2A1 Vector Algebra and Calculus

8 Lectures MT 2008 | Ian Reid
Rev Oct 2008 | ian.reid@eng.ox.ac.uk
Course Page | www.robots.ox.ac.uk/~ian/Teaching/Vectors

Overview

Many of you will know a good deal already about **Vector Algebra** — how to add and subtract vectors, how to take scalar and vector products of vectors, and something of how to describe geometric and physical entities using vectors. This course will remind you about that good stuff, but goes on to introduce you to the subject of **Vector Calculus** which, like it says on the can, combines vector algebra with calculus.

To give you a feeling for the issues, suppose you were interested in the temperature T of water in a river. Temperature T is a *scalar*, and will certainly be a function of a position vector $\mathbf{x} = (x, y, z)$ and may also be a function of time t: $T = T(\mathbf{x}, t)$. It is a scalar field.

Suppose now that you kept y, z, t constant, and asked what is the change in temperature as you move a small amount in x? No doubt you'd be interested in calculating $\partial T/\partial x$. Similarly if you kept the point fixed, and asked how does the temperature change of time, you would be interested in $\partial T/\partial t$.

But why restrict ourselves to movements up-down, left-right, etc? Suppose you wanted to know what the change in temperature along an arbitrary direction. You would be interested in

$$\frac{\partial T}{\partial \mathbf{x}} \quad ,$$

but how would you calculate that? Is $\partial T/\partial \mathbf{x}$ a vector or a scalar?

Now let's dive into the flow. At each point $\mathbf{x}$ in the stream, at each time t, there will be a stream velocity $\mathbf{v}(\mathbf{x}, t)$. The local stream velocity can be viewed directly using modern techniques such as laser Doppler anemometry, or traditional techniques such a throwing twigs in. The point now is that $\mathbf{v}$ is a function that has the same four input variables as temperature did, but its output result is a vector. We may be interested in places $\mathbf{x}$ where the stream suddenly accelerates, or vortices where the stream curls around dangerously. That is, we will be interested in finding the acceleration of the stream, the gradient of its velocity. We may be interested in the magnitude of the acceleration (a scalar). Equally, we may be interested in the acceleration as a vector, so that we can apply Newton's law and figure out the force.

This is the stuff of vector calculus.

Grey book

Vector algebra: scalar and vector products; scalar and vector triple products; geometric applications. Differentiation of a vector function; scalar and vector fields. Gradient, divergence and curl - definitions and physical interpretations; product formulae; curvilinear coordinates. Gauss' and Stokes' theorems and evaluation of integrals over lines, surfaces and volumes. Derivation of continuity equations and Laplace's equation in Cartesian, cylindrical and spherical coordinate systems.

Course Content

- Introduction and revision of elementary concepts, scalar product, vector product.
- Triple products, multiple products, applications to geometry.
- Differentiation and integration of vector functions of a dingle variable.
- Curvilinear coordinate systems. Line, surface and volume integrals.
- Vector operators.
- Vector Identities.
- Gauss' and Stokes' Theorems.
- Engineering Applications.

Learning Outcomes

You should be comfortable with expressing systems (especially those in 2 and 3 dimensions) using vector quantities and manipulating these vectors without necessarily going back to some underlying coordinates.

You should have a sound grasp of the concept of a vector field, and be able to link this idea to descriptions various physical phenomena.

You should have a good intuition of the physical meaning of the various vector calculus operators and the important related theorems. You should be able to interpret the formulae describing physical systems in terms of this intuition.

References

Although these notes cover the material you need to know you should, wider reading is essential. Different explanations and different diagrams in books will give you the perspective to glue everything together, and further worked examples give you the confidence to tackle the tute sheets.

- J Heading, "Mathematical Methods in Science and Engineering", 2nd ed., Ch.13, (Arnold).
- G Stephenson, "Mathematical Methods for Science Students", 2nd ed., Ch.19, (Longman).
- E Kreyszig, "Advanced Engineering Mathematics", 6th ed., Ch.6, (Wiley).
- K F Riley, M. P. Hobson and S. J. Bence, "Mathematical Methods for the Physics and Engineering" Chs.6, 8 and 9, (CUP).
- A J M Spencer, et. al. "Engineering Mathematics", Vol.1, Ch.6, (Van Nostrand Reinhold).
- H M Schey, "Div, Grad, Curl and all that", Norton

Course WWW Pages

Pdf copies of these notes (including larger print versions), tutorial sheets, FAQs etc will be accessible from

$$\boxed{\text{www.robots.ox.ac.uk/}\sim\text{ian/Teaching/Vectors}}$$

Lecture 1

Vector Algebra

1.1 Vectors

Many physical quantities, such a mass, time, temperature are fully specified by one number or magnitude. They are **scalars**. But other quantities require more than one number to describe them. They are **vectors**. You have already met vectors in their more pure mathematical sense in your course on linear algebra (matrices and vectors), but often in the physical world, these numbers specify a **magnitude** and a **direction** — a total of two numbers in a 2D planar world, and three numbers in 3D.

Obvious examples are velocity, acceleration, electric field, and force. Below, probably all our examples will be of these "magnitude and direction" vectors, but we should not forget that many of the results extend to the wider realm of vectors.

There are three slightly different types of vectors:

- **Free vectors:** In many situtations *only* the magnitude and direction of a vector are important, and we can *translate* them at will (with 3 degrees of freedom for a vector in 3-dimensions).
- **Sliding vectors:** In mechanics the line of action of a force is often important for deriving moments. The force vector can slide with 1 degree of freedom.
- **Bound or position vectors:** When describing lines, curves etc in space, it is obviously important that the origin and head of the vector are not translated about arbitrarily. The origins of position vectors all coincide at an overall origin O.

One the advantages of using vectors is that it frees much of the analysis from the restriction of arbitrarily imposed coordinate frames. For example, if two free vectors are equal we need only say that their magnitudes and directions are equal, and that can be done with a drawing that is independent of any coordinate system.

However, coordinate systems are ultimately useful, so it useful to introduce the idea of vector components. Try to spot things in the notes that are independent

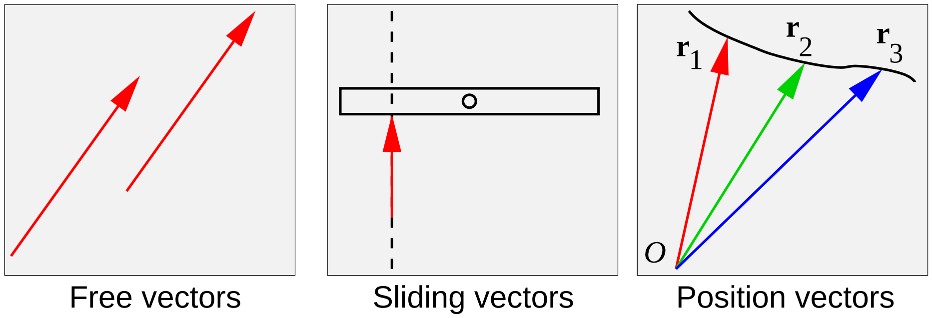


Figure 1.1:

of coordinate system.

1.1.1 Vector elements or components in a coordinate frame

A method of representing a vector is to list the values of its elements or components in a sufficient number of different (preferably mutually perpendicular) directions, depending on the dimension of the vector. These specified directions define a **coordinate frame**. In this course we will mostly restrict our attention to the 3-dimensional Cartesian coordinate frame $O(x, y, z)$. When we come to examine vector fields later in the course you will use curvilinear coordinate frames, especially 3D spherical and cylindrical polars, and 2D plane polar, coordinate systems.

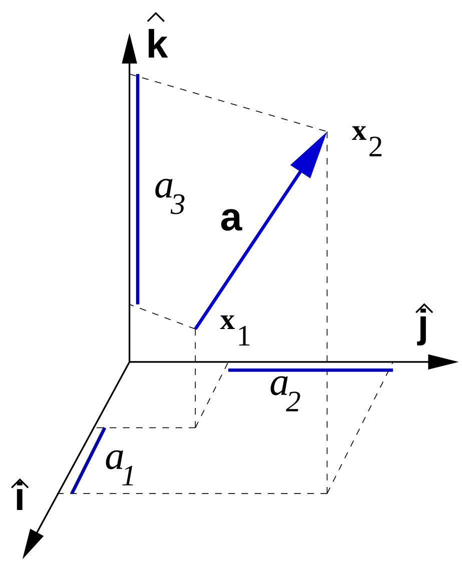


Figure 1.2: Vector components.

In a Cartesian coordinate frame we write

$$\mathbf{a} = [a_1, a_2, a_3] = [x_2 - x_1, y_2 - y_1, z_2 - z_1] \quad \text{or} \quad \mathbf{a} = [a_x, a_y, a_z]$$

as sketched in Figure 1.2. Defining $\hat{\boldsymbol{\imath}}, \hat{\boldsymbol{\jmath}}, \hat{\boldsymbol{k}}$ as unit vectors in the x, y, z directions

$$\hat{\boldsymbol{\imath}} = [1, 0, 0] \quad \hat{\boldsymbol{\jmath}} = [0, 1, 0] \quad \hat{\boldsymbol{k}} = [0, 0, 1]$$

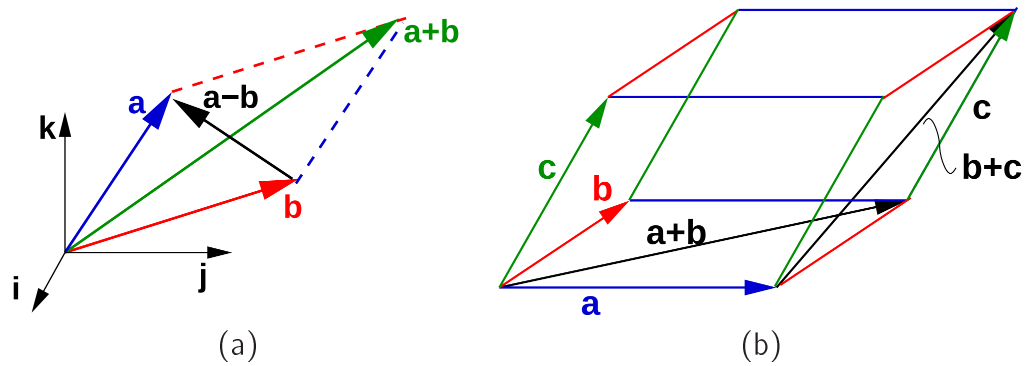


Figure 1.3: (a) Addition of two vectors is commutative. Note that the coordinate frame is irrelevant. (b) subtraction of vectors; (c) Addition of three vectors is associative.

we could also write

$$\mathbf{a} = a_1\hat{\boldsymbol{\imath}} + a_2\hat{\boldsymbol{\jmath}} + a_3\hat{\boldsymbol{k}} \quad .$$

Although we will be most often dealing with vectors in 3-space, you should not think that general vectors are limited to three components.

In these notes we will use bold font to represent vectors $\mathbf{a}, \boldsymbol{\omega}$, In your written work, underline the vector symbol $\underline{a}, \underline{\omega}$ and be **meticulous** about doing so. We shall use the hat to denote a unit vector.

1.1.2 Vector equality

Two free vectors are said to be equal iff their lengths and directions are the same. If we use a coordinate frame, we might say that corresponding components of the two vectors must be equal. This definition of equality will also do for position vectors, but for sliding vectors we must add that the line of action must be identical too.

1.1.3 Vector magnitude and unit vectors

Provided we use an orthogonal coordinate system, the magnitude of a 3-vector is

$$a = |\mathbf{a}| = \sqrt{a_1^2 + a_2^2 + a_3^2}$$

To find the unit vector in the direction of $\mathbf{a}$, simply divide by its magnitude

$$\hat{\mathbf{a}} = \frac{\mathbf{a}}{|\mathbf{a}|} \ .$$

1.1.4 Vector Addition and subtraction

Vectors are added/subtracted by adding/subtracting corresponding components, exactly as for matrices. Thus

$$\mathbf{a} + \mathbf{b} = [a_1 + b_1, \quad a_2 + b_2, \quad a_3 + b_3]$$

Addition follows the parallelogram construction of Figure 1.3(a). Subtraction ($\mathbf{a}-\mathbf{b}$) is defined as the addition ($\mathbf{a}+(-\mathbf{b})$). It is useful to remember that the vector $\mathbf{a}-\mathbf{b}$ goes from $\mathbf{b}$ *to* $\mathbf{a}$.

The following results follow immediately from the above definition of vector addition:

(a) $\mathbf{a}+\mathbf{b}=\mathbf{b}+\mathbf{a}$ (commutativity) (Figure 1.3(a))

(b) $(\mathbf{a}+\mathbf{b})+\mathbf{c}=\mathbf{a}+(\mathbf{b}+\mathbf{c})=\mathbf{a}+\mathbf{b}+\mathbf{c}$ (associativity) (Figure 1.3(b))

(c) $\mathbf{a}+\mathbf{0}=\mathbf{0}+\mathbf{a}=\mathbf{a}$, where the zero vector is $\mathbf{0}=[0,0,0]$.

(d) $\mathbf{a}+(\text{-}\mathbf{a})=\mathbf{0}$

1.1.5 Multiplication of a vector by a scalar. (NOT the scalar product!)

Just as for matrices, multiplication of a vector $\mathbf{a}$ by a scalar c is defined as multiplication of each component by c, so that

$$c\mathbf{a}=[ca_1,ca_2,ca_3].$$

It follows that:

$$|c\mathbf{a}|=\sqrt{(ca_1)^2+(ca_2)^2+(ca_3)^2}=|c||\mathbf{a}|.$$

The direction of the vector will reverse if c is negative, but otherwise is unaffected. (By the way, a vector where the sign is uncertain is called a director.)

♣ Example

Q. Coulomb's law states that the electrostatic force on charged particle Q due to another charged particle q_1 is

$$\mathbf{F}=K\frac{Qq_1}{r^2}\hat{\mathbf{e}}_r$$

where $\mathbf{r}$ is the vector from q_1 to Q and $\hat{\mathbf{r}}$ is the unit vector in that same direction. (Note that the rule "unlike charges attract, like charges repel" is built into this formula.) The force between two particles is not modified by the presence of other charged particles.

Hence write down an expression for the force on Q at $\mathbf{R}$ due to N charges q_i at $\mathbf{r}_i$.

A. The vector from q_i to Q is $\mathbf{R}-\mathbf{r}_i$. The unit vector in that direction is $(\mathbf{R}-\mathbf{r}_i)/|\mathbf{R}-\mathbf{r}_i|$, so the resultant force is

$$\mathbf{F}(\mathbf{R})=\sum_{i=1}^{N}K\frac{Qq_i}{|\mathbf{R}-\mathbf{r}_i|^3}(\mathbf{R}-\mathbf{r}_i)\quad .$$

Note that $\mathbf{F}(\mathbf{R})$ is a vector *field*.

1.2 Scalar, dot, or inner product

This is a product of two vectors results in a scalar quantity and is defined as follows for 3-component vectors:

$$\mathbf{a} \cdot \mathbf{b} = a_1 b_1 + a_2 b_2 + a_3 b_3 \ .$$

Note that

$$\mathbf{a} \cdot \mathbf{a} = a_1^2 + a_2^2 + a_3^2 = |\mathbf{a}|^2 = a^2.$$

The following laws of multiplication follow immediately from the definition:

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

(b) $\mathbf{a} \cdot (\mathbf{b} + \mathbf{c}) = \mathbf{a} \cdot \mathbf{b} + \mathbf{a} \cdot \mathbf{c}$ (distributivity with respect to vector addition)

(c) $(\lambda \mathbf{a}) \cdot \mathbf{b} = \lambda(\mathbf{a} \cdot \mathbf{b}) = \mathbf{a} \cdot (\lambda \mathbf{b})$ scalar multiple of a scalar product of two vectors

1.2.1 Geometrical interpretation of scalar product

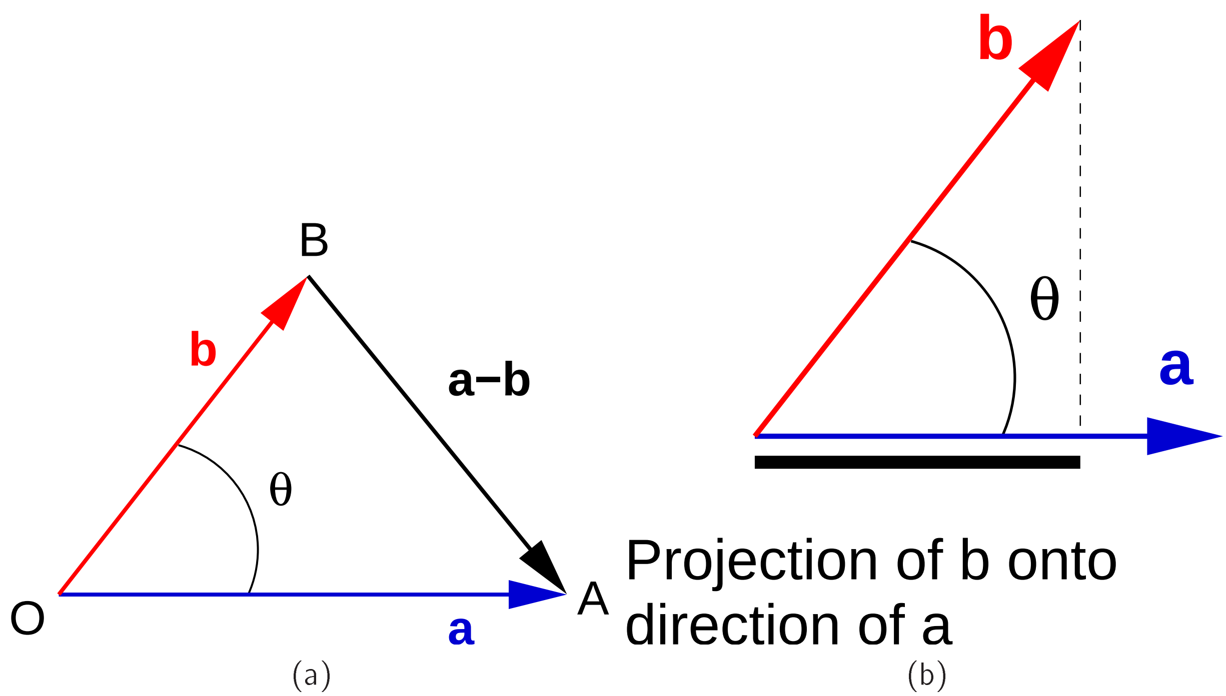


Figure 1.4: (a) Cosine rule. (b) Projection of **b** onto **a**.

Consider the square magnitude of the vector $\mathbf{a} - \mathbf{b}$. By the rules of the scalar product, this is

$$\begin{aligned} |\mathbf{a} - \mathbf{b}|^2 &= (\mathbf{a} - \mathbf{b}) \cdot (\mathbf{a} - \mathbf{b}) \\ &= \mathbf{a} \cdot \mathbf{a} + \mathbf{b} \cdot \mathbf{b} - 2(\mathbf{a} \cdot \mathbf{b}) \\ &= a^2 + b^2 - 2(\mathbf{a} \cdot \mathbf{b}) \end{aligned}$$

But, by the cosine rule for the triangle OAB (Figure 1.4a), the length AB^2 is given by

$$|\mathbf{a} - \mathbf{b}|^2 = a^2 + b^2 - 2ab\cos\theta$$

where θ is the angle between the two vectors. It follows that

$$\mathbf{a} \cdot \mathbf{b} = ab\cos\theta,$$

which is independent of the co-ordinate system used, and that $|\mathbf{a} \cdot \mathbf{b}| \leq ab$. Conversely, the cosine of the angle between vectors **a** and **b** is given by $\cos\theta = \mathbf{a}\cdot\mathbf{b}/ab$.

1.2.2 Projection of one vector onto the other

Another way of describing the scalar product is as the product of the magnitude of one vector and the component of the other in the direction of the first, since $b\cos\theta$ is the component of **b** in the direction of **a** and vice versa (Figure 1.4b).

Projection is particularly useful when the second vector is a unit vector — $\mathbf{a} \cdot \hat{\boldsymbol{\imath}}$ is the component of **a** in the direction of $\hat{\boldsymbol{\imath}}$.

Notice that if we wanted the **vector component** of **b** in the direction of **a** we would write

$$(\mathbf{b} \cdot \hat{\mathbf{a}})\hat{\mathbf{a}} = \frac{(\mathbf{b} \cdot \mathbf{a})\mathbf{a}}{a^2} \quad .$$

In the particular case $\mathbf{a} \cdot \mathbf{b} = 0$, the angle between the two vectors is a right angle and the vectors are said to be mutually orthogonal or perpendicular — neither vector has any component in the direction of the other.

An orthonormal coordinate system is characterised by $\hat{\boldsymbol{\imath}} \cdot \hat{\boldsymbol{\imath}} = \hat{\boldsymbol{\jmath}} \cdot \hat{\boldsymbol{\jmath}} = \hat{\boldsymbol{k}} \cdot \hat{\boldsymbol{k}} = 1$; and $\hat{\boldsymbol{\imath}} \cdot \hat{\boldsymbol{\jmath}} = \hat{\boldsymbol{\jmath}} \cdot \hat{\boldsymbol{k}} = \hat{\boldsymbol{k}} \cdot \hat{\boldsymbol{\imath}} = 0$.

1.2.3 A scalar product is an "inner product"

So far we have been writing our vectors as row vectors $\mathbf{a} = [a_1, \ a_2, \ a_3]$. This is convenient because it takes up less room than writing column vectors

$$\mathbf{a} = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} \quad .$$

In matrix algebra vectors are more usually defined as column vectors, as in

$$\begin{bmatrix} M_{11} & M_{12} & M_{13} \\ M_{21} & M_{22} & M_{23} \\ M_{31} & M_{32} & M_{33} \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix}$$

and a row vector is written as $\mathbf{a}^\top$. Now for most of our work we can be quite relaxed about this minor difference, but here let us be fussy.

Why? Simply to point out at that the scalar product is also the **inner product** more commonly used in linear algebra. Defined as $\mathbf{a}^\top\mathbf{b}$ when vectors are column vectors as

$$\mathbf{a}\cdot\mathbf{b} = \mathbf{a}^\top\mathbf{b} = [a_1,\ a_2,\ a_3]\begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} = a_1b_1 + a_2b_2 + a_3b_3\ .$$

Here we treat a n-dimensional column vector as an $n \times 1$ matrix.

(Remember that if you multiply two matrices $\mathsf{M}_{m\times n}\mathsf{N}_{n\times p}$ then M must have the same columns as N has rows (here denoted by n) and the result has size (rows × columns) of $m \times p$. So for n-dimensional column vectors **a** and **b**, $\mathbf{a}^\top$ is a $1 \times n$ matrix and **b** is $n \times 1$ matrix, so the product $\mathbf{a}^\top\mathbf{b}$ is a 1×1 matrix, which is (at last!) a scalar.)

♣ Examples

Q1. A force **F** is applied to an object as it moves by a small amount $\delta\mathbf{r}$. What work is done on the object by the force?

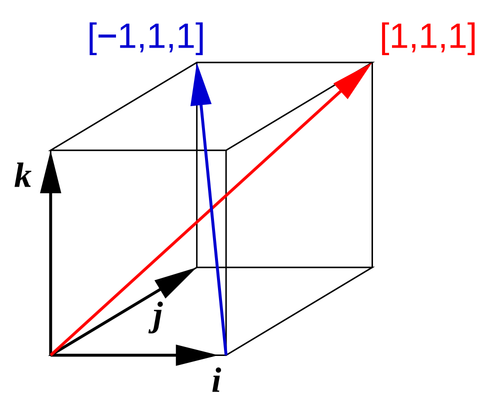
A1. The work done is equal to the component of force in the direction of the displacement multiplied by the displacement itself. This is just a scalar product:

$$\delta W = \mathbf{F}\cdot\delta\mathbf{r}\ .$$

Q2. A cube has four diagonals, connecting opposite vertices. What is the angle between an adjacent pair?

A2. Well, you could plod through using Pythagoras' theorem to find the length of the diagonal from cube vertex to cube centre, and perhaps you should to check the following answer.

The directions of the diagonals are $[\pm1, \pm1, \pm1]$. The ones shown in the figure are $[1, 1, 1]$ and $[-1, 1, 1]$. The angle is thus



$$\theta = \cos^{-1}\frac{[1,1,1]\cdot[-1,1,1]}{\sqrt{1^2+1^2+1^2}\sqrt{-1^2+1^2+1^2}} = \cos^{-1}\frac{1}{3}$$

Q3. A pinball moving in a plane with velocity $\mathbf{s}$ bounces (in a purely elastic impact) from a baffle whose endpoints are $\mathbf{p}$ and $\mathbf{q}$. What is the velocity vector after the bounce?

A3. Best to refer everything to a coordinate frame with principal directions $\hat{\mathbf{u}}$ along and $\hat{\mathbf{v}}$ perpendicular to the baffle:

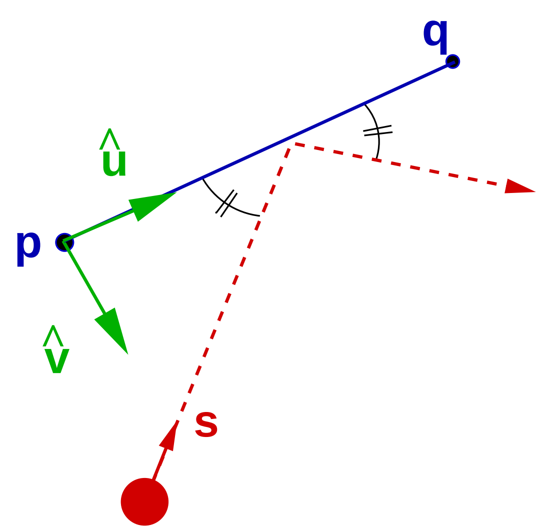
$$\hat{\mathbf{u}} = \frac{\mathbf{q}-\mathbf{p}}{|\mathbf{q}-\mathbf{p}|}$$
$$\hat{\mathbf{v}} = \mathbf{u}^{\perp} = [-u_y, \quad u_x]$$

Thus the velocity before impact is

$$\mathbf{s}_{\text{before}} = (\mathbf{s}.\hat{\mathbf{u}})\hat{\mathbf{u}} + (\mathbf{s}.\hat{\mathbf{v}})\hat{\mathbf{v}}$$

After the impact, the component of velocity in the direction of the baffle is unchanged and the component normal to the baffle is negated:

$$\mathbf{s}_{\text{after}} = (\mathbf{s}.\hat{\mathbf{u}})\hat{\mathbf{u}} - (\mathbf{s}.\hat{\mathbf{v}})\hat{\mathbf{v}}$$



1.2.4 Direction cosines use projection

Direction cosines are commonly used in the field of crystallography. The quantities

$$\lambda = \frac{\mathbf{a}\cdot\hat{\imath}}{a}, \qquad \mu = \frac{\mathbf{a}\cdot\hat{\jmath}}{a}, \qquad \nu = \frac{\mathbf{a}\cdot\hat{\mathbf{k}}}{a}$$

represent the cosines of the angles which the vector $\mathbf{a}$ makes with the co-ordinate vectors $\hat{\imath}$,$\hat{\jmath}$, $\hat{\mathbf{k}}$ and are known as the direction cosines of the vector $\mathbf{a}$. Since $\mathbf{a}\cdot\hat{\imath} = a_1$ etc, it follows immediately that $\mathbf{a} = a(\lambda\hat{\imath} + \mu\hat{\jmath} + \nu\hat{\mathbf{k}})$ and $\lambda^2+\mu^2+\nu^2 = \frac{1}{a^2}[a_1^2+a_2^2+a_3^2] = 1$

1.3 Vector or cross product

The vector product of two vectors $\mathbf{a}$ and $\mathbf{b}$ is denoted by $\mathbf{a}\times\mathbf{b}$ and is defined as follows

$$\mathbf{a}\times\mathbf{b} = (a_2b_3 - a_3b_2)\hat{\imath} + (a_3b_1 - a_1b_3)\hat{\jmath} + (a_1b_2 - a_2b_1)\hat{\mathbf{k}}.$$

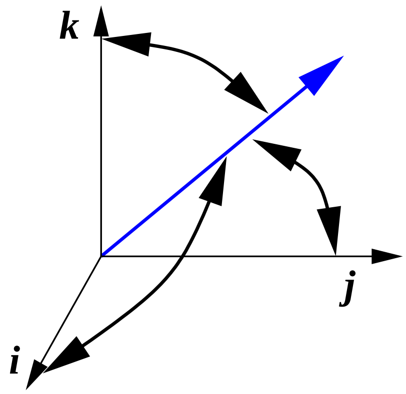


Figure 1.5: The direction cosines are cosines of the angles shown.

It is MUCH more easily remembered in terms of the pseudo-determinant

$$\mathbf{a} \times \mathbf{b} = \begin{vmatrix} \hat{\boldsymbol{\imath}} & \hat{\boldsymbol{\jmath}} & \hat{\boldsymbol{k}} \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix}$$

where the top row consists of the vectors $\hat{\boldsymbol{\imath}}$, $\hat{\boldsymbol{\jmath}}$, $\hat{\boldsymbol{k}}$ rather than scalars.

Since a determinant with two equal rows has value zero, it follows that $\mathbf{a} \times \mathbf{a} = \mathbf{0}$. It is also easily verified that $(\mathbf{a} \times \mathbf{b}) \cdot \mathbf{a} = (\mathbf{a} \times \mathbf{b}) \cdot \mathbf{b} = 0$, so that $\mathbf{a} \times \mathbf{b}$ is orthogonal (perpendicular) to both $\mathbf{a}$ and $\mathbf{b}$, as shown in Figure 1.6.

Note that $\hat{\boldsymbol{\imath}} \times \hat{\boldsymbol{\jmath}} = \hat{\boldsymbol{k}}$, $\hat{\boldsymbol{\jmath}} \times \hat{\boldsymbol{k}} = \hat{\imath}$, and $\hat{\boldsymbol{k}} \times \hat{\boldsymbol{\imath}} = \hat{\boldsymbol{\jmath}}$.

The magnitude of the vector product can be obtained by showing that

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

from which it follows that

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

which is again independent of the co-ordinate system used. This is left as an exercise.

Unlike the scalar product, the vector product does not satisfy commutativity but is in fact anti-commutative, in that $\mathbf{a} \times \mathbf{b} = -\mathbf{b} \times \mathbf{a}$. Moreover the vector product does not satisfy the associative law of multiplication either since, as we shall see later $\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) \neq (\mathbf{a} \times \mathbf{b}) \times \mathbf{c}$.

Since the vector product is known to be orthogonal to both the vectors which form the product, it merely remains to specify its sense with respect to these vectors. Assuming that the co-ordinate vectors form a right-handed set in the order $\hat{\boldsymbol{\imath}}$,$\hat{\boldsymbol{\jmath}}$, $\hat{\boldsymbol{k}}$ it can be seen that the sense of the the vector product is also right handed, i.e the vector product has the same sense as the co-ordinate system used.

$$\hat{\imath} \times \hat{\jmath} = \begin{vmatrix} \hat{\imath} & \hat{\jmath} & \hat{k} \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{vmatrix} = \hat{k} \quad .$$

In practice, figure out the direction from a right-handed screw twisted from the first to second vector as shown in Figure 1.6(a).

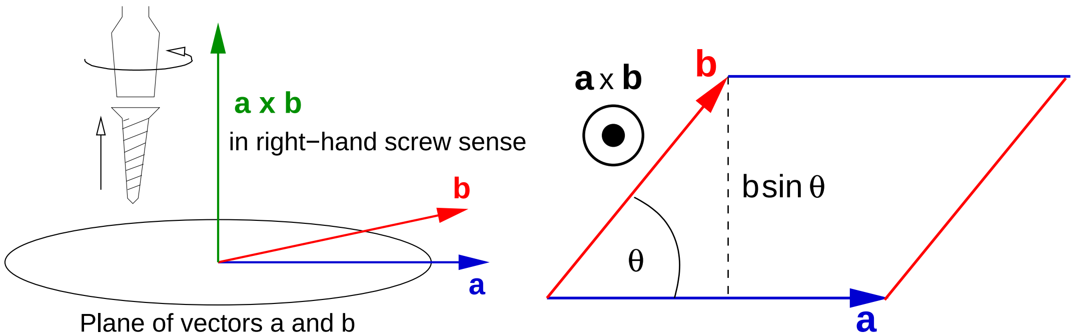


Figure 1.6: (a)The vector product is orthogonal to both **a** and **b**. Twist from first to second and move in the direction of a right-handed screw. (b) Area of parallelogram is $ab \sin\theta$.

1.3.1 Geometrical interpretation of vector product

The magnitude of the vector product ($\mathbf{a} \times \mathbf{b}$) is equal to the area of the parallelogram whose sides are parallel to, and have lengths equal to the magnitudes of, the vectors **a** and **b** (Figure 1.6b). Its direction is perpendicular to the parallelogram.

♣ **Example**

Q. $\mathbf{g}$ is vector from A [1,2,3] to B [3,4,5].
$\hat{\boldsymbol{\ell}}$ is the unit vector in the direction from O to A.
Find $\hat{\mathbf{m}}$, a UNIT vector along $\mathbf{g} \times \hat{\boldsymbol{\ell}}$
Verify that $\hat{\mathbf{m}}$ is is perpendicular to $\hat{\boldsymbol{\ell}}$.
Find $\hat{\mathbf{n}}$, the third member of a right-handed coordinate set $\hat{\boldsymbol{\ell}}, \hat{\mathbf{m}}, \hat{\mathbf{n}}$.

A.

$$\mathbf{g} = [3, 4, 5] - [1, 2, 3] = [2, 2, 2]$$

$$\hat{\boldsymbol{\ell}} = \frac{1}{\sqrt{14}}[1, 2, 3]$$

$$\mathbf{g} \times \hat{\boldsymbol{\ell}} = \frac{1}{\sqrt{14}} \begin{vmatrix} \hat{\imath} & \hat{\jmath} & \hat{k} \\ 2 & 2 & 2 \\ 1 & 2 & 3 \end{vmatrix} = \frac{1}{\sqrt{14}}[2, -4, 2]$$

Hence

$$\hat{\mathbf{m}} = \cdot \frac{1}{\sqrt{24}}[2, -4, 2]$$

and

$$\hat{\mathbf{n}} = \hat{\boldsymbol{\ell}} \times \hat{\mathbf{m}}$$

Lecture 2

Multiple Products. Geometry using Vectors

2.1 Triple and multiple products

Using mixtures of the pairwise scalar product and vector product, it is possible to derive "triple products" between three vectors, and indeed n-products between n vectors.

There is nothing about these that you cannot work out from the definitions of pairwise scalar and vector products already given, but some have interesting geometric interpretations, so it is worth looking at these.

2.1.1 Scalar triple product

This is the scalar product of a vector product and a third vector, i.e. $\mathbf{a}\cdot(\mathbf{b}\times\mathbf{c})$. Using the pseudo-determinant expression for the vector product, we see that the scalar triple product can be represented as the true determinant

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

You will recall that if you swap a pair of rows of a determinant, its sign changes; hence if you swap two pairs, its sign stays the same.

$$\underset{+}{\begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix}} \quad \text{1st SWAP} \quad \underset{-}{\begin{vmatrix} c_1 & c_2 & c_3 \\ b_1 & b_2 & b_3 \\ a_1 & a_2 & a_3 \end{vmatrix}} \quad \text{2nd SWAP} \quad \underset{+}{\begin{vmatrix} c_1 & c_2 & c_3 \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix}}$$

This says that

(i) $\mathbf{a}\cdot(\mathbf{b}\times\mathbf{c}) = \mathbf{b}\cdot(\mathbf{c}\times\mathbf{a}) = \mathbf{c}\cdot(\mathbf{a}\times\mathbf{b})$ (Called cyclic permutation.)

(ii) $\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c}) = -\mathbf{b} \cdot (\mathbf{a} \times \mathbf{c})$ and so on. (Called anti-cyclic permutation.)

(iii) The fact that $\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c}) = (\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c}$ allows the scalar triple product to be written as $[\mathbf{a}, \mathbf{b}, \mathbf{c}]$. This notation is not very helpful, and we will try to avoid it below.

2.1.2 Geometrical interpretation of scalar triple product

The scalar triple product gives the volume of the parallelopiped whose sides are represented by the vectors **a**, **b**, and **c**.

We saw earlier that the vector product $(\mathbf{a} \times \mathbf{b})$ has magnitude equal to the area of the base, and direction perpendicular to the base. The *component* of **c** in this direction is equal to the height of the parallelopiped shown in Figure 2.1(a).

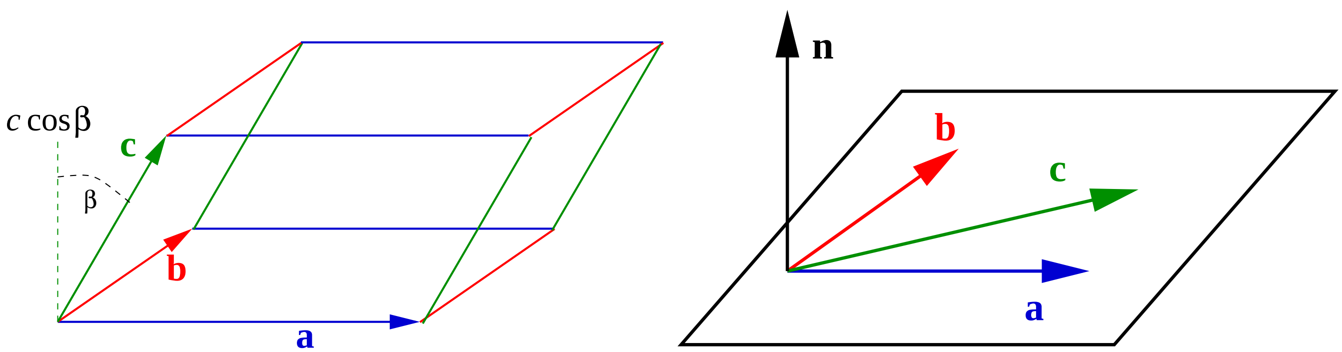


Figure 2.1: (a) Scalar triple product equals volume of parallelopiped. (b) Coplanarity yields zero scalar triple product.

2.1.3 Linearly dependent vectors

If the scalar triple product of three vectors is zero

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

then the vectors are **linearly dependent**. That is, one can be expressed as a linear combination of the others. For example,

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

where λ and μ are scalar coefficients.

You can see this immediately in two ways:

- The determinant would have one row that was a linear combination of the others. You'll remember that by doing row operations, you could reach a row of zeros, and so the determinant is zero.
- The parallelopiped would have zero volume if squashed flat. In this case all three vectors lie in a plane, and so any one is a linear combination of the other two. (Figure 2.1b.)

2.1.4 Vector triple product

This is defined as the vector product of a vector with a vector product, $\mathbf{a}\times(\mathbf{b}\times\mathbf{c})$. Now, the vector triple product $\mathbf{a}\times(\mathbf{b}\times\mathbf{c})$ must be perpendicular to $(\mathbf{b}\times\mathbf{c})$, which in turn is perpendicular to both $\mathbf{b}$ and $\mathbf{c}$. Thus $\mathbf{a}\times(\mathbf{b}\times\mathbf{c})$ can have no component perpendicular to $\mathbf{b}$ and $\mathbf{c}$, and hence must be coplanar with them. It follows that the vector triple product must be expressible as a linear combination of $\mathbf{b}$ and $\mathbf{c}$:

$$\mathbf{a}\times(\mathbf{b}\times\mathbf{c}) = \lambda\mathbf{b} + \mu\mathbf{c} \; .$$

The values of the coefficients can be obtained by multiplying out in component form. Only the first component need be evaluated, the others then being obtained by symmetry. That is

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

The equivalents must be true for the 2nd and 3rd components, so we arrive at the identity

$$\mathbf{a}\times(\mathbf{b}\times\mathbf{c}) = (\mathbf{a}\cdot\mathbf{c})\mathbf{b} - (\mathbf{a}\cdot\mathbf{b})\mathbf{c} \; .$$

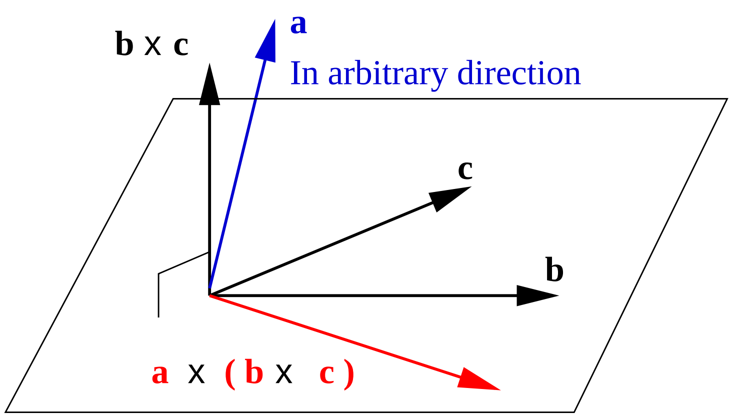


Figure 2.2: Vector triple product.

2.1.5 Projection using vector triple product

An example of the application of this formula is as follows. Suppose $\mathbf{v}$ is a vector and we want its projection into the xy-plane. The component of $\mathbf{v}$ in the z direction is $\mathbf{v}\cdot\hat{\boldsymbol{k}}$, so the projection we seek is $\mathbf{v} - (\mathbf{v}\cdot\hat{\boldsymbol{k}})\hat{\boldsymbol{k}}$. Writing $\hat{\boldsymbol{k}} \leftarrow \mathbf{a}$, $\mathbf{v} \leftarrow \mathbf{b}$,

$\hat{\boldsymbol{k}} \leftarrow \mathbf{c}$,

$$\begin{array}{rcl} \mathbf{a} \times (\mathbf{b} \times \mathbf{c}) & = & (\mathbf{a}\cdot\mathbf{c})\mathbf{b} - (\mathbf{a}\cdot\mathbf{b})\mathbf{c} \\ \downarrow \quad\quad \downarrow \quad \downarrow & & \\ \hat{\boldsymbol{k}} \times (\mathbf{v} \times \hat{\boldsymbol{k}}) & = & (\hat{\boldsymbol{k}}\cdot\hat{\boldsymbol{k}})\mathbf{v} - (\hat{\boldsymbol{k}}\cdot\mathbf{v})\hat{\boldsymbol{k}} \\ & = & \mathbf{v} - (\mathbf{v}\cdot\hat{\boldsymbol{k}})\hat{\boldsymbol{k}} \end{array}$$

So $\mathbf{v} - (\mathbf{v}\cdot\hat{\boldsymbol{k}})\hat{\boldsymbol{k}} = \hat{\boldsymbol{k}} \times (\mathbf{v} \times \hat{\boldsymbol{k}})$.

(Hot stuff! But the expression $\mathbf{v} - (\mathbf{v}\cdot\hat{\boldsymbol{k}})\hat{\boldsymbol{k}}$ is much easier to understand, and cheaper to compute!)

2.1.6 Other repeated products

Many combinations of vector and scalar products are possible, but we consider only one more, namely the vector quadruple product $(\mathbf{a} \times \mathbf{b}) \times (\mathbf{c} \times \mathbf{d})$. By regarding $\mathbf{a} \times \mathbf{b}$ as a single vector, we see that this vector must be representable as a linear combination of $\mathbf{c}$ and $\mathbf{d}$. On the other hand, regarding $\mathbf{c} \times \mathbf{d}$ as a single vector, we see that it must also be a linear combination of $\mathbf{a}$ and $\mathbf{b}$. This provides a means of expressing one of the vectors, say $\mathbf{d}$, as linear combination of the other three, as follows:

$$\begin{array}{rcl} (\mathbf{a} \times \mathbf{b}) \times (\mathbf{c} \times \mathbf{d}) & = & [(\mathbf{a} \times \mathbf{b})\cdot\mathbf{d}]\mathbf{c} - [(\mathbf{a} \times \mathbf{b})\cdot\mathbf{c}]\mathbf{d} \\ & = & [(\mathbf{c} \times \mathbf{d})\cdot\mathbf{a}]\mathbf{b} - [(\mathbf{c} \times \mathbf{d})\cdot\mathbf{b}]\mathbf{a} \end{array}$$

Hence

$$[(\mathbf{a} \times \mathbf{b})\cdot\mathbf{c}]\,\mathbf{d} = [(\mathbf{b} \times \mathbf{c})\cdot\mathbf{d}]\,\mathbf{a} + [(\mathbf{c} \times \mathbf{a})\cdot\mathbf{d}]\,\mathbf{b} + [(\mathbf{a} \times \mathbf{b})\cdot\mathbf{d}]\,\mathbf{c}$$

or

$$\mathbf{d} = \frac{[(\mathbf{b} \times \mathbf{c})\cdot\mathbf{d}]\,\mathbf{a} + [(\mathbf{c} \times \mathbf{a})\cdot\mathbf{d}]\,\mathbf{b} + [(\mathbf{a} \times \mathbf{b})\cdot\mathbf{d}]\,\mathbf{c}}{[(\mathbf{a} \times \mathbf{b})\cdot\mathbf{c}]} \;=\; \alpha\mathbf{a} + \beta\mathbf{b} + \gamma\mathbf{c} \quad .$$

This is not something to remember off by heart, but it is worth remembering that the projection of a vector on any arbitrary basis set is unique.

♣ Example

Q1 Use the quadruple vector product to express the vector $\mathbf{d} = [3, 2, 1]$ in terms of the vectors $\mathbf{a} = [1, 2, 3]$, $\mathbf{b} = [2, 3, 1]$ and $\mathbf{c} = [3, 1, 2]$.

A1 Grinding away at the determinants, we find

$$[(\mathbf{a}\times\mathbf{b})\cdot\mathbf{c}] = -18; \quad [(\mathbf{b}\times\mathbf{c})\cdot\mathbf{d}] = 6; \quad [(\mathbf{c}\times\mathbf{a})\cdot\mathbf{d}] = -12; \quad [(\mathbf{a}\times\mathbf{b})\cdot\mathbf{d}] = -12$$

So, $\mathbf{d} = (-\mathbf{a} + 2\mathbf{b} + 2\mathbf{c})/3$.

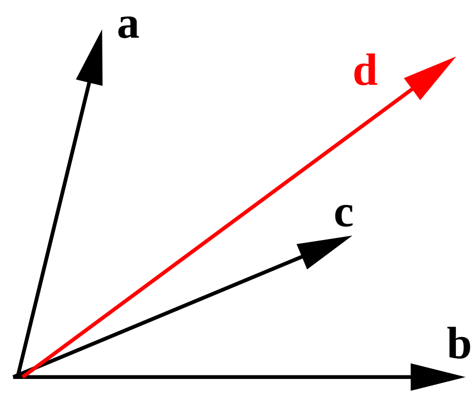


Figure 2.3: The projection of a (3-)vector onto a set of (3) basis vectors is unique. Ie in $\mathbf{d} = \alpha\mathbf{a} + \beta\mathbf{b} + \gamma\mathbf{c}$, the set $\{\alpha, \beta, \gamma\}$ is unique.

2.2 Geometry using vectors: lines, planes

2.2.1 The equation of a line

The equation of the line passing through the point whose position vector is **a** and lying in the direction of vector **b** is

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

where λ is a scalar parameter. If you make **b** a unit vector, $\mathbf{r} = \mathbf{a} + \lambda\hat{\mathbf{b}}$ then λ will represent metric length.

For a line defined by two points $\mathbf{a}_1$ and $\mathbf{a}_2$

$$\mathbf{r} = \mathbf{a}_1 + \lambda(\mathbf{a}_2 - \mathbf{a}_1)$$

or for the unit version

$$\mathbf{r} = \mathbf{a}_1 + \lambda\frac{(\mathbf{a}_2 - \mathbf{a}_1)}{|\mathbf{a}_2 - \mathbf{a}_1|}$$

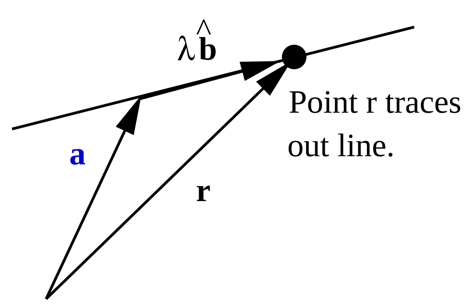


Figure 2.4: Equation of a line. With $\hat{\mathbf{b}}$ a unit vector, λ is in the length units established by the definition of **a**.

2.2.2 The shortest distance from a point to a line

Referring to Figure 2.5(a) the vector $\mathbf{p}$ from $\mathbf{c}$ to any point on the line is $\mathbf{p} = \mathbf{a} + \lambda\hat{\mathbf{b}} - \mathbf{c} = (\mathbf{a}-\mathbf{c}) + \lambda\hat{\mathbf{b}}$ which has length squared $p^2 = (\mathbf{a}-\mathbf{c})^2 + \lambda^2 + 2\lambda(\mathbf{a}-\mathbf{c})\cdot\hat{\mathbf{b}}$. Rather than minimizing length, it is easier to minimize length-squared. The minumum is found when $d\ p^2/d\lambda = 0$, ie when

$$\lambda = -(\mathbf{a}-\mathbf{c})\cdot\hat{\mathbf{b}}\ .$$

So the minimum length vector is

$$\mathbf{p} = (\mathbf{a}-\mathbf{c}) - ((\mathbf{a}-\mathbf{c})\cdot\hat{\mathbf{b}})\hat{\mathbf{b}}.$$

You might spot that is the component of $(\mathbf{a}-\mathbf{c})$ perpendicular to $\hat{\mathbf{b}}$ (as expected!). Furthermore, using the result of Section 2.1.5,

$$\mathbf{p} = \hat{\mathbf{b}} \times [(\mathbf{a}-\mathbf{c})\times\hat{\mathbf{b}}]\ .$$

Because $\hat{\mathbf{b}}$ is a unit vector, and is orthogonal to $[(\mathbf{a}-\mathbf{c})\times\hat{\mathbf{b}}]$, the modulus of the vector can be written rather more simply as just

$$p_{\text{min}} = |(\mathbf{a}-\mathbf{c})\times\hat{\mathbf{b}}|\ \ .$$

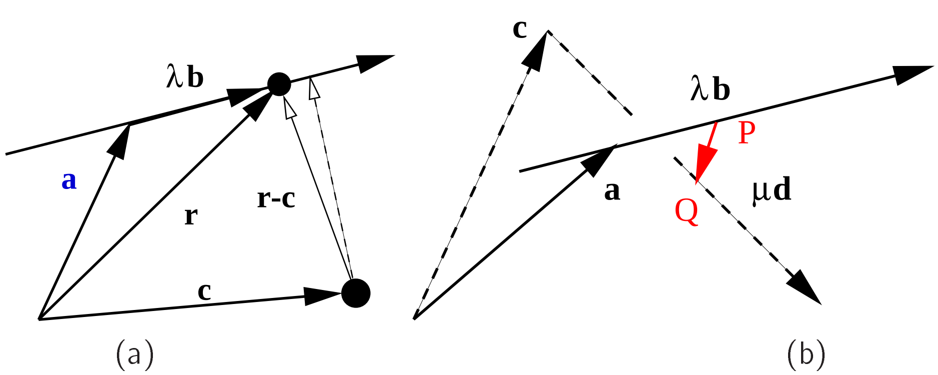


Figure 2.5: (a) Shortest distance point to line. (b) Shortest distance, line to line.

2.2.3 The shortest distance between two straight lines

If the shortest distance between a point and a line is along the perpendicular, then the shortest distance between the two straight lines $\mathbf{r} = \mathbf{a} + \lambda\hat{\mathbf{b}}$ and $\mathbf{r} = \mathbf{c} + \mu\hat{\mathbf{d}}$ must be found as the length of the vector which is mutually perpendicular to the lines.

The unit vector along the mutual perpendicular is

$$\hat{\mathbf{p}} = (\hat{\mathbf{b}}\times\hat{\mathbf{d}})/|\hat{\mathbf{b}}\times\hat{\mathbf{d}}|\ .$$

(Yes, don't forget that $\hat{\mathbf{b}}\times\hat{\mathbf{d}}$ is NOT a unit vector. $\hat{\mathbf{b}}$ and $\hat{\mathbf{d}}$ are not orthogonal, so there is a $\sin\theta$ lurking!)

The minimum length is therefore the component of $\mathbf{a}-\mathbf{c}$ in this direction

$$p_{\text{min}} = \left|(\mathbf{a}-\mathbf{c})\cdot(\hat{\mathbf{b}}\times\hat{\mathbf{d}})/|\hat{\mathbf{b}}\times\hat{\mathbf{d}}|\right|\ .$$

♣ Example

Q Two long straight pipes are specified using Cartesian co-ordinates as follows:
Pipe A has diameter 0.8 and its axis passes through points $(2, 5, 3)$ and $(7, 10, 8)$.
Pipe B has diameter 1.0 and its axis passes through the points $(0, 6, 3)$ and $(-12, 0, 9)$.
Determine whether the pipes need to be realigned to avoid intersection.

A Each pipe axis is defined using two points. The vector equation of the axis of pipe A is

$$\mathbf{r} = [2, 5, 3] + \lambda'[5, 5, 5] = [2, 5, 3] + \lambda[1, 1, 1]/\sqrt{3}$$

The equation of the axis of pipe B is

$$\mathbf{r} = [0, 6, 3] + \mu'[12, 6, 6] = [0, 6, 3] + \mu[-2, -1, 1]/\sqrt{6}$$

The perpendicular to the two axes has direction

$$[1, 1, 1] \times [-2, -1, 1] = \begin{vmatrix} \hat{\boldsymbol{\imath}} & \hat{\boldsymbol{\jmath}} & \hat{\boldsymbol{k}} \\ 1 & 1 & 1 \\ -2 & -1 & 1 \end{vmatrix} = [2, -3, 1] = \mathbf{p}$$

The length of the mutual perpendicular is

$$(\mathbf{a} - \mathbf{c}) \cdot \frac{[2, -3, 1]}{\sqrt{14}} = [2, -1, 0] \cdot \frac{[2, -3, 1]}{\sqrt{14}} = 1.87\ .$$

But the sum of the radii of the two pipes is $0.4 + 0.5 = 0.9$. Hence the pipes do not intersect.

2.2.4 The equation of a plane

There are a number of ways of specifying the equation of a plane.

1. If $\mathbf{b}$ and $\mathbf{c}$ are two non-parallel vectors (ie $\mathbf{b} \times \mathbf{c} \neq \mathbf{0}$), then the equation of the plane passing through the point $\mathbf{a}$ and parallel to the vectors $\mathbf{b}$ and $\mathbf{c}$ may be written in the form

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

   where λ, μ are scalar parameters. Note that $\mathbf{b}$ and $\mathbf{c}$ are free vectors, so don't have to lie in the plane (Figure 2.6(a).)

2. Figure 2.6(b) shows the plane defined by three non-collinear points $\mathbf{a}$, $\mathbf{b}$ and $\mathbf{c}$ in the plane (note that the vectors $\mathbf{b}$ and $\mathbf{c}$ are position vectors, not free vectors as in the previous case). The equation might be written as

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

3. Figure 2.6(c) illustrates another description is in terms of the unit normal to the plane $\hat{\mathbf{n}}$ and a point $\mathbf{a}$ in the plane

   $$\mathbf{r} \cdot \hat{\mathbf{n}} = \mathbf{a} \cdot \hat{\mathbf{n}}\ .$$

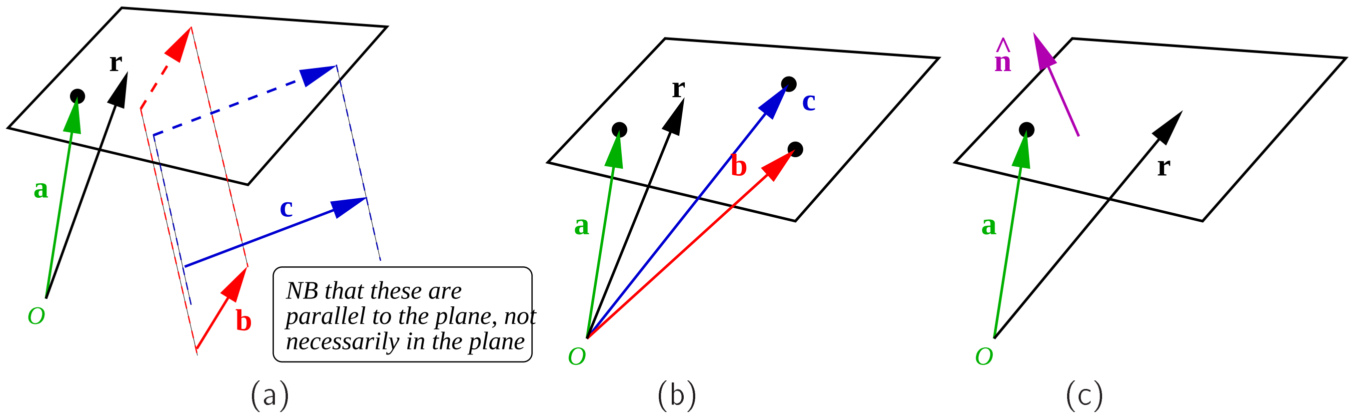


Figure 2.6: (a) Plane defined using point and two lines. (b) Plane defined using three points. (c) Plane defined using point and normal. Vector **r** is the position vector of a general point in the plane.

2.2.5 The shortest distance from a point to a plane

The shortest distance from a point **d** to the plane is along the perpendicular. Depending on how the plane is defined, this can be written as

$$D = |(\mathbf{d} - \mathbf{a}) \cdot \hat{\mathbf{n}}| \qquad \text{or} \qquad D = \frac{|(\mathbf{d} - \mathbf{a}) \cdot (\mathbf{b} \times \mathbf{c})|}{|\mathbf{b} \times \mathbf{c}|} \quad .$$

2.3 Solution of vector equations

It is sometimes required to obtain the most general vector which satisfies a given vector relationship. This is very much like obtaining the locus of a point. The best method of proceeding in such a case is as follows:

(i) Decide upon a system of three co-ordinate vectors using two non-parallel vectors appearing in the vector relationship. These might be **a**, **b** and their vector product $(\mathbf{a} \times \mathbf{b})$.

(ii) Express the unknown vector **x** as a linear combination of these vectors

$$\mathbf{x} = \lambda\mathbf{a} + \mu\mathbf{b} + \nu\mathbf{a} \times \mathbf{b}$$

where λ, μ, ν are scalars to be found.

(iii) Substitute the above expression for **x** into the vector relationship to determine the constraints on λ, μ and ν for the relationship to be satisfied.

♣ Example

Q Solve the vector equation $\mathbf{x} = \mathbf{x} \times \mathbf{a} + \mathbf{b}$.

A Step (i): Set up basis vectors **a**, **b** and their vector product $\mathbf{a} \times \mathbf{b}$.

Step (ii): $\mathbf{x} = \lambda\mathbf{a} + \mu\mathbf{b} + \nu\mathbf{a} \times \mathbf{b}$.

Step (iii): Bung this expression for $\mathbf{x}$ into the equation!

$$\begin{aligned}\lambda\mathbf{a}+\mu\mathbf{b}+\nu\mathbf{a}\times\mathbf{b} &= (\lambda\mathbf{a}+\mu\mathbf{b}+\nu\mathbf{a}\times\mathbf{b})\times\mathbf{a}+\mathbf{b}\\ &= \mathbf{0}+\mu(\mathbf{b}\times\mathbf{a})+\nu(\mathbf{a}\times\mathbf{b})\times\mathbf{a}+\mathbf{b}\\ &= -\nu(\mathbf{a}\cdot\mathbf{b})\mathbf{a}+(\nu a^2+1)\mathbf{b}-\mu(\mathbf{a}\times\mathbf{b})\end{aligned}$$

We have learned that any vector has a unique expression in terms of a basis set, so that the coefficients of $\mathbf{a}$, $\mathbf{b}$ and $\mathbf{a}\times\mathbf{b}$ on either side of the equation much be equal.

$$\begin{aligned}\Rightarrow\lambda &= -\nu(\mathbf{a}\cdot\mathbf{b})\\ \mu &= \nu a^2+1\\ \nu &= -\mu\end{aligned}$$

so that

$$\mu=\frac{1}{1+a^2}\qquad \nu=-\frac{1}{1+a^2}\qquad \lambda=\frac{\mathbf{a}\cdot\mathbf{b}}{1+a^2}\ .$$

So finally the solution is the single point:

$$\mathbf{x}=\frac{1}{1+a^2}\left((\mathbf{a}\cdot\mathbf{b})\mathbf{a}+\mathbf{b}-(\mathbf{a}\times\mathbf{b})\right)$$

2.4 Rotation, angular velocity/acceleration and moments

A rotation can represented by a vector whose direction is along the axis of rotation in the sense of a r-h screw, and whose magnitude is proportional to the size of the rotation (Fig. 2.7). The same idea can be extended to the derivatives, that is, angular velocity $\boldsymbol{\omega}$ and angular acceleration $\dot{\boldsymbol{\omega}}$.

Angular accelerations arise because of a moment (or torque) on a body. In mechanics, the moment of a force $\mathbf{F}$ about a point Q is defined to have magnitude $M = Fd$, where d is the perpendicular distance between Q and the line of action L of $\mathbf{F}$.

The vector equation for moment is

$$\mathbf{M}=\mathbf{r}\times\mathbf{F}$$

where $\mathbf{r}$ is the vector from Q to *any* point on the line of action L of force $\mathbf{F}$. The resulting angular acceleration vector is in the same direction as the moment vector.

The instantaneous velocity of any point P on a rigid body undergoing pure rotation can be defined by a vector product as follows. The angular velocity vector $\boldsymbol{\omega}$ has

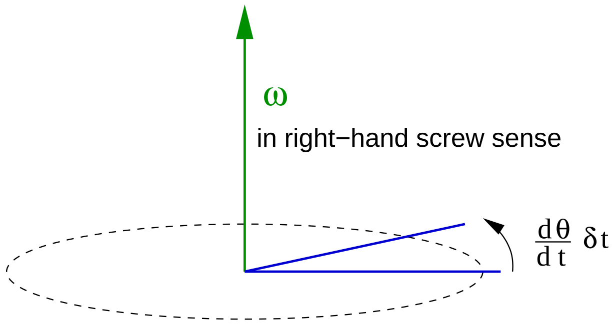


Figure 2.7: The angular velocity vector $\boldsymbol{\omega}$ is along the axis of rotation and has magnitude equal to the rate of rotation.

magnitude equal to the angular speed of rotation of the body and with direction the same as that of the r-h screw. If $\mathbf{r}$ is the vector OP, where the origin O can be taken to be any point on the axis of rotation, then the velocity $\mathbf{v}$ of P due to the rotation is given, in both magnitude and direction, by the vector product

$$\mathbf{v} = \boldsymbol{\omega} \times \mathbf{r}.$$

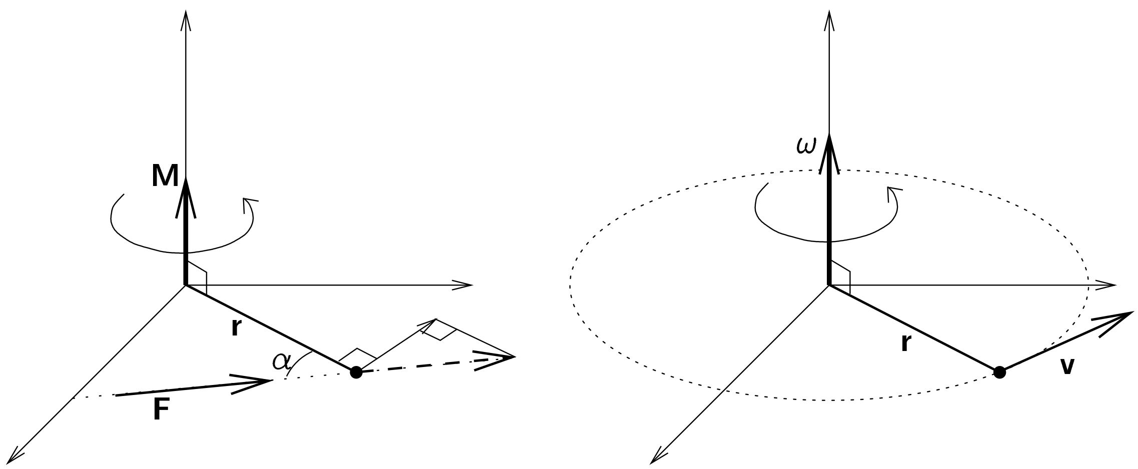


Figure 2.8:

Lecture 3

Differentiating Vector Functions of a Single Variable

Your experience of differentiation and integration has extended as far as *scalar* functions of single and multiple variables — functions like $f(x)$ and $f(x, y, t)$.

It should be no great surprise that we often wish differentiate vector functions. For example, suppose you were driving along a wiggly road with position $\mathbf{r}(t)$ at time t. Differentiating $\mathbf{r}(t)$ wrt time should yield your velocity $\mathbf{v}(t)$, and differentiating $\mathbf{v}(t)$ should yield your acceleration. Let's see how to do this.

3.1 Differentiation of a vector

The derivative of a vector function $\mathbf{a}(p)$ of a single parameter p is

$$\mathbf{a}'(p) = \lim_{\delta p \to 0} \frac{\mathbf{a}(p + \delta p) - \mathbf{a}(p)}{\delta p} \quad .$$

If we write $\mathbf{a}$ in terms of components relative to a FIXED coordinate system ($\hat{\boldsymbol{\imath}}, \hat{\boldsymbol{\jmath}}, \hat{\boldsymbol{k}}$ constant)

$$\mathbf{a}(p) = a_1(p)\hat{\boldsymbol{\imath}} + a_2(p)\hat{\boldsymbol{\jmath}} + a_3(p)\hat{\boldsymbol{k}}$$

then

$$\mathbf{a}'(p) = \frac{da_1}{dp}\hat{\boldsymbol{\imath}} + \frac{da_2}{dp}\hat{\boldsymbol{\jmath}} + \frac{da_3}{dp}\hat{\boldsymbol{k}} \quad .$$

That is, in order to differentiate a vector function, one simply differentiates each component separately. This means that all the familiar rules of differentiation apply, and they don't get altered by vector operations like scalar product and vector products.

Thus, for example:

$$\frac{d}{dp}(\mathbf{a} \times \mathbf{b}) = \frac{d\mathbf{a}}{dp} \times \mathbf{b} + \mathbf{a} \times \frac{d\mathbf{b}}{dp} \qquad \frac{d}{dp}(\mathbf{a} \cdot \mathbf{b}) = \frac{d\mathbf{a}}{dp} \cdot \mathbf{b} + \mathbf{a} \cdot \frac{d\mathbf{b}}{dp} \quad .$$

Note that $d\mathbf{a}/dp$ has a different direction and a different magnitude from $\mathbf{a}$.

Likewise, as you might expect, the chain rule still applies. If $\mathbf{a} = \mathbf{a}(u)$ and $u = u(t)$, say:

$$\frac{d}{dt}\mathbf{a} = \frac{d\mathbf{a}}{du}\frac{du}{dt}$$

♣ Examples

Q A 3D vector $\mathbf{a}$ of constant magnitude is varying over time. What can you say about the direction of $\dot{\mathbf{a}}$?

A Using intuition: if only the direction is changing, then the vector must be tracing out points on the surface of a sphere. We would guess that the derivative $\dot{\mathbf{a}}$ is orthogonal to $\mathbf{a}$.

To prove this write

$$\frac{d}{dt}(\mathbf{a}\cdot\mathbf{a}) = \mathbf{a}\cdot\frac{d\mathbf{a}}{dt} + \frac{d\mathbf{a}}{dt}\cdot\mathbf{a} = 2\mathbf{a}\cdot\frac{d\mathbf{a}}{dt} \quad .$$

But $(\mathbf{a}\cdot\mathbf{a}) = a^2$ which we are told is constant. So

$$\frac{d}{dt}(\mathbf{a}\cdot\mathbf{a}) = 0 \qquad \Rightarrow 2\mathbf{a}\cdot\frac{d\mathbf{a}}{dt} = 0$$

and hence $\mathbf{a}$ and $d\mathbf{a}/dt$ must be perpendicular.

Q The position of a vehicle is $\mathbf{r}(u)$ where u is the amount of fuel consumed by some time t. Write down an expression for the acceleration.n

A The velocity is

$$\mathbf{v} = \frac{d\mathbf{r}}{dt} = \frac{d\mathbf{r}}{du}\frac{du}{dt}$$

$$\mathbf{a} = \frac{d}{dt}\frac{d\mathbf{r}}{dt} = \frac{d^2\mathbf{r}}{du^2}\left(\frac{du}{dt}\right)^2 + \frac{d\mathbf{r}}{du}\frac{d^2u}{dt^2}$$

3.1.1 Geometrical interpretation of vector derivatives

Let $\mathbf{r}(p)$ be a position vector tracing a space curve as some parameter p varies. The vector $\delta\mathbf{r}$ is a secant to the curve, and $\delta\mathbf{r}/\delta p$ lies in the same direction. (See Fig. 3.1.) In the limit as δp tends to zero $\delta\mathbf{r}/\delta p = d\mathbf{r}/dp$ becomes a tangent to the space curve. If the magnitude of this vector is 1 (i.e. a unit tangent), then

$|d\mathbf{r}| = dp$ so the parameter p is arc-length (metric distance). More generally, however, p will not be arc-length and we will have:

$$\frac{d\mathbf{r}}{dp} = \frac{d\mathbf{r}}{ds}\frac{ds}{dp}$$

So, the direction of the derivative is that of a **tangent to the curve**, and its magnitude is $|ds/dp|$, **the rate of change of arc length w.r.t the parameter**. Of course if that parameter p is time, the magnitude $|d\mathbf{r}/dt|$ is the speed.

♣ Example

Q Draw the curve

$$\mathbf{r} = a\cos(\frac{s}{\sqrt{a^2+h^2}})\hat{\boldsymbol{\imath}} \quad + \quad a\sin(\frac{s}{\sqrt{a^2+h^2}})\hat{\boldsymbol{\jmath}} \quad + \quad \frac{hs}{\sqrt{a^2+h^2}}\hat{\boldsymbol{k}}$$

where s is arc length and h, a are constants. Show that the tangent $d\mathbf{r}/ds$ to the curve has a constant elevation angle w.r.t the xy-plane, and determine its magnitude.

A

$$\frac{d\mathbf{r}}{ds} = -\frac{a}{\sqrt{a^2+h^2}}\sin()\,\hat{\boldsymbol{\imath}} + \frac{a}{\sqrt{a^2+h^2}}\cos()\,\hat{\boldsymbol{\jmath}} + \frac{h}{\sqrt{a^2+h^2}}\hat{\boldsymbol{k}}$$

The projection on the xy plane has magnitude $a/\sqrt{a^2+h^2}$ and in the z direction $h/\sqrt{a^2+h^2}$, so the elevation angle is a constant, $\tan^{-1}(h/a)$.

We are expecting $d\mathbf{r}/ds = 1$, and indeed

$$\sqrt{a^2\sin^2() + a^2\cos^2() + h^2}/\sqrt{a^2+h^2} = 1.$$

3.1.2 Arc length is a special parameter!

It might seem that we can be completely relaxed about saying that any old parameter p is arc length, but this is not the case. Why not? The reason is that arc length is special is that, whatever the parameter p,

$$s = \int_{p_0}^{p} \left|\frac{d\mathbf{r}}{dp}\right| dp \quad .$$

Perhaps another way to grasp the significance of this is using Pythagoras' theorem on a short piece of curve: in the limit as dx etc tend to zero,

$$ds^2 = dx^2 + dy^2 + dz^2 \ .$$

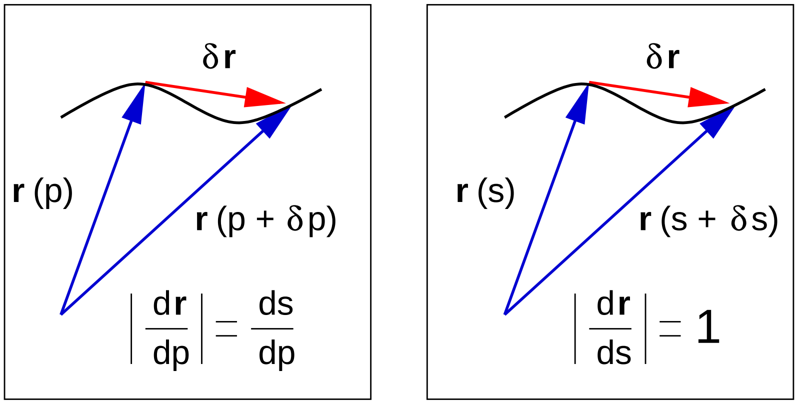


Figure 3.1: Left: $\delta\mathbf{r}$ is a secant to the curve but, in the limit as $\delta p \to 0$, becomes a tangent. Right: if the parameter is arc length s, then $|d\mathbf{r}| = ds$.

So if a curve is parameterized in terms of p

$$\frac{ds}{dp} = \sqrt{\frac{dx}{dp}^2 + \frac{dy}{dp}^2 + \frac{dz}{dp}^2} \, .$$

As an example, suppose in our earlier example we had parameterized our helix as

$$\mathbf{r} = a\cos p\hat{\boldsymbol{\imath}} \quad + \quad a\sin p\hat{\boldsymbol{\jmath}} \quad + \quad hp\hat{\boldsymbol{k}}$$

It would be easy just to *say* that p was arclength, but it would not be correct because

$$\begin{aligned}\frac{ds}{dp} &= \sqrt{\frac{dx}{dp}^2 + \frac{dy}{dp}^2 + \frac{dz}{dp}^2} \\ &= \sqrt{a^2\sin^2 p + a^2\cos^2 p + h^2} = \sqrt{a^2 + h^2}\end{aligned}$$

If p really was arclength, $ds/dp = 1$. So $p/\sqrt{a^2 + h^2}$ **is** arclength, not p.

3.2 Integration of a vector function

The integration of a vector function of a single scalar variable can be regarded simply as the reverse of differentiation. In other words

$$\int_{p_1}^{p_2} \frac{d\mathbf{a}(p)}{dp} dp$$

For example the integral of the acceleration vector of a point over an interval of time is equal to the change in the velocity vector during the same time interval. However, many other, more interesting and useful, types of integral are possible, especially when the vector is a function of more than one variable. This requires the introduction of the concepts of scalar and vector fields. See later!

3.3 Curves in 3 dimensions

In the examples above, parameter p has been either arc length s or time t. It doesn't have to be, but these are the main two of interest. Later we shall look at some important results when differentiating w.r.t. time, but now let use look more closely at 3D curves defined in terms of arc length, s.

Take a piece of wire, and bend it into some arbitrary non-planar curve. This is a *space curve.* We can specify a point on the wire by specifying $\mathbf{r}(s)$ as a function of distance or arc length s along the wire.

3.3.1 The Frénet-Serret relationships

We are now going to introduce a local orthogonal coordinate frame for each point s along the curve, ie one with its origin at $\mathbf{r}(s)$. To specify a coordinate frame we need three mutually perpendicular directions, and these should be *intrinsic* to the curve, not fixed in an external reference frame. The ideas were first suggested by two French mathematicians, F-J. Frénet and J. A. Serret.

1. **Tangent $\hat{\mathbf{t}}$**

   There is an obvious choice for the first direction at the point $\mathbf{r}(s)$, namely the **unit tangent $\hat{\mathbf{t}}$**. We already know that

   $$\hat{\mathbf{t}} = \frac{d\mathbf{r}(s)}{ds}$$

2. **Principal Normal $\hat{\mathbf{n}}$**

   Recall that earlier we proved that if $\mathbf{a}$ was a vector of constant magnitude that varies in direction over time then $d\mathbf{a}/dt$ was perpendicular to it. Because $\hat{\mathbf{t}}$ has constant magnitude but varies over s, $d\hat{\mathbf{t}}/ds$ must be perpendicular to $\hat{\mathbf{t}}$.

   Hence the principal normal $\hat{\mathbf{n}}$ is

   $$\frac{d\hat{\mathbf{t}}}{ds} = \kappa\hat{\mathbf{n}} : \text{ where } \kappa \geq 0 \ .$$

   κ is the **curvature**, and $\kappa = 0$ for a straight line. The plane containing $\hat{\mathbf{t}}$ and $\hat{\mathbf{n}}$ is called the **osculating plane**.

3. **The Binormal $\hat{\mathbf{b}}$**

   The local coordinate frame is completed by defining the binormal

   $$\hat{\mathbf{b}}(s) = \hat{\mathbf{t}}(s) \times \hat{\mathbf{n}}(s) \ .$$

Since $\hat{\mathbf{b}} \cdot \hat{\mathbf{t}} = 0$,

$$\frac{d\hat{\mathbf{b}}}{ds} \cdot \hat{\mathbf{t}} + \hat{\mathbf{b}} \cdot \frac{d\hat{\mathbf{t}}}{ds} = \frac{d\hat{\mathbf{b}}}{ds} \cdot \hat{\mathbf{t}} + \hat{\mathbf{b}} \cdot \kappa \hat{\mathbf{n}} = 0$$

from which

$$\frac{d\hat{\mathbf{b}}}{ds} \cdot \hat{\mathbf{t}} = 0.$$

But this means that $d\hat{\mathbf{b}}/ds$ is along the direction of $\hat{\mathbf{n}}$, or

$$\frac{d\hat{\mathbf{b}}}{ds} = -\tau(s)\hat{\mathbf{n}}(s)$$

where τ is the **torsion**, and the negative sign is a matter of convention.

Differentiating $\hat{\mathbf{n}} \cdot \hat{\mathbf{t}} = 0$ and $\hat{\mathbf{n}} \cdot \hat{\mathbf{b}} = 0$, we find

$$\frac{d\hat{\mathbf{n}}}{ds} = -\kappa(s)\hat{\mathbf{t}}(s) + \tau(s)\hat{\mathbf{b}}(s).$$

The Frénet-Serret relationships:

$$\begin{aligned} d\hat{\mathbf{t}}/ds &= \kappa \hat{\mathbf{n}} \\ d\hat{\mathbf{n}}/ds &= -\kappa(s)\hat{\mathbf{t}}(s) + \tau(s)\hat{\mathbf{b}}(s) \\ d\hat{\mathbf{b}}/ds &= -\tau(s)\hat{\mathbf{n}}(s) \end{aligned}$$

♣ Example

Q Derive $\kappa(s)$ and $\tau(s)$ for the helix

$$\mathbf{r}(s) = a\cos\left(\frac{s}{\beta}\right)\hat{\imath} + a\sin\left(\frac{s}{\beta}\right)\hat{\jmath} + h\left(\frac{s}{\beta}\right)\hat{\boldsymbol{k}}; \qquad \beta = \sqrt{a^2 + h^2}$$

and comment on their values.

A We found the unit tangent earlier as

$$\hat{\mathbf{t}} = \frac{d\mathbf{r}}{ds} = \left[-\frac{a}{\beta}\sin\left(\frac{s}{\beta}\right), \quad \frac{a}{\beta}\cos\left(\frac{s}{\beta}\right), \quad \frac{h}{\beta}\right] .$$

Differentiation gives

$$\kappa\hat{\mathbf{n}} = \frac{d\hat{\mathbf{t}}}{ds} = \left[-\frac{a}{\beta^2}\cos\left(\frac{s}{\beta}\right), \quad -\frac{a}{\beta^2}\sin\left(\frac{s}{\beta}\right), \quad 0\right]$$

Curvature is always positive, so

$$\kappa = \frac{a}{\beta^2} \qquad \hat{\mathbf{n}} = \left[-\cos\left(\frac{s}{\beta}\right), \quad -\sin\left(\frac{s}{\beta}\right), \quad 0 \right] .$$

So the curvature is constant, and the normal is parallel to the xy-plane. Now use

$$\hat{\mathbf{b}} = \hat{\mathbf{t}} \times \hat{\mathbf{n}} = \begin{vmatrix} \hat{\boldsymbol{\imath}} & \hat{\boldsymbol{\jmath}} & \hat{\boldsymbol{k}} \\ (-a/\beta)S & (a/\beta)C & (h/\beta) \\ -C & -S & 0 \end{vmatrix} = \left[\frac{h}{\beta}\sin\left(\frac{s}{\beta}\right), \quad -\frac{h}{\beta}\cos\left(\frac{s}{\beta}\right), \quad \frac{a}{\beta} \right]$$

and differentiate $\hat{\mathbf{b}}$ to find an expression for the torsion

$$\frac{d\hat{\mathbf{b}}}{ds} = \left[\frac{h}{\beta^2}\cos\left(\frac{s}{\beta}\right), \quad \frac{h}{\beta^2}\sin\left(\frac{s}{\beta}\right), \quad 0 \right] = \frac{-h}{\beta^2}\hat{\mathbf{n}}$$

so the torsion is

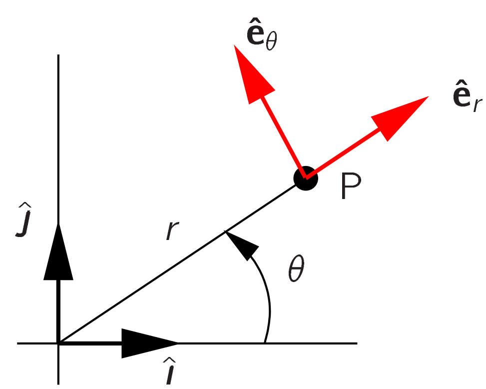
$$\tau = \frac{h}{\beta^2}$$

again a constant.

3.4 Radial and tangential components in plane polars

In plane polar coordinates, the radius vector of any point P is given by

$$\begin{aligned} \mathbf{r} &= r\cos\theta\hat{\boldsymbol{\imath}} + r\sin\theta\hat{\boldsymbol{\jmath}} \\ &= r\hat{\mathbf{e}}_r \end{aligned}$$



where we have introduced the unit radial vector

$$\hat{\mathbf{e}}_r = \cos\theta\hat{\boldsymbol{\imath}} + \sin\theta\hat{\boldsymbol{\jmath}} \quad .$$

The other "natural" (we'll see why in a later lecture) unit vector in plane polars is orthogonal to $\hat{\mathbf{e}}_r$ and is

$$\hat{\mathbf{e}}_\theta = -\sin\theta\hat{\boldsymbol{\imath}} + \cos\theta\hat{\boldsymbol{\jmath}}$$

so that $\hat{\mathbf{e}}_r \cdot \hat{\mathbf{e}}_r = \hat{\mathbf{e}}_\theta \cdot \hat{\mathbf{e}}_\theta = 1$ and $\hat{\mathbf{e}}_r \cdot \hat{\mathbf{e}}_\theta = 0$.

Now suppose P is moving so that $\mathbf{r}$ is a function of time t. Its velocity is

$$\begin{aligned}
\dot{\mathbf{r}} = \frac{d}{dt}(r\hat{\mathbf{e}}_r) &= \frac{dr}{dt}\hat{\mathbf{e}}_r + r\frac{d\hat{\mathbf{e}}_r}{dt} \\
&= \frac{dr}{dt}\hat{\mathbf{e}}_r + r\frac{d\theta}{dt}(-\sin\theta\hat{\boldsymbol{\imath}} + \cos\theta\hat{\boldsymbol{\jmath}}) \\
&= \frac{dr}{dt}\hat{\mathbf{e}}_r + r\frac{d\theta}{dt}\hat{\mathbf{e}}_\theta \\
&= \text{radial} + \text{tangential}
\end{aligned}$$

The radial and tangential components of velocity of P are therefore dr/dt and $rd\theta/dt$, respectively.

Differentiating a second time gives the acceleration of P

$$\begin{aligned}
\ddot{\mathbf{r}} &= \frac{d^2r}{dt^2}\hat{\mathbf{e}}_r + \frac{dr}{dt}\frac{d\theta}{dt}\hat{\mathbf{e}}_\theta + \frac{dr}{dt}\frac{d\theta}{dt}\hat{\mathbf{e}}_\theta + r\frac{d^2\theta}{dt^2}\hat{\mathbf{e}}_\theta - r\frac{d\theta}{dt}\frac{d\theta}{dt}\hat{\mathbf{e}}_r \\
&= \left[\frac{d^2r}{dt^2} - r\left(\frac{d\theta}{dt}\right)^2\right]\hat{\mathbf{e}}_r + \left[2\frac{dr}{dt}\frac{d\theta}{dt} + r\frac{d^2\theta}{dt^2}\right]\hat{\mathbf{e}}_\theta
\end{aligned}$$

3.5 Rotating systems

Consider a body which is rotating with constant angular velocity $\boldsymbol{\omega}$ about some axis passing through the origin. Assume the origin is fixed, and that we are sitting in a fixed coordinate system $Oxyz$.

If $\boldsymbol{\rho}$ is a vector of constant magnitude and constant direction in the rotating system, then its representation $\mathbf{r}$ in the fixed system must be a function of t.

$$\mathbf{r}(t) = \mathrm{R}(t)\boldsymbol{\rho}$$

At any instant as observed in the fixed system

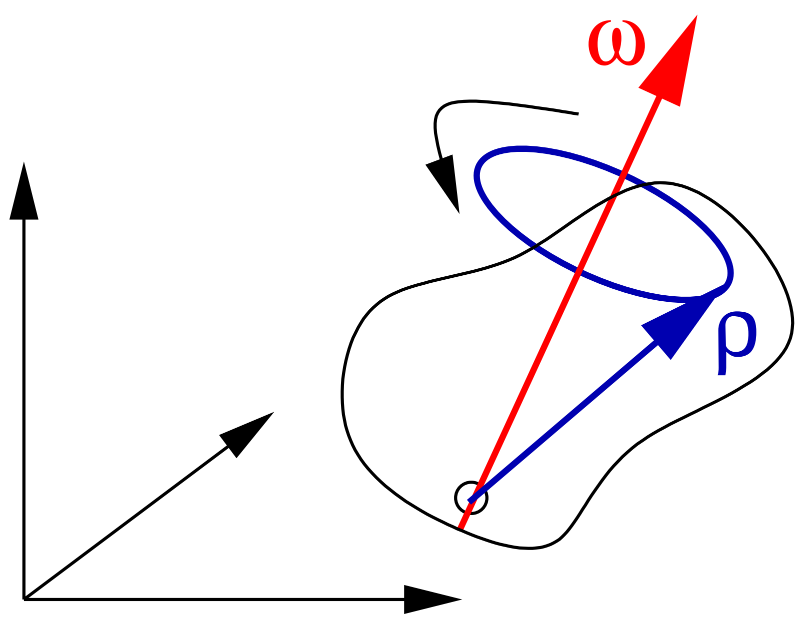
$$\frac{d\mathbf{r}}{dt} = \dot{\mathrm{R}}\boldsymbol{\rho} + \mathrm{R}\dot{\boldsymbol{\rho}}$$

but the second term is zero since we assumed $\boldsymbol{\rho}$ to be constant so we have

$$\frac{d\mathbf{r}}{dt} = \dot{\mathrm{R}}\mathrm{R}^\top\mathbf{r}$$

Note that:

- $d\mathbf{r}/dt$ will have fixed magnitude;
- $d\mathbf{r}/dt$ will always be perpendicular to the axis of rotation;
- $d\mathbf{r}/dt$ will vary in direction within those constraints;
- $\mathbf{r}(t)$ will move in a plane in the fixed system.



Now let's consider the term $\dot{R}R^\top$. First, note that $RR^\top = I$ (the identity), so differentiating both sides yields

$$\begin{aligned} \dot{R}R^\top + R\dot{R}^\top &= 0 \\ \dot{R}R^\top &= -R\dot{R}^\top \end{aligned}$$

Thus $\dot{R}R^\top$ is anti-symmetric:

$$\dot{R}R^\top = \begin{bmatrix} 0 & -z & y \\ z & 0 & -x \\ -y & x & 0 \end{bmatrix}$$

Now you can verify for yourself that application of a matrix of this form to an arbitrary vector has precisely the same effect as the cross product operator, $\boldsymbol{\omega}\times$, where $\boldsymbol{\omega} = [xyz]^\top$. Loh-and-behold, we then we have

$$\dot{\mathbf{r}} = \boldsymbol{\omega} \times \mathbf{r}$$

matching the equation at the end of lecture 2, $\mathbf{v} = \boldsymbol{\omega}\times\mathbf{r}$, as we would hope/expect.

3.5.1 Rotation: Part 2

Now suppose $\boldsymbol{\rho}$ is the position vector of a point P which **moves** in the rotating frame. There will be two contributions to motion with respect to the fixed frame, one due to its motion within the rotating frame, and one due to the rotation itself. So, returning to the equations we derived earlier:

$$\mathbf{r}(t) = R(t)\boldsymbol{\rho}(t)$$

and the instantaenous differential with respect to time:

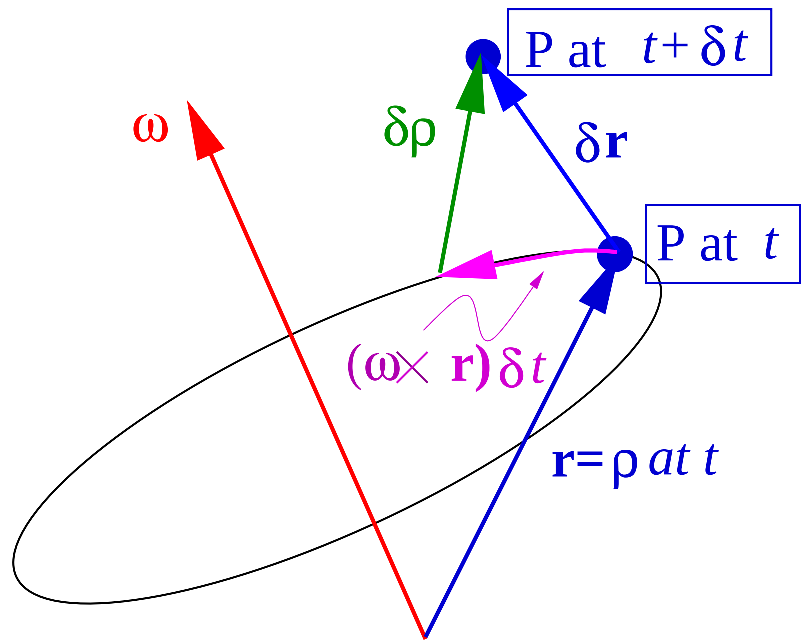
$$\frac{d\mathbf{r}}{dt} = \dot{R}\boldsymbol{\rho} + R\dot{\boldsymbol{\rho}} = \dot{R}R^\top\mathbf{r} + R\dot{\boldsymbol{\rho}}$$

Now $\boldsymbol{\rho}$ is not constant, so its differential is not zero; hence rewriting this last equations we have that

The **instantaneous velocity** of P in the fixed frame is

$$\frac{d\mathbf{r}}{dt} = R\dot{\boldsymbol{\rho}} + \boldsymbol{\omega} \times \mathbf{r}$$

The second term of course, is the contribution from the rotating frame which we saw previously. The first is the linear velocity measured in the rotating frame $\dot{\boldsymbol{\rho}}$, referred to the fixed frame (via the rotation matrix R which aligns the two frames)



3.5.2 Rotation 3: Instantaneous acceleration

Our previous result is a general one relating the time derivatives of any vector in rotating and non-rotating frames. Let us now consider the second differential:

$$\ddot{\mathbf{r}} = \dot{\boldsymbol{\omega}} \times \mathbf{r} + \boldsymbol{\omega} \times \dot{\mathbf{r}} + \dot{\mathsf{R}}\dot{\boldsymbol{\rho}} + \mathsf{R}\ddot{\boldsymbol{\rho}}$$

We shall assume that the angular acceleration is zero, which kills off the first term, and so now, substituting for $\dot{\mathbf{r}}$ we have

$$\begin{aligned}
\ddot{\mathbf{r}} &= \boldsymbol{\omega} \times (\boldsymbol{\omega} \times \mathbf{r} + \mathsf{R}\dot{\boldsymbol{\rho}}) + \dot{\mathsf{R}}\dot{\boldsymbol{\rho}} + \mathsf{R}\ddot{\boldsymbol{\rho}} \\
&= \boldsymbol{\omega} \times (\boldsymbol{\omega} \times \mathbf{r}) + \boldsymbol{\omega} \times \mathsf{R}\dot{\boldsymbol{\rho}} + \dot{\mathsf{R}}\dot{\boldsymbol{\rho}} + \mathsf{R}\ddot{\boldsymbol{\rho}} \\
&= \boldsymbol{\omega} \times (\boldsymbol{\omega} \times \mathbf{r}) + \boldsymbol{\omega} \times \mathsf{R}\dot{\boldsymbol{\rho}} + \dot{\mathsf{R}}(\mathsf{R}^\top\mathsf{R})\dot{\boldsymbol{\rho}} + \mathsf{R}\ddot{\boldsymbol{\rho}} \\
&= \boldsymbol{\omega} \times (\boldsymbol{\omega} \times \mathbf{r}) + 2\boldsymbol{\omega} \times (\mathsf{R}\dot{\boldsymbol{\rho}}) + \mathsf{R}\ddot{\boldsymbol{\rho}}
\end{aligned}$$

The **instantaneous acceleration** is therefore

$$\ddot{\mathbf{r}} = \mathsf{R}\ddot{\boldsymbol{\rho}} + 2\boldsymbol{\omega} \times (\mathsf{R}\dot{\boldsymbol{\rho}}) + \boldsymbol{\omega} \times (\boldsymbol{\omega} \times \mathbf{r})$$

- The first term is the acceleration of the point P in the rotating frame measured in the rotating frame, but referred to the fixed frame by the rotation R

- The last term is the centripetal acceleration to due to the rotation. (Yes! Its magnitude is $\omega^2 r$ and its direction is that of $-\mathbf{r}$. Check it out.)

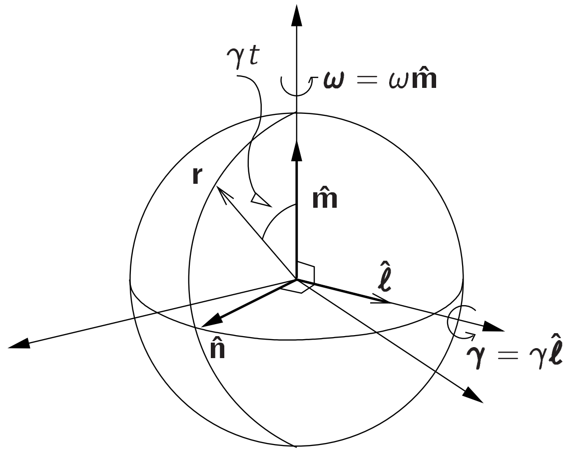


Figure 3.2: Coriolis example.

- The middle term is an extra term which arises because of the velocity of P in the rotating frame. It is known as the **Coriolis acceleration**, named after the French engineer who first identified it.

Because of the rotation of the earth, the Coriolis acceleration is of great importance in meteorology and accounts for the occurrence of high pressure anticyclones and low pressure cyclones in the northern hemisphere, in which the Coriolis acceleration is produced by a pressure gradient. It is also a very important component of the acceleration (hence the force exerted) by a rapidly moving robot arm, whose links whirl rapidly about rotary joints.

♣ Example

Q Find the instantaneous acceleration of a projectile fired along a line of longitude (with angular velocity of $\boldsymbol{\gamma}$ constant relative to the sphere) if the sphere is rotating with angular velocity $\boldsymbol{\omega}$.

A Consider a coordinate frame defined by mutually orthogonal unit vectors, $\hat{\boldsymbol{\ell}}$, $\hat{\mathbf{m}}$ and $\hat{\mathbf{n}}$, as shown in Fig. 3.2. We shall assume, without loss of generality, that the fixed and rotating frames are instantaneously aligned at the moment shown in the diagram, so that $\mathsf{R} = \mathsf{I}$, the identity, and hence $\mathbf{r} = \boldsymbol{\rho}$.

In the rotating frame

$$\dot{\boldsymbol{\rho}} = \boldsymbol{\gamma} \times \boldsymbol{\rho} \qquad \text{and} \qquad \ddot{\boldsymbol{\rho}} = \boldsymbol{\gamma} \times \dot{\boldsymbol{\rho}} = \boldsymbol{\gamma} \times (\boldsymbol{\gamma} \times \boldsymbol{\rho})$$

So the in the fixed reference frame, because these two frames are instantaneously aligned

$$\ddot{\mathbf{r}} \;=\; \boldsymbol{\gamma} \times (\boldsymbol{\gamma} \times \boldsymbol{\rho}) \;+\; 2\boldsymbol{\omega} \times (\boldsymbol{\gamma} \times \boldsymbol{\rho}) \;+\; \boldsymbol{\omega} \times (\boldsymbol{\omega} \times \mathbf{r}) \; .$$

The first term is the centripetal acceleration due to the projectile moving around the sphere — which it does because of the gravitational force. The last term is the centripetal acceleration resulting from the rotation of the sphere. The middle term is the Coriolis acceleration.

Using Fig. 3.2, at some instant t

$$\mathbf{r}(t) = \boldsymbol{\rho}(t) = r\cos(\gamma t)\hat{\mathbf{m}} \quad + \quad r\sin(\gamma t)\hat{\mathbf{n}}$$

and

$$\boldsymbol{\gamma} = \gamma\hat{\boldsymbol{\ell}}$$

Then

$$\boldsymbol{\gamma} \times (\boldsymbol{\gamma} \times \boldsymbol{\rho}) = (\boldsymbol{\gamma} \cdot \boldsymbol{\rho})\boldsymbol{\gamma} - \gamma^2\boldsymbol{\rho} = -\gamma^2\boldsymbol{\rho} = -\gamma^2\mathbf{r},$$

Check the direction — the negative sign means it points *towards* the centre of the sphere, which is as expected.

Likewise the last term can be obtained as

$$\boldsymbol{\omega} \times (\boldsymbol{\omega} \times \mathbf{r}) = -\omega^2 r\sin(\gamma t)\hat{\mathbf{n}}$$

Note that it is perpendicular to the axis of rotation $\hat{\mathbf{m}}$, and because of the minus sign, directed towards the axis)

The Coriolis term is derived as:

$$\begin{aligned} 2\boldsymbol{\omega} \times \dot{\boldsymbol{\rho}} &= 2\boldsymbol{\omega} \times (\boldsymbol{\gamma} \times \boldsymbol{\rho}) \\ &= 2\begin{bmatrix} 0 \\ \omega \\ 0 \end{bmatrix} \times \left(\begin{bmatrix} \gamma \\ 0 \\ 0 \end{bmatrix} \times \begin{bmatrix} 0 \\ r\cos\gamma t \\ r\sin\gamma t \end{bmatrix} \right) \\ &= 2\omega\gamma r\cos\gamma t\hat{\boldsymbol{\ell}} \end{aligned}$$

Instead of a projectile, now consider a rocket on rails which stretch north from the equator. As the rocket travels north it experiences the Coriolis force (exerted by the rails):

$$\begin{array}{ccccc} 2 & \gamma & \omega & R\cos\gamma t & \hat{\boldsymbol{\ell}} \\ +\text{ve} & \text{-ve} & +\text{ve} & +\text{ve} & \end{array}$$

Hence the coriolis force is in the direction opposed to $\hat{\boldsymbol{\ell}}$ (i.e. in the opposite direction to the earth's rotation). In the absence of the rails (or atmosphere) the rocket's tangetial speed (relative to the surface of the earth) is *greater* than the speed of the surface of the earth underneath it (since the radius of successive lines of latitude decreases) so it would (to an observer on the earth) appear to deflect to the east. The rails provide a coriolis force keeping it on the same meridian.

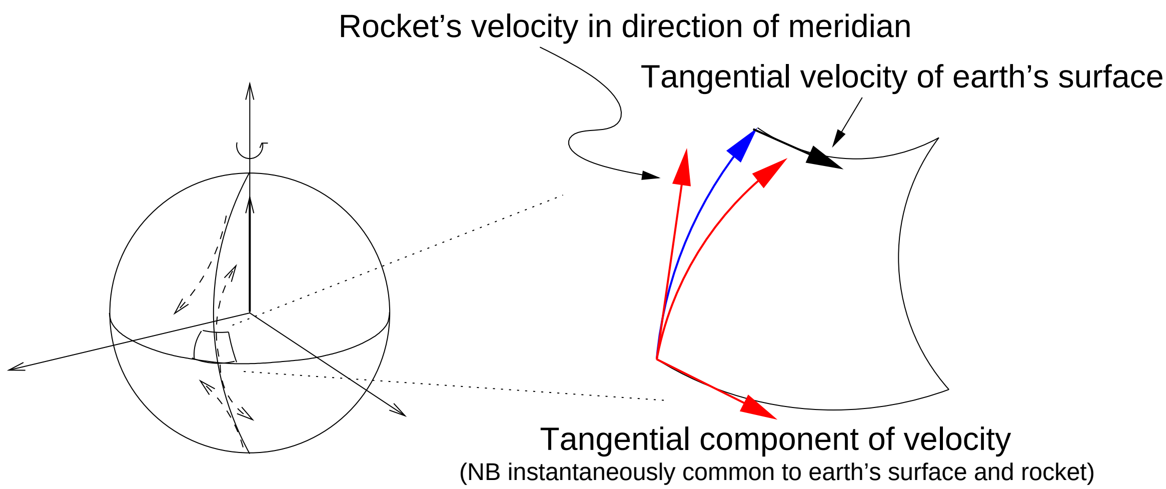


Figure 3.3: Rocket example

Figure 3.4: Coriolis effect giving rise to weather systems