flux out of V takes on a negative value, one says that the surface S contains a *sink* of the vector field $\vec{A}$. In the case of electric charge this corresponds to a negative charge.

Consider the expression $\frac{1}{V}\oint_S \vec{A}\cdot d\vec{S}$ and let the volume V shrink to a point P with coordinates $\vec{r}_0$, then

$$\operatorname{div}\vec{A}(\vec{r}_0) = \lim_{V\to 0}\frac{1}{V}\oint_S \vec{A}\cdot d\vec{S} \tag{49}$$

is called the *divergence* of $\vec{A}$ at the point P. From our preceding discussion we see that the divergence is a local measure of the strength of the sources of the field $\vec{A}$.

It is frequently convenient to use a representation of the divergence in terms of partial derivatives of the components of $\vec{A}$:

$$\operatorname{div}\vec{A}(\vec{r}\,) = \frac{\partial A_x}{\partial x} + \frac{\partial A_y}{\partial y} + \frac{\partial A_z}{\partial z} \tag{50}$$

To derive this formula choose V in the form of a rectangular volume with faces parallel to the coordinate planes, see Fig. NN. The surface integral becomes a sum of integrals over the six faces of V. Consider the two faces parallel to the yz-plane:

$$\int_{(1)} \vec{A}(x,y,z)\cdot(-\hat{x})dy\,dz + \int_{(2)} \vec{A}(x+\delta x,y,z)\cdot\hat{x}dy\,dz \quad =$$

$$= \int\left\{\left[A_x(x,y,z) + \frac{\partial A_x(x,y,z)}{\partial x}\,\delta x + \mathcal{O}(\delta x^2)\right] - A_x(x,y,z)\right\} dy\,dz$$

where we have expanded $A_x(x+\delta x,y,z)$ up to the linear term in δx. The terms $A_x(x,y,z)$ cancel and we are left with

$$\left[\int_y^{y+\delta y} dy \int_z^{z+\delta z} dz \frac{\partial A_x(x,y,z)}{\partial x}\right]\delta x = \frac{\partial A_x(x,y+\eta\delta y,z+\zeta\delta z)}{\partial x}\delta x\delta y\delta z$$

where $\eta\in[0,1]$ and $\zeta\in[0,1]$. If we divide this by $dV=\delta x\delta y\delta z$ and let the volume shrink to a point we are left with $\partial A_x(x,y,z)/\partial x$. Similarly we get from the other two pairs of faces $\partial A_y(x,y,z)/\partial y$ and $\partial A_z(x,y,z)/\partial z$, and hence for the entire surface S

$$\operatorname{div}\vec{A}(\vec{r}\,) = \frac{\partial A_x(x,y,z)}{\partial x} + \frac{\partial A_y(x,y,z)}{\partial y} + \frac{\partial A_z(x,y,z)}{\partial z}$$

Properties of divergence.

(i) The divergence is a *linear* operation, i.e. given the vector fields $\vec{A}$ and $\vec{B}$ and the scalar constant c the following identities hold true:

$$\operatorname{div}(\vec{A}+\vec{B}) = \operatorname{div}\vec{A} + \operatorname{div}\vec{B}, \qquad \text{and} \qquad \operatorname{div}(c\,\vec{A}) = c\operatorname{div}\vec{A} \tag{51}$$

(ii)

$$\operatorname{div}\vec{r} = 3 \tag{52}$$

(iii)

$$\operatorname{div} f(r)\vec{r} = r\frac{df(r)}{dr} + 3f(r) \tag{53}$$

(iv)

$$\operatorname{div} f(\vec{r})\vec{r} = \vec{r}\cdot\operatorname{div} f(\vec{r}) + 3f(\vec{r}) \tag{54}$$

(v)

$$\operatorname{div}\,(r^n\vec{r}) = (n+3)r^n \tag{55}$$

(vi)

$$\operatorname{div}[f(\vec{r})\vec{A}(\vec{r})] = \vec{A}(\vec{r})\cdot\operatorname{grad} f(\vec{r}) + f(\vec{r})\operatorname{div}\vec{A}(\vec{r}) \tag{56}$$

4. **The nabla operator.** In the preceding sections we have encountered expressions in which partial derivatives w.r.t. x, y and z play an important role. One can see that it is convenient to combine the partial derivative *operators* $\partial/\partial x$, $\partial/\partial y$ and $\partial/\partial z$ into a vector operator

$$\nabla \equiv \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\right) \tag{57}$$

called *nabla* (or sometimes "del").

Using the nabla operator the gradient and divergence are written as

$$\operatorname{grad}\Phi(\vec{r}) = \nabla\,\Phi(\vec{r}) \tag{58}$$

and

$$\operatorname{div}\vec{A}(\vec{r}) = \nabla\cdot\vec{A}(\vec{r}) \tag{59}$$

respectively. Thus the nabla operator can be handled like a vector, except that one must remember that it does not commute with field functions like an ordinary vector because it is also a differential operator that operates on the function to the right.

5. **Theorems of Gauss and Green.** Consider Eq. (49)

$$\operatorname{div}\vec{A}(\vec{r}_0) = \lim_{V\to 0}\frac{1}{V}\oint_S \vec{A}\cdot d\vec{S}$$

which we have written down as the definition of the divergence of $\vec{A}(\vec{r})$. Realising that the integral on the r.h.s. and the volume V are infinitesimal in the limit, we can also write

$$\operatorname{div}\vec{A}(\vec{r}_0) = \frac{d}{dV}\oint_S \vec{A}\cdot d\vec{S}$$

or

$$d\oint_S \vec{A}\cdot d\vec{S} = \operatorname{div}\vec{A}(\vec{r}_0)\,dV$$

Integrating this over the a finite volume V enclosed by the surface S we get

$$\oint_S \vec{A} \cdot d\vec{S} = \int_V \operatorname{div} \vec{A}(\vec{r}_0)\, dV \tag{60}$$

which is known as Gauss' theorem or divergence theorem.

The derivation of Gauss' theorem given here may seem rather formal. A more rigorous and more transparent derivation proceeds as follows. Consider a volume V enclosed by the surface S. Divide the volume into infinitesimal volume elements δV_i, $i = 1, 2, \ldots, n$ enclosed by surfaces S_i. Then, by definition of the divergence, we have

$$(\operatorname{div} \vec{A})_i \delta V_i = \oint_{S_i} \vec{A} \cdot d\vec{S}$$

and summing over all volume elements we get

$$\sum_{i=1}^{n} (\operatorname{div} \vec{A})_i \delta V_i = \sum_{i=1}^{n} \oint_{S_i} \vec{A} \cdot d\vec{S}$$

On the r.h.s. each volume element either shares a surface with a neighbouring volume element or has an external surface. The external surface elements together make up the surface S. The contributions of all internal surfaces, on the other hand, cancel since they enter into the sum with opposite signs for the two adjacent volume elements. Therefore, letting $n \to \infty$ and at the same time the maximum linear dimensions of δV_i tend to zero, we get Gauss' theorem.

Of interest is frequently the divergence of vector fields which are the gradients of scalar fields. Thus, if $\Phi(\vec{r})$ is a scalar field, we want to find the divergence of its gradient. Using the nabla operator we write this in the following form:

$$\begin{aligned}
\operatorname{div} \operatorname{grad} \Phi(\vec{r}) &= \nabla \cdot (\nabla \Phi(\vec{r})) = \frac{\partial}{\partial x}\left(\frac{\partial \Phi(\vec{r})}{\partial x}\right) + \frac{\partial}{\partial y}\left(\frac{\partial \Phi(\vec{r})}{\partial y}\right) + \frac{\partial}{\partial z}\left(\frac{\partial \Phi(\vec{r})}{\partial z}\right) \\
&= \frac{\partial^2 \Phi(\vec{r})}{\partial x^2} + \frac{\partial^2 \Phi(\vec{r})}{\partial y^2} + \frac{\partial^2 \Phi(\vec{r})}{\partial z^2} = \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}\right) \Phi(\vec{r})
\end{aligned}$$

where in the last step we have written the expression in the form of a differential operator operating on the scalar function. This differential operator,

$$\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \tag{61}$$

is called the *Laplacian* operator. Formally it is given by the scalar product $\nabla \cdot \nabla$. Therefore we can write the equivalent forms

$$\operatorname{div} \operatorname{grad} \Phi = \nabla \cdot (\nabla\, \Phi) = (\nabla \cdot \nabla)\Phi = \nabla^2 \Phi$$

Consider the scalar fields $\Phi(\vec{r})$ and $\Psi(\vec{r})$. The gradients of Φ and Ψ are vectors, and so is the expression $\vec{A} = [\Phi(\nabla\Psi) - \Psi(\nabla\Phi)]$. Take the divergence of this vector:

$$\mathrm{div}\,\vec{A} = \nabla \cdot [\Phi(\nabla\Psi) - \Psi(\nabla\Phi)] = [\Phi(\nabla^2\Psi) - \Psi(\nabla^2\Phi)]$$

and if we apply the divergence theorem to this equation we get

$$\oint_S [\Phi(\nabla\Psi) - \Psi(\nabla\Phi)] \cdot d\vec{S} = \int_V [\Phi(\nabla^2\Psi) - \Psi(\nabla^2\Phi)] dV \tag{62}$$

which is known as Green's theorem. Thus Green's theorem is a generalization of Gauss' divergence theorem.

6. Curl of a vector field; Stokes' theorem.

The last of the vector operations to consider is the *curl* of a vector field. This concept arises in fluid dynamics in connection with the mathematical description of vortices and in electromagnetism for instance in the relationship between magnetic fields and currents.

Considering a vortex one observes that the fluid rotates about the centre of the vortex like a solid disk. The linear velocity of small volume lements of the fluid increases linearly from the centre and is directed perpendicularly to the radius vector if we put the origin of the coordinate system at the centre of the vortex. If we put the xy-plane normal to the vortex axis, we can write for the linear velocity

$$\vec{v} = \omega r \hat{\varphi}$$

where ω is the angular velocity, which is independent of r, and $\hat{\varphi}$ is a unit vector perpendicular to the radius vector and pointing in the direction of increasing azimuth φ. Thus $\hat{\varphi} = (-\hat{x}\sin\varphi + \hat{y}\cos\varphi)$, and therefore, recalling that $x = r\cos\varphi$ and $y = r\sin\varphi$, we get $\vec{v} = \omega(-y\hat{x} + x\hat{y})$.

Now consider the line integral

$$I = \oint_C \vec{v} \cdot d\vec{r}$$

where C is a closed circular path around the origin (see Fig. NN). Since $r = \text{const.}$ on C, we have $d\vec{r} = r\hat{\varphi}d\varphi$, and hence

$$I = \omega r^2 \int_0^{2\pi} d\varphi = 2\pi r^2 \omega$$

and noting that $A = \pi r^2$ is the area of the circle enclosed by C we finally get

$$I = 2A\omega$$

We see that the integral characterizes the strength of the vortex; it is called the *circulation* of the vector field $\vec{v}$ (Lord Kelvin).

If we take a different closed path in the vortex field we can reproduce the same result even if the path does not enclose the centre of the vortex (see exercise NN).

In the expression for the circulation we have the arbitrary factor A. A better characteristic of the vortex is therefore the following expression

$$\frac{1}{A}\oint_C \vec{v}\cdot d\vec{r}$$

which can be used to get a local measure of the circulation if we let the path C contract into a point. Assuming that the surface A is already small enough to be flat, we can assign a normal $\hat{n}$ to it and define the component of the curl of $\vec{v}$ in the direction of $\hat{n}$:

$$(\mathrm{curl}\,\vec{v})_n = \lim_{A\to 0}\oint_C \vec{v}\cdot d\vec{r} \tag{63}$$

Here we use the convention that, looking down from the unit vector $\hat{n}$, the integration along C goes in anticlockwise direction ("mathematically positive" sense).

In the particular case when C lies in the yz-plane (or in a plane parallel to the yz-plane) and the normal points along the x axis we get $(\mathrm{curl}\,\vec{v})_x$. Let us evaluate this component of the curl choosing the path C in the form of a rectangle whose sides are parallel to the y and z axes, see Fig. NN. The integral then becomes the sum of four integrals along the four sides of the rectangle:

$$\oint_C \vec{v}\cdot d\vec{r} = \int_{(1)} v_z(x,y+\delta y,z)dz - \int_{(2)} v_y(x,y,z+\delta z)dy - \int_{(3)} v_z(x,y,z)dz + \int_{(4)} v_y(x,y,z)dz$$

In the integrals along sides (1) and (2) we expand the argument about δy and δz, respectively, and we get

$$\int_z^{z+\delta z} dz\frac{\partial v_z}{\partial y}\delta y - \int_y^{y+\delta y} dy\frac{\partial v_y}{\partial z}\delta z$$

For sufficiently small δy and δz the integrands are constant and can be taken outside the integrals, hence

$$\oint_C \vec{v}\cdot d\vec{r} = \frac{\partial v_z}{\partial y}\delta y\int_z^{z+\delta z} dz - \frac{\partial v_y}{\partial z}\delta z\int_y^{y+\delta y} dy = \frac{\partial v_z}{\partial y}\delta y\delta z - \frac{\partial v_y}{\partial z}\delta z\delta y$$

but $\delta A = \delta y\delta z$ is the area enclosed by C, hence, dividing by δA and letting all sides of the rectangle tend to zero, we get

$$(\mathrm{curl}\,\vec{v})_x = \frac{\partial v_z}{\partial y} - \frac{\partial v_y}{\partial z}$$

and similarly we get

$$(\mathrm{curl}\,\vec{v})_y = \frac{\partial v_x}{\partial z} - \frac{\partial v_z}{\partial x} \qquad \text{and} \qquad (\mathrm{curl}\,\vec{v})_z = \frac{\partial v_y}{\partial x} - \frac{\partial v_x}{\partial y}$$

and hence, finally,

$$\text{curl}\,\vec{v} = \left(\frac{\partial v_z}{\partial y} - \frac{\partial v_y}{\partial z}, \frac{\partial v_x}{\partial z} - \frac{\partial v_z}{\partial x}, \frac{\partial v_y}{\partial x} - \frac{\partial v_x}{\partial y} \right). \tag{64}$$

Stokes' theorem equates the integral of the curl of a vector field $\vec{v}(\vec{r})$ over a surface S to the line integral of $\vec{v}$ over the boundary ℓ of S:

$$\int_S (\text{curl}\,\vec{v}) \cdot d\vec{S} = \oint_\ell \vec{v} \cdot d\vec{\ell} \tag{65}$$

The surface S need not be a plane surface. To get the correct sign in Stokes' theorem the unit normal vector $\hat{n}$ must satisfy the corkscrew rule with the direction, in which the boundary ℓ is traversed.

The idea of the proof of Stokes' theorem is similar to that of Gauss' theorem. Divide S into n cells $\delta\vec{S}_i = \hat{n}_i \delta S_i$. In the limit of $\delta S_i \to 0$ we get for the ith cell, by definition of the curl,

$$(\text{curl}\,\vec{v})_i = \frac{\hat{n}_i}{\delta S_i} \oint_{\ell_i} \vec{v} \cdot d\vec{\ell}$$

Taking the scalar product with $\hat{n}_i \delta S_i$ and summing over i we get

$$\sum_{i=1}^{n} (\text{curl}\,\vec{v})_i \cdot \delta\vec{S}_i = \sum_{i=1}^{n} \oint_{\ell_i} \vec{v} \cdot d\vec{\ell}$$

If we take the limit of $n \to \infty$ and at the same time let the maximum diameter of δS_i tend to nought, then we get on the left-hand side $\int_S (\text{curl}\,\vec{v}) \cdot d\vec{S}$. On the right-hand side the contributions from all *internal* lines cancel, since they are traversed twice in opposite directions for each of the neighbouring cells, and only the contributions from the external lines remain, which add up to yield the line integral $\oint_\ell \vec{v} \cdot d\vec{\ell}$.

The curl of a vector field takes on a compact form when it is expressed in terms of the nabla operator. Indeed, consider the cross product $\nabla \times \vec{v}$: by definition of nabla we have

$$\nabla \times \vec{v} = \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z} \right) \times (v_x, v_y, v_z) = \left(\frac{\partial v_z}{\partial y} - \frac{\partial v_y}{\partial z}, \frac{\partial v_x}{\partial z} - \frac{\partial v_z}{\partial x}, \frac{\partial v_y}{\partial x} - \frac{\partial v_x}{\partial y} \right)$$

i.e. the curl of $\vec{v}$.

Properties of the curl.

(i)

$$\nabla \times \vec{r} = 0$$

(ii)

$$\nabla \times (\nabla \Phi(\vec{r})) = 0$$

i.e. *a potential field is irrotational.*

(iii)

$$\nabla \cdot (\nabla \times \vec{v}) = 0$$

i.e. *a vortex field has no sources.*

(iv)

$$\nabla \cdot (\vec{u} \times \vec{v}) = \vec{v} \cdot (\nabla \times \vec{u}) - \vec{u} \cdot (\nabla \times \vec{v})$$

(v)

$$\nabla \times (\nabla \times \vec{v}) = \nabla(\nabla \cdot \vec{v}) - \nabla^2 \vec{v}$$

The proof of (i) to (iv) is straightforward. The proof of (v) is also straightforward but lengthy. Here it is done for the x component:

$$\begin{aligned}
[\nabla \times (\nabla \times \vec{v})]_x &= \frac{\partial}{\partial y}(\nabla \times \vec{v})_z - \frac{\partial}{\partial z}(\nabla \times \vec{v})_y \\
&= \frac{\partial}{\partial y}\left(\frac{\partial v_y}{\partial x} - \frac{\partial v_x}{\partial y}\right) - \frac{\partial}{\partial z}\left(\frac{\partial v_x}{\partial z} - \frac{\partial v_z}{\partial x}\right) \\
&= \frac{\partial^2 v_y}{\partial y \partial x} + \frac{\partial^2 v_z}{\partial z \partial x} - \frac{\partial^2 v_x}{\partial y^2} - \frac{\partial^2 v_x}{\partial z^2}
\end{aligned}$$

this expression does not change if we *add* $\partial^2 v_x / \partial x^2$ to the first pair of terms and *subtract* it from the second pair. This gives us

$$\frac{\partial}{\partial x}\left(\frac{\partial v_x}{\partial x} + \frac{\partial v_y}{\partial y} + \frac{\partial v_z}{\partial z}\right) - \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}\right) v_x = \frac{\partial}{\partial x}(\nabla \cdot \vec{v}) - \nabla^2 v_x$$

and similarly for the y and z components.

7. Application: Maxwell equations of Electromagnetism.

Maxwell's equations in SI units[3] are of the following form:

$$\begin{aligned}
\mathrm{curl}\vec{H} &= \vec{\jmath} + \frac{\partial \vec{D}}{\partial t} && \text{Ampére-Maxwell law} && (66) \\
\mathrm{curl}\vec{E} &= -\frac{\partial \vec{B}}{\partial t} && \text{Faraday-Lenz law} && (67) \\
\nabla \cdot \vec{D} &= \rho && \text{Gauss' law} && (68) \\
\nabla \cdot \vec{B} &= 0 && \text{no magnetic charges} && (69) \\
\vec{D} &= \varepsilon \vec{E} && && (70) \\
\vec{B} &= \mu \vec{H} && && (71)
\end{aligned}$$

Here $\vec{H}$ is the *magnetic field strength*, $\vec{B}$ is the *magnetic flux density*, $\vec{E}$ is the *electric field strength* and $\vec{D}$ is the *electric displacement*; these four vectors are functions of space and time. ρ and $\vec{\jmath}$ are the charge density and current density, respectively.

[3]French: Système International d'Unités

$\varepsilon = \varepsilon_r \varepsilon_0$ is the *permittivity*, with ε_r the *relative permittivity* (also called *dielectric constant*) and ε_0 the *permittivity of free space* or *electric constant*; ε_r is dimensionless, $\varepsilon_0 = 8.854\,\mathrm{pF\ m^{-1}}$. $\mu = \mu_r \mu_0$ is the *permeability* with μ_r the *relative permeability* (which is dimensionless) and $\mu_0 = 4\pi \times 10^{-7}$ Henry per meter the *permeability of free space* or *magnetic constant*.

Note that

$$1/\sqrt{\varepsilon_0 \mu_0} = c \tag{72}$$

is the speed of light in free space.

From the Ampére-Maxwell law together with Gauss' law we get immediately the continuity equation: take the divergence of Eq. (66), use the vector identity $\nabla \cdot (\nabla \times \vec{H}) = 0$, and substitute the divergence of $\vec{D}$ from Eq. (68), hence

$$\frac{\partial \rho}{\partial t} + \nabla \cdot \vec{j} = 0 \tag{73}$$

In free space $\varepsilon_r = \mu_r = 1$, $\rho = 0$ and $\vec{j} = 0$. We can therefor rewrite the Maxwell equations in terms of, for instance, $\vec{H}$ and $\vec{D}$, eliminating also ε_0 and μ_0, thus

$$\mathrm{curl}\vec{H} = \frac{\partial \vec{D}}{\partial t} \tag{74}$$

$$\mathrm{curl}\vec{D} = -\frac{1}{c^2}\frac{\partial \vec{H}}{\partial t} \tag{75}$$

$$\nabla \cdot \vec{D} = 0 \tag{76}$$

$$\nabla \cdot \vec{H} = 0 \tag{77}$$

which can be shown to have plane wave solutions

$$\vec{D} = \vec{D}_0\, e^{i(\vec{k}\cdot\vec{r} - \omega t)} \tag{78}$$

and

$$\vec{H} = \vec{H}_0\, e^{i(\vec{k}\cdot\vec{r} - \omega t)} \tag{79}$$

with $\vec{D} \perp \vec{H}$, $\vec{D} \perp \vec{k}$ and $\vec{H} \perp \vec{k}$. The direction of $\vec{D}$ at different times t defines a plane, transverse to the direction of propagation of the wave. This plane is the *plane of polarization*, and the wave is said to be a *transverse wave*. The superposition of two waves travelling in the same direction with vectors $\vec{D}_1$ and $\vec{D}_2$, which are mutually orthogonal and whose phases differ by $\pi/2$, produces an *elliptically polarised* wave; in the particular case of equal amplitudes, $|\vec{D}_{01}| = |\vec{D}_{02}|$, the wave is said to be *circularly polarised*.

Exercise: Assuming that $\vec{D}$ is a plane wave, show that $\vec{H}$ is also a plane wave with the same phase as $\vec{D}$, and that the electric and magnetic fields are orthogonal to each other and to the direction of propagation.

We can deduce wave equations for $\vec{D}$ and $\vec{H}$ from the Maxwell equations, for instance by taking the curl of Eq. (74) and substituting the curl of $\vec{D}$ from (75), hence, applying the identity

$$\nabla \times (\nabla \times \vec{a}(\vec{r})) = \nabla(\nabla \cdot \vec{a}(\vec{r})) - \nabla^2 \vec{a}(\vec{r}) \tag{80}$$

and using Eq. (77), we get the wave equation

$$\left(\frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \nabla^2\right) \vec{H} = 0 \tag{81}$$

substituting into which the plane wave function (79) we get the *dispersion law* of electromagnetic waves in free space,

$$\omega = ck \tag{82}$$

The wave equation for $\vec{D}$ is obtained similarly.