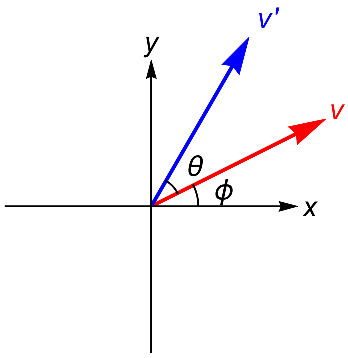
LECTURE 5

Linear transformations of vectors

We have seen how to change the length of a vector by multiplying by a scalar. Often we want to do a more interesting transformation (for example, reflect a vector in some plane). To do this, we'll introduce the idea of a linear transformation — in particular the idea of a rotation transformation — which will lead us on to the subject of matrices.

Rotations in the plane

Given a vector $\mathbf{v}$ in the xy plane, we can generate a new vector $\mathbf{v}'$ by rotating $\mathbf{v}$ through some angle θ. We will use the convention that a positive angle corresponds to an *anti-clockwise* rotation.



It is a matter of trigonometry to work out how the components of $\mathbf{v}'$ are related to those of $\mathbf{v}$. The starting vector has length v $(= |\mathbf{v}|)$ and is at some angle ϕ, relative to the positive x-axis. Hence

$$\begin{aligned}\mathbf{v} &= v_x\,\mathbf{i} + v_y\,\mathbf{j}\\ &= v\cos\phi\,\mathbf{i} + v\sin\phi\,\mathbf{j}.\end{aligned}$$

The rotated vector has the same length, but is at angle $\phi + \theta$, so

$$\mathbf{v}' = v\cos(\phi+\theta)\,\mathbf{i} + v\sin(\phi+\theta)\,\mathbf{j}.$$

Using the angle addition formulae, this becomes

$$\mathbf{v}' = v(\cos\phi\cos\theta - \sin\phi\sin\theta)\,\mathbf{i} + v(\cos\phi\sin\theta + \sin\phi\cos\theta)\,\mathbf{j}.$$

and, by recognising that $v\cos\phi = v_x$, etc., this becomes

$$\mathbf{v}' = (v_x\cos\theta - v_y\sin\theta)\,\mathbf{i} + (v_x\sin\theta + v_y\cos\theta)\,\mathbf{j}.$$

So the components of the rotated vector are related to those of the original vector by

$$v'_x = v_x \cos\theta - v_y \sin\theta \tag{5.1a}$$
$$v'_y = v_x \sin\theta + v_y \cos\theta. \tag{5.1b}$$

This is an example of a *linear transformation*: the new components depend linearly on the old ones, with coefficients that depend only on the form of the transformation itself.

It is a pain to have to write out these equations in full every time we want to work with a rotation. So we introduce the concept of a *matrix*. Doing this has two major advantages: a) it simplifies the notation, and b) it 'separates' the object doing the transformation from the object being transformed.

Linear transformations between two-dimensional vectors

The rotation transformation above is an example of a *linear transformation* of a vector. Every linear transformation from one two-dimensional vector to another takes the following form:

Linear transformations can, in general, transform an n-dimensional vector into an m-dimensional vector. For the most part, the transformations we use in chemistry are between vectors of the same dimension.

$$v'_1 = a_{11}v_1 + a_{12}v_2$$
$$v'_2 = a_{21}v_1 + a_{22}v_2$$

The numbers a_{ij} control the transformation. Notice that we have switched to numbering the vector components, rather than labelling them by x and y, i.e.

$$\mathbf{v} = v_1\,\mathbf{i} + v_2\,\mathbf{j}$$

(and likewise for $\mathbf{v}'$).

Looking at these equations, there is a pattern, as shown by the colour coding below:

$$v'_1 = a_{11}v_1 + a_{12}v_2$$
$$v'_2 = a_{21}v_1 + a_{22}v_2.$$

The two equations can be summarised as

$$v'_i = a_{i1}v_1 + a_{i2}v_2,$$

or, equivalently,

$$v'_i = \sum_{j=1}^{2} a_{ij}v_j. \tag{5.2}$$

It's crucial to understand this expression, in summation notation, because many of the later proofs involving matrices will rely on it.

Equation (5.2) shows how the individual components of the vectors are related by the transformation. But it is a rather ugly expression, and it is fiddly to write out all the subscripts. So we will switch to matrix notation, which shows how the vectors are related as a whole.

Let us start by viewing the vector $\mathbf{v}'$ as a column vector. Using (5.2), it follows that

$$\begin{pmatrix} v'_1 \\ v'_2 \end{pmatrix} = \begin{pmatrix} a_{11}v_1 + a_{12}v_2 \\ a_{21}v_1 + a_{22}v_2 \end{pmatrix}.$$

Then we say that the expression on the right-hand side is the product of a 2×2 *matrix* and the original column vector, i.e.

$$\begin{pmatrix} v'_1 \\ v'_2 \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}. \tag{5.3}$$

In other words, we *define* the matrix-vector product above such that

One way to view the matrix-vector product is as a collection of *dot products* of the rows of the matrix and the vector it is multiplying.

$$\boxed{\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} a_{11}v_1 + a_{12}v_2 \\ a_{21}v_1 + a_{22}v_2 \end{pmatrix}.} \tag{5.4}$$

Then we can write the linear transformation of (5.3) as

$$\mathbf{v}' = \mathsf{A}\mathbf{v} \tag{5.5}$$

where

$$\mathsf{A} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}.$$

Please remember: this is just *an alternative way to denote the linear transformation*, without the hassle of specifying all the subscripts. The definition of the matrix-vector product in (5.4) simply gives us a convenient way to write the sum in (5.2).

Example 1. Find the vector $\mathbf{v}' = \mathsf{A}\mathbf{v}$ where

$$\mathsf{A} = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \qquad \mathbf{v} = \begin{pmatrix} 5 \\ 6 \end{pmatrix}.$$

The v'_1 component is obtained by multiplying row 1 of A by $\mathbf{v}$ and summing the terms:

$$\begin{pmatrix} \boxed{1 \quad 2} \\ 3 \quad 4 \end{pmatrix} \begin{pmatrix} 5 \\ 6 \end{pmatrix} = \begin{pmatrix} \boxed{1 \times 5 + 2 \times 6} \\ \cdot \end{pmatrix}$$

The v'_2 component is obtained in the same way, using row 2 of A:

$$\begin{pmatrix} 1 \quad 2 \\ \boxed{3 \quad 4} \end{pmatrix} \begin{pmatrix} 5 \\ 6 \end{pmatrix} = \begin{pmatrix} 17 \\ \boxed{3 \times 5 + 4 \times 6} \end{pmatrix}$$

So

$$\mathbf{v}' = \begin{pmatrix} 17 \\ 39 \end{pmatrix}.$$

The approach can be generalised in an obvious way to transform three-dimensional vectors. The column vectors have three components, and the matrix A becomes a 3×3

matrix of the form

$$\mathsf{A} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}.$$

Finally, note that a linear transformation has the following key property:

$$\mathsf{A}(\alpha\mathbf{u} + \beta\mathbf{v}) = \alpha\mathsf{A}\mathbf{u} + \beta\mathsf{A}\mathbf{v}.$$

You will show this as an exercise.

Examples of linear transformations used in chemistry

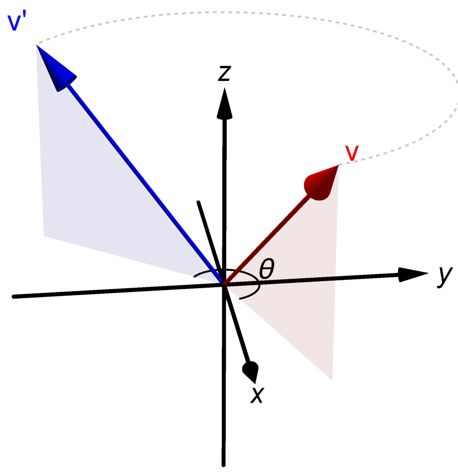
We started by considering the rotation transformation in the xy plane, (5.1). In matrix notation, this reads

$$\mathbf{v}' = \mathsf{R}\mathbf{v}.$$

where

$$\mathsf{R} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

Three dimensional vectors can also be rotated, of course, but the extra dimension makes things rather more complicated. One way to specify a rotation in three dimensions is by giving the *axis* of rotation. For example, we can rotate $\mathbf{v}$ about the z-axis by an angle θ, as shown below.



This rotation is performed by

$$\mathsf{R}_z = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

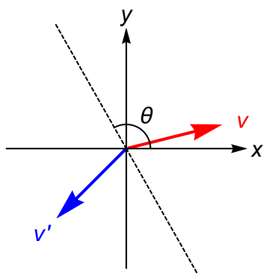
To understand why, look at what happens when it is applied to $\mathbf{v}$:

$$\mathsf{R}_z\mathbf{v} = \begin{pmatrix} v_1\cos\theta - v_2\sin\theta \\ v_1\sin\theta + v_2\cos\theta \\ v_3 \end{pmatrix}.$$

The z-component of the vector is unchanged by a rotation about the z-axis. Only the x- and y-components change, and they do so in the same way as for the two-dimensional rotation [see (5.1)].

Another transformation of importance to chemistry is the *reflection transformation*. As an exercise, you will show that the matrix

$$\begin{pmatrix} \cos 2\theta & \sin 2\theta \\ \sin 2\theta & -\cos 2\theta \end{pmatrix}$$



reflects the two-dimensional vector $\mathbf{v}$ in a line through the origin at an angle θ to the x-axis. So, for example, reflection in the x-axis (i.e. $\theta = 0$) is performed by

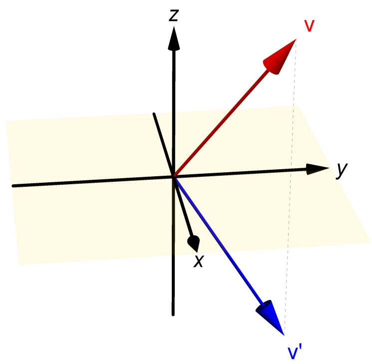
$$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$

while reflection in the line $y = x$ (i.e. $\theta = \pi/4$) is performed by

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

In three dimensions, we reflect vectors in a plane. For example, the matrix that reflects $\mathbf{v}$ in the xy plane is

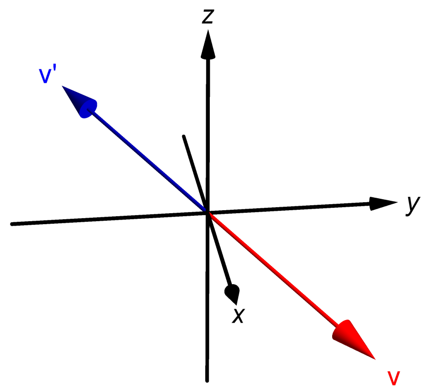
$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$



Of course, this transformation can be performed simply by scalar multiplication by -1, but it is helpful to include it as a matrix transformation too.

A third important matrix is the *inversion* matrix, which changes the signs of all the components of $\mathbf{v}$.

$$\begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$



And, last, but by no means least, it is useful to introduce a transformation that does nothing to $\mathbf{v}$. We define the $n \times n$ **identity matrix** I_n to be the matrix that has 1 along the 'main diagonal', and 0 everywhere else. For example,

$$\mathrm{I}_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Then, $\mathrm{I}_3\mathbf{v} = \mathbf{v}$ for any three-dimensional vector $\mathbf{v}$ (and, of course, similar results hold in all other dimensions).

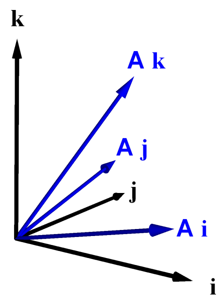
Volume scaling factor of a linear transformation

It is instructive to apply the general 3×3 matrix A to the basis vector $\mathbf{i}$:

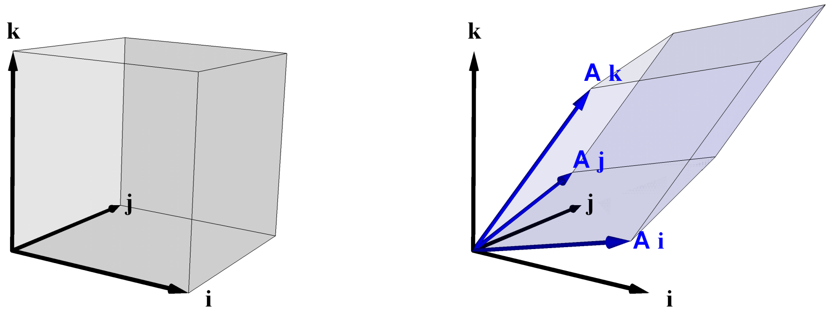
$$\mathsf{A}\,\mathbf{i} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} a_{11} \\ a_{21} \\ a_{31} \end{pmatrix}.$$

Notice that $\mathsf{A}\,\mathbf{i}$ is simply the column vector obtained from column 1 of A. Similarly, $\mathsf{A}\,\mathbf{j}$ and $\mathsf{A}\,\mathbf{k}$ are vectors corresponding to column 2 and 3, respectively, of A.

Viewed geometrically, $\mathbf{A}$ transforms the basis vectors into a set of three new vectors, as illustrated below:



The original basis vectors define a cube of side 1 (and therefore of volume 1). The transformed vectors define a *parallelepiped* of volume V. Hence the two volumes are related by a scaling factor of $V/1 = V$.



As seen previously, the volume of a paralellopiped is the modulus of the triple scalar product

$$V = |(\mathsf{A}\,\mathbf{i}) \cdot (\mathsf{A}\,\mathbf{j}) \times (\mathsf{A}\,\mathbf{k})|\,,$$

which can be written as the determinant constructed from the three vectors as rows. But we argued above that the three vectors are just the columns of A. Hence

We use here that the determinant is unchanged by transposing rows and columns.

$$(\mathsf{A}\,\mathbf{i}) \cdot (\mathsf{A}\,\mathbf{j}) \times (\mathsf{A}\,\mathbf{k}) = \begin{vmatrix} a_{11} & a_{21} & a_{31} \\ a_{12} & a_{22} & a_{32} \\ a_{13} & a_{23} & a_{33} \end{vmatrix} = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = \det \mathsf{A}, \qquad (5.6)$$

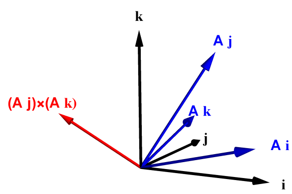
where '$\det \mathsf{A}$' denotes the determinant formed from the matrix A.

Putting all this together, we obtain a geometric interpretation of the determinant of a matrix:

> $|\det \mathsf{A}|$ *is the volume scaling factor of the transformation performed by* A.

Note that $\det \mathsf{A}$ itself can be positive or negative. The sign tells us the *orientation* of the transformation. If the A transforms the right-handed **i**, **j**, **k** basis to a left-handed set of vectors, $\det \mathsf{A}$ will be negative.

To see this, look at the following diagram.



In this example, the transformed vectors are left-handed. The cross product $(\mathsf{A}\,\mathbf{j}) \times (\mathsf{A}\,\mathbf{k})$ points away from $(\mathsf{A}\,\mathbf{i})$ (i.e. the angle between the two is larger than $\pi/2$). Hence, by thinking about the dot product in (5.6), $\det \mathsf{A}$ must be negative.

Example 2. Find the volume scaling factor of the rotation around the z-axis, performed by R_z.

$$\det \mathsf{R}_z = \begin{vmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{vmatrix} = 1 \times (\cos^2\theta + \sin^2\theta) = 1.$$

Therefore the volume scaling factor is the modulus of 1, i.e. 1. This makes sense, of course, given that the transformation is a rotation. Intuitively, we don't expect a rotation of an object to affect its volume.

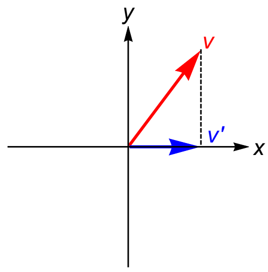
Invertible transformations

An obvious question arises: if A transforms $\mathbf{v}$ into $\mathbf{v}'$, can we find another matrix to transform $\mathbf{v}'$ back to $\mathbf{v}$?

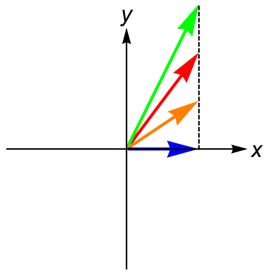
The answer is: sometimes.

It is not always possible to transform back again. A diagram will show us why. Consider the following transformation in two dimensions, which *projects* $\mathbf{v}$ onto the x-axis:

$$\begin{pmatrix} x \\ 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$



It is not possible to reverse this transformation, because *all* vectors $\mathbf{v}$ with the same x component are transformed to the *same* $\mathbf{v}'$. See the diagram below.



So we can't go *backwards* from $\mathbf{v}'$ to a particular $\mathbf{v}$, because we wouldn't know which vector to go back to!

We say that this transformation is not *invertible*.

In a sense, the forward transformation 'loses information'. The original vector has two independent coordinates, whereas the transformed vector only has one. (Its second coordinate is always zero.) So we can't reconstruct the original vector from this single piece of information.

Let's see how this information loss arises algebraically. Suppose that the original vector is

$$\mathbf{v} = v_1\,\mathbf{i} + v_2\,\mathbf{j} + v_3\,\mathbf{k}.$$

It has three separate pieces of information: the three components v_1, v_2 and v_3. Under the transformation, $\mathbf{v}$ turns into

$$\mathbf{v}' = \mathsf{A}\mathbf{v} = v_1\mathsf{A}\,\mathbf{i} + v_2\mathsf{A}\,\mathbf{j} + v_3\mathsf{A}\,\mathbf{k}. \tag{5.7}$$

Now, imagine a scenario where $\mathsf{A}\,\mathbf{j}$ happens to equal $2\mathsf{A}\,\mathbf{k}$. Then we could write

$$\begin{aligned}\mathbf{v}' &= v_1 \mathsf{A}\,\mathbf{i} + v_2 \times 2\mathsf{A}\,\mathbf{k} + v_3 \mathsf{A}\,\mathbf{k} \\ &= v_1 \mathsf{A}\,\mathbf{i} + (v_2 + 2v_3)\mathsf{A}\,\mathbf{k}.\end{aligned}$$

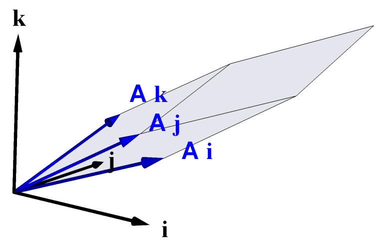
In this scenario, all vectors $\mathbf{v}$ having the same $(v_2 + 2v_3)$ will be transformed onto the same $\mathbf{v}'$. The transformation will therefore not be invertible.

For the transformation to be invertible, therefore, the three vectors $\mathsf{A}\,\mathbf{i}$, $\mathsf{A}\,\mathbf{j}$ and $\mathsf{A}\,\mathbf{k}$ need to be linearly independent. Otherwise, there will be information loss as in the scenario above.

But remember that the transformed vectors $\mathsf{A}\,\mathbf{i}$ etc. are just the columns of A. These columns will be linearly-independent if (and only if) $\det \mathsf{A}$ is non-zero. Hence,

> The transformation by A is invertible if, and only if, $\det \mathbf{A} \neq 0$.

This makes sense geometrically. A transformation with zero determinant will 'squash' the $\mathbf{i}$, $\mathbf{j}$ and $\mathbf{k}$ basis vectors into a space of fewer dimensions (because the parallelepiped has zero volume). Since $\mathbf{v}'$ is a linear combination of these 'squashed' vectors [see (5.7)], all possible vectors $\mathbf{v}'$ will live in this reduced dimensional space.



Once a vector has been transformed into this reduced dimensional space, it loses some of the information needed to transform it back again.

For the projection transformation, discussed at the beginning of the section, all two-dimensional vectors are squashed into one-dimensional vectors aligned with the x-axis.

> **Example 3.** Show that the rotation matrix R_z describes an invertible transformation
>
> We already showed that $\det \mathsf{R}_z = 1$, and hence it follows that the rotation around the z-axis is invertible. In fact, it is fairly obvious that the inverse transformation is simply a rotation by $-\theta$, i.e.
>
> $$\begin{pmatrix} \cos(-\theta) & \sin(-\theta) & 0 \\ -\sin(-\theta) & \cos(-\theta) & 0 \\ 0 & 0 & 1 \end{pmatrix},$$
>
> which simplifies to
>
> $$\begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$