MA10236 – Methods & Applications 1B

Prof Chris Budd OBE, FIMA, CMath, NTF, mascjb@bath.ac.uk

Academic Year **2019/20**

Logistical Information

The lectures are at the following times

- Monday 9.15am in the University Hall
- Wednesday 12.15pm in the University Hall
- Thursday 5.15pm in the University Hall

I will issue an example sheet each week on Monday. The sheet will be a mix of unstarred questions to do for your tutors, starred (*) questions which are harder and for revision, and optional double starred (**) questions which are meant to be a real challenge for those of you who like that sort of thing. Finally there is an optional puzzle each week for a bit of fun.

Warning These typed notes do not cover the whole course. Some additional material will be given in the lectures and on Moodle which is meant to enhance the lecture notes. There will be practical demonstrations and also videos in the lectures, not to mention audience participation! In addition, some parts of these notes indicated by a * (harder) or a ** (really a lot harder) are not on the lectures, they are for general background reading and will not be examined. Also, when the time comes, you may be advised to read parts of Anton (§12 and §13) and make your own notes. You should also make a point of (i) coming to lectures and being prepared to join in (ii) doing the example sheets. Only by doing both will you get the most out of the course, and of course see the demonstrations and hear the jokes (OK so that might put you off). Above all HAVE FUN!!

Acknowledgements These notes have evolved from various units taught in the first year at the University of Bath over several years. Previous lecturers include: Martin Reed, John Willis, Ray Ogden, Keith Walton, Victor Galaktionov, Alastair Spence and Valery Smyshlyaev.

Contents

1 Overview

Applied mathematics is the process of applying mathematics to real world problems, and also learning and developing mathematical tools to solve these problems. Many of the advances in mathematics, such as in calculus, differential equations, vector calculus, Fourier analysis (which you will cover in the Second Year), computation and geometry have come directly from studying such applications. All mathematical descriptions of the real world are to a certain extent an approximation to the truth, but it is quite remarkable how (unreasonably effective) abstract mathematics is at representing the world that we live in, and allowing us to make predictions about how things will behave. When it comes to telling the future Mystic Meg is nothing compared to an applied mathematician. Some mathematicians can give the impression that applied maths is easy, just a matter of applying well known ideas in a routine way to solve routine problems. Nothing could be further from the truth! Firstly, the process of understanding how to describe the world in mathematical terms, and to then interpret the answers in a meaningful way, is very subtle (indeed) and is a skill which we will be teaching you in many of the courses that you will meet at Bath. We will call this process *mathematical modelling*. Secondly, the mathematics that we will use is far from trivial. Often the (for example nonlinear differential) equations that are used to represent reality are very hard to solve, and we need to use a great variety of mathematical techniques to make any progress. I assure you that the mathematics that you will be using in this, and future applied maths courses, is every bit as challenging as the maths that you will meet in all of you other courses. To quote the great Richard Feynmann you have to *think real hard* to do applied mathematics. Thirdly, and wonderfully, it is worth it. The mathematics that you will learn really does tell us what the world is doing, and often reveals surprising truths about the way that things work. If you don't believe me, watch the film *Hidden Figures* which will show you how a bunch of (female) mathematicians got us to the moon, and in the film *Apollo 13* how they got us back again.

The best way to do applied maths is to get a firm mathematical foundation to start with. This course develops the theory of *vectors* – a major mathematical tool for modelling and analysis of various "real world" (physical, mechanical etc.) applications. Vectors have numerous other applications too, for example, in electromagnetic waves which are important in the design of mobile phones, weather and climate forecasting, the design of industrial processes, personality tests and much more besides. We will then combine our understanding of vetors with the main tools of calculus such as integration, differentiation and differential equations. Having done this we can then look at problems in kinematics and mechanics. *Kinematics* is the study of how bodies move. *Mechanics* deals with the action of forces on bodies and includes both statics and the study of motion. We shall consider *Newtonian Mechanics*, including,

- the motion of the planets and satellites,
- projectiles (including a certain drop goal)
- circular orbits
- Some sporting problems

All of these will be illustrated by examples, some of which will involve some practical demonstrations in the lecture. You will also learn the general process of how to approach mathematically modelling a real life situation. In the 19th Century it was thought that Newtonian Mechanics was a well understood and rather dry subject, and that everything was very predictable. We now know that this is far from the case. For example the discovery of *chaos* which I will demonstrate in the lectures, shows that systems governed by Newtonian mechanics can have very bizzare behaviour indeed. In addition, a thorough understanding of mechanics is now *essential* for a career in computer animation with companies such as Pixar, and also in many areas of sport, as well as the space programme. It is also an essential part of the training of pilots, air traffic controllers, forensic scientists (think high speed cars, or bullets) and, as we shall see, in the design of pedestrian crossings, saving the whales and curing cancer!. Be prepared to be both amazed by this course, and highly employable at the end.

2 Vectors

Before looking at the applications we need to understand the tools, namely vectors, and so the first few sections are about *Vector Algebra* and then *Vector Calculus*.

Text book - Anton "Calculus" This is a very readable book, and has lots of exercises. However, we'll not follow Anton all the time, and also we'll cover certain topics differently. You should read both these notes and the relevant sections in Anton.

2.1 Introduction to Vector Algebra

Some quantities in mechanics, physics, or indeed life in general, are characterised by a *single real number*, for example, mass, density, temperature. However other physical quantities require the specification of both a *magnitude* and *direction*, for example, velocity, an acceleration, or a force. We shall assume that we live in a 3-Dimensional world which is part of a (3-D) Euclidean Space $\mathbb{R}^3$, and for this course, vectors are 3-Dimensional. This means that a typical 3-dimensional vector $\mathbf{a} = (x, y, z)$ needs three values to specify it exactly. To give a precise mathematical definition we adopt the axioms of Euclidean geometry, namely, *points, straight lines, planes, parallel lines, distance,* etc. We will then think of vectors as geometrical objects for the main. (Note that this is not the only way to think of a vector. Basically anything which requires several coordinates to specifiy it can be thought of as a vector. For example if I want to go shopping and buy £x of sugar, £y of rice, and £z of strawberries, then my shopping for the day can be summarised by the single vector $\mathbf{a} = (x, y, z)$. However, for most of this course we will stick to the geometric interpretation of a vector which is summarised as follows.)

Definition 2.1 *A vector is a directed line segment (DLS) characterised by two "ordered points" in 3-D Euclidean space, P and Q, and is visualised as an arrow joining P to Q. P is the* initial *point and Q is the* terminal *point. The vector is denoted $\overrightarrow{PQ}$.*

Every vector has a direction and a length associated with it:

Definition 2.2 *The length of $\overrightarrow{PQ}$ is the* distance *between P and Q, and is denoted by $\|\overrightarrow{PQ}\|$. Often we say $\|\overrightarrow{PQ}\|$ is the* norm *of $\overrightarrow{PQ}$.*

Definition 2.3 *If P and Q coincide (that is $P = Q$) then the vector is called the zero segment (or zero vector), and is denoted $\overrightarrow{0}$, $\underline{0}$ or $\mathbf{0}$.*

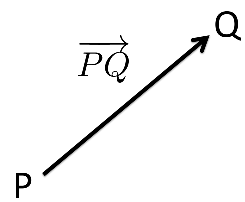


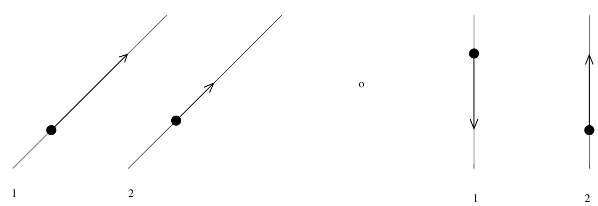
Figure 1: A vector as a directed line segment from the point P to the point Q.

From the Euclidean distance axioms, namely, (a) the distance from a point to itself is zero, and (b) the distance between any two distinct points is strictly positive, we have the following "obvious" result.

Proposition 2.4 $\overrightarrow{PQ} = \overrightarrow{0} \iff \|\overrightarrow{PQ}\| = 0.$

Proof: $\overrightarrow{PQ} = \overrightarrow{0} \iff P = Q \iff \|\overrightarrow{PQ}\| = 0.$ □

Definition 2.5 *Two directed line segments (DLSs)* $\overrightarrow{PQ}$ *and* $\overrightarrow{RS}$ *are said to be* **parallel** *(or collinear) if they lie on parallel straight lines 1 and 2. The zero segment is regarded as being parallel to any vector.*



An important idea is that two DLSs which have the same direction and magnitude are "equal" or "equivalent", and we do not distinguish between them. This leads to the following definition.

Definition 2.6 *Two DLSs* $\overrightarrow{PQ}$ *and* $\overrightarrow{RS}$ *are said to be* equal *(*equivalent*), that is,* $\overrightarrow{PQ} = \overrightarrow{RS}$*, if*

1. *They are collinear.*

2. *They have the same magnitude, that is,* $\|\overrightarrow{PQ}\| = \|\overrightarrow{RS}\|$*.*

3. *They point in the same direction.*

Note that this definition doesn't mention the initial point of a DLS, so a DLS can be "freely" moved, in that, for any DLS $\overrightarrow{PQ}$ and for any other point R, there exists an equal (and unique) DLS with initial point at R. Note that all zero DLSs are regarded as being equivalent.

Notation It is convenient to have a single label to describe this, for example,

$$\mathbf{a} = \overrightarrow{PQ} = \overrightarrow{RS},$$

indicating that we don't distinguish between two equivalent DLSs. Hence, **a** represents a "class" of equivalent DLS's each with the same magnitude and direction as $\overrightarrow{PQ}$.

Finally, $\|\mathbf{a}\|$ will denote the length of the vector **a**.

We will see presently that we can represent **a** as $\mathbf{a} = (a_1, a_2, a_3)$ and that

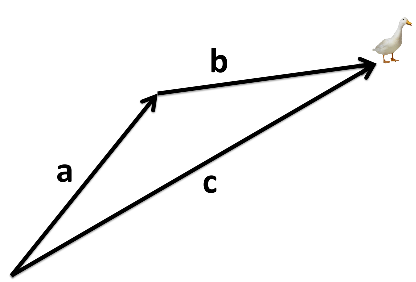
$$\|\mathbf{a}\| = \sqrt{a_1^2 + a_2^2 + a_3^2}.$$

2.2 Basic operations with vectors

In physical problems we often need to evaluate the combined effect of two processes characterised by two vectors, say **a** and **b**, that "follow one after the other" in some sense. This leads us to the first basic operation, namely, **addition** of two vectors.

Example: The wind is blowing with velocity **a**. A bird is flying with velocity **b** with respect to the air. What is the velocity, say **c**, of the bird with respect to an observer on the ground. (As usual the 'velocity' is defined to be the change in position (displacement) per unit of time.)

Answer: The answer is simply that $\mathbf{c} = \mathbf{a} + \mathbf{b}$.



This example motivates the following definition:

Definition 2.7 *Given two vectors* **a** *and* **b**. *Position* **b** *so that its initial point coincides with the terminal point of* **a**. *Then* $\mathbf{c} = \mathbf{a} + \mathbf{b}$ *is defined by the DLS whose initial point coincides with the initial point of* **a** *and whose terminal point coincides with the terminal point of* **b**.

Exercise: Check that this definition is consistent with the following equivalences. If

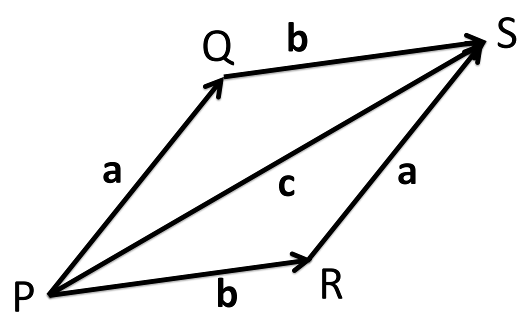
$$\overrightarrow{P_1Q_1} = \overrightarrow{P_2Q_2}\ ,\ \overrightarrow{Q_1S_1} = \overrightarrow{Q_2S_2}$$

then we need to show that

$$\overrightarrow{P_1Q_1} + \overrightarrow{Q_1S_1} = \overrightarrow{P_2Q_2} + \overrightarrow{Q_2S_2}.$$

Property 1 *Commutativity of addition:* $\mathbf{a} + \mathbf{b} = \mathbf{b} + \mathbf{a}$

Proof: Let $\mathbf{a} = \overrightarrow{PQ}$, $\mathbf{b} = \overrightarrow{PR}$. Construct the parallelogram $PQSR$ (see diagram).



Then, from the definition of equivalence of vectors, and the properties of parallelograms

$$\overrightarrow{RS} = \overrightarrow{PQ} = \mathbf{a}, \quad \overrightarrow{QS} = \overrightarrow{PR} = \mathbf{b}.$$

Thus, from the definition of addition, $\overrightarrow{PQ} + \overrightarrow{QS} = \overrightarrow{PS}$, or

$$\mathbf{a} + \mathbf{b} =: \mathbf{c}.$$

On the other hand,

$$\overrightarrow{PR} + \overrightarrow{RS} = \overrightarrow{PS} = \mathbf{c}$$

and so

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

Hence,

$$\mathbf{a} + \mathbf{b} = \mathbf{b} + \mathbf{a} \quad \forall \quad \mathbf{a}, \mathbf{b}.$$

Property 2 *Associativity of addition:* Let $\mathbf{a}, \mathbf{b}$ and $\mathbf{c}$ be three vectors. Then

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

(See Tutorial Sheet 1, Question 2).

Property 3 *The zero vector:* Consider a zero segment, namely, $\overrightarrow{PP}$. All zero segments are regarded as equivalent and we write

$$\overrightarrow{PP} \;=\; \overrightarrow{QQ} \;=\; \mathbf{0}.$$

Since by definition

$$\overrightarrow{PQ} + \overrightarrow{QQ} \;=\; \overrightarrow{PQ} \;=\; \overrightarrow{PP} + \overrightarrow{PQ},$$

it follows that

$$\mathbf{a} + \mathbf{0} = \mathbf{a} = \mathbf{0} + \mathbf{a}.$$

Property 4 *Negation of vectors:* Given a vector $\mathbf{a} = \overrightarrow{PQ}$ we denote the **negative** of $\mathbf{a}$ by $-\mathbf{a}$, and define it by reversing its direction, so that,

$$-\mathbf{a} = \overrightarrow{QP}.$$

Then, clearly,

$$\mathbf{a} + (-\mathbf{a}) = \overrightarrow{PQ} + \overrightarrow{QP} = \overrightarrow{PP} = \mathbf{0}.$$

and so

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

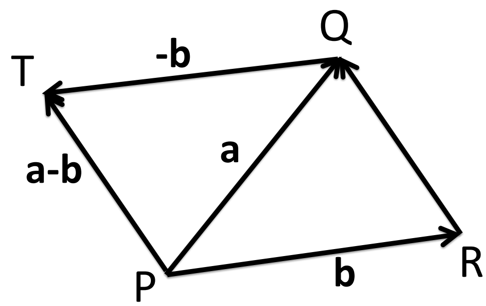
Definition 2.8 (Subtraction)

$$\mathbf{a} - \mathbf{b} \;:=\; \mathbf{a} + (-\mathbf{b}).$$

If $\mathbf{a} = \overrightarrow{PQ}$ *and* $\mathbf{b} = \overrightarrow{PR}$*, then*

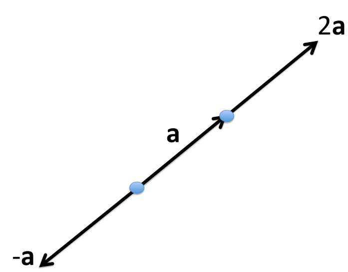
$$\mathbf{a} - \mathbf{b} = \mathbf{a} + (-\mathbf{b}) = \overrightarrow{PQ} + \overrightarrow{QT} = \overrightarrow{PT}.$$

Now, since $PRQT$ *is a parallelogram,* $\overrightarrow{PT} = \overrightarrow{RQ}$ *and hence* $\overrightarrow{RQ} = \mathbf{a} - \mathbf{b}$.



Next we introduce the concept of *multiplication by a scalar (or scalar multiplication)*. Consider $\mathbf{a}+\mathbf{a}$. It is natural to write $\mathbf{a}+\mathbf{a}=2\mathbf{a}$. The vector $2\mathbf{a}$ has the same direction as $\mathbf{a}$ but is twice as long:

$$\|2\mathbf{a}\| = \ 2\ \|\mathbf{a}\|\ .$$



On the other hand, it is natural to write

$$-\mathbf{a}=(-1)\mathbf{a}\ ,\ \overrightarrow{PS}=(-1)\mathbf{a}.$$

Thus, multiplication by -1 results in altering the direction of a vector. This motivates the following definition of scalar multiplication.

Definition 2.9 *For any $\lambda \in \mathbb{R}$, the scalar multiplication of* $\mathbf{a}$ *by* λ *produces the vector* $\mathbf{b}=\lambda\mathbf{a}$ *such that*

1. $\mathbf{b}$ *is collinear to* $\mathbf{a}$*,*

2. $\|\mathbf{b}\| = |\lambda|\ \|\mathbf{a}\|$,

3. $\mathbf{b}$ *and* $\mathbf{a}$ *have the same direction if* $\lambda > 0$*, and opposite if* $\lambda < 0$*. (If* $\lambda = 0$ *then* $\mathbf{b} = 0 \times \mathbf{a} = \mathbf{0}$*.)*

Property 5 *Associativity of scalar multiplication:* $\lambda(\mu\mathbf{a}) = (\lambda\mu)\mathbf{a}$

Property 6 *Distributivity of scalar multiplication with respect to addition of vectors:* $\lambda(\mathbf{a}+\mathbf{b}) = \lambda\mathbf{a} + \lambda\mathbf{b}$

Property 7 *Distributivity of scalar multiplication with respect to addition of scalars:* $(\lambda+\mu)\mathbf{a} = \lambda\mathbf{a} + \mu\mathbf{a}$

Property 8 $1 \cdot \mathbf{a} = \mathbf{a}$

(Properties 5 - 8 can be directly derived from the above definitions.)

Student Exercises:

Advert *You should do these, and later student exercises! They are there to help make sure that you understand the material so far. If you have problems with them, then discuss these with your tutors.*

(a) Draw a figure to show that Property 6 is true. (Hint: Think similar triangles.)

(b) Give arguments to show that Property 5 and Property 7 are true.

Eight Rules of Vector Algebra (or Vector Arithmetic)

1. $\mathbf{a} + \mathbf{b} = \mathbf{b} + \mathbf{a}$
2. $(\mathbf{a} + \mathbf{b}) + \mathbf{c} = \mathbf{a} + (\mathbf{b} + \mathbf{c})$
3. $\mathbf{a} + \mathbf{0} = \mathbf{0} + \mathbf{a} = \mathbf{a}$
4. $\mathbf{a} + (-\mathbf{a}) = \mathbf{0}$
5. $\lambda(\mu\mathbf{a}) = (\lambda\mu)\mathbf{a}$
6. $\lambda(\mathbf{a} + \mathbf{b}) = \lambda\mathbf{a} + \lambda\mathbf{b}$
7. $(\lambda + \mu)\mathbf{a} = \lambda\mathbf{a} + \mu\mathbf{a}$
8. $1 \cdot \mathbf{a} = \mathbf{a}$

These 8 properties are satisfied by directed line segments (with addition, negation and scalar multiplication introduced above). More generally, (1) - (8) may be regarded as “Axioms”, and quantities that satisfy them are called **vectors**. The directed line segment is merely one example of an object satisfying (1) - (8). Other examples are matrices and functions, among others.

Remark Many physical quantities (e.g. velocity, force) are indeed vectors. However not **all** quantities having both a magnitude and a direction satisfy these rules. For example consider the operation of rotation: the axis of rotation gives the direction, the angle is the magnitude. Let “addition” be superposition of rotations. Then the commutative property 1 (namely $\mathbf{a}+\mathbf{b}=\mathbf{b}+\mathbf{a}$) does not hold.

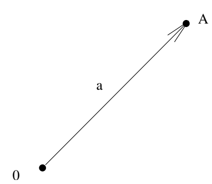
2.3 Unit vectors

Definition 2.10 *A vector whose length is* 1 *is called a unit vector. Given a non-zero vector* $\mathbf{a}$*, that is* $\mathbf{a}\neq\mathbf{0}$*, the unit vector **in the direction of** $\mathbf{a}$ is*

$$\hat{\mathbf{a}}=\frac{1}{\|\mathbf{a}\|}\mathbf{a}.$$

Check: $\|\hat{\mathbf{a}}\|=\left|\frac{1}{\|\mathbf{a}\|}\right|\|\mathbf{a}\|=\frac{\|\mathbf{a}\|}{\|\mathbf{a}\|}=1.$

2.4 Position Vectors:



Let O be the origin (which is a fixed point) in $\mathbb{R}^3$. Any point A can be identified with its **position vector a** by

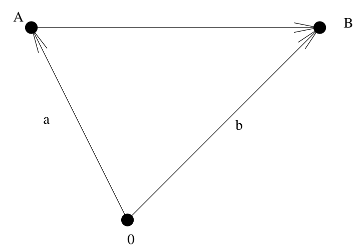
$$\mathbf{a}:=\overrightarrow{OA}.$$

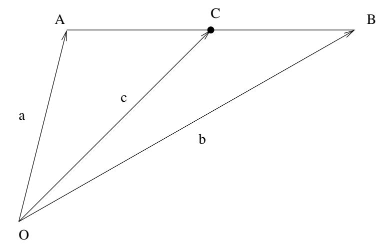
The main concept to remember here is that the initial point of any position vector is always the origin. It follows that for points A and B

$$\overrightarrow{AB}=\overrightarrow{AO}+\overrightarrow{OB}=\overrightarrow{OB}-\overrightarrow{OA}=\mathbf{b}-\mathbf{a}$$

(in the obvious notation). Position vectors are often useful for solving problems in geometry and mechanics.

Useful rules





1. Let the position vectors of A and B be $\mathbf{a}$ and $\mathbf{b}$ respectively.
   Then $\overrightarrow{AB} = \mathbf{b} - \mathbf{a}$

2. Let C be the midpoint of AB and so $\mathbf{c}$ is the position vector of $\overrightarrow{OC}$
   $\therefore \mathbf{c} = \overrightarrow{OA} + \overrightarrow{AC} = \overrightarrow{OA} + \frac{1}{2}\overrightarrow{AB} =$
   $= \mathbf{a} + \frac{1}{2}(\mathbf{b} - \mathbf{a}) = \frac{1}{2}(\mathbf{a} + \mathbf{b})$.

Now let's consider two examples.

Example 1: Assume $\mathbf{a}$ and $\mathbf{b}$ are non-zero and non-collinear, and let λ, μ be real numbers. Then

$$\lambda\mathbf{a} + \mu\mathbf{b} = \mathbf{0} \Rightarrow \lambda = 0, \mu = 0.$$

Proof First, assume $\lambda \neq 0$. Then $\mathbf{a} = -\frac{\mu}{\lambda}\mathbf{b}$, implying $\mathbf{a}$ and $\mathbf{b}$ are collinear: a contradiction. Thus $\lambda = 0$. Hence

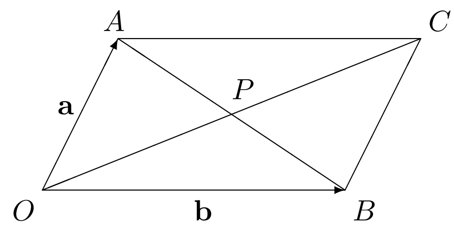
$$\mu\mathbf{b} = \mathbf{0},$$

and since $\mathbf{b} \neq \mathbf{0}$, we must have $\mu = 0$.$\square$

We will now look at a worked example, which is the sort of question that you might be faced with in your Summer exam. This example shows that position vectors are useful for solving geometry problems

Worked example 2: *Prove that the diagonals of a parallelogram bisect each other.*

Answer: **Always** start with a diagram.



The diagonals of the parallelogram are:

$$\overrightarrow{AB} = \mathbf{b} - \mathbf{a}$$
$$\overrightarrow{OC} = \mathbf{b} + \mathbf{a}$$

Let the diagonals intersect at the point P. So

$$\overrightarrow{AP} = \lambda\overrightarrow{AB} = \lambda(\mathbf{b} - \mathbf{a}) \quad \text{for some } \lambda \in \mathbb{R}$$
$$\text{and} \qquad \overrightarrow{OP} = \mu(\mathbf{b} + \mathbf{a}) \quad \text{for some } \mu \in \mathbb{R}.$$

To show that the diagonals bisect each other we need to show that $\lambda = \frac{1}{2}$, $\mu = \frac{1}{2}$. (Check that you understand why this must be so.)

Now

$$\overrightarrow{OA} = \overrightarrow{OP} + \overrightarrow{PA}$$
$$\therefore \qquad \mathbf{a} = \mu(\mathbf{b} + \mathbf{a}) - \lambda(\mathbf{b} - \mathbf{a})$$
$$(1 - \mu - \lambda)\mathbf{a} + (\lambda - \mu)\mathbf{b} = \mathbf{0}$$

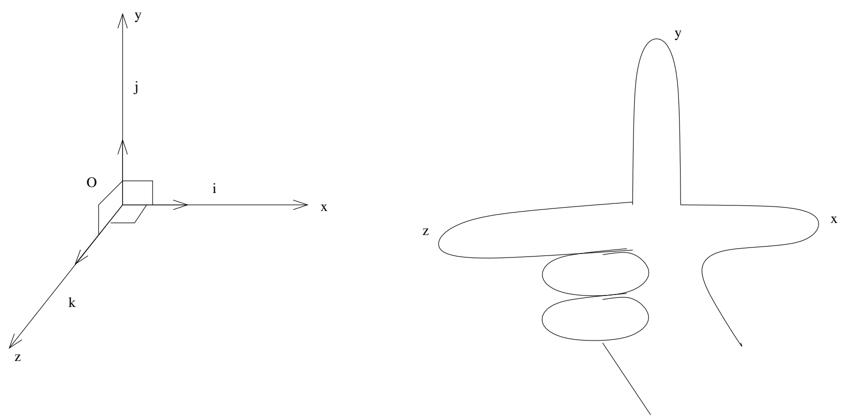
Now, using the previous result it follows that the coefficients of the vectors **a** and **b** must **both** be zero. Hence

$$1 - \mu - \lambda = 0 \quad \text{and} \quad \lambda - \mu = 0.$$

Solving these two equations we see that $\lambda = \mu = \frac{1}{2}$. Hence the diagonals bisect each other. □

2.5 Vectors in Cartesian coordinates

Although vectors are entities in their own right and it is useful for many applications to treat them in this way (also this is how `Matlab` works with vectors, treating them within a coordinate system is often very useful for actual calculations. To do this in three-dimensional Euclidean space we introduce a right-handed mutually orthogonal coordinate frame $(Oxyz)$ with origin O, and with coordinate planes Oxy, Oyz and Ozx. (See diagram following, and Anton §12.1.)



In general, three non-zero vectors $\mathbf{a}, \mathbf{b}, \mathbf{c}$ (which have the same initial point and are not coplanar) form a right-handed co-ordinate system if a right-threaded screw rotated from $\mathbf{a}$ to $\mathbf{b}$ through an angle less than 180^o will advance in the direction $\mathbf{c}$. (See page 4. in "Theory and Problems of Theoretical Mechanics" by Spiegel (in library).)

A **very important set of vectors** are $\mathbf{i}, \mathbf{j}$ and $\mathbf{k}$. These are the unit vectors pointing in the positive directions of Ox, Oy, Oz respectively. Let $\mathbf{a}$ be an arbitrary vector. We can always decompose $\mathbf{a}$ into a "linear combination" of $\mathbf{i}, \mathbf{j}, \mathbf{k}$ as follows.

Proposition 2.11 *Any vector* $\mathbf{a}$*, has components* $a_1, a_2, a_3 \in \mathbb{R}$ *such that* $\mathbf{a} = a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k}$*. The values of the components* a_1, a_2, a_3 *are* unique.

Definition 2.12 *The* a_1, a_2, a_3 *are called* ***components*** *of* $\mathbf{a}$*, and we write* $\mathbf{a} = (a_1, a_2, a_3)$*.*

Proof: In part (1) we prove existence, and then in part (2) show uniqueness.

1. Existence (geometrical):

   Let A be such that $\mathbf{a} = \overrightarrow{OA}$. Construct a "box" (rectangular parallelepiped) with faces parallel to the coordinate planes

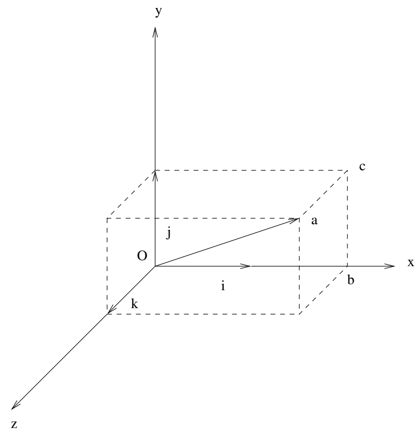
   Let $\mathbf{a} = \overrightarrow{OA}$. Now $\mathbf{a} = \overrightarrow{OA} = \overrightarrow{OB} + \overrightarrow{BC} + \overrightarrow{CA}$, and since $\overrightarrow{OB}$ is parallel to $\mathbf{i}$, there exists a_1 such that $\overrightarrow{OB} = a_1\mathbf{i}$. Similarly $\overrightarrow{BC} = a_2\mathbf{j}$, and $\overrightarrow{CA} = a_3\mathbf{k}$. Therefore

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

2. Uniqueness (algebraic):

   Assume there exists another representation

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



for some $a'_1, a'_2, a'_3 \in \mathbb{R}$. Subtracting gives

$$(a_1 - a'_1)\mathbf{i} + (a_2 - a'_2)\mathbf{j} + (a_3 - a'_3)\mathbf{k} = \mathbf{0}.$$

Assume, without loss of generality, $a_1 \neq a'_1$. Then

$$\mathbf{i} = -\frac{a_2 - a'_2}{a_1 - a'_1}\mathbf{j} - \frac{a_3 - a'_3}{a_1 - a'_1}\mathbf{k}.$$

Therefore **i** is "co-planar" to **j** and **k** (i.e. **i** lies in same plane as **j** & **k**). But **i** is orthogonal to the Oyz plane, a contradiction. □

Examples of the use of components:

Length (this is a really important result):

$$\boxed{\|\mathbf{a}\| = (a_1^2 + a_2^2 + a_3^2)^{\frac{1}{2}}}$$

We obtain this hugely important, and useful, result by applying Pythagoras' theorem twice:

$$\|\mathbf{a}\|^2 = \|\overrightarrow{OA}\|^2 = \|\overrightarrow{OC}\|^2 + \|\overrightarrow{CA}\|^2 = \left(\|\overrightarrow{OB}\|^2 + \|\overrightarrow{BC}\|^2\right) + \|\overrightarrow{CA}\|^2 = a_1^2 + a_2^2 + a_3^2.$$

Addition:

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

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

$$= (a_1 + b_1, a_2 + b_2, a_3 + c_3) \qquad (*)$$

Multiplication by scalar:

$$\lambda\mathbf{a} = \lambda(a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k}) = (\lambda a_1)\mathbf{i} + (\lambda a_2)\mathbf{j} + (\lambda a_3)\mathbf{k} = (\lambda a_1, \lambda a_2, \lambda a_3) \qquad (**)$$

Exercise: Which of the laws of vector algebra have been used in (*) and (**) above?

2.6 The dot product (scalar product)

Definition 2.13 *Given two vectors* $\mathbf{a}$ *and* $\mathbf{b}$*, their* ***dot*** *(or* ***scalar****) product (denoted* $\mathbf{a}.\mathbf{b}$*) is a scalar defined by*

$$\mathbf{a}.\mathbf{b} = \begin{cases} \|\mathbf{a}\|\|\mathbf{b}\|\cos(\theta), & \textit{if } \mathbf{a} \neq \mathbf{0} \textit{ and } \mathbf{b} \neq \mathbf{0} \\ 0 \quad , & \textit{if } \mathbf{a} = \mathbf{0} \textit{ or } \mathbf{b} = \mathbf{0}, \end{cases}$$

where θ *is the angle between* $\mathbf{a}$ *and* $\mathbf{b}$ *(*$0 \leq \theta \leq \pi$*).*

For $\mathbf{a} \neq \mathbf{0}, \mathbf{b} \neq \mathbf{0}$, we have

$$\cos(\theta) = \frac{\mathbf{a}.\mathbf{b}}{||\mathbf{a}||\ ||\mathbf{b}||},$$

and so the dot product tells us about the **angle** between two vectors. In fact, if $\mathbf{a}.\mathbf{b} = 0$, then $\mathbf{a}$ is orthogonal (perpendicular) to $\mathbf{b}$. (See also Anton pp.809-811.) The converse is also true, so we have the **hugely important result** that

Two non-zero vectors $\mathbf{a}$ and $\mathbf{b}$ are **orthogonal** if and only if $\mathbf{a}.\mathbf{b} = 0$.

The dot product is an insanely useful concept. It will be used extensively as a mathematical tool in the manipulation of vectors. However it also has meaning in many physical situations, for example, if $\mathbf{a}$ represents a force and $\mathbf{b}$ represents a displacement, then $W = \mathbf{a}.\mathbf{b}$ is the *work done by the force*. (See the discussion in Anton on work on pp813-814.)

The component (or "algebraic") form of the dot product

From the cosine rule

$$\|\mathbf{a} - \mathbf{b}\|^2 = \|\mathbf{a}\|^2 + \|\mathbf{b}\|^2 - 2\underbrace{\|\mathbf{a}\|\ \|\mathbf{b}\|\cos(\theta)}_{:=\ \mathbf{a}.\mathbf{b}}$$

Therefore,

$$\begin{aligned} \mathbf{a}.\mathbf{b} &= \frac{1}{2}\left(\|\mathbf{a}\|^2 + \|\mathbf{b}\|^2 - \|\mathbf{a} - \mathbf{b}\|^2\right) \\ &= \frac{1}{2}\Big(a_1^2 + a_2^2 + a_3^2 + b_1^2 + b_2^2 + b_3^2 - (a_1 - b_1)^2 - (a_2 - b_2)^2 - (a_3 - b_3)^2\Big) = a_1b_1 + a_2b_2 + a_3b_3 \end{aligned}$$

Thus we have the really **very very** important result

$$\boxed{\mathbf{a}.\mathbf{b} = a_1b_1 + a_2b_2 + a_3b_3.}$$

(Note that we define the dot product in a geometric manner and then deduce the component form. Anton does the reverse - the dot product is defined in Def 12.3.1 and then the geometric form is deduced in Theorem 12.3.3. Matlab does the same as Anton.)

Properties of the dot product: Given the vectors $\mathbf{a}, \mathbf{b}, \mathbf{c}$, and for any scalar λ:

1. $\mathbf{a}.\mathbf{b} = \mathbf{b}.\mathbf{a}$
2. $\mathbf{a}.(\mathbf{b} + \mathbf{c}) = \mathbf{a}.\mathbf{b} + \mathbf{a}.\mathbf{c}$
3. $\lambda(\mathbf{a}.\mathbf{b}) = (\lambda\mathbf{a}).\mathbf{b} = \mathbf{a}.(\lambda\mathbf{b})$
4. $\mathbf{a}.\mathbf{a} = \|\mathbf{a}\|^2$
5. $\mathbf{a}.\mathbf{0} = 0$

Remark These properties can also be used as axioms in an abstract theory of vectors. If we proceed along this line, then we first define a "Linear Space" using the addition and scalar multiplication properties we've seen before, and then introduce an "inner product" based on the rules for the dot product. You'll meet all this in later courses. For example the theory of Hilbert Spaces is developed from these considerations, and this theory is essential in the study of Quantum Mechanics.

Student Exercise (you know what to do!):

(a) Derive (1) and (4) from the definition of $\mathbf{a}.\mathbf{b}$

(b) Derive (2) and (3) from the component form of the dot product.

The dot product is used frequently when solving problems, such as those you will meet in your exam. Here are some examples.

Example 1 (From Part A of a recent exam): (Finding the angle between two vectors) Let $\mathbf{a} = 2\mathbf{i} - \mathbf{j} + \mathbf{k}$, and $\mathbf{b} = \mathbf{i} + \mathbf{j} + 2\mathbf{k}$. Find the angle θ between $\mathbf{a}$ and $\mathbf{b}$.

Answer: We know that

$$\cos(\theta) = \frac{\mathbf{a}.\mathbf{b}}{\|\mathbf{a}\|\|\mathbf{b}\|}.$$

Now, from the component formula for $\mathbf{a}.\mathbf{b}$ we have,

$$\mathbf{a}.\mathbf{b} = 2 \times 1 + (-1) \times 1 + 1 \times 2 = 3$$

and

$$\|\mathbf{a}\| = (2^2 + (-1)^2 + 1^2)^{\frac{1}{2}} = \sqrt{6}, \quad \|\mathbf{b}\| = (1^2 + 1^2 + 2^2)^{\frac{1}{2}} = \sqrt{6}.$$

It follows that

$$\cos(\theta) = \frac{\mathbf{a}.\mathbf{b}}{\|\mathbf{a}\| \cdot \|\mathbf{b}\|} = \frac{3}{\sqrt{6} \cdot \sqrt{6}} = \frac{1}{2},$$

therefore $\theta = \frac{\pi}{3}$. (Exercise: What is the angle in degrees?)

Example 2: The dot product provides an easy way of finding components of vectors. For example, if

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

then

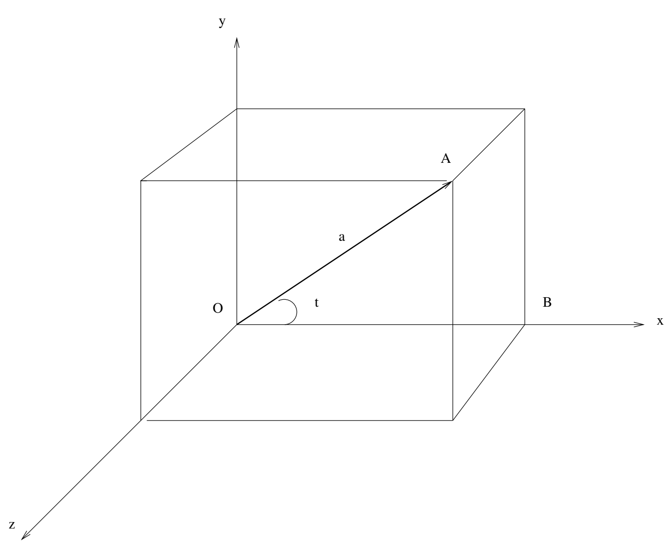
$$\mathbf{a}.\mathbf{i} = a_1, \quad \text{the component of } \mathbf{a} \text{ in the direction of } \mathbf{i}.$$

For the "basis" vectors $\mathbf{i}, \mathbf{j}, \mathbf{k}$ we have

$$\mathbf{i}.\mathbf{i} = \mathbf{j}.\mathbf{j} = \mathbf{k}.\mathbf{k} = 1, \quad \mathbf{i}.\mathbf{j} = \mathbf{j}.\mathbf{k} = \mathbf{k}.\mathbf{i} = 0.$$

□

Example 3 (This is an example of a use of the dot product to do quite a clever piece of geometry) : Find the angle between a diagonal of a cube and one of its edges.

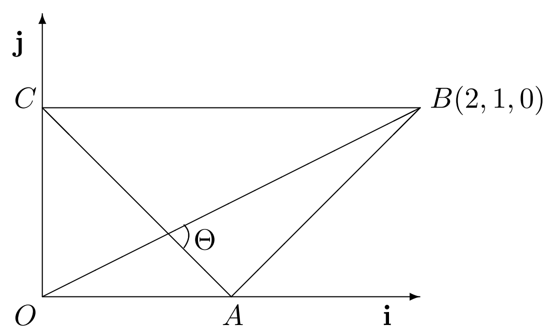


Solution Consider a cube of size d, let $\mathbf{a} = \overrightarrow{OA}$ be the diagonal and let $\mathbf{b} = \overrightarrow{OB}$.

$$\begin{aligned}
\mathbf{a} &= \overrightarrow{OA} \\
&= d\mathbf{i} + d\mathbf{j} + d\mathbf{k} \\
&= (d, d, d) \\
\mathbf{b} = \overrightarrow{OB} = d\mathbf{i} &= (d, 0, 0) \\
\cos\theta = \frac{\mathbf{a}.\mathbf{b}}{\|\mathbf{a}\|\|\mathbf{b}\|} &= \frac{d^2}{\sqrt{3}d.d} = \frac{1}{\sqrt{3}} \\
\therefore \theta &= \cos^{-1}\left(\frac{1}{\sqrt{3}}\right)
\end{aligned}$$

□

* **Example 4:** Find the acute angle between the diagonals of a quadrilateral having vertices $(0, 0, 0)$, $(1, 0, 0)$, $(2, 1, 0)$ and $(0, 1, 0)$.



NB: No **k** direction needed.

$$\overrightarrow{OA} = \mathbf{i}, \qquad \overrightarrow{OB} = 2\mathbf{i} + \mathbf{j},$$
$$\overrightarrow{OC} = \mathbf{j},$$
$$\overrightarrow{CA} = \mathbf{i} - \mathbf{j}.$$

$$\cos\Theta = \frac{\overrightarrow{OB}.\overrightarrow{CA}}{\|\overrightarrow{OB}\|.\|\overrightarrow{CA}\|} = \frac{2.1 + 1.(-1)}{\sqrt{5}.\sqrt{2}} = \frac{1}{\sqrt{10}}$$

$$\Theta = \cos^{-1}\left(\frac{1}{\sqrt{10}}\right).$$

Beware A common error is to say: "If $\mathbf{x}.\mathbf{a} = \mathbf{x}.\mathbf{b}$ then $\mathbf{a} = \mathbf{b}$"

This is clearly incorrect. The correct reasoning goes as follows:

"If $\mathbf{x}.\mathbf{a} = \mathbf{x}.\mathbf{b}$ then $\mathbf{x}.(\mathbf{a} - \mathbf{b}) = 0$. Thus $\mathbf{x}$ is orthogonal to $(\mathbf{a} - \mathbf{b})$."

However, we have a useful Lemma and Corollary, where the key phrase in both is "for all $\mathbf{x}$".

Lemma 2.14 *If* $\mathbf{x}.\mathbf{c} = 0$ *for all* $\mathbf{x}$*, then* $\mathbf{c} = \mathbf{0}$.

Proof Take $\mathbf{x} = \mathbf{c}$. Then $\mathbf{c}.\mathbf{c} = 0$, so $\|\mathbf{c}\|^2 = 0 \Rightarrow \mathbf{c} = \mathbf{0}$.□

Corollary 2.15 *If* $\mathbf{x}.\mathbf{a} = \mathbf{x}.\mathbf{b}$ *for all* $\mathbf{x}$*, then* $\mathbf{a} = \mathbf{b}$.

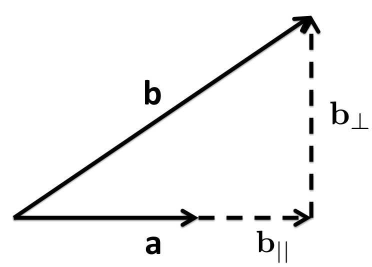
Proof Rearranging gives $\mathbf{x}.(\mathbf{a} - \mathbf{b}) = 0$ for all $\mathbf{x}$, and using Lemma 1.14, we have $\mathbf{a} - \mathbf{b} = \mathbf{0} \Rightarrow \mathbf{a} = \mathbf{b}$.□

Exercise Why is Corollary 2.15 different from the Beware example?

Orthogonal projection Consider $\mathbf{a} \neq \mathbf{0}$ and $\mathbf{b}$. A useful way to represent $\mathbf{b}$ is in the form:

$$\mathbf{b} = \mathbf{b}_{||} + \mathbf{b}_{\perp},$$

where $\mathbf{b}_{||}$ is parallel to $\mathbf{a}$, and $\mathbf{b}_{\perp}$ is orthogonal to $\mathbf{a}$, so that $\mathbf{a}.\mathbf{b}_{\perp} = 0$. (See diagram.)



To show this we let

$$\mathbf{b}_{||} = \lambda \mathbf{a},$$

for some $\lambda \in \mathbb{R}$. Therefore,

$$\begin{aligned} \mathbf{b}.\mathbf{a} &= (\mathbf{b}_{||} + \mathbf{b}_{\perp}).\mathbf{a} \\ &= (\mathbf{b}_{||}).\mathbf{a} \\ &= \lambda \mathbf{a}.\mathbf{a} \\ &= \lambda \|\mathbf{a}\|^2. \end{aligned}$$

Hence, since $\mathbf{a} \neq \mathbf{0}$,

$$\lambda = \frac{\mathbf{b}.\mathbf{a}}{\|\mathbf{a}\|^2}$$

and so we can determine $\mathbf{b}_{||}$. Next, $\mathbf{b}_\perp$ can be found from the equation $\mathbf{b}_\perp = \mathbf{b} - \mathbf{b}_{||}$. Thus,

$$\begin{aligned}
\mathbf{b}_{||} &= \frac{(\mathbf{b}.\mathbf{a})}{\|\mathbf{a}\|^2}\mathbf{a} \\
\mathbf{b}_\perp &= \mathbf{b} - \frac{(\mathbf{b}.\mathbf{a})}{\|\mathbf{a}\|^2}\mathbf{a}
\end{aligned}$$

Definition 2.16 $\mathbf{b}_{||}$ *is called the orthogonal projection of* $\mathbf{b}$ *on* $\mathbf{a}$*: it can also be written as* $(\mathbf{b}.\hat{\mathbf{a}})\hat{\mathbf{a}}$*, where* $\hat{\mathbf{a}}$ *is the unit vector along* $\mathbf{a}$*.* $\mathbf{b}_\perp$ *is the vector component of* $\mathbf{b}$ *orthogonal to* $\mathbf{a}$*.*

Later we shall use the fact (easily deduced from the previous diagram) that $||\mathbf{b}_\perp|| = ||\mathbf{b}|| \sin\theta$, where θ $(0 \leq \theta \leq \pi/2)$ is the angle between $\mathbf{a}$ and $\mathbf{b}$.

***Direction Cosines** Let $\mathbf{a} = \overrightarrow{OA}$ be a vector from O to A, and denote by $\alpha,\ \beta,\ \gamma$ the angles that $\overrightarrow{OA}$ makes with Ox, Oy, Oz, respectively. We define the **direction cosines** of $\mathbf{a}$ to be $\cos(\alpha),\ \cos(\beta),\ \cos(\gamma)$. Now

$$a_1 = \mathbf{i}.\mathbf{a} = \|\mathbf{i}\|\ \|\mathbf{a}\| \cos(\alpha)$$

and so,

$$\cos(\alpha) = \frac{a_1}{\|\mathbf{a}\|}.$$

Similarly, $\cos(\beta) = \frac{a_2}{\|\mathbf{a}\|},\ \cos(\gamma) = \frac{a_3}{\|\mathbf{a}\|}$. It is easily shown that

$$\cos^2(\alpha) + \cos^2(\beta) + \cos^2(\gamma) = 1.$$

***Exercises**

1. What are the direction cosines of $\mathbf{i}$?
2. A line makes angles of 60^o with both the $x-$axis and the $y-$axis and is inclined at an obtuse angle to the $z-$axis. Show that its direction cosines are $\frac{1}{2}, \frac{1}{2}, -\frac{1}{\sqrt{2}}$ and find the angle it makes with the $z-$axis.

Addition of Orthogonal projections and Orthogonal Components: Given $\mathbf{a} \neq \mathbf{0},\ \mathbf{b}, \mathbf{c}$. We can write $\mathbf{b} = \mathbf{b}_{||} + \mathbf{b}_\perp$, and $\mathbf{c} = \mathbf{c}_{||} + \mathbf{c}_\perp$. Similarly, $\mathbf{b} + \mathbf{c} = (\mathbf{b} + \mathbf{c})_{||} + (\mathbf{b} + \mathbf{c})_\perp$. We have the following result:

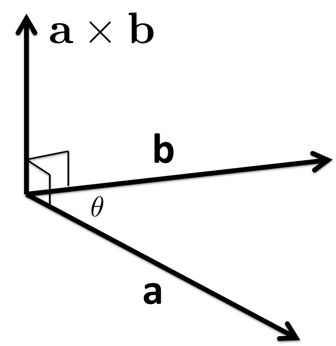
$$(\mathbf{b}+\mathbf{c})_{||} = \mathbf{b}_{||} + \mathbf{c}_{||}, \quad (\mathbf{b}+\mathbf{c})_{\perp} = \mathbf{b}_{\perp} + \mathbf{c}_{\perp}.$$

Proof Now $\mathbf{b}_{||} = \dfrac{(\mathbf{b}.\mathbf{a})}{\|\mathbf{a}\|^2}\mathbf{a}$, and there is a similar expression for $\mathbf{c}_{||}$. Adding gives:

$\mathbf{b}_{||} + \mathbf{c}_{||} = \dfrac{(\mathbf{b}+\mathbf{c}).\mathbf{a}}{\|\mathbf{a}\|^2}\mathbf{a} = (\mathbf{b}+\mathbf{c})_{||}$. Then $\mathbf{b}_{\perp} + \mathbf{c}_{\perp} = \mathbf{b} + \mathbf{c} - \big(\mathbf{b}_{||} + \mathbf{c}_{||}\big) = (\mathbf{b}+\mathbf{c})_{\perp}$ □

2.7 The cross (or vector) product

The cross product gives us a way of 'multiplying' vectors. It was invented/discovered by the great French mathematician Lagrange. Often in applications one needs to find a vector orthogonal to two other vectors. This is accomplished using the cross product. However the cross product has many other uses define.



Definition 2.17 *Given two vectors* $\mathbf{a}$ *and* $\mathbf{b}$*,which are at an angle* θ *apart, their* ***cross*** *(or* **vector** *product)* $\mathbf{c}$ *denoted by* $\mathbf{c} = \mathbf{a} \times \mathbf{b}$ *is a vector*

(i) orthogonal to both $\mathbf{a}$ *and* $\mathbf{b}$

(ii) $\|\mathbf{c}\| = \|\mathbf{a}\| \, \|\mathbf{b}\| \sin(\theta), \quad (0 \le \theta \le \pi)$

(iii) If $\mathbf{a} \parallel \mathbf{b}$*, then* $\mathbf{a} \times \mathbf{b} = \mathbf{0}$*. Otherwise, the direction of* $\mathbf{c}$ *is such that* $\mathbf{a}, \mathbf{b}, \mathbf{c}$ *form a right-handed system (that is, all the angles are less than* π*).*

In fact, it is easy to see that $\mathbf{a} \parallel \mathbf{b}$ if and only if $\mathbf{a} \times \mathbf{b} = \mathbf{0}$.

IMPORTANT NOTE, PLEASE READ In some older text books, exam questions for versions of this course, and possibly at school, you may have seen the notation $\mathbf{c} = \mathbf{a} \wedge \mathbf{b}$ for the cross product. In these notes I will follow the more usual practice now of using $\mathbf{c} = \mathbf{a} \times \mathbf{b}$.

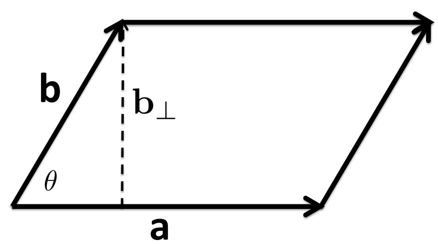
However, these two notations mean exactly the same thing. For this course please use the $\mathbf{c} = \mathbf{a} \times \mathbf{b}$ notation for all of your work.

As well as its numerous applications in geometry (such as calculating volumes and areas), the cross product plays a **VERY** important role in mechanics, especially when rotation is involved. We shall see many examples of this. For example the angular momentum of a body, and the torque/couple generated by a force are all expressed in terms of the cross product. The Coriolis and Magnus forces forces experienced by a spinning body in motion (think of the Earths atmosphere or a football), and the Lorentz force $\mathbf{F}$ experienced by an electron of charge e moving at velocity $\mathbf{v}$ in a magnetic field $\mathbf{B}$ is given by

$$\mathbf{F} = e\, \mathbf{v} \times \mathbf{B}.$$

When coupled to the gradient operator ∇ which we will meet later, we get the 'curl' operator $\nabla\times$ which allows us to define quantities such as vorticity (which is vital for weather forecasting) and links electricity and magnetism together in Maxwell's equations.

As advertised above, there is a simple **geometrical interpretation** of the length of a cross product in terms of an **area**. First, note that $\|\mathbf{a} \times \mathbf{b}\| = ||\mathbf{a}||\ ||\mathbf{b}|| \sin(\theta) = \|\mathbf{a}\|\ \|\mathbf{b}_\perp\|$, since $\|\mathbf{b}_\perp\| = \|\mathbf{b}\|\ \sin(\theta)$. Thus the length of $\|\mathbf{a} \times \mathbf{b}\|$ equals the area of the parallelogram generated by $\mathbf{a}$ and $\mathbf{b}$.



(1) $\mathbf{a}\times\mathbf{b} = -\mathbf{b}\times\mathbf{a}$ So the cross product is **not** commutative. In fact it is **anti-commutative**.

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

(3) $\lambda(\mathbf{a} \times \mathbf{b}) = (\lambda \mathbf{a}) \times \mathbf{b} = \mathbf{a} \times (\lambda \mathbf{b})$

(4) $\mathbf{a} \times \mathbf{a} = \mathbf{0}, \forall \mathbf{a} \quad$ (since $\sin(\theta) = 0$).

Proof of (1): First, $\mathbf{a} \times \mathbf{b}$ and $\mathbf{b} \times \mathbf{a}$ are both orthogonal to $\mathbf{a}$ and $\mathbf{b}$: so they are parallel. Next, $\|\mathbf{a} \times \mathbf{b}\| = \|\mathbf{b} \times \mathbf{a}\| = \|\mathbf{a}\| \, \|\mathbf{b}\| \sin(\theta)$: so that $\mathbf{a} \times \mathbf{b}$ and $\mathbf{b} \times \mathbf{a}$ have the same magnitude. Finally, $\mathbf{a}, \mathbf{b}, \mathbf{a} \times \mathbf{b}$ form a right handed system. Also, $\mathbf{b}, \mathbf{a}, \mathbf{b} \times \mathbf{a}$ form a right handed system. Thus $\mathbf{d}$ is in the opposite direction to $\mathbf{c}$ and so,

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

□

Proof of (2): We'll delay the proof of this till after we introduce the *scalar triple product.*

Important examples:

$$\begin{aligned} \mathbf{i} \times \mathbf{i} &= \mathbf{j} \times \mathbf{j} = \mathbf{k} \times \mathbf{k} = \mathbf{0}. \\ \mathbf{i} \times \mathbf{j} &= \mathbf{k}, \quad \mathbf{j} \times \mathbf{k} = \mathbf{i}, \quad \mathbf{k} \times \mathbf{i} = \mathbf{j}. \\ \mathbf{j} \times \mathbf{i} &= -\mathbf{k}, \quad \mathbf{k} \times \mathbf{j} = -\mathbf{i}, \quad \mathbf{i} \times \mathbf{k} = -\mathbf{j}. \end{aligned}$$

(Proofs from the definition.)

The component form of the vector product: This is a bit messy, but it is an important way to calculate the cross product. The best way to remember this is to use the 'determinant' method which we will come on to shortly Given two vectors

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

then (learn this, you are **expected** to know it for your examination!!!)

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

We establish this by using the rules for the cross products of the unit vectors as follows.

$$\begin{aligned}
\mathbf{a} \times \mathbf{b} &= (a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k}) \times (b_1\mathbf{i} + b_2\mathbf{j} + b_3\mathbf{k}) \\
&= a_1b_1\mathbf{i} \times \mathbf{i} + a_1b_2\mathbf{i} \times \mathbf{j} + a_1b_3\mathbf{i} \times \mathbf{k} + \\
&\quad a_2b_1\mathbf{j} \times \mathbf{i} + a_2b_2\mathbf{j} \times \mathbf{j} + a_2b_3\mathbf{j} \times \mathbf{k} + \\
&\quad a_3b_1\mathbf{k} \times \mathbf{i} + a_3b_2\mathbf{k} \times \mathbf{j} + a_3b_3\mathbf{k} \times \mathbf{k} \\
&= a_1b_2\mathbf{k} - a_1b_3\mathbf{j} - a_2b_1\mathbf{k} + a_2b_3\mathbf{i} + a_3b_1\mathbf{j} - a_3b_2\mathbf{i} \\
&= (a_2b_3 - a_3b_2)\mathbf{i} + (a_3b_1 - a_1b_3)\mathbf{j} + (a_1b_2 - a_2b_1)\mathbf{k}.
\end{aligned}$$

There is a convenient way of calculating, and remembering, the vector product using 3×3 determinants which you will also meet in the algebra courses. This is how I remember the formula myself.

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

DO learn this. I expect you to know it **by heart** for your exam!

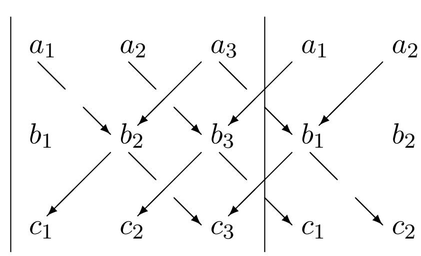
"The rule of Sarrus"

For 3×3 determinants, one can use the "Rule of Sarrus" which goes as follows:

To evaluate

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

the first step is to repeat the first two columns



Now the determinant equals

$$\underbrace{(a_1b_2c_3 + a_2b_3c_1 + a_3b_1c_2) - (a_3b_2c_1 + a_1b_3c_2 + a_2b_1c_3)}$$

"the ↘ terms" − "the ↙ terms"

Exercise: You can easily check that this equals. Make sure that you do this.

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

Remark In these notes we define the vector product geometrically and go on to deduce the component form. Anton first defines the vector product using the determinant (and then deduces the geometric properties. Either approach is good. But I would advise you to get familiar with the vector form as it will make calculations much simpler and quicker later on.'

2.8 The Scalar Triple Product

We saw earlier that if $\mathbf{a}$ and $\mathbf{b}$ are two vectors, then the **area** of the parallelogram spanned by them is given by the magnitude of $\mathbf{a} \times \mathbf{b}$. Now consider *three* vectors $\mathbf{a}, \mathbf{b}, \mathbf{c}$. These can be thought of as the sides of a parallolopiped (think of a squashed cube). The scalar triple product allows us to calculate the **volume** of this solid object.

Definition 2.18 *Given three vectors* $\mathbf{a}, \mathbf{b}, \mathbf{c}$*, their* **scalar triple product***, denoted by* $[\mathbf{a}, \mathbf{b}, \mathbf{c}]$*, is the scalar*

$$\boxed{[\mathbf{a}, \mathbf{b}, \mathbf{c}] := \mathbf{a}.(\mathbf{b} \times \mathbf{c})}$$

The component form of the scalar triple product:

$[\mathbf{a}, \mathbf{b}, \mathbf{c}] = a_1(\mathbf{b} \times \mathbf{c})_1 + a_2(\mathbf{b} \times \mathbf{c})_2 + a_3(\mathbf{b} \times \mathbf{c})_3 =$

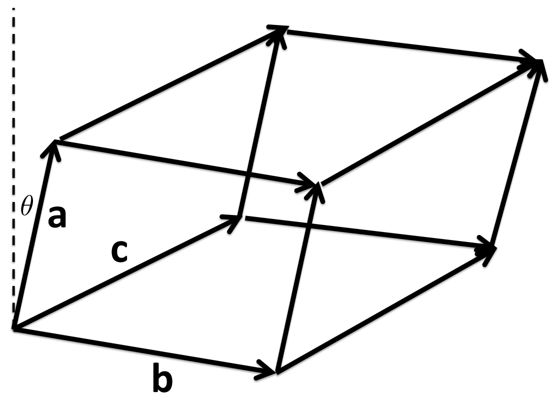
$= a_1(b_2c_3 - b_3c_2) + a_2(b_3c_1 - b_1c_3) + a_3(b_1c_2 - b_2c_1)$

$$= \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix} \text{ - the } 3 \times 3 \text{ determinant again!}$$

Back to the geometric interpretation:

Consider the parallelepiped $\mathcal{P}$ (spanned by $\mathbf{a}, \mathbf{b}, \mathbf{c}$). The area of the base is $\|\mathbf{b} \times \mathbf{c}\|$. The height is $\|\mathbf{a}\| \cos(\theta)$ if $(0 \leq \theta \leq \frac{\pi}{2})$, or $-\|\mathbf{a}\| \cos(\theta)$ if $(\frac{\pi}{2} \leq \theta \leq \pi)$. (Here θ is the angle between $\mathbf{a}$ and $\mathbf{b} \times \mathbf{c}$.) If the volume of $\mathcal{P}$ is denoted by V, then

$$[\mathbf{a}, \mathbf{b}, \mathbf{c}] = \begin{cases} V > 0, \text{ if } \mathbf{a}, \mathbf{b}, \mathbf{c} \text{ is right handed} \\ -V < 0, \text{ if } \mathbf{a}, \mathbf{b}, \mathbf{c} \text{ is left handed.} \end{cases}$$



So $[\mathbf{a}, \mathbf{b}, \mathbf{c}]$ is a "signed" volume of a parallelepiped.

Since the volume of the parallelepiped as well as the right-handedness property are unchanged if the vectors $\mathbf{a}, \mathbf{b}$ and $\mathbf{c}$ are interchanged cyclically we see immediately that

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

Here we have kept the . and the $\times$ fixed and cyclically interchanged $\mathbf{a}, \mathbf{b}$ and $\mathbf{c}$. We give a full list of the properties of the scalar triple product shortly.

This result will become **insanely important** when we want to change variables in multi-variate calculus!

Co-planar vectors: Three vectors are **co-planar** (in other words they all lie in the same plane) if there exist scalars $\lambda, \mu, \nu,$ with $\lambda^2 + \mu^2 + \nu^2 > 0,$ such that $\lambda\mathbf{a} + \mu\mathbf{b} + \nu\mathbf{c} = \mathbf{0}$. (We say that $\mathbf{a}, \mathbf{b}$ and $\mathbf{c}$ are "linearly dependent" vectors.)

$$[\mathbf{a}, \mathbf{b}, \mathbf{c}] = 0 \iff \mathbf{a}, \mathbf{b}, \mathbf{c} \quad \text{are co-planar} \ .$$

Exercise The proof of this result follows immediately from the geometric interpretation. Make sure that you understand why.

Properties of the scalar triple product

(1) $[\mathbf{a}, \mathbf{b}, \mathbf{c}] = [\mathbf{b}, \mathbf{c}, \mathbf{a}] = [\mathbf{c}, \mathbf{a}, \mathbf{b}] =$

$= -[\mathbf{a}, \mathbf{c}, \mathbf{b}] = -[\mathbf{b}, \mathbf{a}, \mathbf{c}] = -[\mathbf{c}, \mathbf{b}, \mathbf{a}]$

(Setting $\mathbf{c} = \mathbf{a}$ gives $[\mathbf{a}, \mathbf{a}, \mathbf{b}] = [\mathbf{a}.\mathbf{b}.\mathbf{a}] = [\mathbf{b}, \mathbf{a}, \mathbf{a}] = 0, \forall \mathbf{a}, \mathbf{b})$

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

(3) $\lambda[\mathbf{a},\mathbf{b},\mathbf{c}] = [\lambda\mathbf{a},\mathbf{b},\mathbf{c}]$

Properties (1) - (3) follow from the properties of the scalar and the vector products.

Example:

$$[\mathbf{i},\mathbf{j},\mathbf{k}] = \mathbf{i}.(\mathbf{j}\times\mathbf{k}) = \mathbf{i}.\mathbf{i} = 1.$$

N.B. $[\mathbf{a},\mathbf{b},\mathbf{c}] = \mathbf{a}.(\mathbf{b}\times\mathbf{c}) = [\mathbf{c},\mathbf{a},\mathbf{b}] = \mathbf{c}.(\mathbf{a}\times\mathbf{b}) = (\mathbf{a}\times\mathbf{b}).\mathbf{c}$. Therefore the order of the . and $\times$ doesn't matter. Also, the brackets can be dropped in $\mathbf{a}.(\mathbf{b}\times\mathbf{c})$, since $(\mathbf{a}.\mathbf{b})\times\mathbf{c}$ makes no sense (a scalar vector product a vector).

Proof of the distributive law, property (2), for the vector product: We wish to prove:

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

Proof To prove this result (which is hard to prove straight from the definitions), we use the cyclic property of the scalar triple product (twice) and the distributive property of the scalar product. Let $\mathbf{x}$ be an arbitrary vector and then consider

$$\begin{aligned}
\mathbf{x}.\mathbf{a}\times(\mathbf{b}+\mathbf{c}) &= (\mathbf{b}+\mathbf{c}).\mathbf{x}\times\mathbf{a} \quad \text{(cyclic STP)}\\
&= \mathbf{b}.\mathbf{x}\times\mathbf{a}+\mathbf{c}.\mathbf{x}\times\mathbf{a} \quad \text{(distributive property of scalar product)}\\
&= \mathbf{x}.\mathbf{a}\times\mathbf{b}+\mathbf{x}.\mathbf{a}\times\mathbf{c} \quad \text{(cyclic STP)}\\
&= \mathbf{x}.(\mathbf{a}\times\mathbf{b}+\mathbf{a}\times\mathbf{c}).
\end{aligned}$$

Therefore

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

for **all** $\mathbf{x}$. Corollary 2.15 then gives us the result that

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

as required.□

2.9 The Triple Vector Product

Now we have defined the scalar triple product it is natural to consider what we might mean by the **vector triple product**. In fact we have two **vector triple products:** defined by

$$\boxed{\mathbf{a}\times(\mathbf{b}\times\mathbf{c}) = (\mathbf{c}.\mathbf{a})\,\mathbf{b}-(\mathbf{b}.\mathbf{a})\,\mathbf{c}}$$

and

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

NOTE These two vector triple products are NOT the same. For those of you who like group theory, the cross product is *not* associative.

How to remember these identities for your exam? (Which you need to do!) One way is note that each of the vectors inside the brackets on the left appears once outside the brackets on the right and the middle' vector $\mathbf{b}$ appears first. Each term contains $\mathbf{a}, \mathbf{b}, \mathbf{c}$ once only.

Another way is to used the mnemonic **CAB - BAC** for the first identity. Think of sitting in the back of a taxi to help remember this. I once shared a taxi with Carol Vorderman, but that is another story.

Proof of the identity $\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) = (\mathbf{a.c})\mathbf{b} - (\mathbf{a.b})\mathbf{c}$.
We use components. First, if $\mathbf{a} = \mathbf{0}$ then the result is true. Assume $\mathbf{a} \neq \mathbf{0}$, and select the axes such that $\mathbf{a} = a\mathbf{i}$.

$$\begin{aligned}
LHS &= a\mathbf{i} \times \{\mathbf{i}(b_2c_3 - b_3c_2) + \mathbf{j}(b_3c_1 - b_1c_3) + \mathbf{k}(b_1c_2 - b_2c_1)\} \\
&= \mathbf{j}(-a(b_1c_2 - b_2c_1)) + \mathbf{k}(a(b_3c_1 - b_1c_3)), \\
RHS &= (a\mathbf{i}.\mathbf{c})\mathbf{b} - (a\mathbf{i}.\mathbf{b})\mathbf{c} \\
&= (ac_1)\mathbf{b} - (ab_1)\mathbf{c} \\
&= (ac_1)(b_1\mathbf{i} + b_2\mathbf{j} + b_3\mathbf{k}) - (ab_1)(c_1\mathbf{i} + c_2\mathbf{j} + c_3\mathbf{k}) \\
&= \mathbf{j}a(c_1b_2 - c_2b_1) + \mathbf{k}a(c_1b_3 - c_3b_1) = LHS. \square
\end{aligned}$$

The equivalent identity for the second vector triple product immediately follows:

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

Important Exercise (you should do this, if only for the real pleasure in seeing a lovely result):

Prove the *beautiful* Jacobi identity:

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

Hint Expand out the 3 vector triple products.

NOTE In the third year of your course (should you choose to do so) we will generalise all of this lovely stuff when we look at *differential geometry* in general and *Lie Algebras* in particular.

2.10 Summary

In this section the main tools for manipulating vectors have been defined and discussed. These tools are, the dot product, the vector product, the scalar triple product and the vector triple product. We shall see how these are used in geometry, in kinematics (especially rotational motion) and in mechanics (where we will bend it like Beckham), in the following Chapters.

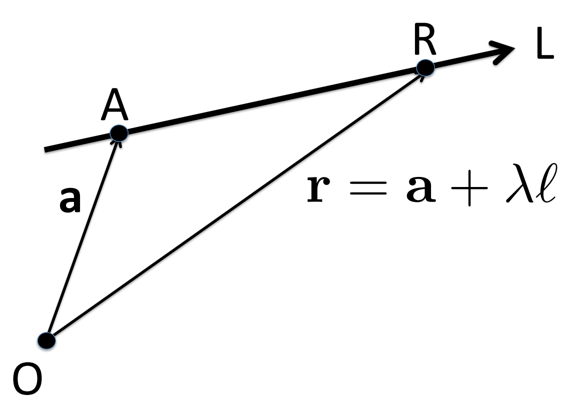
3 Applications of vectors to the geometry of straight lines and planes

In this Chapter we will put the theory of vectors to work, to allow us to represent, and do calculations with, lines and planes and some other surfaces in 3-D. In particular we derive equations for these. We shall answer some typical questions in 3-D geometry, like "what is the equation of the common perpendicular to two given lines", and "find the perpendicular distance from a point to a plane" and "when does a line intersect a plane". These questions are not only interesting and important in their own right, but they lie at the heart of the modern computer graphics industry. (For more information on this see my article *Maths goes to the movies* on Moodle.)

In this Section we will start by looking at lines, then we will look at planes. In the next section we will look at some more exotic surfaces which may change with time.

3.1 The equation of a straight line.

Perhaps the simplest geometric concept (after a point) is that of a straight line. One application of straight lines is in optics, and finding the intersection of straight lines with objects is used in the graphics industry to see how those objects are illuminated by light. To work with lines, we need to know their equations. First we find the equation of a straight line.



A line L in $\mathbb{R}^3$ through a point A, is uniquely specified by the choice of the point A and a direction vector $\mathbf{l} \neq \mathbf{0}$. Let A have position vector $\mathbf{a}$. Then a point R with position vector $\mathbf{r}$ is on the line if and only if there exists $\lambda \in \mathbb{R}$, such that $\mathbf{r} = \mathbf{a} + \lambda \mathbf{l}$. The **vector equation** of the line is given by taking all values of λ so that

$$\mathbf{r} = \mathbf{a} + \lambda \mathbf{l}, \quad (-\infty < \lambda < +\infty)\,. \tag{3.1}$$

Definition 3.1 *Equation (3.1) is the parametric form of the vector equation of a straight line. As the "real parameter" λ varies from $-\infty$ to $+\infty$, the point R traces out the entire line. The **component form** of (3.1) is as follows. For*

$$\mathbf{r} = (x, y, z), \quad \mathbf{a} = (a_1, a_2, a_3), \quad \mathbf{l} = (l_1, l_2, l_3)$$

then

$$L : \begin{cases} x &= a_1 + \lambda l_1 \\ y &= a_2 + \lambda l_2 \\ z &= a_3 + \lambda l_3 \end{cases}$$

□

An **alternative and often very useful form** of the equation of a straight line is as follows. The vector $\overrightarrow{AR}$ is parallel to $\mathbf{l}$ if and only if

$$\boxed{(\mathbf{r} - \mathbf{a}) \times \mathbf{l} = \mathbf{0}} \tag{3.2}$$

We now give several examples which illustrate the use of equations (3.1), (3.2).

Example: Find the vector equation of the line passing through the points A and B with position vectors $\mathbf{a} = \mathbf{i} + \mathbf{j}$, $\mathbf{b} = \mathbf{j} + \mathbf{k}$ respectively.

Solution: Now $\mathbf{l} = \mathbf{b} - \mathbf{a} = \mathbf{k} - \mathbf{i}$, and using the formula $\mathbf{r} = \mathbf{a} + \lambda \mathbf{l}$, we obtain

$$\mathbf{r} = \mathbf{i} + \mathbf{j} + \lambda(\mathbf{k} - \mathbf{i}).$$

In components

$$x = 1 - \lambda,\ y = 1,\ z = \lambda, \qquad (\text{or } y = 1,\ x = 1 - z,\ z \text{ is arbitrary}). \quad \square$$

Exercise Use the alternative form of the equation of the line given by (3.2) to obtain the equation of the line in the previous example. [Hint: $(\mathbf{r}-\mathbf{a})\times\mathbf{l}=\mathbf{0}$ implies

$$(\mathbf{r}-\mathbf{i}-\mathbf{j})\times(\mathbf{k}-\mathbf{i})=\mathbf{0}.$$

Now, with $\mathbf{r}=(x,y,z)$, apply the determinant form of the cross product....] □

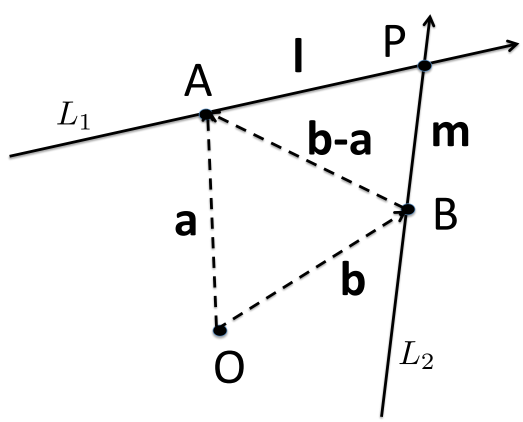
We now consider two important geometrical situations which provide very useful results later on in this Chapter and are are also useful in computer animation applications.

Find the **condition that two lines intersect**. Given two non-parallel lines in $\mathbb{R}^3$

$$\begin{aligned} L_1:\mathbf{r} &= \mathbf{a}+\lambda\mathbf{l}, \quad \lambda\in\mathbb{R} \\ L_2:\mathbf{r} &= \mathbf{b}+\mu\mathbf{m}, \quad \mu\in\mathbb{R}, \end{aligned}$$

show that they intersect if and only if

$$[\mathbf{a},\mathbf{l},\mathbf{m}]=[\mathbf{b},\mathbf{l},\mathbf{m}]\ . \tag{3.3}$$



Solution: From the figure it is clear that if the point A, with position vector $\mathbf{a}$, is on L_1 and B, with position vector $\mathbf{b}$ is on L_2, and there is a point P on the intersection of the two lines so that $P\in L_1\cap L_2$ then the vectors $(\mathbf{a}-\mathbf{b}),\mathbf{l},\mathbf{m}$ are co-planar. Thus $[\mathbf{a}-\mathbf{b},\mathbf{l},\mathbf{m}]=0$. Hence $[\mathbf{a},\mathbf{l},\mathbf{m}]=[\mathbf{b},\mathbf{l},\mathbf{m}]$, using properties of the triple scalar product. □

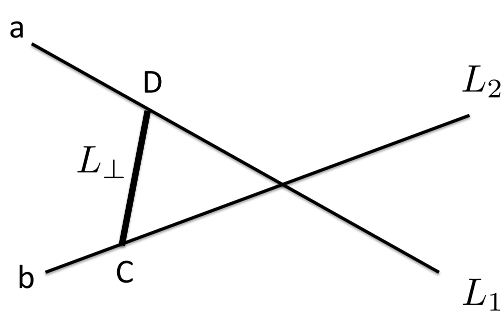
We will now think of a problem faced by an air traffic controller (and by the pilots of the aircraft!). You have two aircraft travelling on straight line paths L_1 and L_2. How close will they get to each other?

To do this calculation we need to work out the distance between the two lines. We do this by calculating the line which is a **common perpendicular** to both lines, and then working out the length of this.

Problem: Let

$$\begin{aligned} L_1 : \mathbf{r} &= \mathbf{a} + \lambda \mathbf{l} \\ L_2 : \mathbf{r} &= \mathbf{b} + \mu \mathbf{m} \end{aligned}$$

be two non-intersecting, non-parallel lines in $\mathbb{R}^3$. What is the equation of the common perpendicular $L_\perp$?



Solution: We seek $L_\perp : \mathbf{r} = \mathbf{c} + \nu\mathbf{n}$ for some direction vector $\mathbf{n}$. Now $L_\perp \perp L_1$ and $L_\perp \perp L_2$. Therefore we can take $\mathbf{n} := \mathbf{l} \times \mathbf{m}$.

Take C to be the point of intersection of $L_2 \cap L_\perp$. Since $C \in L_2$, there exists a real μ_0 such that

$$\mathbf{c} = \mathbf{b} + \mu_0 \mathbf{m}.$$

Therefore

$$L_\perp : \mathbf{r} = \mathbf{b} + \mu_0 \mathbf{m} + \nu(\mathbf{l} \times \mathbf{m}).$$

Now we need to determine μ_0. To do this we use the fact that $L_\perp$ also intersects L_1, and so, using (3.3),

$$[\mathbf{a}, \mathbf{l}, \mathbf{n}] = [\mathbf{c}, \mathbf{l}, \mathbf{n}] = [\mathbf{b} + \mu_0 \mathbf{m}, \mathbf{l}, \mathbf{n}].$$

(Make sure you understand why this is so.) Thus

$$[\mathbf{a},\ \mathbf{l},\ \mathbf{l} \times \mathbf{m}] = [\mathbf{b} + \mu_0 \mathbf{m},\ \mathbf{l},\ \mathbf{l} \times \mathbf{m}].$$

Therefore

$$\mu_0 = \frac{[\mathbf{a} - \mathbf{b},\ \mathbf{l},\ \mathbf{l} \times \mathbf{m}]}{[\mathbf{m},\ \mathbf{l},\ \mathbf{l} \times \mathbf{m}]}.$$

Note that the denominator is

$$[\mathbf{m}, \mathbf{l}, \mathbf{l} \times \mathbf{m}] = [\mathbf{l} \times \mathbf{m}, \mathbf{m}, \mathbf{l}] = (\mathbf{l} \times \mathbf{m}).(\mathbf{m} \times \mathbf{l}) = -(\mathbf{l} \times \mathbf{m}).(\mathbf{l} \times \mathbf{m}) = -\|\mathbf{l} \times \mathbf{m}\|^2,$$

which is nonzero since the lines are non-parallel (hence $\mathbf{l} \times \mathbf{m} \neq \mathbf{0}$). □

We can also find the point D with vector $\mathbf{d}$ where the perpendicular $L_\perp$ intersects the line L_1. We will have

$$\mathbf{d} = \mathbf{a} + \lambda_0 \mathbf{l}.$$

We can find the value of λ_0 by interchanging the two lines in the expression for μ_0. This gives

$$\lambda_0 = \frac{[\mathbf{b} - \mathbf{a},\ \mathbf{m},\ \mathbf{m} \times \mathbf{l}]}{[\mathbf{l},\ \mathbf{m},\ \mathbf{m} \times \mathbf{l}]}.$$

Now we have found the points C and D the length of the shortest line between L_1 and L_2 is given by $\|\mathbf{c} - \mathbf{d}\|$.

NOTE Distances between lines, and between lines and points, is now an important aspect of the rapidly growing field of machine learning.

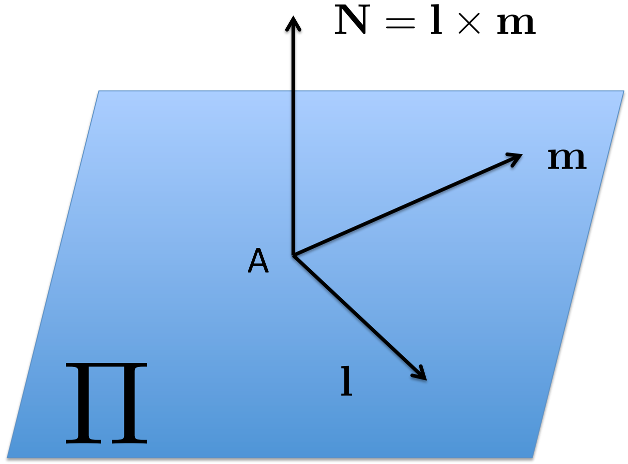
3.2 The equation of the plane.

A plane is the set of points which lie on a flat surface. We usually call a plane a two-dimensional surface.Although we live on a round Earth, because we are small in comparison we experience it as a plane. (Of course if you believe in a flat Earth then it *is* a plane.) In computer graphics we often represent a complicated shape as a set of intersecting planes and look at the way that light reflects from them. All of this makes it worth studying the plane in some detail.

There are several ways that we can represent a plane. In general we will find that it is the set of points (x, y, z) such that

$$ax + by + cz + d = 0,$$

where a, b, c and d are constants.



Let Π be a plane. It can be specified by a point A on the plane with position vector $\mathbf{a}$, and two *non-parallel* vectors $\mathbf{l}$ and $\mathbf{m}$ with in the plane Π. We say $\mathbf{l}$ and $\mathbf{m}$ are parallel to Π. Then for any point $R \in \Pi$ with position vector $\mathbf{r}$ there are a unique pair of *scalars* λ and μ such that

$$\mathbf{r} = \mathbf{a} + \lambda \mathbf{l} + \mu \mathbf{m} \tag{3.4}$$

Equation (3.4) is the *parametric* form of the vector equation of the plane.

Alternative form: There is an alternative, and much more useful, form for the equation of a plane is to think of it as all of the points orthogonal to a given vector. To describe this, let $\mathbf{N}$ be a vector *perpendicular* to Π (a **normal** to Π). For example, one such vector is given by

$$\mathbf{N} = \mathbf{l} \times \mathbf{m}$$

for the above $\mathbf{l}, \mathbf{m}$. Then, $R \in \Pi \Leftrightarrow$

$$(\mathbf{r} - \mathbf{a}).\mathbf{N} = 0. \tag{3.5}$$

We can re-write (3.5) as $\mathbf{r}.\mathbf{N} = \mathbf{a}.\mathbf{N} =: D$, say. In components with

$$N = (A, B, C), \quad \text{and } \mathbf{r} = (x, y, z),$$

then

$$Ax + By + Cz = D\ . \tag{3.6}$$

Example: Find the equation of the plane passing through $A = (1, 2, 3)$ and perpendicular to $\mathbf{N} = (0, 1, 1)$.

Solution: Since $\mathbf{a}.\mathbf{N} = 5$ we have $\mathbf{r}.\mathbf{N} = \mathbf{a}.\mathbf{N} = 5$. Therefore the equation of the plane is

$$y + z = 5.$$

□

A special case is when $\mathbf{N}$ is a **unit** normal, which we denote by $\mathbf{n}$. Then we have

$$(\mathbf{r} - \mathbf{a}) \cdot \mathbf{n} = 0\ . \tag{3.7}$$

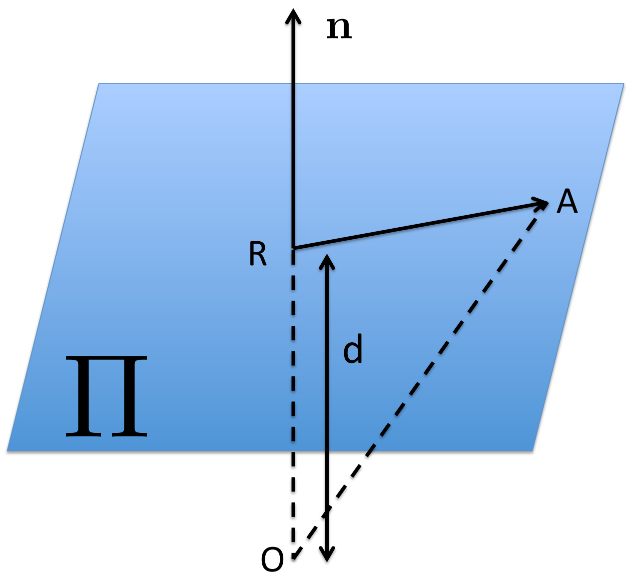
In terms of the above $\mathbf{l}, \mathbf{m}$, see (3.4), we can take the unit normal to be

$$\mathbf{n} = \pm \frac{\mathbf{l} \times \mathbf{m}}{||\mathbf{l} \times \mathbf{m}||}.$$

Equations (3.4), (3.5), (3.6) and (3.7) are equivalent forms of the equation of a plane.
Note that (3.7) may also be written

$$\mathbf{r}.\mathbf{n} = d \tag{3.8}$$

with $d := \mathbf{a}.\mathbf{n}$.



The number d in the above formula has a special geometric meaning. To see this, let $R \in \Pi$ be a point in the plane, with position vector $\mathbf{r}$ so that $\mathbf{r}$ is orthogonal to the plane, and thus $\mathbf{r}$ is parallel to $\mathbf{n}$. It follows that $\mathbf{r} = \lambda \mathbf{n}$ with $\lambda = \pm \|r\|$. So

$$d = \mathbf{r}.\mathbf{n} = \lambda \mathbf{n}.\mathbf{n} = \lambda = \pm \|r\|.$$

Hence d is $\pm$ the **perpendicular distance** from O to Π. $\square$

The component form of (3.8) is, with $\mathbf{n} = (n_1, n_2, n_3)$,

$$n_1 x + n_2 y + n_3 z = d \; . \tag{3.9}$$

Let us consider some examples.

Example 1: Let Π be the plane given by $x + 3y + 4z = -2$. What is the perpendicular distance from O to Π?

Solution: From the equation of the plane we see that the normal $\mathbf{N}$ is given by

$$\mathbf{N} = (1, 3, 4),$$

and the unit normal is

$$\mathbf{n} = \frac{\mathbf{N}}{\|\mathbf{N}\|} = \frac{\mathbf{N}}{\sqrt{26}}.$$

Hence,

$$\mathbf{r}\cdot\mathbf{n} = \frac{\mathbf{r}\cdot\mathbf{N}}{\sqrt{26}} = -\frac{2}{\sqrt{26}}$$

Thus, the distance from O to Π is $\dfrac{2}{\sqrt{26}}$. □

Example 2: Find the intersection of a line

$$L : \mathbf{r} = \mathbf{a} + \lambda\mathbf{l},\ \lambda \in \mathbb{R}$$

and a plane

$$\Pi :\ \mathbf{r}.\mathbf{N} = D.$$

Let S be the point of intersection, with position vector $\mathbf{s} = \mathbf{a} + \lambda_0\mathbf{l}$. (Draw a diagram here if you wish.) Now S lies both on the line and in the plane, so

$$(\mathbf{a} + \lambda_0\mathbf{l}).\mathbf{N} = D \ \ \Rightarrow\ \ \lambda_0 = \frac{D - \mathbf{a}.\mathbf{N}}{\mathbf{l}.\mathbf{N}}. \tag{3.10}$$

□

Example 3: Find the perpendicular distance between the point B with position vector $\mathbf{b}$ and the plane $\Pi : \mathbf{r}.\mathbf{N} = D$.

Solution: Let the perpendicular distance be h.

Draw the perpendicular from B to Π and let it meet Π at B'. Thus, $\overrightarrow{BB'} \perp \Pi$. Therefore $\overrightarrow{BB'} \parallel \mathbf{N}$, and so

$$\overrightarrow{BB'} = \lambda_0\mathbf{N},\ \text{for some } \lambda_0 \in \mathbb{R}.$$

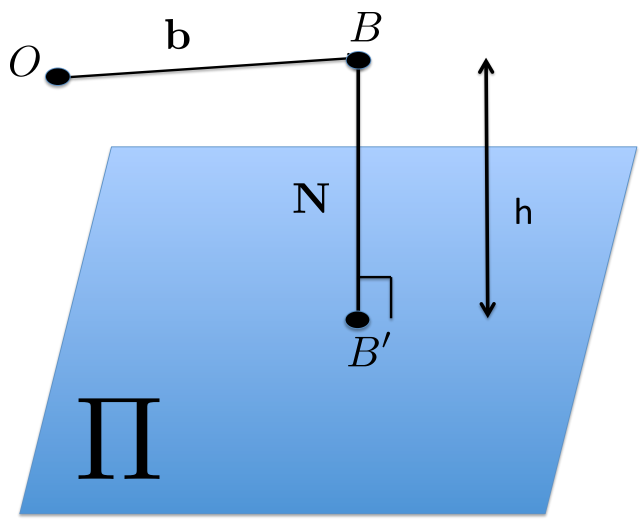
Then $(\mathbf{b} + \lambda_0\mathbf{N})$ is the position vector of B'. Hence

$$(\mathbf{b} + \lambda_0\mathbf{N})\cdot\mathbf{N} = D.$$

So $\mathbf{b}.\mathbf{N} + \lambda_0\|\mathbf{N}\|^2 = D$, and hence

$$\lambda_0 \ = \ \frac{D - \mathbf{b}\cdot\mathbf{N}}{\|\mathbf{N}\|^2}.$$

Therefore, $h = \|\overrightarrow{BB'}\| = |\lambda_0|\ \|\mathbf{N}\| = \dfrac{|D - \mathbf{b}.\mathbf{N}|}{\|\mathbf{N}\|}$. □



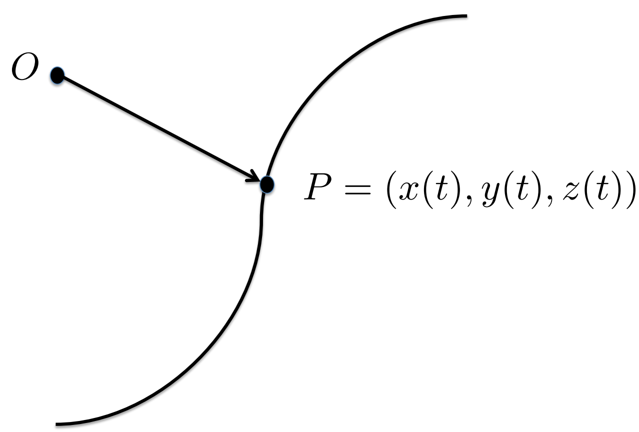
3.3 Summary

In this Chapter we have derived the basic equations of a straight line and a plane in 3-D . In addition, several useful results were obtained, for example, (3.3) gives a condition for two lines to intersect (which is useful in air traffic contriol and also in computer graphics). This result was used to obtain the equation of the common perpendicular to two straight lines. Planes and normals go hand in hand, see (3.5) and (3.7). Important techniques to learn and remember are how to find the normal to a plane, and how to find the perpendicular distance from a point to a plane. These will all come in handy later when we look at how objects move around.

4 Vector functions, line integrals and directional derivatives.

Whilst lines and planes are very important in geometry, real life is full of much more exotic shapes. Fortunately, vector calculus gives us a way of representing these, and also seeing how they might change. This is vital information if you are trying to represent an image in a computer graphics simulation, or to plot the path of a space craft or an aircraft. In this Chapter we will show how we can do this. We will start by looking at general parametric curves. That is, curves which are described by a parameter (usually the variable t). We will then look at what happens when we differentiate these with respect to t to find velocities and accelerations. We will look at integrating with respect to t. This allows us to work out quantities such as energy.

4.1 Examples of parametric curves in 3-D



Let $P = (x(t), y(t), z(t))$ be a point whose position vector $\mathbf{r}$ varies with "time" t, so that,

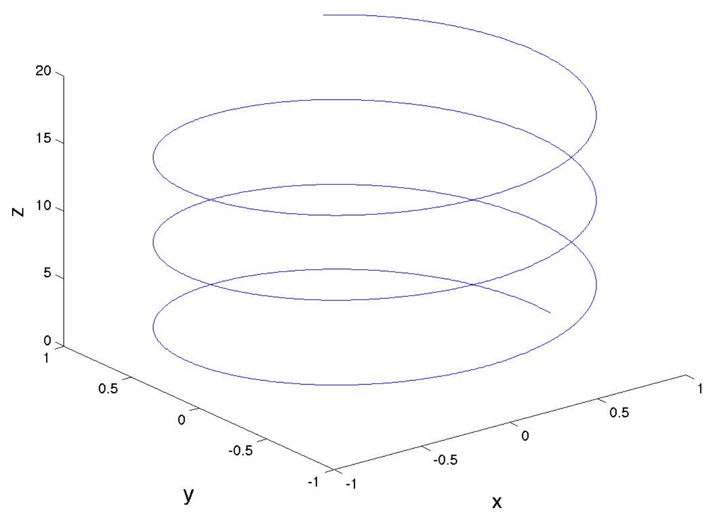
$$\mathbf{r} = \mathbf{r}(t) = x(t)\mathbf{i} + y(t)\mathbf{j} + z(t)\mathbf{k}.$$

The components $x(t)$, $y(t)$, $z(t)$ are real-valued (scalar) functions of the real variable t. So, a vector-valued function $\mathbf{r}(t)$ may be interpreted as a **triple** of real-valued functions. As t varies, the point P traces out a curve C.

Example 1: *The straight line* If $\mathbf{a}$ and $\mathbf{l}$ are constant vectors then $\mathbf{r}(t) = \mathbf{a} + t\mathbf{l}$ represents the equation of a straight line, L say.

As t varies from $-\infty$ to $+\infty$, P traces out the straight line in the direction of $\mathbf{l}$, through the point A.

Example 2: *The Circular helix*



Let

$$\mathbf{r}(t) \;=\; a\cos(t)\,\mathbf{i} + a\sin(t)\,\mathbf{j} \,+\, c\,t\,\mathbf{k} \quad (a,\, c \text{ positive constants })$$

Then

$$x(t) \;=\; a\cos(t), \quad y(t) \,=\, a\sin(t), \quad z(t) \,=\, ct.$$

Note that then $x^2 + y^2 \,=\, a^2$, and therefore the curve C lies on this cylinder $x^2 + y^2 = a^2$ but increases in z with t. The resulting curve C is called a **circular helix**. $\square$

We see the helix in the design of a spiral staircase and also in the shape of the DNA molecule (which is two helices intertwined).

4.2 The calculus of vector-valued functions

If t is thought of as time, then we can think of the points on the helix above as moving along the helix with time. To get the picture, imagine yourself climbing up a circular staircase, with the end of your nose a point. As you move your position changes with time, and therefore we can differentiate it to give you your velocity. This velocity will be itself a vector. We are going

to define differentiation of vector functions in the obvious way, but first, to recall what you are learning in the analysis courses, we need to know what a limit of a vector valued function is.

Definition 4.1 (Limit) $\mathbf{r}(t) \to \mathbf{r}^0$ *as* $t \to t_0$ *iff*

$$\|\mathbf{r}(t) - \mathbf{r}^0\| \to 0 \ \text{as} \ t \to t_0.$$

(Recall that the latter means that for all positive ε *there exists* $\delta > 0$ *such that* $\|\mathbf{r}(t) - \mathbf{r}^0\| < \varepsilon$ *as long as* $|t - t_0| < \delta$*.)*

If we write $\mathbf{r}(t) = (x(t), y(t), z(t))$ *and* $\mathbf{r}^0 = (x^0, y^0, z^0)$ *then obviously* $\mathbf{r}(t) \to \mathbf{r}^0$ *as* $t \to t_0$ *iff*

$$x(t) \to x^0, \ y(t) \to y^0, \ z(t) \to z^0.$$

Here we use the fact that a vector valued function is a triple of real valued scalar functions, and we know, from your analysis courses, how to take the limits of real functions.

Definition 4.2 (Continuity) *We say that* $\mathbf{r}(t)$ *is continuous if for all* t_0, $\mathbf{r}(t) \to \mathbf{r}(t_0)$ *as* $t \to t_0$*. Obviously,* $\mathbf{r}(t)$ *is continuous if and only if all its components* $x(t)$, $y(t)$ *and* $z(t)$ *are continuous.*

Definition 4.3 (Derivative) *Given a time dependent vector* $\mathbf{r}(t)$ *the* ***derivative*** $\mathbf{r}'(t)$ *is*

$$\frac{d\mathbf{r}}{dt} = \mathbf{r}'(t) := \lim_{h \to 0} \left[\frac{\mathbf{r}(t+h) - \mathbf{r}(t)}{h} \right], \tag{4.1}$$

provided this limit exists.

It follows that, provided the scalar derivatives $x'(t)$, $y'(t)$ and $z'(t)$ all exist,

$$\mathbf{r}'(t) = x'(t)\mathbf{i} + y'(t)\mathbf{j} + z'(t)\mathbf{k}, \tag{4.2}$$

so we differentiate the components. From now on, we assume that all $\mathbf{r}(t)$ we consider are "good enough" for the derivatives to exist. We then say $\mathbf{r}(t)$ is *differentiable*.

Notation: There are many ways of writing a derivative. To avoid confusion, in these notes

$$\mathbf{r}'(t), \mathbf{r}', \frac{d\mathbf{r}}{dt}, \frac{d}{dt}\Big(\mathbf{r}(t)\Big)$$

are all equivalent.

A geometrical interpretation of the derivative:

As t varies $\mathbf{r}$ traces out the curve C (think of the example of the helix). Now if P and Q are the end points of the position vectors $\mathbf{r}(t)$ and $\mathbf{r}(t+h)$ respectively,

$$\frac{\mathbf{r}(t+h)\ -\ \mathbf{r}(t)}{h} = \frac{\overrightarrow{PQ}}{h}$$

and so, if $h \to 0$, $Q \to P$, and $\dfrac{\mathbf{r}(t+h) - \mathbf{r}(t)}{h}$ tends to a vector in the direction of the tangent to C at P. Hence, we have the following definition

Assuming $\mathbf{r}'(t) \neq \mathbf{0}$, $\mathbf{r}'(t)$ is called a **tangent vector** to the curve C at point P. The tangent points in the direction of increasing t. The tanjent vector is precisely the **velocity** of the point at that moment.

Example: For the helix

$$\mathbf{r}(t) \ = \ a\cos(t)\,\mathbf{i} + a\sin(t)\,\mathbf{j} \ + \ c\,t\,\mathbf{k}$$

the tangent vector $\mathbf{r}'(t)$, is

$$\mathbf{r}'(t) \ = \ -\,a\sin(t)\,\mathbf{i} + a\cos(t)\,\mathbf{j} \ + \ c\,\mathbf{k}. \qquad \square$$

Rules of differentiation

Let $\mathbf{r}(t), \mathbf{r}_1(t), \mathbf{r}_2(t)$ be differentiable vector-valued functions, $\lambda(t)$ a differentiable real-valued function, and $\mathbf{c}$ a constant vector. We can differentiate sums, dot products and cross products of vectors as follows.

Then:

1. $\dfrac{d}{dt}(\mathbf{c}) = \mathbf{0}$

2. $\dfrac{d}{dt}\Big(\mathbf{r}_1(t) + \mathbf{r}_2(t)\Big) = \dfrac{d}{dt}\Big(\mathbf{r}_1(t)\Big) + \dfrac{d}{dt}\Big(\mathbf{r}_2(t)\Big)$

3. $\dfrac{d}{dt}\Big(\lambda(t)\mathbf{r}(t)\Big) = \dfrac{d}{dt}\Big(\lambda(t)\Big)\mathbf{r}(t) \;+\; \lambda(t)\dfrac{d}{dt}\Big(\mathbf{r}(t)\Big)$

4. For the dot product: $\dfrac{d}{dt}\Big(\mathbf{r}_1 \,.\, \mathbf{r}_2\Big) \;=\; \dfrac{d\mathbf{r}_1}{dt} \,.\, \mathbf{r}_2 \;+\; \mathbf{r}_1 \,.\, \dfrac{d\mathbf{r}_2}{dt}$

5. For the cross product, $\dfrac{d}{dt}\Big(\mathbf{r}_1 \times \mathbf{r}_2\Big) \;=\; \dfrac{d\mathbf{r}_1}{dt} \times \mathbf{r}_2 + \mathbf{r}_1 \wedge \dfrac{d\mathbf{r}_2}{dt}$ (N.B. Keep the correct ordering!)

Thus the usual rules of differentiation apply. Proofs of (1) - (5) directly follow from definition (4.1), or component form (4.2). Verify this as an exercise.

For a position vector $\mathbf{r}(t)$, $\|\mathbf{r}(t)\|$ is a scalar function of t and we can differentiate it.

Example 1: Show that for $\mathbf{r} \neq \mathbf{0}$

$$\frac{d}{dt}\Big(\|\mathbf{r}\|\Big) \;=\; \frac{1}{\|\mathbf{r}\|}\mathbf{r} \,.\, \mathbf{r}'$$

We deduce that $\mathbf{r}.\mathbf{r}' = 0$ iff $\|\mathbf{r}\|$ = constant.

Solution: (i) Recall, $\|\mathbf{r}\|^2 = \mathbf{r}.\mathbf{r}$. Differentiate both sides with respect to t

$$\frac{d}{dt}\Big(\|\mathbf{r}\|^2\Big) \;=\; \frac{d}{dt}\,(\mathbf{r} \,.\, \mathbf{r})$$

$$\therefore\;\; 2\|\mathbf{r}\|\frac{d}{dt}\Big(\|\mathbf{r}\|\Big) \;=\; \frac{d\mathbf{r}}{dt} \,.\, \mathbf{r} + \mathbf{r}\,\frac{d\mathbf{r}}{dt} \;=\; 2\;\mathbf{r}.\frac{d\mathbf{r}}{dt}$$

$$\therefore\;\; \frac{d}{dt}\Big(\|\mathbf{r}\|\Big) = \frac{1}{\|\mathbf{r}\|}\mathbf{r}.\mathbf{r}'. \tag{4.3}$$

We will use this result often.

$$(ii)\;\; \text{If } \|\mathbf{r}\| = \text{ constant } \therefore \frac{d}{dt}\Big(\|\mathbf{r}\|\Big) \;=\; 0$$

$$\therefore \mathbf{r}.\mathbf{r}' = \|\mathbf{r}\|\underbrace{\frac{d}{dt}\Big(\|\mathbf{r}\|\Big)}_{=0} \;=\; 0 \qquad \square$$

Integration By analogy with differentiation, we have

$$\int_a^b \mathbf{r}(t)dt \;:=\; \left(\int_a^b x(t)dt\right)\mathbf{i} + \left(\int_a^b y(t)dt\right)\mathbf{j} + \left(\int_a^b z(t)dt\right)\mathbf{k}.$$

Also, the usual integration rules apply.

Anti-derivatives

$$\int \mathbf{r}'(t)dt \;=\; \mathbf{r}(t) \;+\; \mathbf{c}, \qquad \frac{d}{dt}\left\{\int \mathbf{r}(t)dt\right\} \;=\; \mathbf{r}(t).$$

Example 4: If $\mathbf{r} : \mathbb{R} \to \mathbb{R}^3$ is a twice differentiable function such that $\mathbf{r}''(t) \equiv \mathbf{0}$, show that $\mathbf{r}(t)$ represents a straight line.

Solution: In this case,

$$\begin{aligned}\mathbf{r}''(t) \equiv \mathbf{0} &\Longrightarrow \frac{d}{dt}\big(\mathbf{r}'(t)\big) = \mathbf{0}\\ &\Longrightarrow \mathbf{r}'(t) = \mathbf{c},\end{aligned}$$

where $\mathbf{c}$ is a constant vector (analogous to the constant of integration). Therefore,

$$\int \mathbf{r}'(t)dt = \mathbf{c}t + \mathbf{d},$$

where $\mathbf{d}$ is a constant vector, and thus

$$\mathbf{r}(t) = t\mathbf{c} + \mathbf{d}, \quad \forall t \in \mathbb{R},$$

which is the equation of a straight line through arbitrary point $\mathbf{d}$ in an arbitrary direction $\mathbf{c}$, as required. □

4.3 Arc-length

Imagine that you are an ant crawling along a curve. You may well be interested in how far you have travelled. To do this calculation (even if you are not an ant) we need to find the *arc-length* of the curve. The arc-length of a section of a curve is defined as a limit of lengths of piecewise straight lines. Selecting a point P_0 on a curve C corresponding to $t = t_0$, we introduce the arc length, s say, from P as a *signed length*: s is positive for $t > t_0$ and negative for $t < t_0$.

NOTE It is often convenient to use the arc-length s as a parameter to describe C (instead of t).

Change of variable: (The Chain Rule)

If $\mathbf{r} = \mathbf{r}(t)$ and $t = t(s)$ then

$$\frac{d\mathbf{r}}{ds} = \frac{d\mathbf{r}}{dt}\,\frac{dt}{ds} \tag{4.4}$$

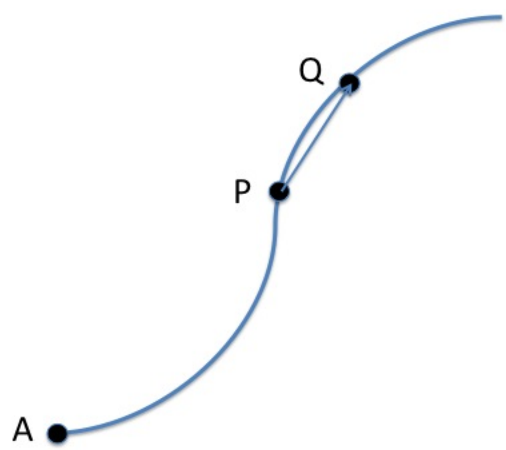
(Prove using components).

Example: For a circle of radius a in the xy-plane, $\mathbf{r} = (x, y, 0)$, where $x(t) = a\cos t$, $y(t) = a\sin t$, and for the arc-length $s(t) = at$. Hence

$$\frac{d\mathbf{r}}{ds} = \frac{d\mathbf{r}}{dt}\,\frac{dt}{ds} = \frac{d\mathbf{r}}{dt}\Big/\frac{ds}{dt} = \frac{1}{a}\mathbf{r}' = -\,\mathbf{i}\sin\left(\frac{s}{a}\right) + \mathbf{j}\cos\left(\frac{s}{a}\right) \quad \square$$

How to compute the arc-length

Warning! Even simple curves can have very hard to compute arc-length. With this warning we proceed to find a formula for arc-length. Let $\mathbf{r}(t)$ be a position vector of P on $\mathcal{C}$, a smooth curve in 3-D.



Let $A \in \mathcal{C}$ be a given reference point and let $s(t)$ be the arc length along $\mathcal{C}$ from A to a point P having position vector $\mathbf{r}(t)$. Now consider a point Q which is close to P for which $s(t+h)$ is length from A to Q, and the line PQ lies close to the part of $\mathcal{C}$ between P and Q.

Then, for small h,

$$\frac{s(t+h)-s(t)}{h} \simeq \frac{\|\overrightarrow{PQ}\|}{h} = \frac{\|\mathbf{r}(t+h)-\mathbf{r}(t)\|}{h} = \left\|\frac{\mathbf{r}(t+h)-\mathbf{r}(t)}{h}\right\|.$$

Now taking the limit as $h \to 0$,

$$\frac{ds}{dt} = \left\|\frac{d\mathbf{r}}{dt}\right\| = \|\mathbf{r}'\| \tag{4.5}$$

Integrating (4.5) from t_0 to t

$$s(t) = \int_{t_0}^{t} \|\mathbf{r}'(\tau)\| \, d\tau = \int_{t_0}^{t} \sqrt{\left(\frac{dx}{d\tau}\right)^2 + \left(\frac{dy}{d\tau}\right)^2 + \left(\frac{dz}{d\tau}\right)^2} \; d\tau$$

(Here τ is a dummy variable of integration which, strictly speaking, could be anything but t which is the upper limit; we nevertheless often use in similar context t also as the integration variable, for notational simplicity.)

See an example in lectures on the arc-length for a circular helix; see also the example in the end of the next subsection. However, be warned (again), this integral is often impossible to calculate!

Warning! In general

$$\left\|\frac{d\mathbf{r}}{dt}\right\| \neq \frac{d}{dt}\Big(\|\mathbf{r}(t)\|\Big).$$

(The length of a derivative is generally different from the derivative of a length.)

4.4 Tanjents, principal unit normal vectors and curvature.

We have seen that $\mathbf{r}'(t)$ is a vector tangent to C at $\mathbf{r}(t)$, assuming $\mathbf{r}'(t) \neq \mathbf{0}$. We apply the usual recipe to construct a vector of unit length pointing in the tangent direction:

Definition 4.4 *With* $\mathbf{r}'(t) \neq \mathbf{0}$,

$$\hat{\mathbf{T}}(t) := \frac{\mathbf{r}'(t)}{\|\mathbf{r}'(t)\|} \tag{4.6}$$

is called the unit vector tangent to C at t (or unit tangent vector).

If we choose the arc length s as the parameter for the curve C, then using the Chain Rule and (4.6)

$$\mathbf{r}' = \frac{d\mathbf{r}}{dt} = \frac{d\mathbf{r}}{ds}\frac{ds}{dt} = \frac{d\mathbf{r}}{ds}\|\mathbf{r}'\| \quad \therefore \quad \text{from (4.6)}$$

$$\hat{\mathbf{T}} = \frac{d\mathbf{r}}{ds} \tag{4.7}$$

$\Big($Note that $\left\|\frac{d\mathbf{r}}{ds}\right\| = 1.\Big)$

We might want to see how rapidly a curve changes. For example if we fold a piece of paper then any curve drawn on that paper changes rapidly at the fold point. We can measure this by defining the *curvature* of the curve.

We have

$$\|\hat{\mathbf{T}}\| = 1 \quad \therefore \quad \|\hat{\mathbf{T}}\|^2 = \hat{\mathbf{T}} \cdot \hat{\mathbf{T}} \equiv 1$$

$$\text{and differentiating the latter,} \quad \hat{\mathbf{T}} \cdot \frac{d\hat{\mathbf{T}}}{dt} = 0$$

Thus $\frac{d\hat{\mathbf{T}}}{dt}$ is orthogonal to the tangent vector $\hat{\mathbf{T}}$ and so is **normal** to the curve C.

Definition 4.5 *Assume* $\hat{\mathbf{T}}' := \frac{d\hat{\mathbf{T}}}{dt} \neq \mathbf{0}$*, then*

$$\hat{\mathbf{N}} := \frac{\hat{\mathbf{T}}'}{\|\hat{\mathbf{T}}'\|} \tag{4.8}$$

is called the ***principal unit normal****.*

(The word "principal" means that this is "the" unit normal to C pointing in the direction in which the curve actually curves, or in other words lying in the plane "best approximating" the curve C near the considered point P. More precise statements would require use of more advanced Analysis.)

In the s-parametrisation, replacing in (4.8) t by s (and recalling that $\frac{ds}{dt} > 0$)

$$\hat{\mathbf{N}} = \frac{\frac{d\hat{\mathbf{T}}}{ds}}{\left\|\frac{d\hat{\mathbf{T}}}{ds}\right\|},$$

or, re-arranging,

$$\frac{d\hat{\mathbf{T}}}{ds} = \kappa\,\hat{\mathbf{N}}, \quad \kappa := \left\|\frac{d\hat{\mathbf{T}}}{ds}\right\|. \tag{4.9}$$

This leads us to the following definitions

The length κ of the vector $\frac{d\hat{\mathbf{T}}}{ds}$ is called **the curvature** of curve C at the given point t.
The quantity $a = 1/\kappa$ is called the *radius of curvature*.

Note that $\hat{\mathbf{T}} = d\mathbf{r}/ds$. It follows that

$$\kappa = \|\frac{d^2\mathbf{r}}{ds^2}\|.$$

The larger that κ is (or the smaller that a is) the more that the curve is changing in its direction over a short distance

By the chain rule and using also (4.5),

$$\kappa = \left\| \frac{d\hat{\mathbf{T}}}{ds} \right\| = \left\| \frac{d\hat{\mathbf{T}}}{dt} \Big/ \frac{ds}{dt} \right\| = \frac{\|\hat{\mathbf{T}}'(t)\|}{\|\mathbf{r}'(t)\|} \tag{4.10}$$

Example: Consider again the circular helix. Then

$$\begin{aligned}
\mathbf{r}(t) &= (a\cos t)\mathbf{i} + (a\sin t)\mathbf{j} + (ct)\mathbf{k} \quad (a,\ c \text{ positive constants }) \\
\mathbf{r}'(t) &= -a\sin t\,\mathbf{i} + a\cos t\,\mathbf{j} + c\mathbf{k} \\
\|\mathbf{r}'\| &= \left(a^2(\sin^2 t + \cos^2 t) + c^2\right)^{\frac{1}{2}} = \sqrt{a^2+c^2}
\end{aligned}$$

So

$$\frac{ds}{dt} = \|\mathbf{r}'\| = \sqrt{a^2+c^2}$$

Take our start point as $t = 0$ i.e. $\mathbf{r}(0) = (a, 0, 0)$.

(1) The **arc-length** from $(a, 0, 0)$ to $\mathbf{r}(t)$ is $s(t) = \int_0^t \|\mathbf{r}'\| dt = \sqrt{a^2+c^2}\, t.$

(2) The **unit tangent vector:**

$$\hat{\mathbf{T}} = \frac{\mathbf{r}'}{\|\mathbf{r}'\|} = \frac{-(a\sin t)\mathbf{i} + (a\cos t)\mathbf{j} + c\mathbf{k}}{\sqrt{a^2+c^2}}.$$

(3) To find the **principal unit normal**, now

$$\hat{\mathbf{T}}' = \frac{d\hat{\mathbf{T}}}{dt} = \frac{-(a\cos t)\mathbf{i} - (a\sin t)\mathbf{j}}{\sqrt{a^2+c^2}}$$

$$\therefore \left\|\hat{\mathbf{T}}'\right\| = \frac{a}{\sqrt{a^2+c^2}} \quad \therefore \quad \hat{\mathbf{N}} = \frac{\hat{\mathbf{T}}'}{\left\|\hat{\mathbf{T}}'\right\|} = -\cos t\,\mathbf{i} - \sin t\,\mathbf{j}.$$

(4) Hence, from (4.9)-(4.15), for the **curvature**,

$$\kappa = \frac{\left\|\frac{d}{dt}\hat{\mathbf{T}}\right\|}{\|\mathbf{r}'\|} = \frac{a}{\sqrt{a^2+c^2}} \frac{1}{\sqrt{a^2+c^2}} = \frac{a}{a^2+c^2},$$

with

$$\frac{d\hat{\mathbf{T}}}{ds} = \frac{d\mathbf{T}}{dt} \Big/ \frac{ds}{dt} = \frac{\frac{1}{\sqrt{a^2+c^2}}(-a\cos t\mathbf{i} - a\sin t\mathbf{j})}{\sqrt{a^2+c^2}} = \frac{-(a\cos t)\mathbf{i} - (a\sin t)\mathbf{j}}{a^2+c^2} = \kappa\,\hat{\mathbf{N}}.$$

Note that, in particular, for $c = 0$ (so that C is a circle) $\kappa = \frac{1}{a}$, the inverse radius. □

More examples of curvature can found in the example sheets. BUt have some fun with origami to give yourselves a feeling for what curvature is all about.

4.5 Line integrals

Imagine that you are carrying a heavy suit of armour as you go up the spiral staircase of a castle. To do this you need to do some work. As well as working against gravity, every step that you take takes effort due to friction, so that the work involves taking into account the arc-length along the curve. To work out the total amount of work that you have to do involves integrating all of this along the whole length of the curve. To this we have to do a *line integral* along the curve. Line integrals are a very important aspect of vector calculus and play a central role in calculations of work and energy. Another example of a line integral is the integration of the velocity vector of the flow of air along a closed loop around the aerofoil of a wing on an aircraft. This line integral called the *circulation* is used to find the lift on the wing. You will meet them not only in this course but also in courses on partial differential equations, multi-variate calculus, fluid mechanics, electromagnetics and in courses involving complex variables.

Definition 4.6 *Suppose that we have a parametrised curve $\mathcal{C}$ in three dimensions from the points A to P given by $\mathbf{r} = (x(t), y(t), z(t))$ for $a < t < b$. This has a tangent vector $\mathbf{r}' = (dx/dt, dy/dt, dz/dt)$. Now suppose that we have a vector quantity $\mathbf{F}(x, y, z)$ which is defined at all points along the curve. Then the* vector line integral *I of the* vector *function $\mathbf{F}(\mathbf{r})$ is given by*

$$I = \int_{t=a}^{b} \mathbf{F}.\mathbf{r}' \, dt = \int_{t=a}^{b} \mathbf{F}(x(t), y(t), z(t)).(dx/dt, dy/dt, dz/dt) \, dt. \tag{4.11}$$

Example 1: Let's, consider a circular staircase with $(x(t), y(t), z(t)) = (\cos(2t), \sin(2t), 3t), \quad 0 < t < 1$. We carry a heavy weight of mass m up this staircase and the force acting on this weight is $\mathbf{F} = (0, 0, -mg)$. The work done in doing this is given by $|I|$ where I is the line integral. Now $\mathbf{r}' = (-2\sin(2t), 2\cos(2t), 3)$. Thus

$$I = \int_0^1 (0, 0, -mg).(-2\sin(2t), 2\cos(2t), 3) \, dt = \int_0^1 -3mg \, dt = -3mg.$$

Example 2: In the same problem above an additional amount of work, proportional to μ is done against friction, so that now $\mathbf{F} = (0, 0, -mg) - \mu\mathbf{r}'$. The total work done is now $|I|$ where

$$I = \int_0^1 (2\mu\sin(2t), -2\mu\cos(2t), -mg - 3\mu).(-2\sin(2t), 2\cos(2t), 3) \, dt = \int_0^1 -3mg - 13\mu \, dt = -3mg - 13\mu.$$

The above definition of a line integral is very useful for finding the integral of vector valued quantities and determining energy. However, we also often want to find the line integral of a scalar quantity. For example, if we have a wire bent into a shape with a varying density $\rho(\mathbf{r})$ along the wire, then the integral of ρ with respect to arc-length s along the wire gives its total mass. We now make this precise

We saw in Section 4.3 that the total arc-length s along the curve $\mathcal{C}$ defined by $\mathbf{r} = \mathbf{r}(t)$, from A to P, is given by:

$$s = \int_{t=a}^{b} ||\mathbf{r}'(t)||\ dt$$

since $ds/dt = ||\mathbf{r}'||$.

This can be generalised to integrate a general scalar function $\rho(\mathbf{r}(t))$ along the curve $\mathcal{C}$:

Definition 4.7 *The* scalar line integral *of the* scalar *function $\rho(\mathbf{r})$ is the scalar I given by*

$$I = \int_{\mathcal{C}} \rho(\mathbf{r})\ ds = \int_{t=a}^{b} \rho(\mathbf{r}(t))\ ||\mathbf{r}'(t)||\ dt. \tag{4.12}$$

Example A wire is described by the curve $\mathbf{r} = (t, t^2, t^3)$, $0 < t < 1$. Its density increases along its length and is given by $\rho(s) = s^2$. Find an expression involving an integral for its total mass m.

We have that

$$m = \int_{\mathcal{C}} \rho(s)\ ds = \int s^2\ ds = \frac{S^3}{3},$$

where S is the total arc-length of the wire. In turn we have that

$$S = \int_{t=0}^{1} \frac{ds}{dt}\ dt = \int_0^1 \sqrt{1 + 4t^2 + 9t^4}\ dt.$$

Hence

$$m = \frac{\left(\int_0^1 \sqrt{1 + 4t^2 + 9t^4}\ dt\right)^3}{3}.$$

If you wish, you can evaluate this integral. Alternatively, you may not.

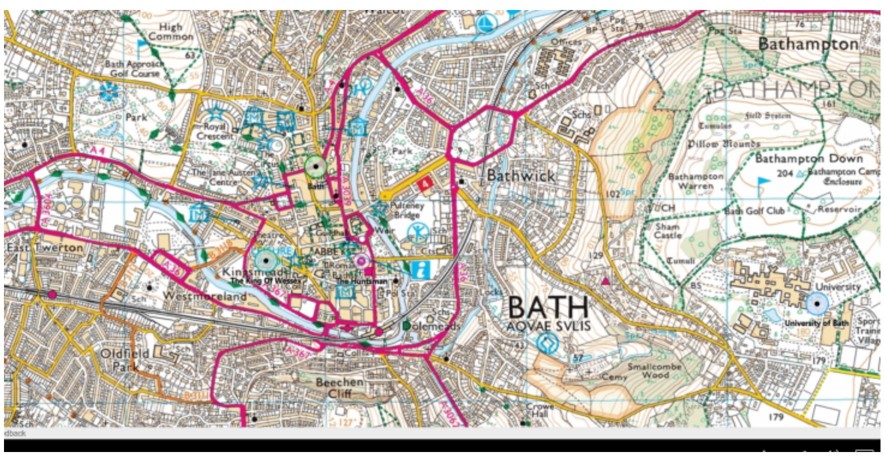
4.6 Gradients and directional derivatives

Suppose that we are walking on a mountain. We sit down to rest and get out an apple to eat. Carelessly we put the apple on the ground and it starts to roll away. What direction will it roll in? Similarly, we walk up the mountain, but on a footpath which keeps changing direction. How fast are we moving up hill? These, and many other important questions, can be answered by using the gradient operator ∇. This gives us a way of finding the derivative of a scalar function f of a vector.

Take the horizontal xy-plane horizontal, and define a differentiable function $f(x, y) : \mathbb{R}^2 \to \mathbb{R}$. Then $z = f(x, y)$ describes a surface in 3D above the xy-plane: $f(a, b)$ is the *height* of the surface above the point (a, b) on the plane. This is called the **graph** of $f(x, y)$. It is the set of points $(x, y, z) \in \mathbb{R}^3$ where $z = f(x, y)$.

Definition If $c \in \mathbb{R}$, then we define a *level curve* of points (x, y) for which $f(x, y) = c$. On the surface, the points (x, y, c) form the **contour lines** of points at the height c.

Back to our mountaineering example. A standard OS map of a mountainous region will depict heights using contour lines. The University of Bath is in just such a (mountainous) region and the 1:25 000 scale OS map of Bath is shown below in which the height contours are all indicated.



At any point (x, y) we can find the partial derivatives of f given by $\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}$, also written as $f_x(x, y)$ and $f_y(x, y)$. These give the rates of change of f with respect to distance in the x and y directions respectively. Geometrically, $f_x(a, b)$ and $f_y(a, b)$ are the slopes of the surface in the positive x and y directions at (a, b). That is, in the directions of the unit vectors $\mathbf{i}, \mathbf{j}$.

We now look at how the value of f changes as we move along a curve. Suppose that we define the curve

$$\mathcal{C} = (x(t), y(t)).$$

We can then define the function g along this curve by

$$g(t) = f(x(t), y(t)).$$

If g is a path on a hill, and f is the height of the hill, then $g(t)$ is our percpetion of how out height changes with time t.

Definition 4.8 *The* directional derivative *of the function f along the curve $\mathcal{C}$ is given by dg/dt.*

By using the *chain rule* we can evaluate the directional derivative. In particular

$$\frac{dg}{dt} = \frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt} \equiv f_x \frac{dx}{dt} + f_y \frac{dy}{dt}. \tag{4.13}$$

A special case of the directional derivative is given when the curve $\mathcal{C}$ is a straight line in the direction of the vector $\mathbf{u} = (u, v)$ passing through the point (a, b). We then have

$$x(t) = a + ut, \quad y(t) = b + vt, \quad g(t) = f(a + ut, b + vt), \quad dx/dt = u, \quad dy/dt = v.$$

The following definition then follows directly from the expression (4.13).

Definition 4.9 *The* directional derivative $D_{\mathbf{u}}f$ *of the function $f(x, y)$ at the point (a, b) and in the direction of the vector $\mathbf{u} = (u, v)$ is given by*

$$D_{\mathbf{u}}f = uf_x + vf_y \tag{4.14}$$

Example Suppose that $f(x, y) = x^2 - 2y^2$. What is the directional derivative of f in the direction of $\mathbf{u} = (1, 2)$ at the point $(a, b) = (1, 1)$?.

Solution At the point $(1, 1)$ we have

$$f_x = 2x = 2, \quad \text{and} \quad f_y = -4y = -4.$$

Therefore

$$D_{\mathbf{u}}f = 2 * 1 - 4 * 2 = -6.$$

Now. If we look at the equation (4.14) it has the form

$$D_{\mathbf{u}}f = \mathbf{u}.(f_x, f_y).$$

There seems to be something special about the vector $f_x\ \mathbf{i} + f_y\ \mathbf{j}$. This observation motivates the following defintion

Definition 4.10 *At any point (x, y) where the first partial derivatives exist, we define the vector $\nabla f(x, y)$, also written as* **grad** *$f(x, y)$, by the expression*

$$\nabla f(x, y) = f_x(x, y)\,\mathbf{i} + f_y(x, y)\,\mathbf{j} \tag{4.15}$$

The symbol ∇, called "grad", "del" or "nabla" is a **vector** differential operator which acts on a *scalar* valued function and is given by

$$\nabla = \mathbf{i}\frac{\partial}{\partial x} \;+\; \mathbf{j}\frac{\partial}{\partial y} \tag{4.16}$$

In later courses you will also meet the *divergence* operator $\nabla.$ and the *curl* operator $\nabla\times$ which act on vector valued functions.

Example If $f(x, y) = (x^2y - 3y^2x)$ then $\nabla f = (2xy, -6xy)$.

The following result follows *immediately* from the definitions

Theorem 4.11 *The directional derivative of the function $f(x, y)$ in the direction* **u** *is given by*

$$D_{\mathbf{u}}f = \mathbf{u}.\nabla f.$$

The following theorem gives a *geometrical* interpretation of ∇f.

Theorem 4.12 *If $f(x, y)$ is differentiable at (a, b) and $\nabla f(a, b) \neq 0$, then $\nabla f(a, b)$ is a vector normal to the level curve (contour) of $f(x, y)$ at (a, b).*

Proof On a contour we have the identity $f(x, y) = c$. If we suppose that the contour is given by the curve $(x(t), y(t))$ for some parameter t, then differentiating the identity with respect to t and applying the chain rule we have that

$$f_x\, dx/dt + f_y\, dy/dt = 0.$$

As the tanjent $\mathbf{T} = (dx/dt, dy/dt)$ and also $\nabla f = (f_x, f_y)$ we see immediately that

$$\nabla f.\mathbf{T} = 0.$$

In other words the vector ∇f is orthogonal to the tanjent vector $\mathbf{T}$. Hence ∇f is in the direction of the normal to the contour.

Now we return to our example of an apple rolling down hill. The apple will roll in the path of *steepest descent* in which its height changes most rapidly. The question we now ask is, what is the direction of steepest descent?

The answer to this is given by the follwing very important theorem

Theorem 4.13 *The vector* $\nabla f(a,b)$ *is in the direction of* **the maximum increase** *of* f *at the point* (a,b)*. If* $\mathbf{u}$ *is a* unit vector *in the direction of* ∇f *then*

$$D_{\mathbf{u}} f = \|\nabla f\|. \tag{4.17}$$

Proof Let $\mathbf{u}$ be an arbitrary unit vector. Then

$$D_{\mathbf{u}} f = \mathbf{u}.\nabla f = \|\nabla f\| \cos(\theta),$$

where θ is the angle between $\mathbf{u}$ and ∇f. Now, the expression above is clearly *maximised* when $\cos(\theta) = 1$ so that $\theta = 0$. It follows that $\mathbf{u}$ must be parallel to, and in the same direction as, ∇f. For this vector $D_{\mathbf{u}} f = \|\nabla f\|$.

The vector ∇f is in the direction of the *steepest ascent* of f. Conversely the vector $-\nabla f$ is in the direction of the *steepest descent*. We can now answer the question of the mountaineers lunch. The apple will roll in the direction of the steepest descent, in other words in the direction of $-\nabla f$.

READ THIS It is hard to over emphasise the importance of the above theorem. It has implications that go far beyond having lunch on a mountain. For example, suppose that you are in charge of a company which makes a profit $f(\mathbf{x})$ which depends on a number of parameters summarised by the vector $\mathbf{x}$. If you are currently operating with the parameters set to $\mathbf{x} = a$ and you want to increase your profit, then then way to do this most rapidly is to move the parameters in the direction of ∇f. Conversely, if you want to minimise a function f then you do this most rapidly by moving in the direction of $-\nabla f$. If $\nabla f = 0$ then the function is at a maximum or at a minimum. These observations lie at the heart of the mathematical field of optimisation. They are used countless times a day by computer codes trying to improve how things operate. Indeed they are at the very heart of the ultra modern subject of *machine learning*.

We will now look at some examples to illustrate the use of these results.

Example 1 The height of the surface close to the top of a mountain is described by the function

$$f(x, y) = 1 - 2x^2 - 3y^2.$$

Question 1 At the point $(1, 1)$ what is the line of steepest descent?

Answer We have

$$\nabla f = (-4x, -6y).$$

At the point $(x, y) = (1, 1)$ it follows that $-\nabla f = (4, 6)$ so the line of steepest descent is in the direction $(4, 6)$.

Question 2 What are the contours of the surface?

Answer The contours are given by the set of curves

$$2x^2 + 3y^2 = c.$$

These are all ellipses centred on the origin.

Question 3 If we are at the point $(1, 1)$ and then move in the direction of $\mathbf{v} = (2, -1)$ do we go uphill or downhill?

Answer The rate of change of f in the direction of $\mathbf{v}$ is given by

$$D_{\mathbf{v}} f = \mathbf{v}.\nabla f = (2, -1).(-4, -6) = -2.$$

IAs this is negative we are walking down hill. This is illustrated in the figure below which shows the contours of the function $f(x, y)$, the point $(1, 1)$ and a vector in the direction $(2, -1)$.

Example 2: The height of a surface close to a pass at $(x, y) = (0, 0)$ between two mountain peaks is described by the function

$$f(x, y) = (x^2 - y^2).$$

In this case we have

$$\nabla f = (2x, -2y).$$

The contours are the set of *hyperbolae* centred on the origin, given by the curves

$$x^2 - y^2 = c.$$

Student exercise Draw the contours of the function $f(x, y)$. Calculate ∇f and hence work out the line of steepest descent at the point $(1, 2)$.

NOTE This shape is also that of a Pringles Crisp, see the figure.

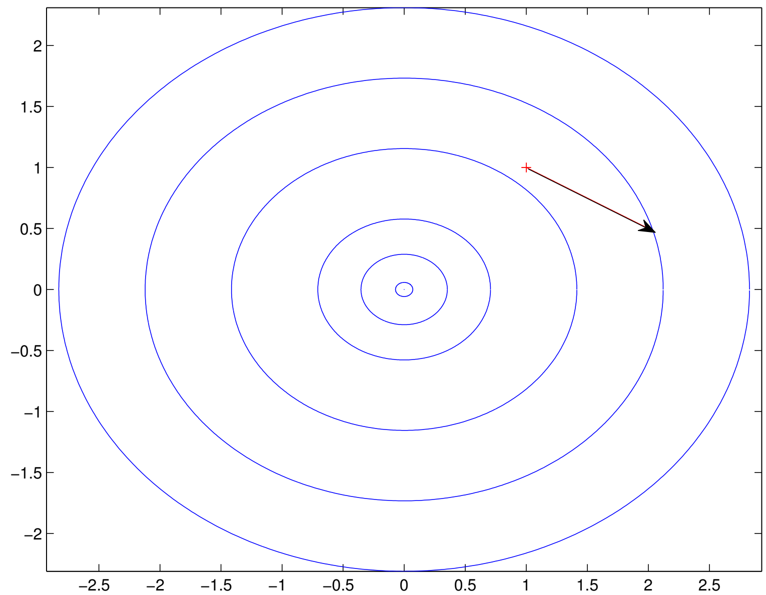


Figure 2: Contours of the function f and the line in the direction $(2, -1)$.

Figure 3: Have a pringle

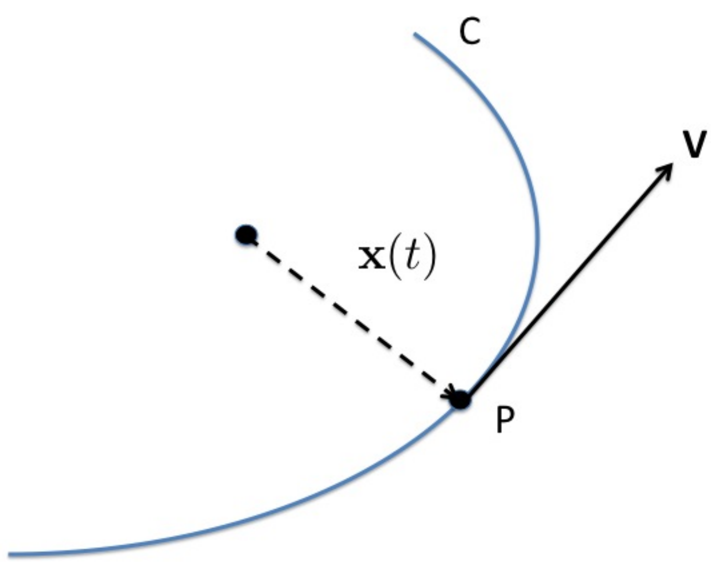
5 Kinematics

Kinematics is the study of the motion of particles along curves, or of particles moving in free space or on some surface. Here we start apply the vector algebra and vector calculus introduced in the previous sections to physical problems to see how things move. We will be looking both at particles moving in straight lines, and also spinning around in circles. To do this we not only need to represent the position of particle using a vector, but we will want to differentiate this position once (to find its velocity), twice (to find its acceleration) and three times (to find its jerk). Yes, I did say jerk, this is an important concept in road traffic collisions. We will mostly be think of the motion of a point particle. In reality moving objects have substance and they both move and spin (think of the Earth). Our intuition for the motion of a point may often not work when thinking of a massive mving and spinning object such as a football. So be warned!

5.1 Velocity and acceleration

We will start then by looking at a particle moving in free space. Consider such a particle P moving in space along a curve C. Let the position vector of P be $\mathbf{x}(t)$, where t represents time.

We will write $r(t) = \|\mathbf{x}(t)\|$ or, dropping the 't', $r = \|\mathbf{x}\|$.



Definition 5.1 *The velocity* $\mathbf{v}(t)$ *is the* rate of change of position *and is defined by:*

$$\mathbf{v}(t) = \lim_{\triangle t \to 0} \frac{\mathbf{x}(t+\triangle t) - \mathbf{x}(t)}{\triangle t} = \frac{d\mathbf{x}(t)}{dt} = \dot{\mathbf{x}}(t).$$

The velocity vector is tangential to C (remember that $\mathbf{r}'$ is the tangent vector defioned earlier.)

Notation Usually we use "$\dot{} = \frac{d}{dt}$", so we write $\dot{\mathbf{x}} = \frac{d\mathbf{x}}{dt}$, and $\ddot{\mathbf{x}} = \frac{d\dot{\mathbf{x}}}{dt} = \frac{d^2\mathbf{x}}{dt^2}$.

Definition 5.2 *The speed* v *is the magnitude of* $\mathbf{v}$*. So*

$$v = \|\mathbf{v}\| = \left\| \frac{d\mathbf{x}}{dt} \right\| \left(= \frac{ds}{dt} = \dot{s} \right)$$

where s *denotes arc-length.*

Definition 5.3 *The acceleration vector* $\mathbf{a}$ *is the rate of change of velocity:*

$$\begin{aligned} \mathbf{a} &= \frac{d\mathbf{v}}{dt} = \frac{d^2\mathbf{x}}{dt^2} \\ &= \dot{\mathbf{v}} = \ddot{\mathbf{x}} \end{aligned}$$

Useful identities

(1) $r\dot{r} = \mathbf{x}.\dot{\mathbf{x}}$

(2) $v\dot{v} = \dot{\mathbf{x}}.\ddot{\mathbf{x}} = \mathbf{v}.\dot{\mathbf{v}}$

Proof of (1)

$$r^2 = \|\mathbf{x}\|^2 = \mathbf{x}.\mathbf{x},$$

therefore, differentiating,

$$2r\dot{r} = 2\mathbf{x}.\dot{\mathbf{x}}.$$

Proof of (2) Similar to the proof of (1) - see Example sheet.

5.2 Equations of Motion and Vector Differential Equations

The equations of motion in two and three dimensions are often described by using vector valued differential equations. The usual rules for solving differential equations apply. To show this we'll look at two examples.

Example 1: Consider the motion of a charged particle moving in a magnetic field which is described by the vector differential equation:

$$\ddot{\mathbf{x}} = k\, \mathbf{x} \times \dot{\mathbf{x}}, \tag{5.1}$$

where k is a constant.

(i) Show that the speed v is constant throughout the motion.

(ii) Find a differential equation for $r(t) = \|\mathbf{x}(t)\|$, and hence solve for $r(t)$.

Solution

(i) Take the dot product of (5.1) with $\dot{\mathbf{x}}$ to give:

$$\dot{\mathbf{x}}.\ddot{\mathbf{x}} = \dot{\mathbf{x}}.(k\mathbf{x} \times \dot{\mathbf{x}}) = 0.$$

Thus, via the second "useful identity" above,

$$v\dot{v} = 0.$$

Therefore $\frac{1}{2}\frac{d}{dt}\left(v^2\right) = 0$, and so v is a constant.

(ii) Remember that $r^2 = \mathbf{x}.\mathbf{x}$. To make use of the differential equation (5.1) we need to differentiate twice to get a term containing $\ddot{\mathbf{x}}$.

$$\begin{aligned}
\frac{d^2}{dt^2}(r^2) &= \frac{d}{dt}\left(\frac{d}{dt}(\mathbf{x}.\mathbf{x})\right) \\
&= \frac{d}{dt}(2\mathbf{x}.\dot{\mathbf{x}}) \\
&= 2\dot{\mathbf{x}}.\dot{\mathbf{x}} + 2\mathbf{x}.\ddot{\mathbf{x}} \\
&= 2\|\mathbf{v}\|^2 + 2\mathbf{x}.(k\mathbf{x} \wedge \dot{\mathbf{x}}) \\
&= 2v^2.
\end{aligned}$$

But v is constant by (i), so solving this scalar ODE gives

$$r(t)^2 = v^2t^2 + At + B,$$

or

$$r(t) = \sqrt{v^2t^2 + At + B},$$

where A, B are constants of integration to be determined e.g. from the initial conditions. □

Example 2 The motion of a particle in three dimensions at the end of a perfect spring, in the absence of air resistance and friction, is described by the vector differential equation

$$\ddot{\mathbf{x}} + \omega^2\mathbf{x} = \mathbf{0}, \tag{5.2}$$

where ω is constant.

(i) Show that $\dot{\mathbf{x}} \times \mathbf{x} = \mathbf{h}$, a constant vector.

(ii) Determine $\mathbf{x}(t)$ given that $\mathbf{x}(0) = \mathbf{x}_0$ and $\dot{\mathbf{x}}(0) = \mathbf{v}_0$.

(iii) Describe the motion geometrically.

VERY IMPORTANT. READ THIS FIRST

The differential equation (5.2), called the harmonic oscilator equation, is one of the most important equations in the whole of mathematics. It describes the periodic solutions that you get to most problems in vibration and wave mechanics. It is *essential* for this and later courses that you are *completely familiar with it* and its solution. There are a number of ways to solve it:

1. *Look up the answer.* This method can be used if you have access to a formula book. However, this is not an option during your exam.

2. *Using the characteristic equation.* This is the usual method. First think of (5.2) as a scalar, rather than a vector equation. It is a very special example of a second order ordinary differential equation with constant coefficients. To solve it you pose a solution of the form

$$x(t) = e^{\lambda t}.$$

Substituting into the differential equation we have

$$\lambda^2 e^{\lambda t} + \omega^2 e^{\lambda t} = 0.$$

The unknown λ therefore satisfies the *characteristic (quadratic) equation*

$$\lambda^2 + \omega^2 = 0.$$

Thus we have $\lambda = \pm i\omega$. The general solution of (5.2) then has the form

$$x(t) = ce^{i\omega t} + de^{-i\omega t},$$

for general c and d. Using the result that $e^{i\omega t} = \cos(\omega t) + i\sin(\omega t)$ we then have that there are general constants a and b so that

$$x(t) = a\cos(\omega t) + b\sin(\omega t). \tag{5.3}$$

Now, to extend this to the vector case you simply replace the scalars a and b by vectors. This is because (5.2) is basically three scalar equations in each of the three coordinate directions. The solution of (5.2) then becomes

$$\mathbf{x}(t) = \mathbf{a}\cos(\omega t) + \mathbf{b}\sin(\omega t). \tag{5.4}$$

This funtion is *periodic* with period

$$T = \frac{2\pi}{\omega}.$$

3. *Know the answer.* Whilst the above method is completely correct and general, I do not expect you to go through this process every time that you meet the equation (5.2). Instead it is *completely satisfactory* for the purposes of this course to simply write down the answer (5.3). If you will do this you will get full marks. You should simply learn that (5.3) is the general solution to (5.2). You will meet the equation (5.2) so often in examples, that you really simply need to know what it's answer is, without having to rederive it every time.

Now we consider the solution of the original problem

(i) To show a quantity is constant it is usually easiest to prove that its derivative is zero. Now,

$$\frac{d}{dt}\left(\dot{\mathbf{x}} \times \mathbf{x}\right) = \ddot{\mathbf{x}} \times \mathbf{x} + \dot{\mathbf{x}} \times \dot{\mathbf{x}} = -\omega^2 \mathbf{x} \times \mathbf{x} + \mathbf{0} = \mathbf{0}.$$

Therefore, $\dot{\mathbf{x}} \times \mathbf{x} = \mathbf{h}$, for some constant vector $\mathbf{h}$.

(ii) From the above discussion you can now write down the result that the equation (5.2) has the general solution

$$\mathbf{x}(t) = \mathbf{a}\cos(\omega t) \ + \ \mathbf{b}\sin(\omega t), \qquad \mathbf{a}, \mathbf{b} \text{ are arbitrary constant vectors.}$$

We now need to find the values of $\mathbf{a}$ and $\mathbf{b}$. From the initial conditions

$$\begin{aligned} \mathbf{x}_0 &= \mathbf{a}\cos(0) + \mathbf{b}\sin(0) = \mathbf{a}, \\ \mathbf{v}_0 &= -\mathbf{a}\,\omega\ \sin(0) + \mathbf{b}\,\omega\cos(0) = \mathbf{b}\,\omega. \quad \therefore\ \mathbf{b} = \frac{\mathbf{v}_0}{\omega}. \end{aligned}$$

Thus the solution is

$$\mathbf{x}(t) = \mathbf{x}_0\cos(\omega t) + \frac{1}{\omega}\mathbf{v}_0\sin(\omega t)).$$

We see that the motion is *oscillatory* in time and it has period $2\pi/\omega$. This is reasonable given that it is describing the motion of a mass which is bouncing up and down at the end of a spring.

If $\mathbf{x}(t) = (x(t), y(t), z(t))$ then we can plot $x(t)$ as a function of t, or we can plot (x, y, z) together. Taking representative vectors $\mathbf{a}$ and $\mathbf{b}$ we then get the following graphs.

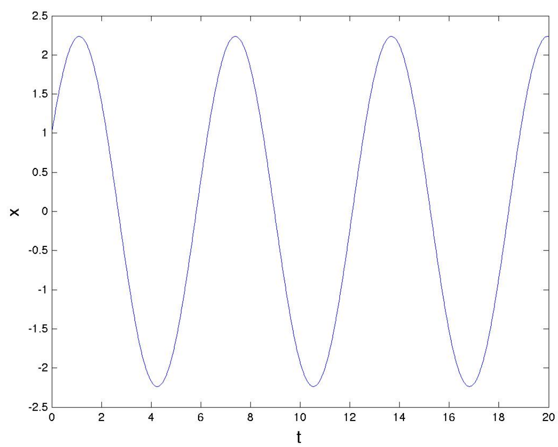


Figure 4: The motion (t, x) which is oscillatory.

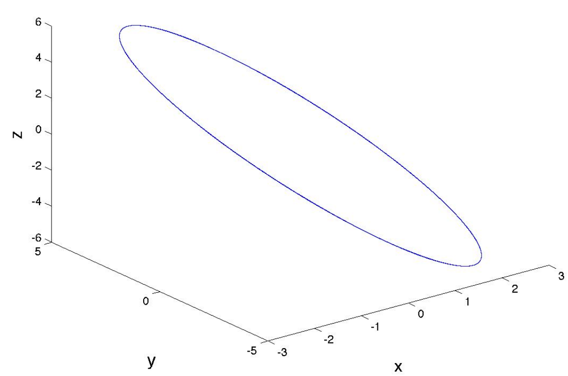


Figure 5: The motion (x, y, z) which is all in one plane.

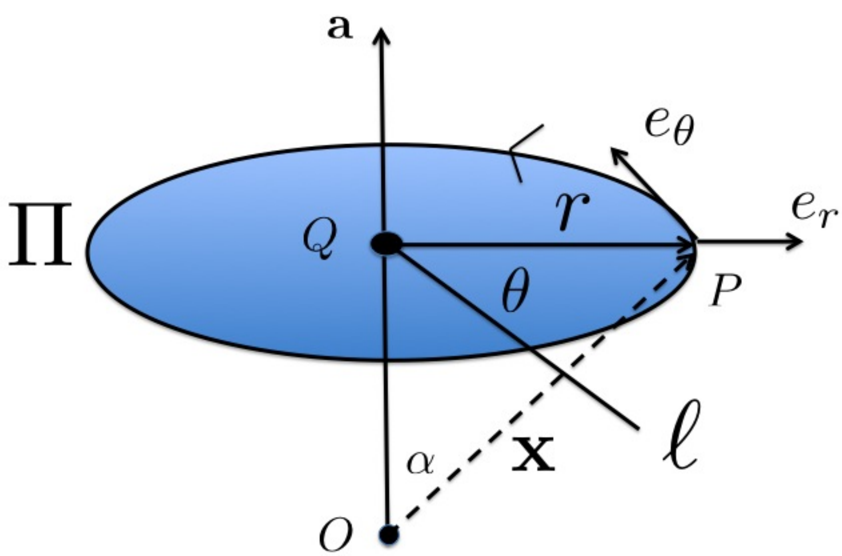
(iii) Since $\mathbf{x}_0, \mathbf{v}_0$ are constant vectors, it follows that $\mathbf{x}(t)$ remains in the plane generated by $\mathbf{x}_0$ and $\mathbf{v}_0$. □

This is a VERY important example. Motion such as this occurs everywhere in vibrating systems.

The quantity **h** is called the *angular momentum*. This is conserved in this example. In fact it is conserved in any system acted on by a (central) force **F** which is parallel to the position vector **r**.

5.3 Rotating systems

Rotating systems arise very naturally in lots of applications, for example in machinery (such as car engines), the motion of the planets, and even the motion of certain types of bacteria. If you spin around on the ice then you too are in circular motion. In this section we will take our first look at rotating systems. This will then motivate a more general approach in the next section using polar coordinates. The figure below will motivate our discussions.



When we think of rotation we naturally think of an *axis*. This will be a line through the origin O in the direction of the unit vector **a** around which things rotate. We also have an *angle* measuring the amount of the rotation. Consider the rotation of a point P about such a fixed axis. The point will move around a circle of radius r in a plane Π which is perpendicular to **a**.

Now we will introduce *polar coordinates* on the plane Π. To do this we will let Q be the origin, and we then set $r = \|\overrightarrow{QP}\|$. We then draw a fixed line ℓ on Π through Q. If we take the line QP then there an angle θ between this line and the line ℓ measured in an anti-clockwise sense.

The angle θ will change as the vector $\overrightarrow{QP}$ rotates about the axis.

We will now introduce two important vectors. We define $\mathbf{e}_r$ to be the *unit vector* in the direction $\overrightarrow{QP}$, and $\mathbf{e}_\theta$ to be the unit vector in Π perpendicular to $\mathbf{e}_r$ pointing in the direction of increasing θ. These are illustrated in the figure. We will look at these vectors in much more detail in the next section.

If $\mathbf{x} = \overrightarrow{OP}$ then it can be seen from the diagram that:

$$\begin{aligned} \mathbf{x}(t) &= \overrightarrow{OQ} + \overrightarrow{QP} \\ &= \|OQ\|\mathbf{a} + r\,\mathbf{e}_r. \end{aligned}$$

If r is constant, the point at P will move in the direction of the vector $\mathbf{e}_\theta$ and we have

$$\dot{\mathbf{x}} = r\dot{\theta}\,\mathbf{e}_\theta$$

This result will be fully justified in the next section.

We see from the diagram that the unit vectors $\mathbf{a}$, $\mathbf{e}_r$ and $\mathbf{e}_\theta$ are orthogonal and form a right-handed orthonormal system. Therefore,

$$\mathbf{e}_\theta = \mathbf{a} \times \mathbf{e}_r.$$

Hence,

$$\mathbf{a} \times \mathbf{x} = \mathbf{a} \times (\|OQ\|\mathbf{a} + r\,\mathbf{e}_r) = r\,\mathbf{e}_\theta$$

Therefore

$$\dot{\mathbf{x}} = r\dot{\theta}\,\mathbf{e}_\theta = \dot{\theta}\mathbf{a} \times \mathbf{x}.$$

If we denote

$$\omega = \dot{\theta}\,\mathbf{a} \tag{5.5}$$

it follows that

$$\dot{\mathbf{x}} = \omega \times \mathbf{x}. \tag{5.6}$$

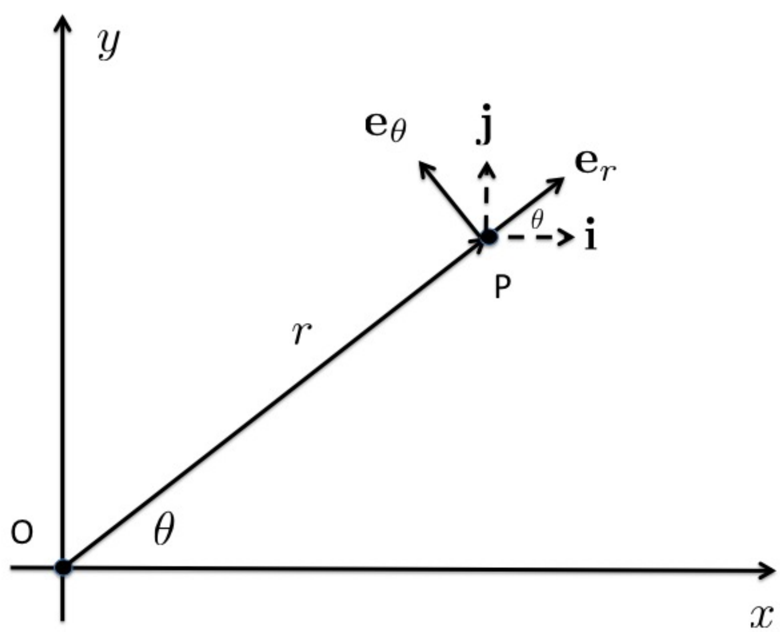
This shows that the cross product can be used to calculate velocities in rotating systems

5.4 Velocity and Acceleration in polar coordinates in 2-D

We will now generalise these ideas. So far we have always used the unit vectors $\mathbf{i}, \mathbf{j}, \mathbf{k}$ to represent vectors. However, it is often convenient to deal with "moving" or "rotating" axes, such as the vectors $\mathbf{e}_r$ and $\mathbf{e}_\theta$ described above. In this section we will look at how we can calculate velocity and acceleration in terms of these.

To do this we take the plane $\Pi = OXY$ and $P = (x, y)$ a point on the plane with position vector $\mathbf{x}(t) = (x, y)$. We now introduce polar coordinates (r, θ) for the point P, where $r := \|\mathbf{x}\|$, θ is the angle between $\mathbf{x}$ and the OX-axis as on the diagram. We have

$$x = r\cos(\theta), \quad y = r\sin(\theta).$$



The unit vectors that we considered in the last section now become

$$\mathbf{e}_r = \frac{\mathbf{x}}{r} \text{ (the radial direction),}$$

$\mathbf{e}_\theta$ orthogonal to $\mathbf{e}_r$, in the direction of increasing θ (the angular direction).

Hence $\mathbf{e}_r$, $\mathbf{e}_\theta$ are unit vectors pointing in the direction of increasing r and θ, respectively. In terms of $\mathbf{i}$ and $\mathbf{j}$, by considering the diagram, it follows from trigonometry that

$$\left.\begin{aligned} \mathbf{e}_r &= \cos(\theta)\,\mathbf{i} + \sin(\theta)\,\mathbf{j}, \\ \mathbf{e}_\theta &= -\sin(\theta)\,\mathbf{i} + \cos(\theta)\,\mathbf{j}. \end{aligned}\right\} \tag{5.7}$$

Note that $\mathbf{e}_r$, $\mathbf{e}_\theta$ are independent of r. Now, differentiate (5.7) with regard to θ:

$$\left.\begin{aligned} \frac{d\mathbf{e}_r}{d\theta} &= -\sin(\theta)\,\mathbf{i} + \cos(\theta)\,\mathbf{j} &&= \mathbf{e}_\theta, \\ \frac{d\mathbf{e}_\theta}{d\theta} &= -\cos(\theta)\mathbf{i} - \sin(\theta)\mathbf{j} &&= -\,\mathbf{e}_r. \end{aligned}\right\} \tag{5.8}$$

So, for a moving point $\mathbf{x}(t)$, the chain rule gives,

$$\begin{aligned} \dot{\mathbf{e}}_r &= \frac{d\mathbf{e}_r}{dt} = \frac{d\mathbf{e}_r}{d\theta}\frac{d\theta}{dt} = \dot{\theta}\,\mathbf{e}_\theta. \\ \dot{\mathbf{e}}_\theta &= \frac{d\mathbf{e}_\theta}{dt} = \frac{d\mathbf{e}_\theta}{d\theta}\frac{d\theta}{dt} = -\,\dot{\theta}\,\mathbf{e}_r. \end{aligned}$$

The resulting important equations, **that you should remember**, are then

$$\boxed{\begin{aligned} \dot{\mathbf{e}}_r &= \dot{\theta}\,\mathbf{e}_\theta, \\ \dot{\mathbf{e}}_\theta &= -\dot{\theta}\,\mathbf{e}_r. \end{aligned}}$$

Velocity in polar coordinates.

From the definition of $\mathbf{e}_r$, the position vector of P is given by

$$\mathbf{x}(t) = r(t)\mathbf{e}_r(\theta(t)), \tag{5.9}$$

and differentiating with respect to t gives,

$$\mathbf{v} = \dot{\mathbf{x}} = \dot{r}\mathbf{e}_r + r\dot{\mathbf{e}}_r = \dot{r}\mathbf{e}_r + r\dot{\theta}\mathbf{e}_\theta \tag{5.10}$$

So, $\dot{r}$ (called the radial velocity) and $r\dot{\theta}$ (called the angular velocity) are the components of the velocity $\mathbf{v}$ in the $\mathbf{e}_r$ and $\mathbf{e}_\theta$ directions, respectively.

The speed is given by:

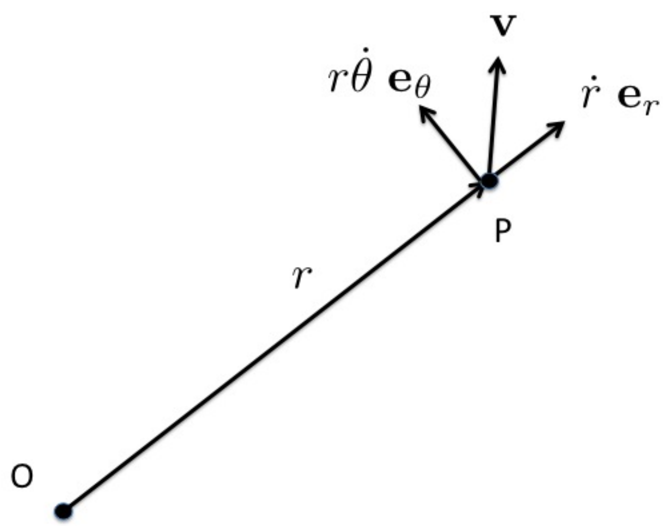
$$v = \|\mathbf{v}\| = \sqrt{\dot{r}^2 + r^2\dot{\theta}^2}\,, \tag{5.11}$$

(using Pythagoras' theorem, or directly via $v^2 = \mathbf{v}\cdot\mathbf{v}$ etc.).

Acceleration in polar coordinates

Differentiating the expression for the velocity, and using again the identity in the box we have,

$$\begin{aligned} \ddot{\mathbf{x}} &= \frac{d}{dt}(\dot{\mathbf{x}}) = \frac{d}{dt}\left(\dot{r}\mathbf{e}_r + r\dot{\theta}\mathbf{e}_\theta\right) \\ &= \ddot{r}\mathbf{e}_r + \dot{r}\dot{\mathbf{e}}_r + \dot{r}\dot{\theta}\mathbf{e}_\theta + r\ddot{\theta}\mathbf{e}_\theta + r\dot{\theta}\dot{\mathbf{e}}_\theta \\ &= \ddot{r}\mathbf{e}_r + \dot{r}\dot{\theta}\mathbf{e}_\theta + \dot{r}\dot{\theta}\mathbf{e}_\theta + r\ddot{\theta}\mathbf{e}_\theta - r\,\dot{\theta}^2\mathbf{e}_r. \end{aligned}$$



Thus we have the very (very) important expression **which you must learn**

$$\boxed{\ddot{\mathbf{x}} = \left(\ddot{r} - r\dot{\theta}^2\right)\mathbf{e}_r + \left(r\ddot{\theta} + 2\dot{r}\dot{\theta}\right)\mathbf{e}_\theta}$$

Hence, $(\ddot{r} - r\dot{\theta}^2)$ and $(r\ddot{\theta} + 2\dot{r}\dot{\theta})$ are the acceleration components in the $\mathbf{e}_r$ and $\mathbf{e}_\theta$ directions, respectively.

Note that for the angular component of the acceleration we have

$$r\ddot{\theta} + 2\dot{r}\dot{\theta} = \frac{1}{r}\frac{d}{dt}\left(r^2\dot{\theta}\right).$$

Definition 5.4 *$\dot{\theta}$ is the rate of change of the angle θ, often called "angular speed" ω. So $\dot{\theta} = \omega$.*

Funky stuff When we are in a rotating frame we pick up the two additional acceleration components given by $r\dot{\theta}^2$ and $r\ddot{\theta} + 2\dot{r}\dot{\theta}$. These are called the centrifugal and coriolis terms respectively. In such a rotating frame they play an identical role to forces and it was thinking about this which helped Einstein to formulate his General Theory of Relativity. The Earth is such a rotating frame. The resulting Coriolis Force acting on the atmosphere is a prime driver for our weather systems.

Example 1: A particle moves on a circle of (constant) radius R with a constant angular acceleration α (i.e. $\ddot{\theta} = \alpha$). If the particle starts from rest, show that after time t, (a) its angular speed, ω, is given by $\omega = \alpha t$, and (b) the arc length covered is $\frac{R}{2}\omega t$.

Solution

$$
\begin{aligned}
\dot{\mathbf{x}} &= \dot{r}\mathbf{e}_r + r\dot{\theta}\,\mathbf{e}_\theta = R\dot{\theta}\mathbf{e}_\theta \quad (\text{since } \dot{r} = 0) \\
\ddot{\mathbf{x}} &= (\ddot{r} - r\dot{\theta}^2)\mathbf{e}_r + (2\dot{r}\dot{\theta} + r\ddot{\theta})\mathbf{e}_\theta = -R\dot{\theta}^2\mathbf{e}_r + R\ddot{\theta}\mathbf{e}_\theta .
\end{aligned}
$$

(a) Angular acceleration

$$
\begin{aligned}
\ddot{\theta} &= \alpha \\
\text{hence} \quad \dot{\theta} &= \alpha t + A.
\end{aligned}
$$

But $A = 0$ since $\dot{\theta}(0) = 0$ (that is, the particle starts from rest).

$$\text{hence} \quad \omega = \dot{\theta} = \alpha t \ .$$

(b)

$$
\begin{aligned}
& \theta = \frac{1}{2}\alpha t^2 \ (\text{we can take } \theta(0) = 0) \\
\therefore \quad & R\theta = \frac{R}{2}\alpha t^2 \\
& \text{hence the arclength travelled} = R\theta = \frac{R}{2}\alpha t^2 = \frac{R}{2}\omega t.
\end{aligned}
$$

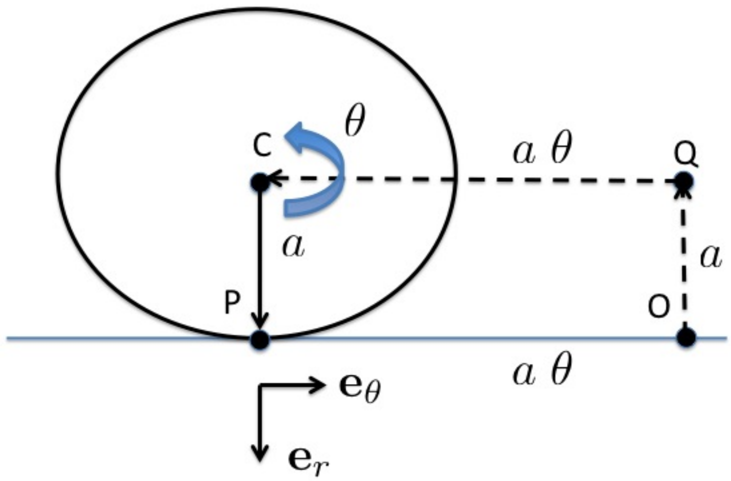
□

Next we give an example to show how the fixed $\mathbf{i}, \mathbf{j}$ frame and the moving $\mathbf{e}_r, \mathbf{e}_\theta$ frame can be used together.

Example 2:

A wheel of (constant) radius a and centre C rolls without slipping at a constant angular speed ω along a horizontal plane which contains the origin O (which is the starting point of the rolling motion). At $t = 0$ the spoke CP is vertically downwards, with the point P in contact with the plane. Find the acceleration of the point P.

Solution Suppose that the wheel has rotated an angle θ. The no-slipping condition implies that the arc length from O to P is given by $a\,\theta = a\omega\, t$. If Q is a point a distance a vertically above



the origin O then, after a time t, the position vector of P, namely $\mathbf{x}(t)$, is given by

$$\begin{aligned}\mathbf{x}(t) &= \overrightarrow{OP} = \overrightarrow{OQ} + \overrightarrow{QC} + \overrightarrow{CP} \\ &= a\mathbf{j} - a\,\theta\mathbf{i} + a\mathbf{e}_r.\end{aligned}$$

Thus, the velocity of P is given by (using $\dot\theta = \omega$ and the the fact that $\dot a = 0$)

$$\begin{aligned}\mathbf{v} &= \dot{\mathbf{x}} = -\,a\dot\theta\,\mathbf{i} + a\dot{\mathbf{e}}_r = -\,a\omega\mathbf{i} + a\dot\theta\mathbf{e}_\theta \\ &= a\omega(\mathbf{e}_\theta - \mathbf{i}).\end{aligned}$$

Note that as $\mathbf{e}_\theta = \mathbf{i}$ the velocity of P is *zero*. This is the no-slip condition in action. The bottom of a wheel has zero velocity, this is why wheels and tyres work.

Hence, the acceleration of P is given by,

$$\mathbf{a} = \ddot{\mathbf{x}} = a\omega\dot{\mathbf{e}}_\theta = -\,a\omega\dot\theta\mathbf{e}_r = -\,a\omega^2\mathbf{e}_r,$$

and so $\mathbf{a}$ acts only in the radial direction. □

5.5 Some curves in polar coordinates (r, θ)

Having defined polar coordinates, it is interesting to look at some well-known curves in 2-D expressed in terms of these. We will meet these again in the dynamics part of this course.

Example 1 The curve given by $r = l,\ l > 0$, is a circle of radius l, centred on $(0, 0)$.

Example 2 The curve given by

$$r = \frac{l}{1 + e\cos(\theta)},$$

with e, l constants, arises frequently in celestial mechanics. It is a conic section and was discovered by the Greek mathematician Appolonius about 3000 years ago. If $l > 0$, and $0 < e < 1$, it is an ellipse, with focus at the origin O. See the figure below.

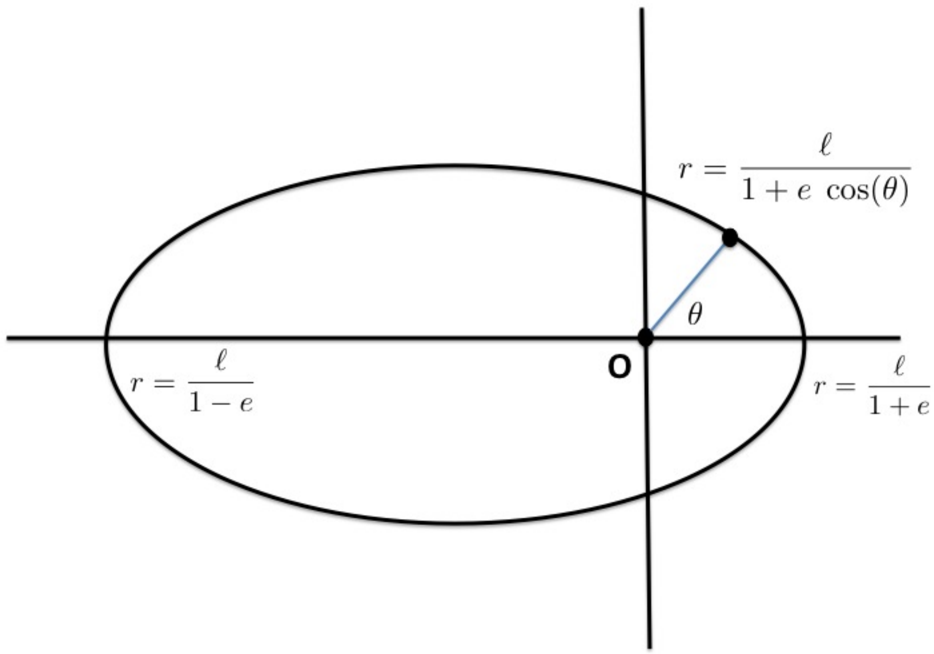


Figure 6: An elliptical orbit with the origin O at the focus.

Since we need to know about elliptical orbits of planets later on, we give a derivation in terms of Cartesian coordinates.

Derivation: $l = r + e\underbrace{r\cos(\theta)}_{=x}$

$\therefore \quad l - ex = r$

$\therefore \quad (l - ex)^2 = r^2 = x^2 + y^2$

$\therefore \quad (1 - e^2)x^2 + 2lex + y^2 = l^2$

$$\therefore \quad (1 - e^2)\left(x + \frac{le}{1 - e^2}\right)^2 + y^2 = \frac{l^2}{1 - e^2}$$

$$\therefore \qquad \frac{\left(x+\frac{le}{l-e^2}\right)^2}{\left(\frac{l}{1-e^2}\right)^2}+\frac{y^2}{\left(\frac{l}{\sqrt{1-e^2}}\right)^2}=1.$$

This is the usual equation of an ellipse in Cartesian coordinates where,

$$\frac{\hat{x}^2}{a^2}+\frac{\hat{y}^2}{b^2}=1,$$

$$\hat{x} = x+\frac{le}{1-e^2}, \quad \hat{y}=y,$$

$$a = \frac{l}{1-e^2}, \quad b=\frac{l}{\sqrt{1-e^2}}.$$

Here, e is called the "eccentricity" and l is the "semi-latus rectum". If $e=0$ then we have a circle. The ellipse is the orbit of a planet going around the Sun. This iss Kepler's first law of planetary motion and was one of the starting points of modern science. Later on we will prove this result for an inverse square law of gravity.

Example 3

$$r=\frac{l}{1+e\cos\theta}, \quad e>1,$$

is the equation of a hyperbola. This curve is the orbit of a fast moving comet passing by the Sun, or of an Alpha particle passing by the nucleus of an atom. (This was discovered by Ernest Rutherford in the early part of the 20th Century.)

The hyperbola has *asymptotes* which are the lines at the two angles for which

$$\cos(\theta)=-\frac{1}{e}.$$

Hyperbolas play an important role in sending satellites to distant parts of the Solar system, when they pass by planets using hyperbolic 'sling-shot' orbits.

Example 4

$$r=\frac{l}{1+\cos(\theta)},$$

is the equation of a parabola. The parabola is the cross-sectional shape of a satellite TV dish and also of a car headlight.

5.6 Kepler's laws of motion

One of the greatest breakthroughs in the history of science came from the subject of Kinematics! The ancients thought that the Sun and the planets went round the Earth. As the paths of these

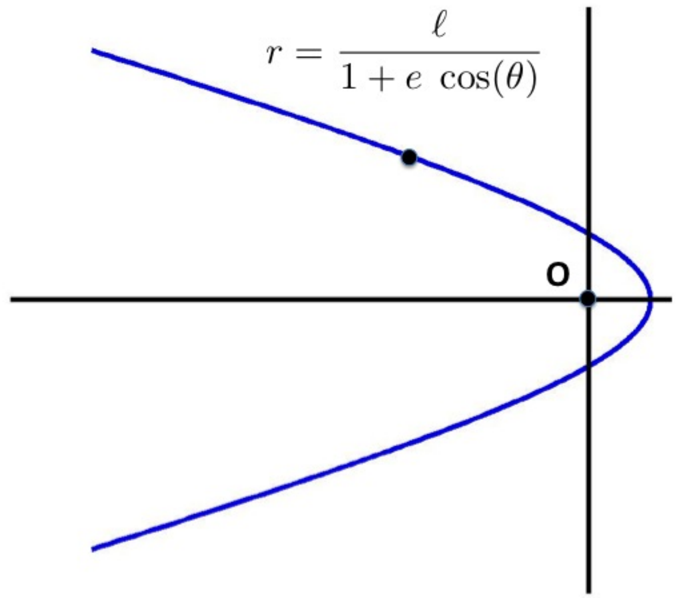


Figure 7: A hyperbolic orbit, with the origin O at the focus.

heavenly bodies take complicated paths when viewed from the Earth the original Ptolmeic view of the solar system had planets moving on small circles (epicycles) on larger circles around the Earth. This was a very complicated was of describing the motion. A dramatic simplification (of the science of not of the philosophy) was then proposed by Copernicus who suggested that instead to the planets and Sun going round the Earth, that instead the Earth and the planets went around the Sun. The orbits of all of these would be circles with the Sun at the centre. This caused a huge controversy as it no longer placed the Earth at the centre of the universe. Galileo was famously put under house arrest for supporting the Copernican view of things. However there still remained a problem. The Copernican view of the solar system did not agree very well with the experimental measurements. In particular in the 16^{th} century, the Danish astronomer Tycho Brahe made detailed observations of the motion of the planets and showed that they did not agree with the predications of the Copernican model. Brahe in fact advocated a different model in which the planets went around the Sun, and the Sun then went around the Earth. The problem was that Copernicus had thought that the planets had to go around the Sun in circles as these were the 'perfect' orbit. Fortunately there were other possible curves which were almost as perfect. These were ellipses, which had be known about for at least 2000 years. Brahe's student, Johannes Kepler (illustrated) analysed his data and realised that if he replaced circles by ellipses

Figure 8: Johannes Kepler

then everything worked perfectly, with brilliant fit between theory and observations. This was a huge breakthrough in science and the dawn of the modern scientific revolution! In this sense humanity was incredibly lucky in that the planets went around the Sun along curves which had already been studied by mathematicians. It was a classic example of a mathematical discovery having to wait 2000 years before it found an application. Politicians who fund mathematics should take not of this. In 1609 (i.e. before the invention of calculus) Kepler published his **three laws** which are valid for all planets in the solar system and which fitted the observations perfectly. These laws then led Newton directly to the invention of calculus. The three laws of the *kinematics* of a planet are as follows:

K1: Each planet moves in an ellipse with the Sun at a *focus.*

K2: Areas swept out by the radius vector from the Sun to a planet in equal times are equal.

K3: The square of the period of revolution T of a planet is proportional to the cube of the semi-major axis,

$$T^2 = ca^3, \tag{5.12}$$

where the constant of proportionality c is *independent* of the planet in the Solar System.

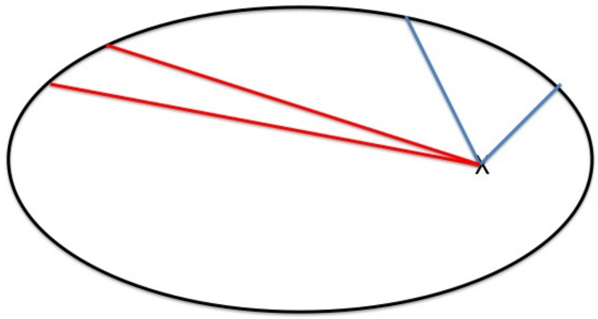


Figure 9: An illustration of K1 (with the planet going round the Sun which as the focus of an ellipse) and K2 (which is that the area swept out between the red lines is the same as that between the blue lines, in the same period of time).

Some implications of Kepler's laws: We now look at some of the kinematic implications of these laws. Later in the dynamics part of this course we will derive them directly from Newton's law of gravitation.

Take polar axes in the plane of the orbit with the Sun at the origin.

K1: As we have already seen, in polar coordinates an ellipse is given by

$$r = \frac{l}{1 + e\cos(\theta)}, \quad (0 < e < 1), \tag{5.13}$$

where the constants l and e take different values for different planets. Recall from before that the semi-major axis, a, and the semi-minor axis, b, are given by

$$a = \frac{l}{1 - e^2}, \qquad b = \frac{l}{\sqrt{1 - e^2}}. \tag{5.14}$$

K2: Suppose in a small time increment δt, the angle θ increases by an amount $\delta\theta$ (see the figure following). Let $A(t)$ be the area swept by the radius vector from time t_0 to t. The area swept out by the radius vector in time δt is

$$\delta A = \frac{1}{2} r^2 \delta\theta + O\left((\delta\theta)^2\right)$$

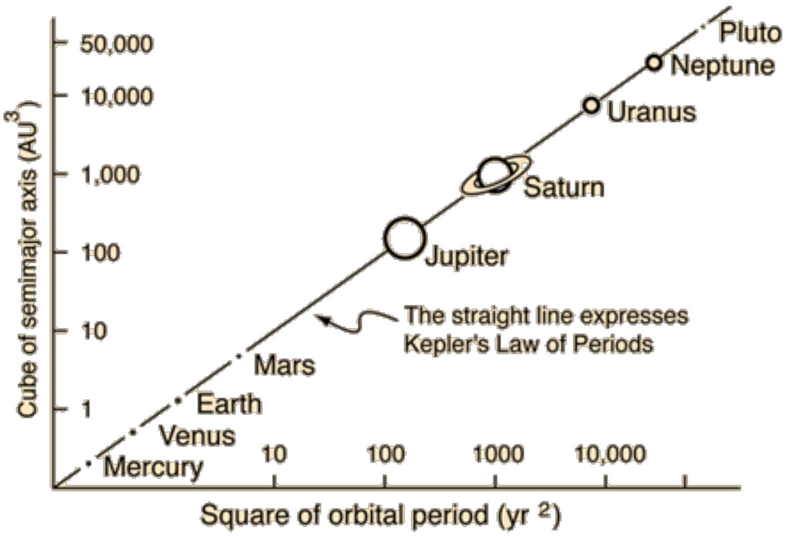


Figure 10: The period of a planet squared compared against the semi-major axis cubed for the planets in the Solar System

.

(i.e., to first order with respect to small $\delta\theta$, δA is approximated by $\frac{1}{2}r^2\delta\theta$). Divide by δt and let $\delta t \to 0$, to give

$$\frac{dA}{dt} = \lim_{\delta t \to 0} \frac{\delta A}{\delta t} = \lim_{\delta t \to 0} \frac{1}{2} r^2 \frac{\delta\theta}{\delta t} = \frac{1}{2} r^2 \dot{\theta}.$$

K2 implies that the rate of change of the swept out area $\frac{dA}{dt}$ is constant, so dropping the factor $\frac{1}{2}$ we have

$$r^2\dot{\theta} = h, \text{ a constant .} \tag{5.15}$$

Note that the constant h takes different values for different planets.

K3: First, note that the period of revolution T is the time taken to go from $\theta = 0$ to $\theta = 2\pi$, i.e. for the swept out area to coincide with the area of the ellipse. Since, from the above,

$$\frac{dA}{dt} = \frac{1}{2} h$$

is constant, by the fundamental theorem of calculus,

$$A(T) - A(0) = \frac{1}{2} h T = \text{area of the ellipse.}$$

The area of an ellipse with semi-axes a and b is known to equal πab.

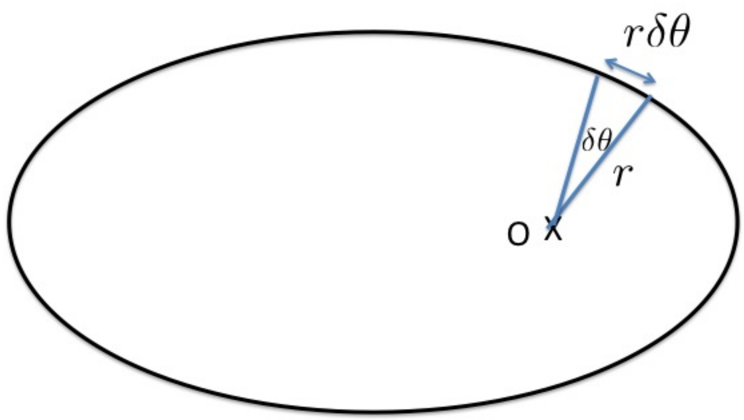


Figure 11: The implications of K2. The area swept out by the radius vector is (to leading order) given by $\delta A = r^2\delta\theta/2$.

[1]

Hence

$$T = \frac{2\pi ab}{h}.$$

[1]* As a non-examinable exercise, you can also evaluate the area $\mathcal{A}$ of an ellipse in polar coordinates, as follows:

$$\mathcal{A} = \int_{-\pi}^{\pi} \frac{1}{2} r^2 d\theta = \frac{l^2}{2} \int_{-\pi}^{\pi} \frac{1}{(1 + e\cos\theta)^2} d\theta$$

Now use the substitution $s = \tan\theta/2$, so $d\theta = \dfrac{2ds}{1+s^2}$, and $\cos\theta = \dfrac{1-s^2}{1+s^2}$, to obtain

$$\mathcal{A} = l^2 \int_{-\infty}^{+\infty} \frac{(1+s^2)ds}{((1+e)+(1-e)s^2)^2} = l^2 \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \frac{\left(1 + \frac{1+e}{1-e}\tan^2\phi\right)}{(1+e)^2\sec^4\phi} \frac{(1+e)^{\frac{1}{2}}}{(1-e)^{\frac{1}{2}}} \sec^2\phi \, d\phi$$

where we have used the further substitution $(1-e)^{\frac{1}{2}} s = (1+e)^{\frac{1}{2}} \tan\phi$ to get the second integral. After simplifying the integrand,

$$\mathcal{A} = \frac{l^2}{(1-e^2)^{\frac{3}{2}}} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \left((1-e)\cos^2\phi + (1+e)\sin^2\phi\right) d\phi = \frac{\pi l^2}{(1-e^2)^{\frac{3}{2}}} = \pi\, a\, b.$$

□

Finally, using (5.14),

$$T^2 = \frac{4\pi^2 a^2 b^2}{h^2} = \frac{4\pi^2 l^4}{h^2(1-e^2)^3} = \left(\frac{4\pi^2 l}{h^2}\right) a^3. \tag{5.16}$$

K3 states that $T^2 = c\,a^3$, with the constant of proportionality c being independent of the planet. Thus $\dfrac{4\pi^2 l}{h^2}$ is a constant, which is the same for all the planets. Therefore K3 implies that

$$\mu := \frac{h^2}{l}, \tag{5.17}$$

is independent of the particular planet under consideration.

In Chapter 8 we will show why this follows from Newton's inverse square law of gravitation.

6 Dynamics

In the last section we looked at how bodies *move* and how vector calculus allows us to study this motion. In general such bodies move under the action of *forces*. It is understanding this which allows us to predict their motions, and introduces us to the great subject of *dynamics*.

6.1 Forces and Newton's Laws

We start our discussion of dynamics with some definitions and physical concepts.

(a) An **inertial** frame of reference (or coordinate system) is one which is **fixed** or is moving with uniform velocity compared to a "universal" fixed frame of reference. We shall assume that the sun (and sometimes the earth too) forms an inertial system. (This is a reasonable assumption as long as speeds aren't too large.) Newton's Laws hold under the assumption that all measurements or observations are made with respect to an inertial system.

(b) Any body (lump of matter) has associated with it a numerical value, called its **mass**. It is an intrinsic property of the body, and is a measure of its resistance to change in its motion (or a measure of its **inertia**). We'll refer to all bodies whose size does not matter in the considered context and which can therefore be regarded as points as **particles**, for example, tennis balls, apples, satellites, planets....

(c) The **linear momentum** of a particle is defined to be the vector $\mathbf{p} = m\mathbf{v} = m\dot{\mathbf{x}}$, where m is the mass. The linear momentum is a measure of how hard it is to stop the body.

(d) **Forces** cause particles to move (as experience suggests). Experimental evidence suggests that there are fundamental and non-fundamental forces (see Feynmann, 'Lectures on Physics I' for a physical discussion). **Fundamental forces** are gravitational (long range), electromagnetic (long range), and nuclear (short range). **Non-fundamental forces** include molecular, elastic, and frictional forces. We will be mostly considering in this course the gravitational and frictional forces.

Newton's laws of motion:
In an inertial frame (which is thereby postulated to exist, but see Einstein's General Theory of Relativity for an extension of this idea):

N1: Any freely moving particle moves with uniform velocity.

N2: A particle of mass m subject to a total force $\mathbf{F}$ undergoes an acceleration $\mathbf{a}$ given by

$$\mathbf{F} = m\,\ddot{\mathbf{x}} = m\,\mathbf{a}.$$

N3: Action and reaction are equal in magnitude and opposite in direction. So if $\mathbf{F}_{12}$ denotes the force exerted by particle 2 on particle 1, and $\mathbf{F}_{21}$ denotes the force exerted by particle 1 on particle 2, then

$$\mathbf{F}_{12} = -\mathbf{F}_{21}.$$

6.2 Examples of forces

There are many different types of force. Here is a list of the most common ones, which we will all consider in this course.

1. **Constant forces** These are the simplest types of force. We have

$$\mathbf{F} = \mathbf{C}$$

   for some constant vector $\mathbf{C}$. An example of this (as we shall see) is the gravitational force of the Earth on a particle close to the Earth's surface.

2. **Frictional and viscous forces** These always act to *oppose* the motion of a particle and generally take the form

$$\mathbf{F} = -k(\mathbf{x}, \dot{\mathbf{x}})\,\dot{\mathbf{x}}.$$

   If the force arises from having the particle in contact with a surface, then it is usually due to *friction* (which can be a static or a dynamic force). If it arises from the motion of the particle through a fluid (such as air or water) then this force is due to *viscosity* and is often also called *drag*.

3. **Central forces** These are forces which act in the same direction as $\mathbf{x}$. Examples of these are the gravitational force of the Sun on the Earth, or the force of an atomic nucleus on an electron.

6.3 Inverse square law forces

These type of forces are very common in three dimensional spaces. We consider a body at position $\mathbf{x}_1$ acting on another body at position $\mathbf{x}_2$. The resulting force $\mathbf{F}$ between them is then given by

$$\mathbf{F} = K\frac{\mathbf{x}_1 - \mathbf{x}_2}{\|\mathbf{x}_1 - \mathbf{x}_2\|^3}. \tag{6.1}$$

If one of the bodies if fixed at the origin (for example the Sun) then this is a central force acting on the other body.

A very important example of such a force is given by Newton's law ofGgravitation. In this case we consider the two bodies to have mass m_1 and m_2. The resulting force is then

$$\mathbf{F} = -Gm_1m_2\frac{\mathbf{x}_1 - \mathbf{x}_2}{\|\mathbf{x}_1 - \mathbf{x}_2\|^3}, \tag{6.2}$$

where G is the Gravitational Constant.

Equally if the two bodies are charged with charges e_1 and e_2 then the Coulomb electrostatic force between them is

$$\mathbf{F} = -ke_1e_2\frac{\mathbf{x}_1 - \mathbf{x}_2}{\|\mathbf{x}_1 - \mathbf{x}_2\|^3}. \tag{6.3}$$

where k is the Coulomb constant.

If $r = \|\mathbf{x}_1 - \mathbf{x}_2\|$ then it follows from (6.3) that

$$\|\mathbf{F}\| = \frac{|K|}{r^2}. \tag{6.4}$$

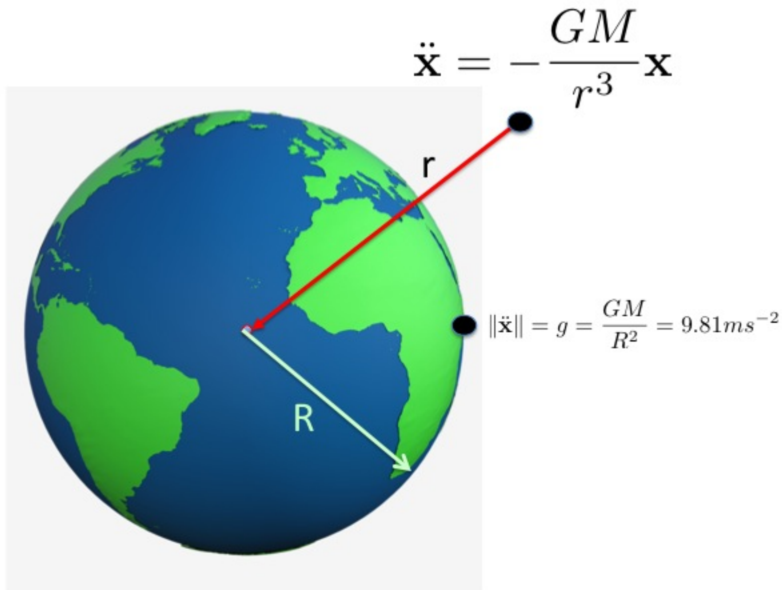
For this reason, these forces are called *inverse square-law forces.* Newton derived his inverse square law for Gravity by considering Kepler's laws of motion. It took him 20 years and he had to invent calculus to do it. However, we now know that inverse square laws arise very naturally in three dimensional space. Very loosely speaking it is because the strength of the forca at the centre has to be spread over a the surface of a sphere of radius r. This surface has area $4\pi r^2$. So the force reduces by a factor of $1/r^2$ as we move away from the centre. This rather vague statement will be proved in the second year course on *Vectors and PDEs*.

6.4 Particle motion in the neighbourhood of the Earth

Close to the surface of the Earth, the gravitational force becomes (close to) constant.

Consider a two-particle system with M (Earth) $\gg m$ (particle, for example, a satellite). Then we may regard the Earth as fixed and the equation of motion of the particle is

$$m\ddot{\mathbf{x}} = \mathbf{F} = -\frac{GMm}{r^3}\mathbf{x}, \quad r = \|\mathbf{x}\|.$$



$$\therefore \ddot{\mathbf{x}} = -\frac{GM}{r^3}\mathbf{x}, \quad \|\ddot{\mathbf{x}}\| = \frac{GM}{r^2}.$$

Particle motion very close to the Earth

Assume now that the earth is a sphere of radius R, and that the particle is sufficiently close to the surface of the Earth so we may take $\|\mathbf{x}\| = R$. So, to a good approximation,

$$\mathbf{F} = m\ddot{\mathbf{x}} = -\frac{GMm}{R^3}\mathbf{x} = -\frac{GMm}{R^2}\hat{\mathbf{x}},$$

where $\hat{\mathbf{x}} = \frac{\mathbf{x}}{\|\mathbf{x}\|} = \frac{\mathbf{x}}{R}$, is the unit vector pointing 'away' from the centre of the earth. Introduce the scalar $g := \frac{GM}{R^2}$, and the vector $\mathbf{g} := -g\hat{\mathbf{x}}$.

Funky stuff Note that the mass of the particle m cancels out in this equation and that $\mathbf{g}$ is *independent of* m. So the *acceleration* of a mass m in a gravitational field does not depend on m. This was one of the starting points of Einstein's General Theory of Relativity.

For a particle very close to the earth, the acceleration is given by

$$\ddot{\mathbf{x}} = \mathbf{g} \quad \text{('free fall' acceleration).}$$

Definition 6.1 *The **weight** of a particle close to the surface of the earth is defined to be* $\mathbf{F} = m\mathbf{g}$.

Important Note: The above discussion regards the earth as a particle. This seems reasonable for planetary motion (that is, when $\|\mathbf{x}\| \gg R$). However, Newton showed that even near the surface, the gravitational attraction is the same as if all Earth's mass were concentrated at its centre, assuming the Earth has a spherically symmetric mass distribution.

7 Motion under a constant gravitational force: projectiles, pendulums and friction

In this section we will use the Newton's second law to derive the differential equations of motion for projectiles, pendulums and motion under the action of friction. We first analyse the motion of projectiles in a non-resisting medium in §7.1, and then we consider the more realistic case of motion in a resisting medium in §7.2. We will then look at the rigid pendulum. Finally we will look at problems with friction.

7.1 Motion of a projectile under gravity in non-resisting medium.

Consider the motion of a particle very close to the Earth, subject to the (constant) Earth's gravitational attraction $m\mathbf{g}$ and assuming no air resistance. From Newton's Second law we have

$$m\ddot{\mathbf{x}} = m\mathbf{g} \quad \text{or} \quad \ddot{\mathbf{x}} = \mathbf{g}. \tag{7.1}$$

Thus

$$\ddot{\mathbf{x}} = \mathbf{g}$$

We can solve (7.1) by integrating twice to give

$$\begin{aligned} \mathbf{x}(t) &= \mathbf{x}_0 + \mathbf{v}_0 t + \frac{1}{2}\mathbf{g}t^2, \qquad (7.2) \\ \text{where} \quad \mathbf{x}_0 &= \mathbf{x}(0), \;\; \mathbf{v}_0 = \dot{\mathbf{x}}(0). \end{aligned}$$

Example 1: Assume that at time $t = 0$, a particle (projectile) is launched from the surface of the earth at an angle α and with speed v_0. Find the horizontal distance travelled and the path of the particle in (x, y) coordinates (ie. the trajectory).

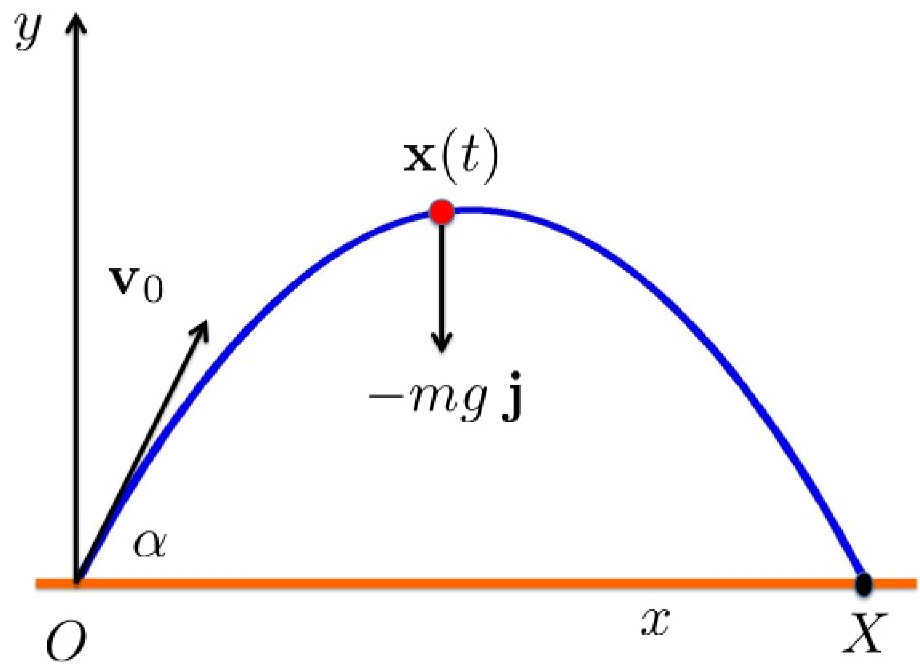
Solution: To find solution to this problem we will follow general strategy which is applicable to most of the problems from now on. This is a general approach to the great subject of *mathematical modelling*. For the purposes of this course it will consist of the following **five steps**

Step 1: *Draw a careful diagram, showing the origin, the axes, the particle and the forces acting on it, and any initial data that we have*

For this problem the particle has the position vector

$$\mathbf{x}(t) = x(t)\,\mathbf{i} + y(t)\,\mathbf{j}.$$

The diagram of its motion then has the following form:



Step 2 *Resolve the forces and apply Newton's second law linking forces and accelerations.*

Only the gravity force $m\mathbf{g}$ applies where $\mathbf{g} = -g\mathbf{j}$.
Newton's second law them implies that the particle satisfies the Ordinary Differential Equation (ODE) (7.1).

Step 3: *Determine the Initial conditions, I.C.*

As the particle starts from the origin we have $\mathbf{x}_0 = 0$.

Considering the diagram and using a bit of trigonometry we have

$$\dot{\mathbf{x}}(0) \;=\; \mathbf{v}_0 \;=\; v_0\cos(\alpha)\,\mathbf{i} + v_0\sin(\alpha)\,\mathbf{j}.$$

Step 4 *Solve the ODE together with the I.C.*

The solution is given by (7.2).

Step 5 *Analyse the solution to answer the question and find the motion of the particle*

For $\mathbf{x}(t) = x(t)\,\mathbf{i} + y(t)\,\mathbf{j}$, from (7.2) we have, in components,

$$\mathbf{i} \text{ component: } x(t) = (v_0 \cos(\alpha))t, \tag{7.3}$$

and,

$$\mathbf{j} \text{ component: } y(t) = (v_0 \sin(\alpha))t - \frac{1}{2}\,g\,t^2. \tag{7.4}$$

To eliminate t note that from (7.3), $t = \dfrac{x}{v_0 \cos(\alpha)}$, so that,

$$y = v_0 \sin(\alpha)\frac{x}{v_0 \cos(\alpha)} - \frac{1}{2}g\frac{x^2}{v_0^2 \cos^2(\alpha)},$$

and so we have the important result

$$\boxed{y = x\ \tan(\alpha) - \frac{g}{2\,v_0^2 \cos^2(\alpha)}\ x^2.} \tag{7.5}$$

This curve is a *parabola* in the (x, y) plane. This important result was discovered by Galileo in the 16th Century.

To find the distance travelled, set $y = 0$. Thus, $x = 0$ (the starting point) or

$$x = X = \frac{2v_0^2 \sin(\alpha)\cos(\alpha)}{g} = \frac{v_0^2 \sin(2\alpha)}{g}.$$

Definition 7.1 *X, the horizontal distance travelled, is called the* ***range****.*

The maximum value for X is $\dfrac{v_0^2}{g}$ and this occurs when $\sin(2\alpha) = 1$, that is, $\alpha = \dfrac{\pi}{4}$.

Example 2:

A rugby ball is kicked at an angle α to the horizontal from a point at ground level with a speed v_0. It is aimed at a goal which is a distance d away from the kicker. The height of the crossbar on the goal is h.

Figure 12: Jonny Wilkinson winning the 2003 Rugby World Cup for England with the drop goal of all time.

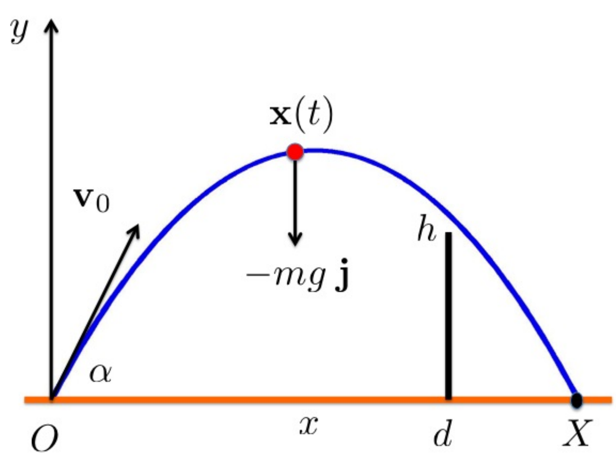
(i) Find the angle α which maximises the ball's height y when it gets to the goal

(ii) Show that the ball will clear the crossbar (and the player will score a goal) if

$$v_0^2 > g\left\{(d^2+h^2)^{\frac{1}{2}}+h\right\}.$$

Solution:

Step 1 The diagram for this situation is given below.



Steps 2–4 These are the same as in the previous example.

Step 5 From (7.5), at $x = d$,

$$y(d) \;=\; d\tan(\alpha) - \frac{gd^2}{2\,v_0^2\cos^2(\alpha)}.$$

To find the maximum value for $y(d)$, we seek α such that $\frac{dy}{d\alpha}(d) = 0$. Therefore

$$\frac{dy}{d\alpha} = d\sec^2(\alpha) \;-\; \frac{gd^2}{v_0^2}\,\frac{\sin(\alpha)}{\cos^3(\alpha)} = 0.$$

Thus,

$$\sec^2(\alpha)\left(d - \frac{gd^2}{v_0^2}\tan(\alpha)\right) = 0,$$

which gives

$$\tan(\alpha) = \frac{v_0^2}{gd}.$$

Therefore

$$\frac{1}{\cos^2(\alpha)} = \tan^2(\alpha) + 1 \;=\; \frac{v_0^4}{g^2d^2} \;+\; 1.$$

So, from the expressions for $y(d)$ and $\tan(\alpha)$ etc,

$$\begin{aligned} y_{\max} \;&=\; d\frac{v_0^2}{gd} \;-\; \frac{gd^2}{2v_0^2}\left(\frac{v_0^4}{g^2d^2} \;+\; 1\right) \\ &=\; \frac{v_0^2}{g} - \frac{v_0^2}{2g} - \frac{gd^2}{2v_0^2} \\ &=\; \frac{v_0^2}{2g} \;-\; \frac{gd^2}{2v_0^2}\,. \end{aligned}$$

(ii) The ball will clear the crossbar if and only if $y_{\max} > h$. Now,

$$\begin{aligned} & y_{\max} > h \\ \iff\;& \frac{v_0^2}{2g} \;-\; \frac{gd^2}{2v_0^2} > h \\ \iff\;& v_0^4 - g^2d^2 > 2ghv_0^2 \\ \iff\;& (v_0^2 - gh)^2 > g^2(d^2+h^2) \\ \iff\;& v_0^2 \;>\; gh + g(d^2+h^2)^{\frac{1}{2}} \\ & \text{or } v_0^2 \;<\; gh - g(d^2+h^2)^{\frac{1}{2}} \end{aligned}$$

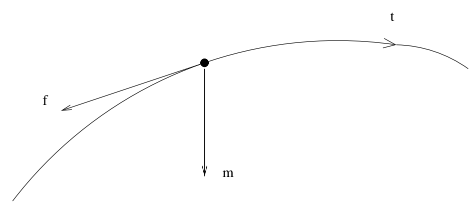
Now the term on the right-hand side of the last inequality is negative, so the inequality is never satisfied. Therefore

$$y_{\max} > h \;\iff\; v_0^2 \;>\; gh + g(d^2+h^2)^{\frac{1}{2}}.$$

□

7.2 Motion of a projectile under gravity and air resistance.

In reality, when particles move through the air a resistive force, $\mathbf{f}$ say, is exerted on the particle. This is often called the *drag* on the particle.



Experiments suggest that provided the particle is not spinning then this force acts in a direction opposite to its velocity, i.e.

$$\mathbf{f} = -f(v)\mathbf{v}$$

where $v = \|\mathbf{v}\|$, $\mathbf{v} = \dot{\mathbf{x}}$, and where $f(v)$ is some positive function. For v not very large $f(v)$ is found to be

$$f(v) \;=\; \mu$$

where $\mu > 0$ is a physical constant called the *drag coefficient*. Thus

$$\mathbf{f} \;=\; -\,\mu\mathbf{v} \;=\; -\,\mu\dot{\mathbf{x}}, \tag{7.6}$$

and N2 implies that $m\ddot{\mathbf{x}} = -\mu\dot{\mathbf{x}} + m\mathbf{g}$. Rearranging we obtain

$$\ddot{\mathbf{x}} + k\dot{\mathbf{x}} = \mathbf{g}, \quad k = \mu/m, \tag{7.7}$$

a second order linear vector ODE with initial conditions:

$$\mathbf{x}(0) = \mathbf{x}_0, \quad \dot{\mathbf{x}}(0) = \mathbf{v}_0.$$

Note that k reduces as m increases. This means that large bodies, such as Elephants, are much less affected by air resistance than small bodies such as mice. This leads to the conclusion that large bodies fall faster than smaller ones. (Although in a vacuum all bodies fall at exactly the same rate.)

If we set $\mathbf{v} = \dot{\mathbf{x}}$ then we have the vector ODE

$$\dot{\mathbf{v}} + k\mathbf{v} = \mathbf{g}.$$

You should be able to solve this first order ODE using the techniques that you learned last semester. To do this you look at the separate ODEs for the different components of the solution, and then using an integrating factor to solve each ODE in turn. If you do this you will get:

$$\mathbf{v} = \frac{\mathbf{g}}{k} + \left(\mathbf{v}_0 - \frac{1}{k}\mathbf{g}\right) e^{-kt}. \tag{7.8}$$

Now integrate this equation to get

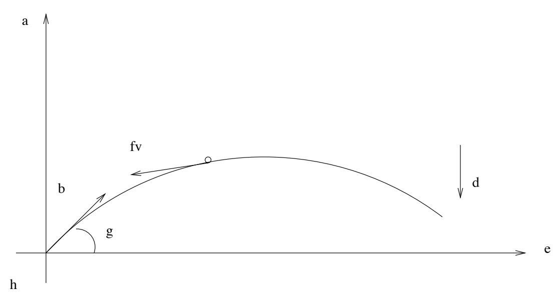
$$\mathbf{x}(t) = \mathbf{x}_0 + \mathbf{v}_0 \frac{1}{k}\left(1 - e^{-kt}\right) + \mathbf{g}\frac{1}{k^2}\left(e^{-kt} - 1 + kt\right). \tag{7.9}$$

For a comparison, the "no air resistance" case which we studied in §7.1 has solution:

$$\mathbf{x}(t) = \mathbf{x}_0 + \mathbf{v}_0 t + \frac{1}{2}\mathbf{g}t^2. \tag{7.10}$$

** Challenge. Show that (7.9) reduces to (7.10) as $k \to 0$.

Example: At the time $t = 0$, a golf ball is launched from the surface of the Earth at an angle α and speed v_0. Assuming air resistance as described above, describe the trajectory of the ball.



Solution:

Step 1 The situation is illustrated below.

Syeps 2–4 are the same as before.

Step 5 As we have shown, the ODE (7.7) holds and has solution (7.9). Using (7.9), in components we have

$$x(t) = \frac{(1 - e^{-kt})}{k} v_0 \cos(\alpha) \tag{7.11}$$

$$y(t) = \frac{(1 - e^{-kt})}{k} v_0 \sin(\alpha) - \frac{g}{k^2}\left(e^{-kt} - 1 + kt\right) \tag{7.12}$$

To find the trajectory in terms of y and x we eliminate t. From (7.11)

$$1 - e^{-kt} = \frac{kx}{v_0 \cos(\alpha)},$$

and so

$$kt = -\ln\left(1 - \frac{kx}{v_0\cos(\alpha)}\right).$$

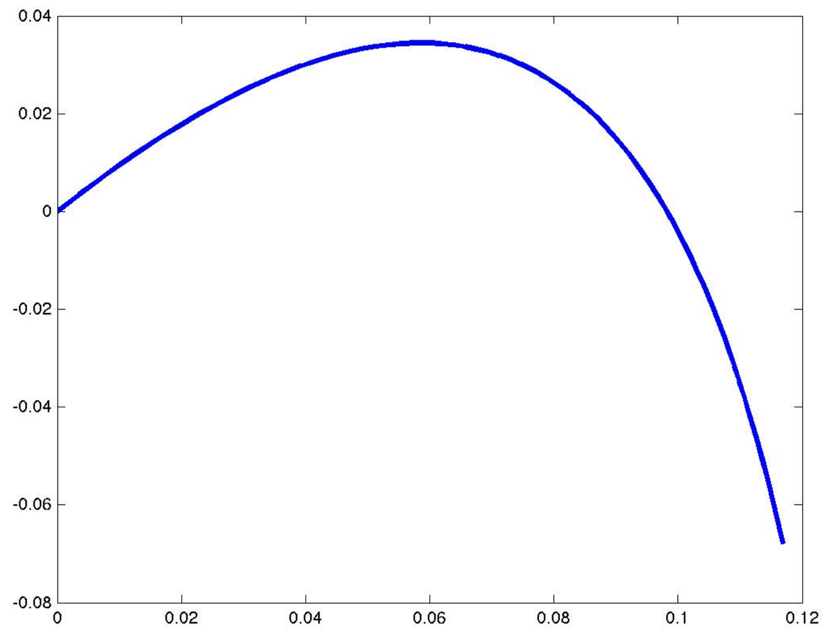
Hence the trajectory is from (7.12)

$$y \;=\; x\tan(\alpha) \;+\; \frac{g}{k^2}\left[\frac{kx}{v_0\cos(\alpha)} + \ln\left(1 - \frac{kx}{v_0\cos(\alpha)}\right)\right]. \qquad (7.13)$$

This can be compared with the trajectory with no resistance (7.5) given by

$$y \;=\; x\tan(\alpha) \;-\; \frac{gx^2}{2v_0^2\cos^2(\alpha)}.$$

The shape of this trajectory is illustrated below. Anyone who has watched a golf ball in flight will recognise its shape.



Implications:

1) It follows from from (7.11), that as $0 < e^{-kt} \to 0$ as $t \to \infty$ then

$$x(t) < \frac{v_0\cos(\alpha)}{k} \quad \text{and} \quad x(t) \longrightarrow \frac{v_0\cos(\alpha)}{k} \quad \text{as } t \to \infty.$$

2) From (7.12) $\quad y(t) \longrightarrow -\infty,$ as $t \to \infty$.

3) From (7.8) $\quad \mathbf{v}(t) \longrightarrow \dfrac{1}{k}\mathbf{g}$ as $t \longrightarrow \infty$. Hence the velocity tends to a constant. This is called the *terminal velocity.*

7.3 The simple pendulum

We will now look at the problem of a swinging pendulum. This will comprise of a particle of mass m is fixed to the end of either a light rod (the rigid pendulum) or an inextensible light string (the non-rigid pendulum). In both cases the rod and the string are of length l. The other end is pivoted at the origin O. The rod/string swings in a vertical plane, so the motion is in two dimensional. We will assume that the angle that the rod/string makes to the vertical is θ. Physically, the tension in the rod/string exerts a force $\mathbf{T}$ on the particle acting towards the origin. In the case of the rigid rod this tension can be either positive or negative and the particle is always constrained to move at the end of the rod. In the case of a string, it has to be positive for the strong to remain taut. If it drops to zero then the string goes slack and the system changes, with the particle moving freely in space

Example: A rigid pendulum comprises a mass at the end of a light rod of length l. The rod is freely pivoted at the origin and makes an angle θ with the vertical. At the time $t = 0$, we have $\theta(0) = 0$, and the horizontal speed is $l\, v_0$.

(1) Prove that (in the absence of friction and air resistance) the angle θ satisfies the nonlinear second order ODE

$$\ddot{\theta} + \frac{g}{l}\sin(\theta) = 0.$$

(Do not attempt to solve this ODE).

(2) If the energy E of the pendulum is defined by

$$E = \frac{\dot{\theta}^2}{2} - \frac{g}{l}\cos(\theta).$$

Prove that E is constant. Hence show that

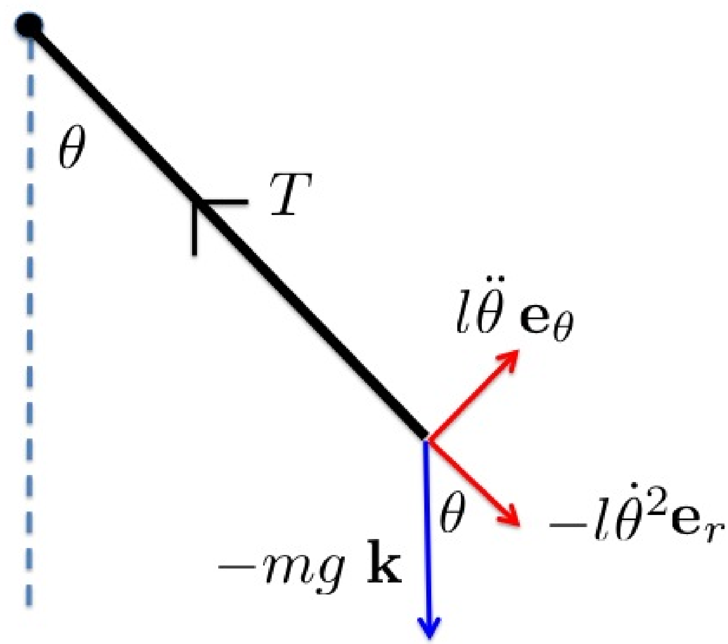
$$\dot{\theta}^2 = v_0^2 - \frac{2g}{l}(1 - \cos(\theta)).$$

(3) Calculate the tension T in the rod.

(4) Hence describe the possible motions of the rigid pendulum.

Solution We take the five point approach to solving this problem

Step 1: Draw a picture.



We take the origin to be the fixed end of the rod. In polar coordinates, with the unit vectors, $\mathbf{e}_r, \mathbf{e}_\theta$ we have from before that the position vector is $\mathbf{x} = r\mathbf{e}_r$, The position, velocity and acceleration are then given by

$$\begin{aligned}
\mathbf{x} &= r\,\mathbf{e}_r, \text{ and,} \\
\dot{\mathbf{x}} &= \dot{r}\,\mathbf{e}_r + r\dot{\theta}\,\mathbf{e}_\theta, \\
\ddot{\mathbf{x}} &= (\ddot{r} - r\dot{\theta}^2)\,\mathbf{e}_r + (r\ddot{\theta} + 2\dot{r}\dot{\theta})\,\mathbf{e}_\theta.
\end{aligned}$$

But the rod is rigid and has constant length l. It follows that $r = l$, $\dot{r} = \ddot{r} = 0$. Thus

$$\ddot{\mathbf{x}} = -l\,\dot{\theta}^2\,\mathbf{e}_r + l\,\ddot{\theta}\,\mathbf{e}_\theta.$$

Step 2: Resolve all of the forces.

There are two forces acting on the mass. The tension T in the rod, which acts along the rod in the direction $\mathbf{e}_r$, and the gravitational force which acts in the vertical direction $-\mathbf{k}$. It is convenient to resolve all of these forces in the directions $\mathbf{e}_r$ and $\mathbf{e}_\theta$. This gives the following

(which you should check)

$$
\begin{aligned}
m\mathbf{g} &= mg\cos(\theta)\,\mathbf{e}_r - mg\sin(\theta)\,\mathbf{e}_\theta, \\
\mathbf{T} &= -T\,\mathbf{e}_r.
\end{aligned}
$$

Hence, if we apply Newton' Second Law, comparing forces and mass times accelerations, we have:

$$m\ddot{\mathbf{x}} = m\mathbf{g} + \mathbf{T},$$

or,

$$-m\,l\,\dot{\theta}^2\,\mathbf{e}_r + m\,l\,\ddot{\theta}\,\mathbf{e}_\theta = (m\,g\,\cos(\theta) - T)\,\mathbf{e}_r - m\,g\,\sin(\theta)\,\mathbf{e}_\theta. \tag{7.14}$$

Now consider the components in the direction of $\mathbf{e}_\theta$ we have the second order differential equation

$$m\,l\,\ddot{\theta} = -m\,g\,\sin(\theta).$$

Hence, θ satisfies the second order ODE

$$l\,\ddot{\theta} + g\,\sin(\theta) = 0.$$

as required. **Warning** this equation is very (very) hard to solve analytically.

If we next resolve in the direction of $\mathbf{e}_r$ we find that the tension T satisfies the equation

$$m\,g\,\cos(\theta) - T = -m\,l\,\dot{\theta}^2. \tag{7.15}$$

To find T we must calculate $\dot{\theta}$, which we do next.

Step 3: Find the initial conditions.
These are given by: $\theta(0) = 0$ and the horizontal speed at $t = 0$ is $l\,v_0$. Since, at $t = 0$, $r = l$, a constant, then $\dot{r}(0) = 0$. So

$$
\begin{aligned}
\dot{\mathbf{x}} &= \underbrace{\dot{r}\,\mathbf{e}_r}_{=\mathbf{0}} + r\dot{\theta}\,\mathbf{e}_\theta = l\dot{\theta}\,\mathbf{e}_\theta \\
&\Rightarrow\ l\,\dot{\theta}(0) = l\,v_0 \text{ (setting } t = 0). \\
&\Rightarrow\ \dot{\theta}(0) = v_0.
\end{aligned}
$$

Step 4: Integrate the motion to calculate the energy E and hence find the tension T .

The equation (7.15) is a nonlinear ODE for $\theta(t)$. We cannot solve this ODE in general (honest, I did warn you) BUT we can find a first integral. In fact this is important as it allows us to find the *Energy* of the whole system and then the tension in the rod.

To do this we multiply all terms in the expression (7.15) by $\dot{\theta}$ to give

$$\dot{\theta}\ddot{\theta} + \frac{g}{l}\sin(\theta)\dot{\theta} = 0. \tag{7.16}$$

Next we use the two results that

$$\frac{d}{dt}\dot{\theta}^2 = 2\dot{\theta}\ddot{\theta} \quad \text{and} \quad \frac{d}{dt}\cos(\theta) = -\sin(\theta)\dot{\theta}.$$

Hence we can integrate (7.16) to give

$$E = \frac{1}{2}\dot{\theta}^2 - \frac{g}{l}\cos(\theta), \tag{7.17}$$

for some constant, E.

E is the *energy* of the system, which is *conserved* during the motion.

Setting $t = 0$ in (7.17) and using the fact that $\dot{\theta}(0) = v_0$ (from Step 3), we have

$$C = \frac{v_0^2}{2} - \frac{g}{l}$$

Hence

$$\dot{\theta}^2 = v_0^2 - \frac{2g}{l}\left(1 - \cos(\theta)\right)\ . \tag{7.18}$$

This is the result in (2).

It follows that θ satisfies the *first order differential equation*

$$\frac{d\theta}{dt} = \sqrt{v_0^2 - \frac{2g}{l}\left(1 - \cos(\theta)\right)}.$$

This ODE cannot be solved analytically (without using Elliptic Functions), but could be solved numerically using techniques discussed in later Units. However, we can obtain quite a bit of information about the solution with further analysis.

Finding the tension. Recall from (7.15) that the tension T is given by

$$T = mg\cos(\theta) + ml\dot{\theta}^2.$$

Thus, using (7.18), we have that the tension is explicitly given by

$$T = m\,l\,v_0^2 + m\,g\,\left(3\cos(\theta) - 2\right)\ . \tag{7.19}$$

Step 5: Find the resulting motion.

We notice that as θ increases from zero, $\cos(\theta)$ decreases and so

(a) T decreases (from (7.19)), and

(b) $\dot{\theta}$ decreases (from (7.18)), so the pendulum slows down.

In fact, from (7.18), $\dot{\theta}^2 = 0$ (ie. zero speed) $\iff$

$$v_0^2 = \frac{2g}{l}\left(1 - \cos(\theta)\right),$$

or,

$$\cos(\theta) = 1 - \frac{lv_0^2}{2g} \;\; =:\; P,\ \text{say.}$$

If $-1 \le P \le 1$ then there is a *critical angle* θ_1 where $\dot{\theta}_1 = 0$. At this angle the pendulum comes momentarily to rest. The critical angle is given by

$$\theta_1 = \cos^{-1}(P).$$

Now we can work out the possible types of motion.

Case 1. We first, assume that $v_0^2 < V^2 \equiv \frac{4g}{l}$.

In this case, $P > -1$ and so $\theta_1 < \pi$. Thus $\dot{\theta} = 0$ occurs before $\theta = \pi$. At this point we have (from the underlying differential equation)

$$\ddot{\theta} = -\frac{g}{l}\sin(\theta_1) < 0.$$

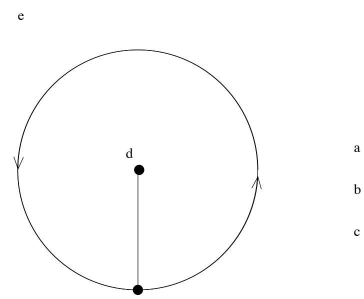
Hence, θ reaches a maximum value at $\theta = \theta_1$ and *then decreases.* The pendulum swings back to the origin, where now (from considering the energy) $\dot{\theta} = -v_0$ (Prove this). It continues swinging until $\theta = -\theta_1$. When is momentarily comes to rest again, before swinging back. Thus the pendulum exhibits *periodic motion.*

Case 2. Now consider the case $v_0^2 = V^2$.

In this case $\theta_1 = \pi$ and we have $\dot{\theta} = \ddot{\theta} = \pi$. The pendulum eventually (in fact it takes an infinite time!) comes to rest at the vertically upright position with $\theta = \pi$. This is in fact *unstable* (you can test this by trying to balance a pencil on its end).

Case 3. Finally consider the case $v_0^2 > V^2$.

In this case there is no value of θ for which $\dot{\theta} = 0$. In this case the pendulum completes whole revolutions and continues to swing like this for ever, with θ constantly increasing with time.



To summarise. The rigid pendulum has

(a) periodic oscillations if $v_0^2 < 4g/l \quad \Leftrightarrow \quad P > -1$, and

(b) full revolutions if $v_0^2 > 4g/l \quad \Leftrightarrow \quad P < -1$.

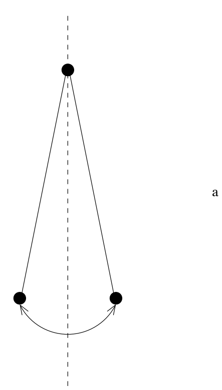
The borderline case $v_0^2 = 4g/l$ corresponds to a *bifurcation*: which is a switch between the two qualitatively different regimes as v_0 is varied.

7.4 Small oscillations

7.4.1 Oscillations when there is no air resistance

Let v_0 be small, we expect that a small initial velocity will lead to small oscillations. More precisely, we assume:

$$\frac{v_0^2 l}{2g} \ll 1.$$



Then we are obviously in the case $\frac{v_0^2 l}{2g} < 1$ in the previous section. Now

$$\begin{aligned}|\theta(t)| &\leq \theta_1 = \cos^{-1} P \\ &= \cos^{-1}\left(1 - \frac{v_0^2}{2gl}\right) \ll 1 \\ \Rightarrow \sin(\theta) &= \theta - \frac{\theta^3}{6} + \cdots \approx \theta, \text{ to a good approximation.}\end{aligned}$$

Then the equation $\ddot{\theta} + \frac{g}{l}\sin(\theta) = 0$ is approximated by the linear second order ODE

$$\begin{aligned}\ddot{\theta} + \frac{g}{l}\theta &= 0 \\ \Leftrightarrow \ddot{\theta} + \omega^2\theta &= 0, \quad \omega := \left(\frac{g}{l}\right)^{\frac{1}{2}},\end{aligned}$$

which gives rise to **simple harmonic motion**. in fact it has the exact solution, which satisfies the I.C given by

$$\theta(t) = v_0\sqrt{\frac{l}{g}}\sin\left(\sqrt{\frac{g}{l}}t\right).$$

Exercise: Check this

The period T of the pendulum with small oscillations is then given by

$$T = 2\pi\sqrt{\frac{l}{g}}.$$

NOTE 1 that T is independent of v_0 if v_0 is small. This result was discovered by Galileo. As the size of the oscillations increases so does T, but slowly.

NOTE2 In SI Units $g \approx \pi^2$ to a very good approximation. Thus we have the very good approximation

$$T \approx 2\sqrt{l}.$$

Hence a pendulum which is one metre long, will have a half swing (from side to side) whihc is almost exactly one second. As far as I know this is a coiincidence.

7.4.2 Small oscillations with air resistance

This is called the damped pendulum equation. If we include air resistance this will always act in an opposite direction to that of motion i.e. in the direction $-\mathbf{e}_\theta$. If we set the air resistance as $-2kl\dot{\theta}\mathbf{e}_\theta$ and resolve in the direction $\mathbf{e}_\theta$ then the resulting ODE for the mass in the case of small oscillations is given by

$$\ddot{\theta} + 2k\,\dot{\theta} + \frac{g}{l}\theta = 0.$$

We supoose that initially we start the pendulum moving from a state of rest so that:

$$\theta = \theta_0 \quad \text{and that} \quad \dot{\theta} = 0.$$

We solve this ODE by posing a solution of the form

$$\theta = e^{\lambda t}$$

and substituting into the ODE. This leads to the following quadratic (characteristic) equation for θ

$$\lambda^2 + 2k\lambda + \frac{g}{l} = 0.$$

Assuming that k is *small*, in particular that $k^2 < g/l$, this quadratic equation has the complex roots

$$\lambda_{\pm} = -k \pm i\sqrt{g/l - k^2} \equiv -k \pm i\omega,$$

where

$$\omega^2 = g/l - k^2.$$

Thus the general solution to the damped pendulum equation is:

$$\theta(t) = e^{-kt}\left[A\cos(\omega t) + B\sin(\omega t)\right].$$

Here A and B are constants.

To find A and B we need to make use of the initial conditions. It follows immediately that

$$\theta_0 = \theta(0) = A.$$

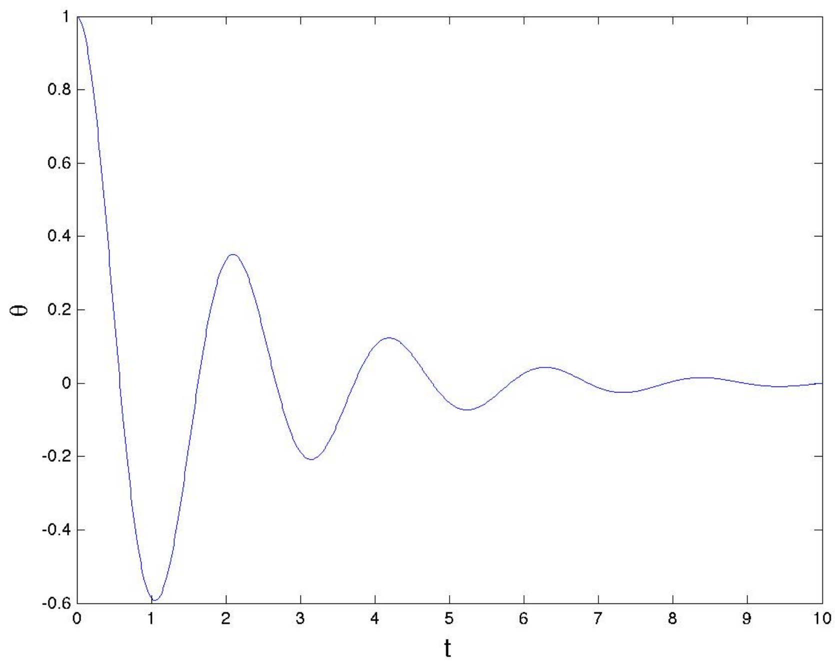
Similarly (with a bit more work)

$$0 = \dot{\theta}(0) = -kA + \omega B.$$

Thus

$$A = \theta_0 \quad \text{and} \quad B = \frac{k\theta_0}{\omega}.$$

The resulting motion is illustrated below with $k = 0.5, \omega = 3$ and $\theta_0 = 1$. It takes the form of oscillations which gradually damp to zero. You can try this at home if you make your self a simple pendulum using a piece of string and a weight of some sort.



7.5 * Optional and non-examinable. A mass on a string eg. a lion

If the mass is on a string rather than a rigid rod, then the motion is exactly the same provided that the string remains taught i.e. if the tension $T > 0$. This arises provided that

$$T = m\ l\ v_0^2 + m\ g\ \ (3\cos(\theta) - 2) > 0.$$

It is clear that T decreases as θ increases. If $v_0 < V$ then at $\theta = \theta_1$ we have $\cos(\theta_1) = 1 - l v_0^2/2g$, thus after some manipulation we have

$$T(\theta_1) = mg\left(1 - \frac{v_0^2 l}{2g}\right).$$

There are then three cases

1. $v_0^2 < 2g/l$. In this case the string remains taut during the motion and the system behaves like the rigid pendulum.

2. $2g/l < v_0^2 < 5g/l$. In this case the string goes slack at some angle $\theta_2 < \theta_1$.

3. $v_0^2 > 5g/l$. The string never comes to rest and the tension T is always positive. The mass swirls around the pivot on the end of a taut string (Think of the lion here).

In Case 2, at the point when $\theta = \theta_2$ where

$$m\ l\ v_0^2 + m\ g\,(3\cos(\theta_2) - 2)\,,$$

the mass will move off as a free projectile.

The limit of $v_0^2 = 5g/l$ comes the expression

$$T = m\ l\ v_0^2 + m\ g\ \ (3\cos(\theta) - 2) > m\ l\ v_0^2 - 5\ m\ g$$

so that $T > 0$ for all θ in this case, with the minimum value of T arising when $\theta = \pi$.

7.6 ** Very optional and non-examinable. The chaotic double pendulum

The double pendulum comprises two rigid pendulums coupled together through a smooth joint with respective masses $m_{1,2}$ lengths $l_{1,2}$ and angles $\theta_{1,2}$ to the vertical. This is illustrated below.

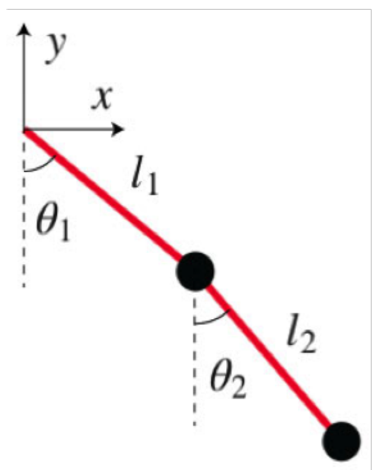


Figure 13: A schematic of the double pendulum.

The differential equations for the double pendulum, are given (in the absence of friction) by

$$(m_1 + m_2)l_1\ddot{\theta}_1 + m_2 l_2\ddot{\theta}_2\cos(\theta_1 - \theta_2) + m_2 l_2\dot{\theta}_2^{\,2}\sin(\theta_1 - \theta_2) + g(m_1 + m_2)\sin(\theta_1) = 0,$$

$$m_2 l_2\ddot{\theta}_2 + m_2 l_1\ddot{\theta}_1\cos(\theta_1 - \theta_2) - m_2 l_1\dot{\theta}_1^2\sin(\theta_1 - \theta_2) + m_2 g\sin(\theta_2) = 0.$$

By taking θ_1 and θ_2 to be *small*, we can show that there are two distinct types of periodic oscillation with either θ_1 and θ_2 in phase, or with θ_1 and θ_2 out of phase.

If, for example (as in the lectures) we take $m_1 = 2m_2 = 2$ and $l_1 = l_2 = 1$, then for larger values of θ_1 and θ_2 *chaotic motion* is possible. This is very complex and apparently random

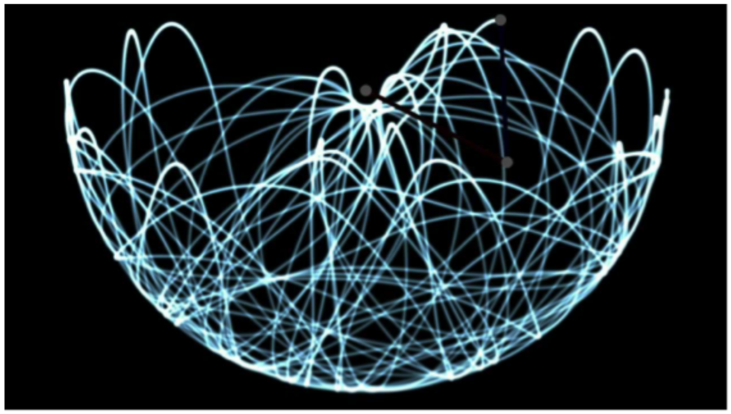
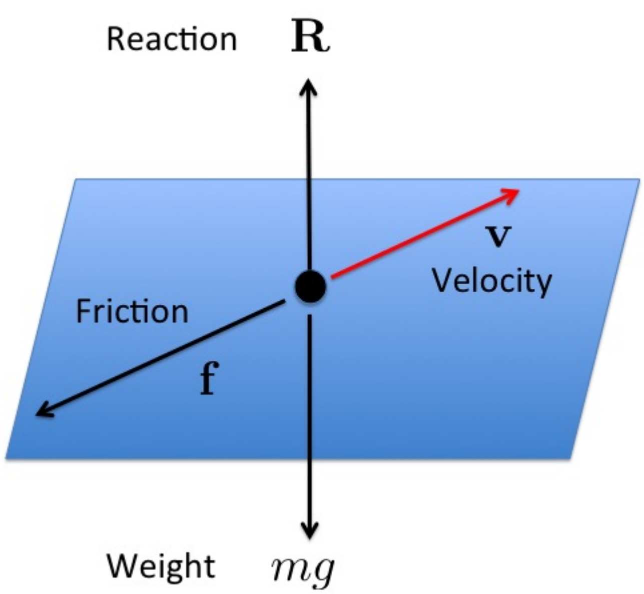


Figure 14: Chaotic motion of the double pendulum. In this figure a light is shone from the bottom of the bottom part of the pendulum.

motion. We illustrate the chaotic motion in Figure 14. Chaotic motion cannot be random as it is described by differential equations, but it certainly appears to be. You can check this by solving the above equations using Matlab. If you go onto the Moodle site you will find much more information about the double pendulum. It shows how complex the motion of even a simple mechanical system can be.

7.7 Motion along surfaces with friction



In this sub-section we will consider that most mysterious of all forces, namely friction. Friction is the force that arises when two bodies are in contact. It a weird force because it can take many differ values when the two objects are stationary with respect to each other (static friction), and a different value when they are moving relative to each other (dynamic friction). Friction is an example of is what is called an *emergent prpoerty*, as we see, on the large scale, what emerges from the interaction of a lot of things on a small scale. It is vital to understand friction because much of our technology relies on it in some way. An example of relevance to all of us is the technology behind shoes which are designed not to slip on wet ground.

Consider a particle *sliding* on a surface $\mathcal{S}$ and acted upon by gravity, so that motion is essentially in 2D. The surface exerts a force on the particle which has two components:

1. A **normal reaction**, $\mathbf{R}$ which is orthogonal to the surface, so $\mathbf{R} = R\mathbf{n}$, where $\mathbf{n}$ is the normal to the surface and $R \geq 0$ on physical grounds. This force acts to prevent the particle moving through the surface (and is a function of the electronic bonds in the surface matertial).

2. The tangential component is the **dynamic frictional force**, $\mathbf{f}$, which acts in the direction opposite to the direction of motion.

Experiments show that to a good approximation, that whilst the particle is *moving* we have:

$$f = \mu R \tag{7.20}$$

where $f = \|\mathbf{f}\|, R = \|\mathbf{R}\|$. The constant $\mu \geq 0$ is the coefficient of (dynamic) friction and is a constant.

If the body is *at rest* relative to a surface, and a force F is applied to it, then the frictional force will oppose its motion. The maximum force F that can be applied until the particle moves is given by

$$F = \mu_S R, \tag{7.21}$$

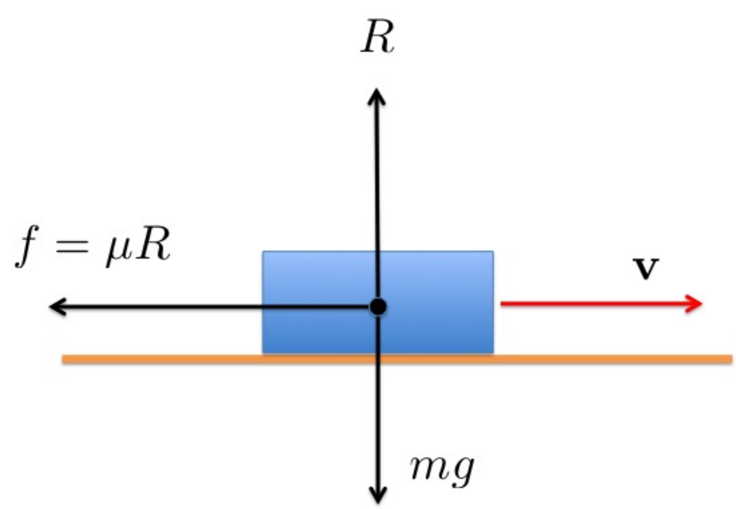
The resistance to the applied force F is called the *static friction* and this can take any value between 0 and $\mu_S R$. Weird! What is even wierder is that in most cases $\mu_S > \mu$. What we see is that if an increasing force $F(t)$ is applied to the particle then it stays at rest until $F = \mu_S R$. It then starts to slide, and at this point the friction force decreases, so that the force on the particle increases and it accelerates. Try this for your self by pushing an object with your finger.

Notice that (7.21) only specifies *the magnitude* f of the static friction force $\mathbf{f}$. Its *direction* is such that maintains the particle's static position at the surface. In most physical models of friction, the coefficient of static friction μ_s in (7.21) is larger than the coefficient of the dynamic friction μ in (7.20). The two equations (7.20) and (7.21) constitute *Coulomb's Law of Friction*, due to C.A. Coulomb (1736 - 1806).

Example: A brick of mass m slides along a horizontal plane. At time $t = 0$ its speed is v_0. Let μ be the coefficient of dynamic friction. Show that the particle comes to rest at time $T = \dfrac{v_0}{g\mu}$.

Solution We find the solution by using the familiar 5-point plan

Step 1: (Draw a diagram showing the forces and the motion)



The position vector is $\mathbf{x} = x\mathbf{i}$, and the velocity $\mathbf{v} = v\mathbf{i}$.

Step 2: (Resolve the forces and apply Newton' s second law)

The forces acting on the brick are: the normal reaction $\mathbf{R} = R\ \mathbf{j}$, gravity $m\mathbf{g} = -mg\ \mathbf{j}$ and dynamic friction $\mathbf{f} = -f\ \mathbf{i}$. So, from Newton's Second Law:

$$m\ddot{\mathbf{x}} = \mathbf{F} = \mathbf{R} + \mathbf{f} + m\mathbf{g},$$

or

$$m\ddot{x}\,\mathbf{i} = R\,\mathbf{j} - f\,\mathbf{i} - mg\,\mathbf{j}. \tag{7.22}$$

Step 3: (I.C.)

These are $\mathbf{x}(0) = x(0)\,\mathbf{i} = \mathbf{0}$, so $x(0) = 0$. Also $\dot{\mathbf{x}}(0) = \dot{x}(0)\,\mathbf{i} = v_0\mathbf{i}$, so $\dot{x}(0) = v_0$.

Step 4: (Solve the system arising from applying Newton's Second Law)

(i) First, equate the y-components of (7.22)), to give the reaction via

$$0 = R - mg \quad \text{so that we have} \quad R = mg.$$

(ii) Next, resolve in the x-direction to give

$$m\ddot{x} = -f = -\mu R \;\; \text{(using Coulomb's Law)}$$

Hence, the motion is described by

$$\ddot{x}(t) = -\mu g. \tag{7.23}$$

With the initial condition $\dot{x}(0) = v_0$, we can integrate (7.23) to get

$$\dot{x}(t) = -\mu g t + v_0.$$

If the particle comes to rest at $t = T$, then $\dot{x}(T) = 0$. Hence $0 = -\mu g T + v_0$, or $T = \dfrac{v_0}{\mu g}$. □

At this point the *friction force immediately drops to zero* and the brick remains at rest. Isn't friction fun! You too can be a *tribologist*. .

8 Motion under Central Forces.

8.1 Central forces

A central force is a generalisation of the gravitational force which we considered in Chapter 6. In this Chapter we will consider the types of motion that we are likely to see under a central force. The most commonly encountered such motion is a periodic *elliptical orbit* (as described by Kepler's laws of motion), however we can also see motion along a *hyperbola* and motion in a straight line. Indeed one of the great breakthroughs in mechanics (usually attributed to Newton, although others such as Hooke and Halley can claim credit for it) was realising that the same force which pulled objects down to the Earth in what seemed to be straight lines when dropped, could also produce elliptical orbits when looking at planets going around the Sun. It is essential to study central forces if you want to be a rocket scientist (see Moodle for several ways to be a rocket scientist).

Consider a particle of mass m at P with nonzero position vector $\mathbf{x} = \overrightarrow{OP}$ relative to the origin at O. (For planetary motion we will assume that O is the location of the Sun.)

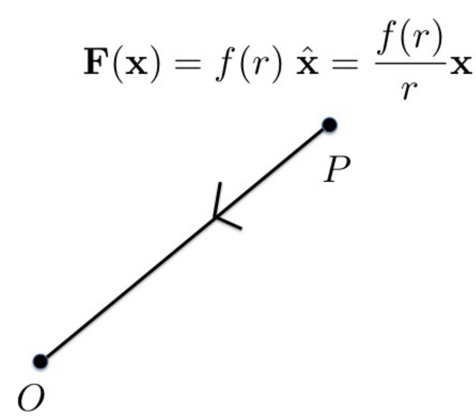


Figure 15: An attracting central force

Definition 8.1 $\mathbf{F}(\mathbf{x})$ *is a* ***central force*** *if it is always directed towards (or away from) O, and its magnitude depends* only *on the distance $r = \|\mathbf{x}\| = \|\overrightarrow{OP}\|$.*

All central forces can be written in the form

$$\mathbf{F}(\mathbf{x}) = f(r)\,\hat{\mathbf{x}} = \frac{f(r)}{r}\mathbf{x},$$

where $\hat{\mathbf{x}} = \mathbf{x}/r$ is a unit vector in the direction of $\mathbf{x}$, and $f(r)$ a scalar function.

The central force is *attractive towards* O if $f(r) < 0$, or *repulsive away from* O if $f(r) > 0$.

Example 1: If a particle of mass m is in orbit about a planet of mass M with its centre at O, then gravitational force between them is an attractive central force with

$$f(r) = -\frac{\mu\, m}{r^2}, \quad \mu = GM.$$

Example 2:. If a mass at point P is on the end of a spring which is fastened at one end to O and at the other end to P (think of P being a bungee jumper at the end of a bungee) then there is a central *elastic* force between O and P. This is given by Hooke's law and we have.

$$f(r) = -kr.$$

The elastic force is an example of a *derived* force and is a result of the combined molecular forces in the spring. Robert Hooke was a contemporary of Newton and (like me) was a Gresham Professor of Geometry.

Example 3: The electrostatic force between two charged particles at points P_1 and P_2 has

$$f(r) = \frac{K\ q_1\ mq_2}{r^2}$$

where q_1 and q_2 are the electric charges of the points at P_1 and P_2, and $K > 0$ is a physical constant. In this force particles which have the same charge repel (which is quite unlike gravity), and particles which have opposite charges attract. The above is called Coulomb's law of electrostatics, and was discovered by the same Coulomb as for the friction law. This force is *much* stronger than the gravitational force, and it is the force which keeps molecules (and therefore our bodies) together. It is one of the three *fundamental* forces of nature.

8.2 Motion under the action of central forces

Applying Newton's Second law, the equation of motion of the point at P is of the form:

$$m\ddot{\mathbf{x}} = \mathbf{F}(\mathbf{x}) = \frac{f(r)}{r}\mathbf{x}. \tag{8.1}$$

Notation: In this section we assume that the particle has mass $m = 1$, and so from now on the term m won't explicitly enter into the analysis. (An alternative way of thinking of this is that we assume the central force is a force "per unit mass". There is no loss in generality, since in our analysis $f(r)$ is arbitrary and so can be replaced by $m\, f(r)$ to this effect.) So

$$\ddot{\mathbf{x}} = \frac{f(r)}{r}\mathbf{x}. \tag{8.2}$$

8.3 Properties of the motion:

Property 1: Equation (8.2) describes motion in a plane.

Proof: This is a consequence of conservation of the angular momentum $\mathbf{h} = \mathbf{x} \times \dot{\mathbf{x}}$ in a central force. This follows because

$$\dot{\mathbf{h}} = \frac{d}{dt}(\mathbf{x} \times \dot{\mathbf{x}}) = \dot{\mathbf{x}} \times \dot{\mathbf{x}} + \mathbf{x} \times \ddot{\mathbf{x}} = \mathbf{x} \times \ddot{\mathbf{x}}.$$

But

$$\mathbf{x} \times \ddot{\mathbf{x}} = \mathbf{x} \times \frac{f(r)}{r}\mathbf{x} = \mathbf{0}.$$

Thus

$$\mathbf{h} \quad \text{is a constant.} \tag{8.3}$$

If we take the dot product of (8.3) with $\mathbf{x}$ then

$$\begin{aligned} \mathbf{x}.\mathbf{h} &= \mathbf{x}.\mathbf{x} \times \dot{\mathbf{x}} = [\mathbf{x}, \mathbf{x}, \dot{\mathbf{x}}] = 0 \\ \therefore \ \mathbf{x}.\mathbf{h} &= 0. \end{aligned}$$

If $\mathbf{h} \neq \mathbf{0}$ this is the equation of plane, Π say, perpendicular to $\mathbf{h}$ and containing the origin. Then $\mathbf{x}(t) \in \Pi$, for all t i.e. the motion is in the plane Π.

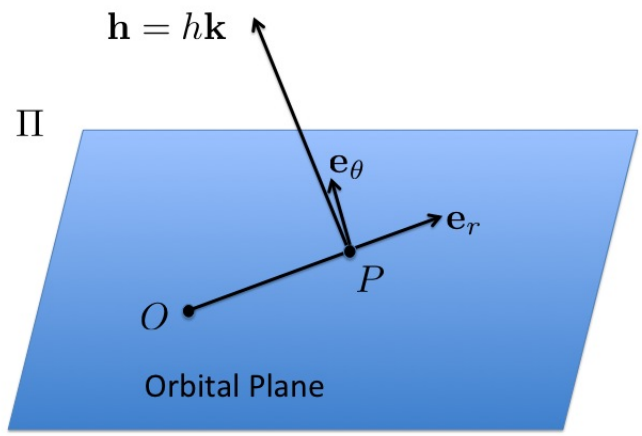
In the special case of $\mathbf{h} = \mathbf{0}$, (8.3) $\mathbf{x} \times \dot{\mathbf{x}} = \mathbf{0}$, and so $\dot{\mathbf{x}}$ is parallel to $\mathbf{x}$, and we have motion along a line (and hence in any plane containing this line).

NOTE In the case of the Solar system all of the planets orbit in a plane. This plane is called the *ecliptic*

Motion in the plane

Since motion under the action of a central force is in a plane (the *orbital plane* Π) it is 2-Dimensional, and we can therefor we can use polar coordinates in the plane to describe the way that it moves. With $\mathbf{k}$, the unit vector in the $\mathbf{h}$ direction, choose the polar coordinates on Π so that $\mathbf{e}_r$, $\mathbf{e}_\theta$, $\mathbf{k}$ form a right-handed system. Hence (8.3) can be written as

$$\mathbf{x} \times \dot{\mathbf{x}} = h\,\mathbf{k}, \quad h =: \|\mathbf{h}\|.$$



Crucially a central force only acts in the direction of $\mathbf{e}_r$.

There is NO component of the force in the direction of $\mathbf{e}_\theta$

Hence there is *no acceleration in the direction of* $\mathbf{e}_\theta$.

Property 2: For motion under any central force, the quantity $h = \ r^2\dot{\theta} = \text{constant}$.

First, recall that in general

$$\begin{aligned}
\mathbf{x} &= r\mathbf{e}_r \\
\dot{\mathbf{x}} &= \dot{r}\mathbf{e}_r + r\dot{\theta}\mathbf{e}_\theta \\
\ddot{\mathbf{x}} &= \left(\ddot{r} - r\dot{\theta}^2\right)\mathbf{e}_r + \left(r\ddot{\theta} + 2\dot{r}\dot{\theta}\right)\mathbf{e}_\theta.
\end{aligned}$$

But from the above, we have that there is no acceleration in the $\mathbf{e}_\theta$ direction. It follows that

$$r\ddot{\theta} + 2\dot{r}\dot{\theta} = 0.$$

But, a bit of calculus shows us that

$$r\ddot{\theta} + 2\dot{r}\dot{\theta} = \frac{1}{r}\frac{d}{dt}\left(r^2\dot{\theta}\right).$$

Hence we have

$$\frac{d}{dt}\left(r^2\dot{\theta}\right) = 0.$$

It follows that

$$r^2\dot{\theta} = h \tag{8.4}$$

where h (the scalar angular momentum) is a constant.

> As we have seen earlier, this is precisely K2 (Kepler's equal areas law).
>
> We can now see that this law is true for **any** central force, not only for gravitation.

Property 3 : The orbit equation.

With $\ddot{\mathbf{x}} = \left(\ddot{r} - r\dot{\theta}^2\right)\mathbf{e}_r$ and $\dfrac{f(r)}{r}\mathbf{x} = f(r)\mathbf{e}_r$, then $\ddot{\mathbf{x}} = \dfrac{f(r)}{r}\mathbf{x}$, implies

$$\ddot{r} - r\dot{\theta}^2 \;=\; f(r),$$

or, since $r^2\dot{\theta} = h$, we have,

$$\ddot{r} - \frac{h^2}{r^3} = f(r)\,. \tag{8.5}$$

This is a 2nd order nonlinear ODE for $r(t)$ which we can combine with the first order ODE for θ in (8.4) to find the orbit $(r(t), \theta(t))$. In general we cannot solve this, except for special forms of the function $f(r)$. Fortunately the inverse square law force is one of these special cases.

Property 4: Conservation of Energy

If we multiply the orbit equation (8.5) by $\dot{r}$ we have

$$\dot{r}\ddot{r} - \frac{h^2\dot{r}}{r^3} = f(r)\dot{r}.$$

Now

$$\frac{d}{dt}\left(\dot{r}^2\right) = 2\dot{r}\ddot{r}, \quad \frac{d}{dt}\left(\frac{1}{r^2}\right) = -2\frac{\dot{r}}{r^3} \quad \text{and} \quad \frac{d}{dt}\int f(r)\,dt = f(r)\dot{r}.$$

Hence, we can integrate (8.5) once to obtain

$$\frac{1}{2}\,\dot{r}^2 + \frac{h^2}{2r^2} - \int f(r)\;dr = E, \tag{8.6}$$

for some constant E. This can be regarded as an equation describing the conservation of the energy E.

8.4 Finding the orbit

Now we return to the question of finding the **orbit** (= the trajectory of the particle). We seek to eliminate t in the differential equations (8.4,8.5) to obtain $r = r(\theta)$. To do this we make the following change of variable:

$$r = \frac{1}{u}, \quad u = u(\theta),$$

which makes sense only if

$$\dot{\theta} \neq 0,$$

since otherwise θ is constant. Now,

$$\begin{aligned} \dot{r} &= \frac{d}{dt}\left(u^{-1}\right) = \frac{d}{d\theta}\left(u^{-1}\right)\frac{d\theta}{dt} \\ &= -\frac{1}{u^2}\frac{du}{d\theta}\,\dot{\theta} = -\frac{du}{d\theta}\underbrace{r^2\dot{\theta}}_{=h} = -\,h\frac{du}{d\theta}. \end{aligned}$$

So, again using $\dot{\theta} = \dfrac{h}{r^2} = hu^2$,

$$\ddot{r} = \frac{d}{dt}\dot{r} = \dot{\theta}\frac{d\dot{r}}{d\theta} = -\,\dot{\theta}\;h\frac{d}{d\theta}\left(\frac{du}{d\theta}\right) = -\,h^2u^2\,\frac{d^2u}{d\theta^2}.$$

Hence (8.5) becomes

$$-h^2u^2\,\frac{d^2u}{d\theta^2}\; - \;h^2u^3 = f(u^{-1}),$$

or,

$$\frac{d^2u}{d\theta^2} \;+\; u \;=\; -\,\frac{f(u^{-1})}{h^2u^2}\,, \qquad \text{assuming } (h \neq 0). \tag{8.7}$$

Definition 8.2 *Equation (8.7) is called the* ***orbit*** *equation. It holds provided $\dot{\theta} \neq 0$ (or, equivalently, $h \neq 0$).*

8.5 Orbits under an inverse square law

Usually, (8.7) will be nonlinear and essentially impossible to solve, but in the case of an inverse square law force it reduces to a linear equation which we can solve easily. **This is a major fluke. We didn't deserve to be so lucky!!** From this we can derive Kepler's three laws from first principles.

Consider, for example, the Gravitational inverse square law force

$$f(r) = -\frac{\mu}{r^2}, \quad \text{so that} \quad f(u^{-1}) = -\mu u^2.$$

Equation (8.7) then becomes

$$\frac{d^2u}{d\theta^2} + u = \frac{\mu}{h^2}. \tag{8.8}$$

This is a 2nd order *linear* (yes linear!!!) ODE with exact (yes exact!!!) solution (student exercise)

$$u(\theta) = A\cos(\theta) + B\sin(\theta) + \frac{\mu}{h^2},$$

where A, B are arbitrary constants. Funky man! Thus

$$\frac{1}{r} = u = C\cos(\theta - \theta_0) + \frac{\mu}{h^2}, \tag{8.9}$$

where C, θ_0 are alternative arbitrary constants. (We can choose $C \geq 0$, with $A = C\cos(\theta_0), B = C\sin(\theta_0)$.) We are also free to choose the line $\theta = 0$ (that is, the x-axis), so without loss of generality, we take $\theta_0 = 0$. Also, we can choose

$$C = \frac{\mu e}{h^2}, \quad e > 0,$$

so that $e = Ch^2/\mu$. Therefore,

$$\frac{h^2/\mu}{r} = 1 + e\cos\theta$$

thus we have

$$r(\theta) = \frac{l}{1 + e\cos\theta}, \quad \text{with} \quad l = \frac{h^2}{\mu}. \tag{8.10}$$

Take a look at this. Does it remind you of something? We looked at this equation in Section 5. It is *exactly the polar equation of a conic section.* Another **gigantic fluke**.

If $e = 0$ the orbit is a *circle*, which we look at next, If $0 < e < 1$ the orbit is an *ellipse*. If $e > 1$ the orbit is a *hyperbola*. Only the ellipse (and the circle) lead to bounded orbits in which the planet goes round the Sun periodically.

Thus we have *proven* Kepler's first law for an inverse square law central force.

Yay!!. Aren't we clever!

NOTE Newton was the first person to prove this (with a bit of help from Halley and Hooke). In his book the *Principia* he gave a purely geometrical proof (even though he has used calculus to do it originally).

Having proven Kepler's First and Second laws from first principles, it remains for us to prove Kepler's Third law.

We saw in Chapter 5 equation (5.16) that if T is the period of the planet motion and a is the size of the semi-major axis, h is the angular momentum and l is as above then

$$T^2 = \left(\frac{4\pi^2 l}{h^2}\right) a^3.$$

But from (8.10) we now know that $l = h^2/\mu$. Substituting we have

$$T^2 = \frac{4\pi^2}{\mu} a^3. \tag{8.11}$$

Thus T^2 is *exactly proportional to* a^3 and the constant of proportionality is $4\pi^2/\mu$. Hence we have proved Kepler's Third Law. Yay again!

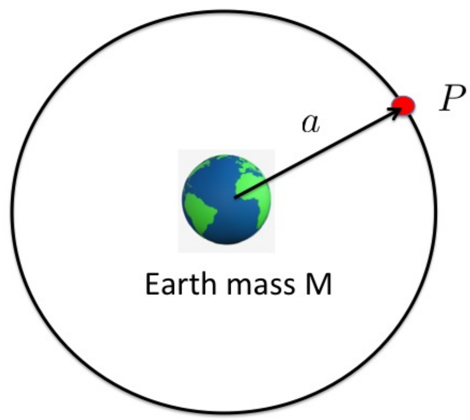
NOTE In astronomical units based around the Earth, $T = 1$ (year) and $a = 1$. In these units we see that $\mu = 4\pi^2$.

For the example of motion under an *inverse cube law* which *does not* obey Keplers Third Law, see example Sheet 10.

8.6 Circular orbits

For motion under a central force it is possible to have circular orbits. For a particle P to be in a circular orbit about a point at O (such as the centre of the Earth), we must have $r \equiv a$, a constant, and this simplifies the equations considerably.
Recall that the basic equations of motion under a central force are (8.4) and (8.5):

$$r^2\dot{\theta} = h, \quad \ddot{r} - \frac{h^2}{r^3} = f(r),$$



with h and $f(r)$ given. If $r = a$ =constant, then $\dot{r} = \ddot{r} = 0$. Hence

$$\dot{\theta} = \frac{h}{a^2} = \text{ a constant, } \omega \text{ say,}$$

and,

$$h^2 + a^3 f(a) = 0, \quad h = a^2\omega. \tag{8.12}$$

This implies

$$a\omega^2 \, + \, f(a) \, = \, 0, \tag{8.13}$$

which can only have a solution if $f(a) < 0$ (i.e. in the case of attraction). In this case,

$$\omega \, = \, \left(\frac{-f(a)}{a}\right)^{1/2},$$

i.e. a circular orbit does exist for any a with $f(a) < 0$ with the above given ω.

For example, for gravity

$$f(r) = -\mu/r^2, \quad \mu > 0$$

there exists a unique circular orbit for any a with

$$\dot{\theta} \equiv \omega = \left(\frac{\mu}{a^3}\right)^{1/2}.$$

The velocity v of a particle on this orbit is given by

$$v = a\dot{\theta} = \left(\frac{\mu}{a}\right)^{1/2} \tag{8.14}$$

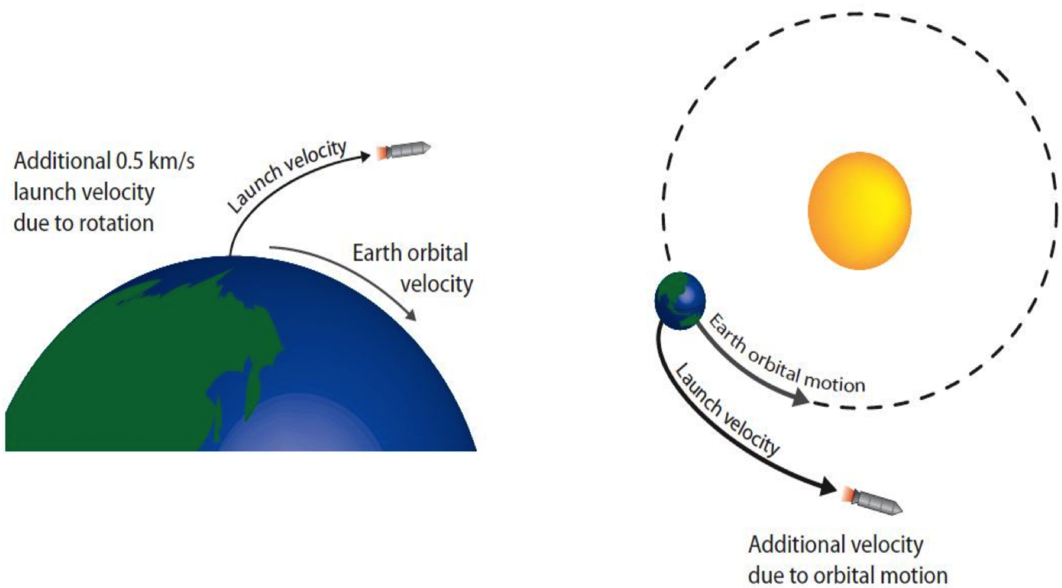
Example: *A Low Earth Orbit* Suppose that a satellite orbits close to the Earth so that a is close to $R = 6371\ km$ the radius of the Earth. From the previous study of the gravitational force, we know that if M is the mass of the Earth, G is the Gravitational constant, and $g = 9.81\ ms^{-2}$ is the acceleration due to gravity at the Earth's surface, then

$$\mu = GM \quad \text{and} \quad g = \frac{\mu}{R^2} \quad \text{so that} \quad \mu = gR^2.$$

Substituting $a = R$ into (8.14) we have

$$v = \sqrt{gR} = 7905\ ms^{-1} = 7.9\ kms^{-1}.$$

Such a satellite takes 1.4 hours to orbit the Earth. Typically to achieve this orbit, a rocket will

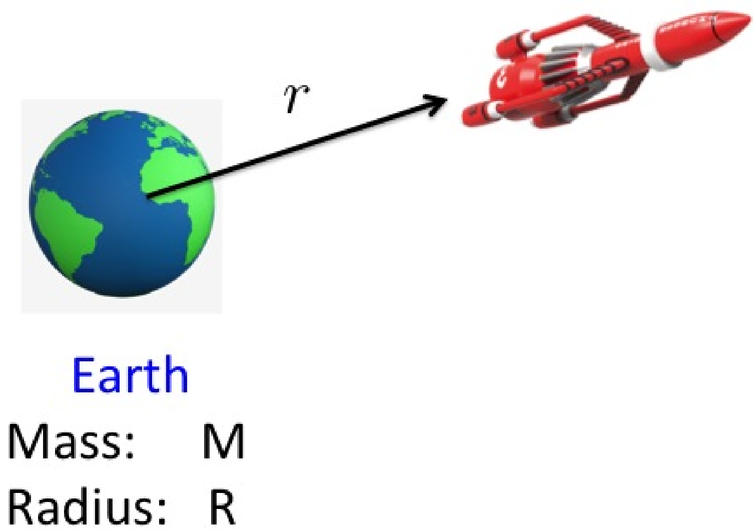


launch vertically from the Earth and then turn over to insert the satellite into the orbit. The launch will make use of the boost of $0.5\ kms^{-1}$ given by the Earth's rotation. See Figure 8.6. The very high orbital speed shows you why space craft on reentry into the Earth's atmosphere have to be specially protected to avoid burning up.

8.7 Escape velocity

We will now study the exmaple of a rocket (of unit mass $m = 1$) which takes off *vertically* (ie. **radially** in the direction $\mathbf{e}_r$ from the surface of a planet (e.g. the Earth) of radius R with speed v_0. Thus, $\dot{\mathbf{x}}(0) = v_0\mathbf{e}_r$, and from $\dot{\mathbf{x}} = \dot{r}\mathbf{e}_r + r\dot{\theta}\mathbf{e}_\theta$ the initial conditions are:

$$r = R,\ \dot{r} = v_0,\ \theta = 0,\ \dot{\theta} = 0, \quad \text{at } t = 0.$$



For the case of Newtonian gravitation,

$$\ddot{\mathbf{x}} = \frac{f(r)}{r}\mathbf{x} = -\frac{\mu}{r^3}\mathbf{x}$$

with, as before

$$f(r) = -\frac{\mu}{r^2}, \;\; \mu = GM = gR^2.$$

In polar coordinates $r^2\dot{\theta} = h$, and the orbit equation (8.5) becomes

$$\ddot{r} - \frac{h^2}{r^3} = -\frac{\mu}{r^2}. \tag{8.15}$$

The initial conditions give:

$$h = r^2(0)\,\dot{\theta}(0) = R^2 \times 0 = 0,$$

and so $r^2(t)\dot{\theta}(t) = 0$. Therefore, for $r \neq 0$,

$$\theta(t) = \text{constant} = 0,$$

since we take $\theta(0) = 0$. Hence motion is always in the radial direction. Since $h = 0$, (8.15) gives, the following differential equation for r.

$$\ddot{r} = -\frac{\mu}{r^2}.$$

We have already seen this equation. To solve it, we multiply both sides by $\dot{r}$ and integrate. This gives the energy equation

$$\frac{\dot{r}^2}{2} = \frac{\mu}{r} + E,$$

where the energy E is a constant. From the initial conditions

$$E = \frac{1}{2}v_0^2 - \frac{\mu}{R},$$

so, if $v = \dot{r}$ we have

$$v^2(r) = \frac{2\mu}{r} + v_0^2 - \frac{2\mu}{R}. \tag{8.16}$$

Note that $v(r)$ decreases with r.

We no consider the following two cases:

(i) $v_0^2 < \dfrac{2\mu}{R}$.

Equation (8.16) is

$$0 \le v^2(r) = \underbrace{\frac{2\mu}{r}}_{\to 0 \text{ as } r\to\infty} + \underbrace{v_0^2 - \frac{2\mu}{R}}_{<0}.$$

This implies

$$r \le \frac{2\mu R}{2\mu - v_0^2 R} =: r_{\max},$$

with $v^2 = 0$ at $r = r_{\max}$ and $v^2 > 0$ for $r < r_{\max}$.

Thus r cannot exceed $r_{\max}$, since then v^2 would be negative. Thus, the rocket travels out to a maximum distance of $r = r_{\max}$ and then returns to the planet. It cannot therefore escape the planet's gravitational attraction.

(ii) $v_0^2 \ge \dfrac{2\mu}{R}$.

In this case

$$v^2 = \underbrace{\frac{2\mu}{r}}_{>0} + \underbrace{v_0^2 - \frac{2\mu}{R}}_{\ge 0} > 0, \ \forall r,$$

and v is never zero. Thus the rocket never stops and r keeps increasing indefinitely, i.e. the rocket escapes the planet and can reach the distant stars. There the rocket crew can seek out new life and new civilisations, to boldly go where no Bath mathematics student has gone before.

Conclusion: To escape the planet the initial speed v_0 must be greater than (or equal to)

$$v_{\text{esc}} := \sqrt{\frac{2\mu}{R}} = \sqrt{\frac{2MG}{R}}.$$

Definition 8.3 *The* ***escape speed*** *is the speed v_{esc}, such that for $v_0 \geq v_{esc}$ a particle escapes the gravitational attraction of the planet. Similarly, the* ***escape velocity*** *is $\mathbf{v}_{esc} = v_{esc}\,\mathbf{e}_r$.*

For the earth $v_{\text{esc}} \approx 11.2 km\, s^{-1}$. This is a factor of $\sqrt{2}$ greater than the speed for a low Earth Orbit.

8.8 * Black Holes

Consider a star and assume that its radius satisfies

$$R < \frac{2MG}{c^2},$$

where c is the speed of light ($c \simeq 3 \times 10^8 km\, s^{-1}$). Since G and c are universal constants, this is a condition relating the radius of the star to its mass. Rearranging, we have

$$c < \sqrt{\frac{2MG}{R}},$$

so the escape speed is greater than the speed of light. Since not even light can escape from such a star, it will appear invisible at a distance greater than $r_{\max}$. Such a star is called a **black hole**. The value $\frac{2MG}{c^2}$ is called the "Schwarzschild radius". [For a planet with the mass of our sun to be a black hole R must be less than 1cm!]

A rough definition of a black hole is as follows: A black hole is an object in space that has collapsed under its own gravitational forces to such an extent that its escape speed is equal to the speed of light. Black holes are believed to be formed in the gravitational collapse of massive stars at the end of their life. They seem to be very common in the Universe, and massive black hols lie at the centre of galxies such as our own. The existence of black holes was first postulated by Laplace in 1798! (The above simplified analysis assumes of course that light also satisfies the Newton's laws of mechanics: the actual physical mechanism is more subtle, referring to Einstein's Theory of General Relativity and cosmology which we do not discuss here.)

8.9 *Hyperbolic Orbits

If a satellite approaches a planet at a velocity higher than the escape speed then it will go past the planet on a hyperbolic orbit. This is described by the polar equation

$$r = \frac{l}{1 + e\ \cos(\theta)}, \quad \text{with} \quad e > 1, \quad l = \frac{h^2}{\mu}.$$

A typical motion has the satellite approaching from $r = \infty$, having a closest approach when $r = l/(1 + e)$, and then departing to $r = \infty$. The lines of approach and departure are the *asymptotes* of the hyperbola which are lines at the angle $\theta = \pm\theta_\infty$ where

$$\cos(\theta_\infty) = -1/e.$$

It follows from the orbit formula that

$$\dot{r} = \frac{l\ e\ \sin(\theta)\ \dot{\theta}}{(1 + e\cos(\theta))^2} = \frac{e\ r^2\ \dot{\theta}\ \sin(\theta)}{l} = \frac{\mu\ e\ \sin(\theta)}{h}.$$

Thus at $r = \infty$ we have a speed $v_\infty = |\dot{r}|$ where

$$v_\infty = \frac{\mu}{h} \tan\left(\theta_\infty\right).$$

If we project the asymptote from infinity then its closest perpendicular distance from the planet is b. See Figure 16. If follows from the definition of the angular momentum (Exercise, check

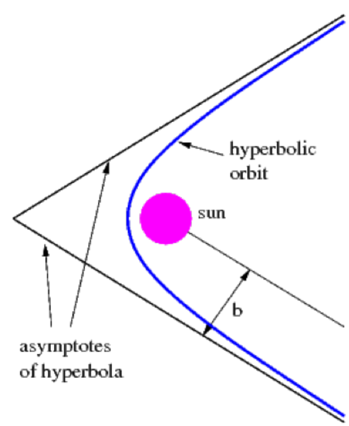


Figure 16: A typical hyperbolic orbit of a body (such as a comet or a satellite) around the Sun

this) that $h = bv_\infty$. Hence

$$\tan(\theta_\infty) = \frac{b\ v_\infty^2}{\mu}.$$

Thus if we know v_∞ and b we can calculate θ_∞. The satellite is then deflected around the planet through an angle of $2\theta_\infty$.

This effect is used to great use on long range satellite missions to distant planets in the Solar system. In particular, the satellites have *sling-shot orbits* when they are defelected around bodies such as Jupiter, and use the gravitational pull of this body to accelerate them on their orbit. An example is given by Voyager probes to Uranus and Neptune which were deflected around Jupiter and Saturn, before ending up on Star Trek the Motion Picture.

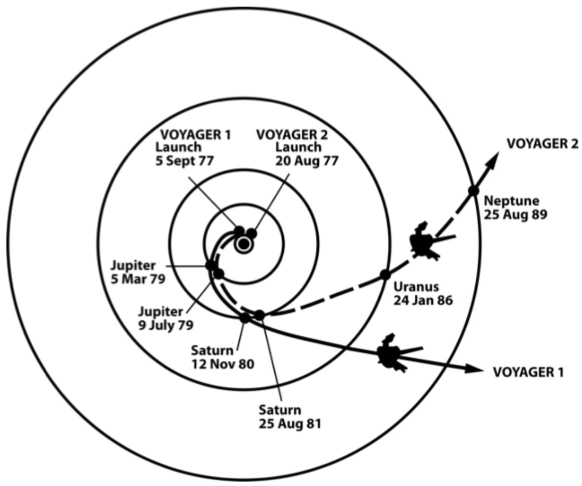


Figure 17: The sling-shot orbits of the Voyager probes.