Tomi Johnson[1]

Keble College - Michaelmas 2014
CP3&4: Mathematical methods I&II
Tutorial 3 - Vectors and matrices I

Prepare full solutions to the 'problems' with a self assessment of your progress on a cover page.
Leave these at Keble lodge by 5pm on Monday of 2nd week.
Look at the 'class problems' in preparation for the tutorial session.
Suggested reading: RHB 7 and the lecturer's problem sets.

Goals

- Start to think about vectors more abstractly than merely as arrows pointing towards points in co-ordinate space.
- Identify and close gaps in your understanding of basic operations involving three dimensional vectors, and their relation to points in co-ordinate space.
- Learn how to describe lines, planes and surfaces using vectors and vector operations.
- Become proficient in using vectors and vector operations, in advance of their use in formulating geometric and kinematic problems.

Problems

Let us start by examining mathematically what we mean by a real vector $|v\rangle$, such as could be used to describe a velocity, force, spatial translation, or magnetic field. Intuitively, a vector $|v\rangle$, e.g. a specific velocity, is one of a set $\mathcal{V}$ of many vectors, e.g. the set of all possible velocities, which we call a vector space. Again, matching our intuitions, this vector space can be supposed to satisfy a few reasonable requirements: adding two vectors gives another vector in the vector space, multiplying a vector by a real constant gives another vector in the vector space, there is a zero vector $|0\rangle$ such that $|v\rangle + |0\rangle = |v\rangle$, and there are also the usual properties of associativity, commutativity and distributivity. That's all we need to begin talking about vectors abstractly.

From its featuring in the definition of a vector space, the notion of adding some vectors to obtain another is an important one. It leads us to the idea of a complete basis $\mathcal{B}$, a set of basis vectors $|v_i\rangle$ from which all other vectors in the vector space $\mathcal{V}$ can be built via linear superposition $|v\rangle = \sum_i c_i |v_i\rangle$ (a more fancy way of saying it: the basis vectors span the vector space). This idea builds on every day intuition e.g. to get to the Sheldonian from Keble, go South down Parks Rd for 500m then West along Broad St for 50m. To avoid redundancy, we can insist that no basis vector can be built from the others via linear superposition, i.e. they are linearly independent. This is equivalent to there being no solution to $|0\rangle = \sum_i c_i |v_i\rangle$ other than the solution $c_i = 0$ and it implies that a decomposition $|v\rangle = \sum_i c_i |v_i\rangle$ is unique. A vector space $\mathcal{V}$ will permit a maximum number N of linearly independent vectors and this number will equal the minimum number of vectors required to form a basis i.e. the exact number of vectors required to form a linearly independent complete basis (all of this can be proved without introducing any more assumptions). N is called the dimension of the vector space.

[1]These problems were compiled by Prof. D. Jaksch based on problem sets by Prof. N. Harnew and past Oxford Prelims exam questions.

Given some basis $\mathcal{B}$ it is then possible to represent a vector $|v\rangle$ by its components c_i, written as a column vector

$$\mathbf{v} = \begin{pmatrix} c_1 \\ \vdots \\ c_N \end{pmatrix} .$$

To save space, column vectors are usually typeset as $\mathbf{v} = (c_1, \ldots, c_N)^T$ where ^{T} indicates the transpose. Even the transpose sign ^{T} is often left out when it is obvious the object is a column vector. Due to the difficulty of writing in bold font by hand, column vectors $\mathbf{v}$ are also written $\underline{v}$ or $\vec{v}$. It doesn't matter which you use, just be consistent. If two vectors are added, i.e. $|w\rangle = z_u |u\rangle + z_v |v\rangle$ with $\mathbf{u} = (b_1, \ldots, b_N)^T$ and $\mathbf{v} = (c_1, \ldots, c_N)^T$ in basis $\mathcal{B}$, then the column vector representing the vector sum in the same basis $\mathcal{B}$ is that with the components of the original column vectors added $\mathbf{w} = (z_u b_1 + z_v c_1, \ldots, z_u b_N + z_v c_N)^T$.

The difference between a vector $|v\rangle$ and a column vector $\mathbf{v}$ is quite subtle and often, for notational convenience, they are treated as the same thing and used interchangeably. The vector $|v\rangle$ is the more fundamental object, while the column vector $\mathbf{v}$ is a representation of that object for a particular choice of basis $\mathcal{B}$. The vector $|v\rangle$ doesn't care about which basis $\mathcal{B}$ you have decided on, but this matters to the column vector $\mathbf{v}$. Change basis $\mathcal{B}$ and your column vector representation $\mathbf{v}$ of $|v\rangle$ must change also. When writing a column vector representation $\mathbf{v}$ you should always state (if not obvious) the basis that you are using, otherwise the vector $|v\rangle$ being represented is not clear. The lecturer prefers to use the same notation for both vectors and column vectors, making it clear at all times what basis is being used.

So far we have defined a vector and this is enough to use it. However, we are used to assigning some values of length to a single vector or angle to pairs of vectors. It is clear that to do this we need an extra ingredient in addition to the vector space, so let's introduce a scalar function of two vectors $f(|v\rangle, |u\rangle) = \langle u| v\rangle$, called the inner product, and see what this enables us to do. The only single-vector quantity available is $\langle v| v\rangle$ and since we want this to give us a length we require it to be non-negative $\langle v| v\rangle \geq 0$ with equality only for $|v\rangle = |0\rangle$. This enables us to think of $||\,|v\rangle\,|| = \sqrt{\langle v| v\rangle}$ as the length (or norm) of a vector, with $|0\rangle$ having zero length but every other vector a positive length.

Let's see if we can also use this function to define an angle between two vectors. In familiar real vector spaces, the angle between vectors $|v\rangle$ and $|u\rangle$ is the same as that between $|u\rangle$ and $|v\rangle$, so the inner product might be expected to be symmetric $\langle u| v\rangle = \langle v| u\rangle$. Also, if we want length to have the basic property that two copies of a vector added have twice the length of the original vector, then we require that $f(|v\rangle, |u\rangle)$ is linear in its arguments. Amazingly, after only this, we are already in a position to define an angle between two vectors, a fact that is revealed by answering the following question.

1. Prove the Cauchy-Schwarz inequality $|\,\langle u| v\rangle\,| \leq ||\,|v\rangle\,||\;||\,|u\rangle\,||$, and determine when the equality holds.

   Solution: Divide up $|v\rangle = |v_\parallel\rangle + |v_\perp\rangle$, where $|v_\parallel\rangle = \langle v| u\rangle\,|u\rangle / ||\,|u\rangle\,||^2$ and $|v_\perp\rangle = |v\rangle - |v_\parallel\rangle$, assuming $||\,|u\rangle\,||^2 \neq 0$. Later we'll think of these as the projection of $|v\rangle$ onto the direction of $|u\rangle$ and onto the space orthogonal to $|u\rangle$, respectively. The linearity and symmetry of the scalar product means that $||\,|v\rangle\,||^2 = ||\,|v_\parallel\rangle\,||^2 + ||\,|v_\perp\rangle\,||^2 + 2\,\langle v_\parallel| v_\perp\rangle$. The same two properties also enable us to show $\langle v_\parallel| v_\perp\rangle = 0$ (later we'll come to think of this as their being orthogonal) and so $||\,|v\rangle\,||^2 = ||\,|v_\parallel\rangle\,||^2 + ||\,|v_\perp\rangle\,||^2 \geq ||\,|v_\parallel\rangle\,||^2$, where we have used the non-negativity of length and equality holds only when $|v_\perp\rangle = 0$. This rearranges to the Cauchy-Schwarz inequality. $|v_\perp\rangle = 0$ and thus the equality holds when $|v\rangle = |0\rangle$ or when $|v\rangle \propto |u\rangle$. It is also simple to show directly using linearity that it holds for the yet unexamined case $|u\rangle = |0\rangle$.

We see that, following only from the basic assumptions we have made about it, the inner product $\langle u|\, v\rangle$ is constrained to be between $-||\,|v\rangle\,||\,||\,|u\rangle\,||$, obtained when $|u\rangle = c\,|v\rangle$ with c non-positive, and $||\,|v\rangle\,||\,||\,|u\rangle\,||$, obtained when $|u\rangle = c\,|v\rangle$ with c non-negative. It follows therefore that we can think of a quantity $\langle u|\, v\rangle\,/||\,|v\rangle\,||\,||\,|u\rangle\,|| = \cos(\theta_{vu})$ as being the fractional projection of $|u\rangle$ onto $|v\rangle$ and defining an angle θ_{vu} between any two non-zero vectors that is entirely in keeping with our intuition. For example, the angle between a vector and itself is 0 and the angle between a vector and its negative is π.

In fact, all of the usual properties of vectors follow from the few simple assumptions we have made above. Here is another example of a notion we intuitively expect to be true.

2. Prove the Pythagorean theorem, i.e. $||\,|v\rangle + |u\rangle\,||^2 = ||\,|v\rangle\,||^2 + ||\,|u\rangle\,||^2$ iff $|v\rangle$ and $|u\rangle$ are orthogonal.

Solution: By the linearity of the inner product, $||\,|v\rangle + |u\rangle\,||^2 = ||\,|v\rangle\,||^2 + ||\,|u\rangle\,||^2 + 2\,\langle u|\, v\rangle$. The last term is zero iff the vectors are orthogonal, leaving Pythagoras' theorem.

3. Prove the triangle inequality $||\,|v\rangle + |u\rangle\,|| \leq ||\,|v\rangle\,|| + ||\,|u\rangle\,||$ and determine when the equality holds.

Solution: By the linearity of the inner product, $||\,|v\rangle + |u\rangle\,||^2 = ||\,|v\rangle\,||^2 + ||\,|u\rangle\,||^2 + 2\,\langle u|\, v\rangle$. Combining this with the Cauchy-Schwarz inequality gives $||\,|v\rangle + |u\rangle\,||^2 \leq ||\,|v\rangle\,||^2 + ||\,|u\rangle\,||^2 + 2||\,|v\rangle\,||\,||\,|u\rangle\,|| = (||\,|v\rangle\,|| + ||\,|u\rangle\,||)^2$. The triangle inequality follows. Equality occurs when $\langle u|\, v\rangle = ||\,|v\rangle\,||\,||\,|u\rangle\,||$, which is when $\theta_{vu} = 0$.

The notion of angle and therefore orthogonality brought about by introducing an inner product also allows us to impose an incredibly useful structure to vectors. Decomposing a vector into its orthogonal projections simplifies the mathematics and allows a clear interpretation of the parts making up this decomposition.

For this reason, when discussing a basis, it is helpful and usual to restrict ourselves to a basis comprising orthonormal basis vectors $\langle v_j|\, v_i\rangle = \delta_{ij}$. As before, any vector is decomposed as $|v\rangle = \sum_i c_i\,|v_i\rangle$, but there is a much more natural interpretation of this decomposition in the orthonormal case. Evaluating $\langle v_i|\, v\rangle$, we find it equal to the component c_i, and thus $c_i\,|v_i\rangle$ equal to $\langle v_i|\, v\rangle\,|v_i\rangle$, the projection of $|v\rangle$ onto the vector $|v_i\rangle$. The decomposition $|v\rangle = \sum_i c_i\,|v_i\rangle$ then reveals the vector is equal to the sum of its projections onto orthogonal vectors $|v_i\rangle$, with each projection having value c_i. The column vector $\mathbf{v} = (c_1, \ldots, c_N)^T$ representing $|v\rangle$ in this basis is made up of components corresponding to the various projections. This interpretation is not possible without orthonormality.

Decomposing two vectors $|v\rangle = \sum_i c_i\,|v_i\rangle$, $|u\rangle = \sum_i b_i\,|v_i\rangle$ in the same orthonormal basis, it follows that the inner product takes on the simple form $\langle u|\, v\rangle = \sum_i b_i c_i$. It is remarkable that it is only for an orthonormal basis that the inner product can be solely expressed in terms of the components in a basis (in other cases we would need to also know the inner products of the basis vectors). Thus only for orthonormal bases are we able to deduce all the properties (including lengths and angles) of vectors $|v\rangle$ solely from their column vector representations $\mathbf{v}$. This fact suggests the introduction of an operation, called $\mathbf{u}\cdot\mathbf{v}$, the scalar (or dot) product, between column vectors $\mathbf{u}$ and $\mathbf{v}$ that, for an orthonormal basis, equals the inner product $\mathbf{u}\cdot\mathbf{v} = \langle u|\, v\rangle$ of the two vectors being represented. This operation is defined by $\mathbf{u}\cdot\mathbf{v} = (b_1, \ldots, b_N)^T \cdot (c_1, \ldots, c_N)^T = \mathbf{u}^T\mathbf{v} = \sum_i b_i c_i$. Similarly, we can introduce the length (magnitude) of a column vector as $v = |\mathbf{v}| = \sqrt{\mathbf{v}^T\mathbf{v}} = ||\,|v\rangle\,||$ and obtain the angle θ_{uv} from $\cos(\theta_{uv}) = \mathbf{u}^T\mathbf{v}/uv$. To enable the obtaining of vector properties from column vector properties is a reason why we always choose orthonormal bases if possible.

Having established that, provided we use a fixed orthonormal basis, we may deal solely with column vectors $\mathbf{v}$ *representing vectors* $|v\rangle$*, we now spend the rest of this problem set using such a representation. You are probably very used to thinking of a vector, e.g. velocity, force, spatial translation, or magnetic field, by its representation in an orthonormal basis, and you may feel the previous three pages it took to arrive at the decision to use such a representation were a bit of a waste of a time. However, the notion of a vector, as used in physics, is a very general one. During your course you'll come to describe functions as vectors, and you'll deal with vectors with complex components to represent quantum mechanical states. Normal intuition will fail you, but the mathematical intuition gained through the pedantic exploration of the last few pages will hold up. More immediately, when considering changes of basis, you will appreciate the ability to retreat from the column vector representation* $\mathbf{v}$ *in a particular basis to the more fundamental vector* $|v\rangle$ *itself.*

More specifically, for the remainder of the problem set we will consider three-dimensional vector spaces. These are standard in mechanics and electromagnetism, for obvious reasons. It is always possible to write the choice of orthonormal basis as $\mathcal{B} = \{|e_1\rangle, |e_2\rangle, |e_3\rangle\}$ *so the basis vectors are represented by column vectors* $\hat{\mathbf{e}}_1 = (1,0,0)^T$*,* $\hat{\mathbf{e}}_2 = (0,1,0)^T$*,* $\hat{\mathbf{e}}_3 = (0,0,1)^T$*, where the* $\hat{}$ *indicates that the vector has unit length (is a unit vector).*

For the particular case of 3D there is a function of a pair of column vectors that is of particular usefulness. The vector (cross) product is defined in the column vector representation as

$$\mathbf{u} \times \mathbf{v} = \mathbf{u} \wedge \mathbf{v} = \begin{pmatrix} u_2 v_3 - u_3 v_2 \\ u_3 v_1 - u_1 v_3 \\ u_1 v_2 - u_2 v_1 \end{pmatrix} = \begin{vmatrix} \hat{\mathbf{e}}_1 & \hat{\mathbf{e}}_2 & \hat{\mathbf{e}}_3 \\ u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \end{vmatrix}.$$

The vector product is not associative $(\mathbf{u} \times \mathbf{v}) \times \mathbf{w} \neq \mathbf{u} \times (\mathbf{v} \times \mathbf{w})$ *and is anti-commutative* $\mathbf{u} \times \mathbf{v} = -\mathbf{v} \times \mathbf{u}$*. It is distributive* $(\mathbf{u}+\mathbf{v}) \times \mathbf{w} = \mathbf{u} \times \mathbf{w} + \mathbf{v} \times \mathbf{w}$*. The vector product can also be written as* $\mathbf{u} \times \mathbf{v} = uv|\sin(\theta_{uv})|\hat{\mathbf{n}}$ *where* $\hat{\mathbf{n}}$ *is orthogonal to both* $\mathbf{u}$ *and* $\mathbf{v}$*. Finally, also specific to 3D, the triple scalar product is given by* $\{\mathbf{uvw}\} = \mathbf{u} \cdot (\mathbf{v} \times \mathbf{w})$ *and has the useful property that* $\{\mathbf{uvw}\} = -\{\mathbf{vuw}\} = \{\mathbf{wuv}\}$*. It can be written as a determinant*

$$\{\mathbf{uvw}\} = \begin{vmatrix} u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \end{vmatrix}.$$

The problem with introducing the vector and triple scalar products above is that their meaning is not entirely independent of the choice of basis $\mathcal{B}$*. In particular, the sign of the result of the product depends on something called the handedness of* $\mathcal{B}$*. To use these products to represent physical quantities unambiguously, one needs to fix the handedness of the basis being used. We explore this in the next question.*

4. Think of describing a 3D vector space with reference to an x, y and z axis. We write the unit vectors along the positive directions of the x, y and z axis as $|i\rangle$, $|j\rangle$ and $|k\rangle$, and represent them by $\hat{\mathbf{i}}$, $\hat{\mathbf{j}}$ and $\hat{\mathbf{k}}$, respectively. One possible choice is to set $\mathcal{B} = \{|i\rangle, |j\rangle, |k\rangle\}$, respectively. Another choice is to set $\mathcal{B} = \{|j\rangle, |i\rangle, |k\rangle\}$. Show that these choices lead to different physical interpretations of the vector represented by $\mathbf{u} \times \mathbf{v}$, where $\mathbf{u} = u_x\hat{\mathbf{i}} + u_y\hat{\mathbf{j}} + u_z\hat{\mathbf{k}}$ and $\mathbf{v} = v_x\hat{\mathbf{i}} + v_y\hat{\mathbf{j}} + v_z\hat{\mathbf{k}}$. Discuss why this leads one to define two types of basis, right and left-handed, and give definitions for these types. Which orderings of $|i\rangle$, $|j\rangle$ and $|k\rangle$ are conventionally said to give a right or left-handed set?

Solution: For the two choices of basis orderings, $\mathbf{u} \times \mathbf{v}$, written in terms of $\hat{\mathbf{i}}$, $\hat{\mathbf{j}}$ and $\hat{\mathbf{k}}$, differs by a minus sign and thus has a different physical interpretation. All ordered basis choices related by rotation have the same sign and are therefore said to have the same handedness, those by reflection and rotation are of opposite handedness. All ordered bases whose ordered scalar product is 1 have the same handedness and all resulting in -1 have the opposite handedness. Convention chooses $\mathcal{B} = \{|i\rangle, |j\rangle, |k\rangle\}$ to be called right-handed and $\{\hat{\mathbf{i}}\hat{\mathbf{j}}\hat{\mathbf{k}}\} = 1$. Thus $\{|k\rangle, |i\rangle, |j\rangle\}$ and $\{|j\rangle, |k\rangle, |i\rangle\}$ are also right-handed, with the other three orderings left-handed.

It boils down to this: if you represent physics in a 3D vector space and you wish to use the vector or triple scalar product then you need to fix the handedness of your coordinate system. It matters. The usual way to resolve the handedness ambiguity is to fix/always choose a right-handed basis. Many physical relationships involving the vector or triple scalar product, e.g. $m\ddot{\mathbf{r}} = q\dot{\mathbf{r}} \times \mathbf{B}$ *for the motion of a particle in a magnetic field, assume a right-handed basis is being used, often without mentioning it, and would be incorrect if a left-handed basis was used. Basically, never use a left-handed coordinate system and you can forget about this complication!*

Let us now specialise to thinking about the particular case of three-dimensional vectors represented by column vectors (bold roman letters) in a fixed right-handed basis $\mathcal{B}$. *We will think of each column vector corresponding to a point (capital roman letter) in 3D coordinate space or, equivalently, arrows (drawn) from the origin O to the point. As is common, we will often use slightly imprecise language and treat the vector, column vector, point and arrow as equivalent for the ease of discussion. By answering the next two questions you will reveal the usefulness of the language of vectors in solving problems in coordinate space.*

5. Points A and B correspond to column vectors $\mathbf{a}$ and $\mathbf{b}$. Find the column vector $\mathbf{g}$ corresponding to the midpoint G of the straight line connecting the two vectors.

Solution: $\mathbf{g} = (\mathbf{a} + \mathbf{b})/2$.

6. The vertices of triangle ABC correspond to column vectors $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$. Find the column vector $\mathbf{g}$ corresponding to the centroid G of the triangle. The centroid is the intersection of the three lines from each vertex to the centre of the opposite side.

Solution: $\mathbf{g} = (\mathbf{a} + \mathbf{b} + \mathbf{c})/3$.

Let's go through two of the common objects, lines and planes, we deal with in coordinate space and explore how they are represented and analysed using the language of vectors. Firstly, consider a line. All points $\mathbf{r}$ *on a line in 3D space fulfill* $\mathbf{r} = \mathbf{a} + \lambda\mathbf{b}$ *where* $\mathbf{a}$ *is a point on the line,* $\mathbf{b}$ *points in the direction of the line, and* λ *is a real parameter. Implicitly this equation can be rewritten as* $(r_x - a_x)/b_x = (r_y - a_y)/b_y = (r_y - a_z)/b_z$. *Get to grips with lines in the next three questions.*

7. Show that the points $(1, 0, 1)$, $(1, 1, 0)$ and $(1, -3, 4)$ lie on a line. Give the equation of the line in the form $\mathbf{r} = \mathbf{a} + \lambda\mathbf{b}$.

Solution: The line through the first two points is given by $\mathbf{a} = (1, 1, 0)$, $\mathbf{b} = (0, 1, -1)$. It goes through the third point for $\lambda = -4$.

8. Two small objects travel with equal speed. The first starts from the point $(2, 3, 3)$ and travels in the direction of $(-1, -1, 0)$, while the second starts at the same time from $(3, 2, 1)$ and travels in the direction of $(-2, 0, 2)$. Determine whether or not they collide.

Solution: Their paths intersect but they do not collide.

9. Derive an expression for the shortest distance ℓ between the two non-parallel lines $\mathbf{r}_i = \mathbf{a}_i + \lambda_i\mathbf{b}_i$, for $i = 1, 2$. Hence find the shortest distance between the lines

$$\frac{x-2}{2} = y - 3 = \frac{z+1}{2} \quad \text{and} \quad x + 2 = \frac{y+1}{2} = z - 1\,.$$

Solution: $\ell = |\{(\mathbf{a}_1 - \mathbf{a}_2)\mathbf{b}_1\mathbf{b}_2\}|/|\mathbf{b}_1 \times \mathbf{b}_2|$. For the given values $\ell = \sqrt{18}$.

The points $\mathbf{r}$ on a plane in 3D are given by $\mathbf{r} = \mathbf{a} + \lambda\mathbf{b} + \mu\mathbf{c}$ with $\mathbf{a}$ a point on the plane, $\mathbf{b}$ and $\mathbf{c}$ vectors parallel to the plane, and λ and μ real parameters. The implicit form of this equation is $\mathbf{r} \cdot \hat{\mathbf{n}} = d$ with $\hat{\mathbf{n}}$ a unit vector normal to the plane and d the shortest distance from the plane to the origin O.

10. Find the shortest distance d_{P} from a plane to a point P corresponding to column vector $\mathbf{p}$.

Solution: $d_{\mathrm{P}} = \mathbf{p} \cdot \hat{\mathbf{n}} - d$.

11. Find the equation of the line passing through $\mathbf{a} = (1, 2, 3)$ perpendicular to the plane $x - 2y + z = 1$.

Solution: $\mathbf{r} = (1, 2, 3) + \lambda(1, -2, 1)$.

The question of whether two vectors are linearly dependent is simple. Are they identical up to a multiplicative factor? Further, in three dimensions, it is impossible to have four linearly independent vectors. Hence, for 3D coordinate space, it is usual to discuss linear independence for sets of three vectors and it boils down to a question of coplanarity, are all the vectors parallel to the same plane? Let's explore this in the next two questions.

12. Show that three vectors $\mathbf{a}$, $\mathbf{b}$ and $\mathbf{c}$ are linearly dependent iff they are coplanar and thus cannot form a basis.

Solution: If $\mathbf{b} = \mu\mathbf{c}$, with some constant μ, the vectors are linearly dependent and all three vectors are parallel to the plane defined by plane vectors $\mathbf{a}$ and $\mathbf{b}$ (or $\mathbf{a}$ and $\mathbf{c}$), i.e. coplanar. If instead $\mathbf{b} \neq \mu\mathbf{c}$, then every vector parallel to the plane defined by plane vectors $\mathbf{b}$ and $\mathbf{c}$ can be written as a linear combination of the form $\mu\mathbf{b} + \lambda\mathbf{c}$. Since the three vectors are co-planar iff $\mathbf{a}$ is parallel to the plane, we can thus write that the vectors are coplanar iff $\mathbf{a} = \mu\mathbf{b} + \lambda\mathbf{c}$ i.e. they are linearly dependent.

13. Prove that the three vectors $\mathbf{a} = \hat{\mathbf{i}} + \hat{\mathbf{j}} + 2\hat{\mathbf{k}}$, $\mathbf{b} = 2\hat{\mathbf{i}} + 3\hat{\mathbf{j}} - \hat{\mathbf{k}}$, $\mathbf{c} = \hat{\mathbf{i}} + 2\hat{\mathbf{j}} - 3\hat{\mathbf{k}}$ are coplanar.

Solution: The triple scalar product is $\{\mathbf{abc}\} = 0$ and hence they all lie in a plane (also containing the origin).

Finally, away from the physical discussion of lines and planes etc. sometime it is worth just getting proficient at solving abstract vector equations. This and some of the class problems are good examples.

14. If $\mathbf{a} \times \mathbf{b} = \mathbf{a} - \mathbf{b}$, prove that $\mathbf{a} = \mathbf{b}$.

Class problems

15. Show that the three vectors $\hat{\mathbf{i}}$, $\hat{\mathbf{j}}$ and $\hat{\mathbf{k}}$ form an orthogonal basis for all points in 3D co-ordinate space.

16. Find a vector along the line of intersection of the planes $x + 3y - z = 5$ and $2x - 2y + 4z = 3$.

Solution: Any multiple of $\mathbf{b} = (10, -6, -8)^T$.

17. A line intersects a plane at an angle $\alpha = \pi/6$. The line is defined by $\mathbf{r} = \mu\hat{\mathbf{n}}$ and the plane by $\mathbf{r} \cdot \hat{\mathbf{m}} = 0$, with $\hat{\mathbf{n}}$ and $\hat{\mathbf{m}}$ unit vectors. Calculate the shortest distance ℓ from the plane to the point on the line with $\mu = 2$.

Solution: $\ell = 1$ if the angle is taken between line and plane and $\ell = \sqrt{3}$ if the angle is taken between the line and the normal to the plane.

18. Calculate the shortest distance between the planes $x + 2y + 3z = 1$ and $x + 2y + 3z = 5$.

19. Identify the following surfaces:

(a) $|\mathbf{r}| = k$

(b) $\mathbf{r} \cdot \hat{\mathbf{u}} = \ell$

(c) $\mathbf{r} \cdot \hat{\mathbf{u}} = m|\mathbf{r}|$ for $-1 \leq m \leq +1$.

(d) $|\mathbf{r} - (\mathbf{r} \cdot \hat{\mathbf{u}})\hat{\mathbf{u}}| = n$.

Here k, ℓ, m and n are fixed scalars and $\hat{\mathbf{u}}$ is a fixed unit vector.

20. Which of the following set of vectors are i) linearly independent and ii) orthogonal (or both) and explain why:

(a) $\mathbf{a} = (0, 1, 0)^T$, $\mathbf{b} = (1, 0, 0)^T$, and $\mathbf{c} = (0, 0, 1)^T$,

(b) $\mathbf{a} = (0, 1, 1)^T$, $\mathbf{b} = (1, 1, 1)^T$, and $\mathbf{c} = (0, 0, 1)^T$,

(c) $\mathbf{a} = (1, 1, 1)^T$, $\mathbf{b} = (1, -1, 1)^T$, and $\mathbf{c} = (1, 1, -2)^T$,

(d) $\mathbf{a} = (1, 0, 1)^T$, $\mathbf{b} = (2, 3, 1)^T$, and $\mathbf{c} = (1, 6, -1)^T$.

21. Show which of the following statements about general vectors $\mathbf{a}$, $\mathbf{b}$ and $\mathbf{c}$ are true and which are false (with $c = |\mathbf{c}|$):

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

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

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

(d) $\mathbf{d} = \lambda\mathbf{a} + \mu\mathbf{b}$ implies $(\mathbf{a} \times \mathbf{b}) \cdot \mathbf{d} = 0$,

(e) $\mathbf{a} \times \mathbf{c} = \mathbf{b} \times \mathbf{c}$ implies $\mathbf{c} \cdot \mathbf{a} - \mathbf{c} \cdot \mathbf{b} = c|\mathbf{a} - \mathbf{b}|$,

(f) $(\mathbf{a} \times \mathbf{b}) \times (\mathbf{c} \times \mathbf{b}) = \mathbf{b}[\mathbf{b} \cdot (\mathbf{c} \times \mathbf{a})]$.

22. For a given ordered set of vectors $\{\mathbf{a}, \mathbf{b}, \mathbf{c}\}$ cyclic permutations are $\{\mathbf{a}, \mathbf{b}, \mathbf{c}\}$, $\{\mathbf{b}, \mathbf{c}, \mathbf{a}\}$ and $\{\mathbf{c}, \mathbf{a}, \mathbf{b}\}$ while anti-cyclic permutations are $\{\mathbf{a}, \mathbf{c}, \mathbf{b}\}$, $\{\mathbf{b}, \mathbf{a}, \mathbf{c}\}$, and $\{\mathbf{c}, \mathbf{b}, \mathbf{a}\}$. How many permutations are needed to get from the initial set of vectors to a cyclic permutation and how many to get to an anti-cyclic permutation? How can we make use of these definitions in the above examples?

In the near future you'll be studying solid state physics, where atoms are located at a regular translationally symmetric lattice points $\mathbf{R} = l\mathbf{a} + m\mathbf{b} + n\mathbf{c}$, *for integer* l, m *and* n. *This structure is revealed by scattering waves off the the lattice and observing peaks when waves are scattered by momenta* $\mathbf{K}$ *satisfying* $\exp(\mathrm{i}\mathbf{K} \cdot \mathbf{R}) = 1$. *It is possible to write* $\mathbf{K} = h\mathbf{a}' + k\mathbf{b}' + l\mathbf{c}'$, *for integer* h, k *and* l, *where* $\mathbf{a}'$, $\mathbf{b}'$ *and* $\mathbf{c}'$ *are called reciprocal lattice vectors. It's worth spending one question getting to grips with the definition of these.*

23. The vectors $\mathbf{a}$, $\mathbf{b}$ and $\mathbf{c}$ are not coplanar. Verify that the expressions

$$\mathbf{a}' = \frac{2\pi\mathbf{b} \times \mathbf{c}}{\{\mathbf{abc}\}}, \qquad \mathbf{b}' = \frac{2\pi\mathbf{c} \times \mathbf{a}}{\{\mathbf{abc}\}}, \qquad \mathbf{c}' = \frac{2\pi\mathbf{a} \times \mathbf{b}}{\{\mathbf{abc}\}},$$

define a valid set of reciprocal vectors $\mathbf{a}'$, $\mathbf{b}'$ and $\mathbf{c}'$ by confirming the following properties:

(a) $\mathbf{a}' \cdot \mathbf{a} = \mathbf{b}' \cdot \mathbf{b} = \mathbf{c}' \cdot \mathbf{c} = 2\pi$,

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

(c) $\{\mathbf{a}'\mathbf{b}'\mathbf{c}'\} = 2\pi/\{\mathbf{abc}\}$,

(d) $\mathbf{a} = (2\pi \mathbf{b}' \times \mathbf{c}')/\{\mathbf{a}'\mathbf{b}'\mathbf{c}'\}$.

24. Figure 1 shows a Mathematica script attempting to calculate of the length of a column vector with complex components. The built-in Mathematica function `Norm` gives a different result for its length than naively applying definition above for real column vectors. Explore the difference between the two ways of calculating the length of a column vector and discuss which is a better approach.

```
In[1]:= a = {i, 2, 4 + 2 i};
```

We wish to work out the length of this vector

```
In[2]:= Print["|a| = ", Sqrt[a.a] // Simplify]

        |a| = √(15 + 16 i)
```

When using the built in *Mathematica* function the result is different

```
In[3]:= Print["|a| = ", Norm[a] // Simplify]

        |a| = 5
```

Figure 1: Calculating the length of a complex vector in Mathematica.

There's nothing in our definition of a vector space $\mathcal{V}$ that stops us from allowing them to be built using complex rather than real numbers: letting $z|v\rangle$ be a vector in the space if $|v\rangle$ is, where z may be complex. In this case we'll be dealing with components c_i in $|v\rangle = \sum_i c_i |v_i\rangle$ that could be complex. This last question highlights that if we allow complex numbers then something needs to change in our definition of an inner product $f(|v\rangle, |u\rangle) = \langle u|\, v\rangle$. The first thing we wanted from an inner product was a concept of length. Nothing changes here: The only single-vector quantity available is $\langle v|\, v\rangle$ and since we want this to give us a length we require it to be non-negative $\langle v|\, v\rangle \geq 0$ with equality only for $|v\rangle = |0\rangle$. This enables us to think of $|| \, |v\rangle \, || = \sqrt{\langle v|\, v\rangle}$ as the length (or norm) of a vector, with $|0\rangle$ having zero length but every other vector a positive length. However, it now becomes impossible to insist that $f(|v\rangle, |u\rangle)$ is linear in both its first and second arguments as then $f(z|v\rangle, z|v\rangle) = z^2 f(|v\rangle, |v\rangle)$, i.e. $||z|v\rangle|| = z|| \, |v\rangle \, ||$, would mean that both could not be real for arbitrary complex z. The simplest thing to do to fix this is let the inner product be linear in its first argument but conjugate linear in its second $f(|v\rangle, z_u|u\rangle + z_w|w\rangle) = z_u^* f(|v\rangle, |u\rangle) + z_w^* f(|v\rangle, |w\rangle)$. Alternatively one can say that $f(|v\rangle, |u\rangle)$ is linear in its first argument and $f(|u\rangle, |v\rangle) = (f(|v\rangle, |u\rangle))^*$, meaning it is not symmetric. For real numbers this makes no difference, so we tend to use the complex vector space notation anyway. If we do have complex numbers it means that $f(z|v\rangle, z|v\rangle) = |z|^2 f(|v\rangle, |u\rangle)$, i.e. $||z|v\rangle|| = |z|| \, |v\rangle \, ||$, thus solving our problem. The only other thing it changes is that decomposing two vectors $|v\rangle = \sum_i c_i |v_i\rangle$, $|u\rangle = \sum_i b_i |v_i\rangle$ in the same orthonormal basis, it follows that the inner product is $\langle u|\, v\rangle = \sum_i b_i^* c_i$. We can let this define anew what we mean by the scalar product $\mathbf{u} \cdot \mathbf{v} = (b_1, \ldots, b_N)^T \cdot (c_1, \ldots, c_N)^T = \mathbf{u}^\dagger \mathbf{v} = \sum_i b_i^* c_i$, where we have defined the symbol $\dagger$ to meant the conjugate transpose (transpose the elements and then conjugate them).

In the next vectors and matrices problem set we'll work entirely with complex vector spaces. Why bother generalising in this way? It turns out that using complex numbers make things more concise and some mathematics, e.g. of quantum mechanics, is most easily expressed by making use of them (it is possible to represent quantum mechanics using only real numbers, since all measurable results will be real, but it's a total nightmare that is never attempted other than by those who just wanted to check it could be done).