Chapter 1

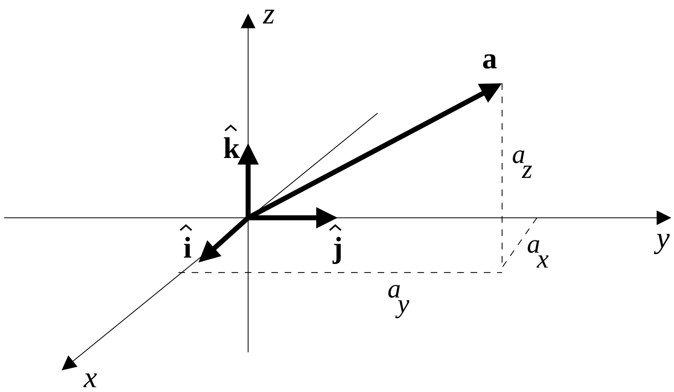
Vectors

One of the simplest and most easily understood ways we have to model the universe is by using vectors. Usually we say that something is a vector if it has both a *magnitude* and a *direction*. In physics, vectors are used to describe forces, flows of fluids, electric and magnetic fields, and in fact any other quantity with an associated direction.

Because we live in a three-dimensional world, a vector typically consists of three numbers, which correspond to how much the vector protrudes in the x, y or z directions. Usually we write these numbers next to each other, separated by commas:

$$\mathbf{a} = \langle 1.3, 3.0, -2.1 \rangle$$

An equivalent notation is to write the vector in terms of the *unit axis vectors* $\hat{\mathbf{i}}, \hat{\mathbf{j}}$ and $\hat{\mathbf{k}}$.



The vector above can be written as

$$\mathbf{a} = 1.3\hat{\mathbf{i}} + 3.0\hat{\mathbf{j}} - 2.1\hat{\mathbf{k}} \quad ,$$

where

$$\begin{aligned} \hat{\mathbf{i}} &= \langle 1, 0, 0 \rangle \\ \hat{\mathbf{j}} &= \langle 0, 1, 0 \rangle \\ \hat{\mathbf{k}} &= \langle 0, 0, 1 \rangle \ . \end{aligned}$$

The three numbers which describe a vector are called the *components* of the vector, and are often written using a subscript according to exactly which component they are, a_x for the x-component of the vector $\mathbf{a}$, a_y for the y component, and a_z for the z component. This means that we can write

$$\mathbf{a} = a_x\hat{\mathbf{i}} + a_y\hat{\mathbf{j}} + a_z\hat{\mathbf{k}} \tag{1.1}$$

and so in the example given above, $a_x = 1.3$, $a_y = 3.0$ and $a_z = -2.1$.

The *magnitude*, or 'length', or 'norm' of a vector is usually written with square brackets on either side of the vector, such as $\|\mathbf{a}\|$. It can be calculated using Pythagoras' theorem. The result is

$$\|\mathbf{a}\| = \sqrt{a_x^2 + a_y^2 + a_z^2} \tag{1.2}$$

Any vector with a magnitude equal to 1 is known as a *unit vector*, and is often written with a hat ($\hat{}$) above the symbol, such as for $\hat{\mathbf{i}}$,$\hat{\mathbf{j}}$ and $\hat{\mathbf{k}}$.

The particular vector which points to a *position* in space is usually denoted by the symbol $\mathbf{r}$, and is defined

$$\mathbf{r} = x\hat{\mathbf{i}} + y\hat{\mathbf{j}} + z\hat{\mathbf{k}} \tag{1.3}$$

Note that $\mathbf{r}$ is (of course) a function of x, y, and z, and always points directly outwards from the origin.

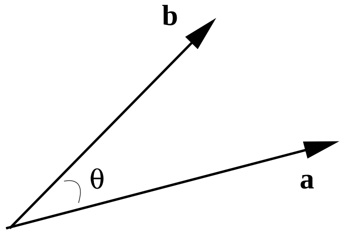
Vector multiplication

Because each vector is a collection of three numbers, we find that there are various ways of combining them to form scalars or other vectors. These combinations are called 'multiplications', even their connection with the ordinary multiplication which we learned in primary school is somewhat tenuous. The first sort of multiplication which we consider here is called the *dot product*:

1. The dot product

The dot product between two vectors $\mathbf{a} = a_x\hat{\mathbf{i}} + a_y\hat{\mathbf{j}} + a_z\hat{\mathbf{k}}$ and $\mathbf{b} = b_x\hat{\mathbf{i}} + b_y\hat{\mathbf{j}} + b_z\hat{\mathbf{k}}$ is written

$$\mathbf{a} \cdot \mathbf{b}$$



and is defined to be

$$\mathbf{a} \cdot \mathbf{b} = a_x b_x + a_y b_y + a_z b_z \ . \tag{1.4}$$

Note that this way of combining the two vectors has resulted in a *scalar quantity*, that is, in a regular number. For this reason, the dot product is often called the *scalar product* of the two vectors.

The scalar product has the following important properties, which can be derived directly from the definition (1.4) above:

(i) $\mathbf{a} \cdot \mathbf{a} = \|\mathbf{a}\|^2$

(ii) $\mathbf{a} \cdot \mathbf{0} = 0$

(iii) $\mathbf{a} \cdot \mathbf{b} = \mathbf{b} \cdot \mathbf{a}$

(iv) $\mathbf{a} \cdot \mathbf{b} = |\mathbf{a}||\mathbf{b}| \cos\theta$

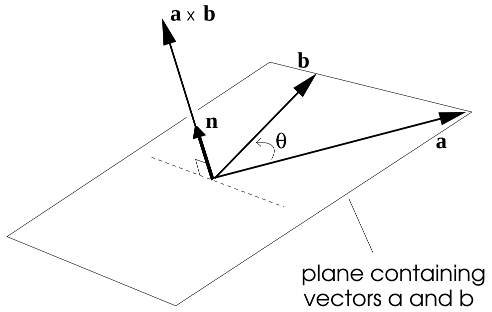
where θ is the angle between the two vectors (as shown above). Property (iv) is an interesting equation - it says that the dot product is a measure of how much the two vectors point in the same direction. If the vectors are orthogonal then the angle θ between them is ninety degrees, and so $\cos\theta$ is zero, therefore the dot product of the two vectors is also zero. This goes the other way around; if the dot product of any two vectors is equal to zero, then the vectors are perpendicular. Mathematically we can write this

$$\mathbf{a} \perp \mathbf{b} \ \text{ if and only if } \ \mathbf{a} \cdot \mathbf{b} = 0 \ . \tag{1.5}$$

Because the range of $\cos\theta$ is limited so that $-1 \leq \cos\theta \leq 1$, we immediately have the result

$$\|\mathbf{a} \cdot \mathbf{b}\| \leq \|\mathbf{a}\|\|\mathbf{b}\| \tag{1.6}$$

which is known as *Schwarz's inequality*.



2. The cross product

The cross product between two vectors $\mathbf{a} = a_x\hat{\mathbf{i}} + a_y\hat{\mathbf{j}} + a_z\hat{\mathbf{k}}$ and $\mathbf{b} = b_x\hat{\mathbf{i}} + b_y\hat{\mathbf{j}} + b_z\hat{\mathbf{k}}$ is written

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

and is defined to be

$$\mathbf{a} \times \mathbf{b} := \begin{vmatrix} \hat{\mathbf{i}} & \hat{\mathbf{j}} & \hat{\mathbf{k}} \\ a_x & a_y & a_z \\ b_x & b_y & b_z \end{vmatrix} \tag{1.7}$$

where the square brackets on either side of the three-by-three matrix are shorthand for taking the determinant; if we go ahead and work this out, the full result is

$$\mathbf{a} \times \mathbf{b} = \hat{\mathbf{i}}(a_y b_z - a_z b_y) - \hat{\mathbf{j}}(a_x b_z - a_z b_x) + \hat{\mathbf{k}}(a_x b_y - a_y b_x) \tag{1.8}$$

Note that the cross product $\mathbf{a} \times \mathbf{b}$ has both a magnitude and a direction, and for this reason it is sometimes called the *vector product*.

From the definition (1.7) we can prove that

$$\mathbf{a} \times \mathbf{b} := \|\mathbf{a}\|\|\mathbf{b}\| \sin\theta \ \hat{\mathbf{n}} \tag{1.9}$$

where θ is the angle between the two vectors and $\hat{\mathbf{n}}$ is a *unit vector* which is *perpendicular* to both $\mathbf{a}$ and $\mathbf{b}$ (see the previous figure). There is a bit of ambiguity in this: there are *two* vectors which are perpendicular to both $\mathbf{a}$ and $\mathbf{b}$. Which one should we choose?

The convention is to adopt the so-called 'right-hand rule'. Hold your right hand flat so that the fingers all point in the direction of the vector $\mathbf{a}$. Then curl your fingers around in the direction of the vector $\mathbf{b}$, as if you were trying to turn $\mathbf{a}$ so that it points in $\mathbf{b}$'s direction. If you keep your thumb sticking straight out, then the direction your thumb is pointing in is the direction of the cross product.

Notice that this 'choice' depends crucially on the *order* of $\mathbf{a}$ and $\mathbf{b}$. If you pick $\mathbf{b}$ first in your right-hand rule, and form the cross product $\mathbf{b} \times \mathbf{a}$, then the resulting vector (given by your thumb) ends up pointing in the opposite direction! This is a fundamental property of cross products, and formally we can write that

$$\mathbf{a} \times \mathbf{b} = -\mathbf{b} \times \mathbf{a} \quad . \tag{1.10}$$

This identity is easily proved from (1.7).

To summarize, the cross product has the following important properties:

(i) $\mathbf{a} \times \mathbf{b}$ is perpendicular to both $\mathbf{a}$ and $\mathbf{b}$, i.e. $\mathbf{a} \cdot (\mathbf{a} \times \mathbf{b}) = \mathbf{0}$ and $\mathbf{b} \cdot (\mathbf{a} \times \mathbf{b}) = \mathbf{0}$. The direction is given by the right-hand rule.

(ii) $\|\mathbf{a} \times \mathbf{b}\| := \|\mathbf{a}\|\|\mathbf{b}\| \sin\theta$

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

3. Multiplication by a scalar

Any vector can be multiplied by a scalar quantity. For any vector $\mathbf{a} = a_x\hat{\mathbf{i}} + a_y\hat{\mathbf{j}} + a_z\hat{\mathbf{k}}$ and scalar c,

$$c\mathbf{a} = ca_x\hat{\mathbf{i}} + ca_y\hat{\mathbf{j}} + ca_z\hat{\mathbf{k}} \quad . \tag{1.11}$$

The effect of multiplication by a scalar is simply to change the length of the vector, or, if the scalar happens to be negative, to cause it to point in the opposite direction.

Vector addition

Happily, vector addition is a lot simpler than vector multiplication. Vectors can be added together just like regular numbers, provided we remember to keep the different components separate. If we represent a vector $\mathbf{a}$ by

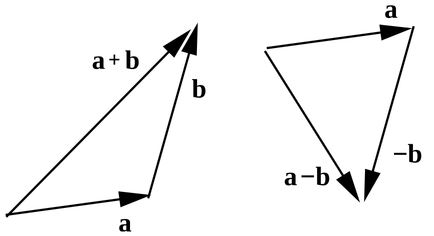
$$\mathbf{a} = a_x\hat{\mathbf{i}} + a_y\hat{\mathbf{j}} + a_z\hat{\mathbf{k}}$$

and a vector $\mathbf{b}$ by

$$\mathbf{b} = b_x\hat{\mathbf{i}} + b_y\hat{\mathbf{j}} + b_z\hat{\mathbf{k}}$$

then the vector sum of $\mathbf{a}$ and $\mathbf{b}$ is straightforward:

$$\mathbf{a} + \mathbf{b} = (a_x + b_x)\hat{\mathbf{i}} + (a_y + b_y)\hat{\mathbf{j}} + (a_z + b_z)\hat{\mathbf{k}} \quad . \tag{1.12}$$



We can also subtract vectors in the same way:

$$\begin{aligned}\mathbf{a}-\mathbf{b} &= \mathbf{a}+(-\mathbf{b}) \\ &= (a_x-b_x)\hat{\mathbf{i}}+(a_y-b_y)\hat{\mathbf{j}}+(a_z-b_z)\hat{\mathbf{k}} \quad .\end{aligned}$$

Problems

1. If $\mathbf{a} = \langle 1.0, 2.2, -1.4\rangle$ and $\mathbf{b} = \langle 0.0, 1.0, 0.0\rangle$ then calculate:

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

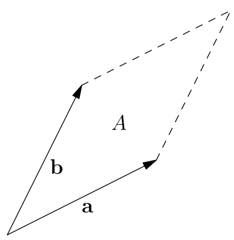
(b) $3\mathbf{a}-2\mathbf{b}$

(c) $\mathbf{a}\cdot\mathbf{b}$

(d) $\mathbf{a}\times\mathbf{b}$

2. Show using equation (1.9) or otherwise that the area of the parallelogram formed by two vectors **a** and **b** is

$$A = \|\mathbf{a}\times\mathbf{b}\| \ . \tag{1.13}$$



3. Using equations (1.7) and (1.4), prove the identity

$$\mathbf{a}\cdot(\mathbf{a}\times\mathbf{b}) = \mathbf{0}$$

4*. Prove equation (1.10) from the definition (1.7).

5*. Prove equation (1.9) from the definition (1.7).

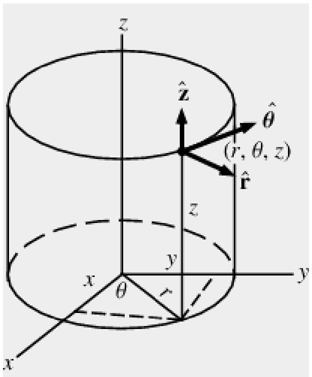
[SOLUTIONS: 1. a) $\langle 1.0, 3.2, -1.4\rangle$, b) $\langle 3.0, 4.6, -4.2\rangle$, c) 2.2, d) $\langle 1.4, 0.0, 1.0\rangle$.]

Curvilinear coordinates

The unit vectors $\hat{\mathbf{i}}$, $\hat{\mathbf{j}}$ and $\hat{\mathbf{k}}$ are not the only possible way of describing vectors in three dimensions. In fact, if the system being studied possesses some sort of symmetry then it is often much more convenient to work in *curvilinear coordinates*. These coordinate systems remain orthogonal, but depend implicitly on the position vector.

1. Cylindrical Polar Coordinates

With the z axis defined as the axis of symmetry, the cylindrical polar coordinate system is defined according to the diagram below:



In terms of the Cartesian coordinates x, y and z, we have

$$\begin{aligned} r &= \sqrt{x^2+y^2} \\ \theta &= \tan^{-1}(y/x) \\ z &= z \quad \text{(z coordinate unchanged)} \; . \end{aligned} \tag{1.14}$$

The inverse operations are

$$\begin{aligned} x &= r\cos\theta \\ y &= r\sin\theta \\ z &= z \end{aligned}$$

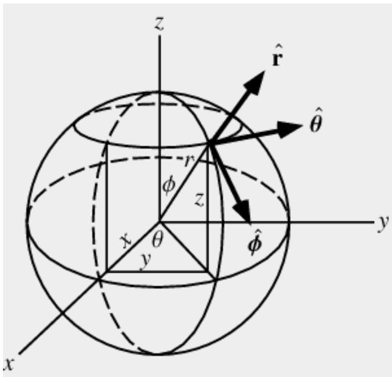
As in Cartesian coordinates, any vector can be expanded in terms of the unit vectors $\hat{\mathbf{r}}$,$\hat{\theta}$ and $\hat{\mathbf{z}}$:

$$\mathbf{a} = a_r\hat{\mathbf{r}} \; + a_\theta\hat{\theta} \; + a_z\hat{\mathbf{z}} \quad .$$

It is important to remember that the unit vectors for cylindrical polar coordinates are *dependent on position*, that is, they change direction as you move around the three-dimensional space. This is in contrast to the Cartesian coordinates, which don't change. For example, on the x-axis, the coordinate vector $\hat{\theta}$ points in the positive y direction, i.e. it is parallel to the $\hat{\mathbf{j}}$ coordinate vector. However along the y axis the vector $\hat{\theta}$ points in the negative x-direction.

2. Spherical Polar Coordinates

We introduce the spherical polar coordinates r, θ and ϕ as outlined in the diagram below:



The coordinate r is always positive. We define θ to lie within the range $0 \leq \theta \leq 2\pi$, whereas ϕ must lie between $0 \leq \phi \leq \pi$, i.e. between the north and south 'poles' of our coordinate system.

In terms of the Cartesian coordinates x, y and z, we have

$$\begin{aligned} r &= \sqrt{x^2 + y^2 + z^2} \\ \theta &= \tan^{-1}(y/x) \\ \phi &= \cos^{-1}(z/r) \ . \end{aligned} \tag{1.15}$$

The inverse operations are

$$\begin{aligned} x &= r\cos\theta\sin\phi \\ y &= r\sin\theta\sin\phi \\ z &= r\cos\phi \end{aligned}$$

Instead of the three cartesian coordinates $\hat{\mathbf{i}}, \hat{\mathbf{j}}$ and $\hat{\mathbf{k}}$ we now use the coordinate system $\hat{\mathbf{r}}$, $\hat{\theta}$ and $\hat{\phi}$. These coordinate vectors are orthogonal (i.e. they are at right-angles to one another) and any vector **a** in three dimensions can be expanded in a linear combination, using these three as a basis:

$$\mathbf{a} = a_r \hat{\mathbf{r}} + a_\theta \hat{\theta} + a_\phi \hat{\phi} \tag{1.16}$$

The position vector, which points to a general point in space, is

$$\mathbf{r} = r\ \hat{\mathbf{r}} \ . \tag{1.17}$$

Note that since a position vector always points *away* from the origin, the vector **r** doesn't have any components in the $\hat{\theta}$ or $\hat{\phi}$ directions.

Problems

1. Perform each of the following conversions.

(a) Convert the point $\langle \sqrt{2}, \pi/3, 1 \rangle$ from cylindrical to cartesian coordinates.

(b) Convert the point $\langle 1, 1, 1 \rangle$ from Cartesian to spherical coordinates.

2. Sketch or describe the surface resulting from the following equations in spherical coordinates:

(a) $r = 2$

(b) $\theta = \pi/3$

(c) $\phi = \pi/4$.