2 Linear Vector Spaces and Matrices

2.1 Revision of determinants

2.1.1 Two-by-Two Determinants

A 2×2 determinant is an object with two rows and columns, sandwiched between two vertical lines. It just represents an ordinary scalar quantity; numerical value given by

$$\Delta = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21} \, . \tag{1}$$

for any set of four numbers a_{ij}. The notation with the indices is conventional and is important for matrix multiplication. An example:

$$\Delta = \begin{vmatrix} 1 & 3 \\ 4 & 2 \end{vmatrix} = 1 \times 2 - 3 \times 4 = -10 \, .$$

Rule 1

Interchanging rows and columns leaves a determinant unchanged.

$$\Delta' = \begin{vmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{21}a_{12} = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} \, .$$

Rule 2

A determinant vanishes if one of the rows or columns contains only zeroes.

Rule 3

If we multiply a row (or column) by a constant, then the value of the determinant is multiplied by that constant.

$$\Delta' = \begin{vmatrix} \alpha\, a_{11} & \alpha\, a_{12} \\ a_{21} & a_{22} \end{vmatrix} = \alpha\, a_{11}a_{22} - \alpha\, a_{12}a_{21} = \alpha \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} \, .$$

Rule 4

A determinant vanishes if two rows (or columns) are multiples of each other. For example, if $a_{i2} = \alpha\, a_{i1}$ for $i = 1, 2$, then $\Delta = \alpha\, a_{11}a_{21} - \alpha\, a_{11}a_{21} = 0$.

Rule 5

If we interchange a pair of rows or columns, the determinant changes sign.

$$\Delta' = \begin{vmatrix} a_{12} & a_{11} \\ a_{22} & a_{21} \end{vmatrix} = a_{12}a_{21} - a_{11}a_{22} = - \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} \, .$$

Rule 6

Adding a multiple of one row to another (or a multiple of one column to another) does not change the value of a determinant.

$$\Delta' = \begin{vmatrix} (a_{11} + \alpha\, a_{12}) & a_{12} \\ (a_{21} + \alpha\, a_{22}) & a_{22} \end{vmatrix} = (a_{11} + \alpha\, a_{12})a_{22} - a_{12}(a_{21} + \alpha\, a_{22})$$

$$= [a_{11}a_{22} - a_{12}a_{21}] + \alpha\,[a_{12}a_{22} - a_{12}a_{22}] = \begin{vmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \end{vmatrix} + 0\,.$$

This is a very useful rule to help simplify higher order determinants. In our 2×2 example, take 4 times row 1 from row 2 to give

$$\Delta = \begin{vmatrix} 1 & 3 \\ 4 & 2 \end{vmatrix} = \begin{vmatrix} 1 & 3 \\ 0 & -10 \end{vmatrix} = 1 \times (-10) - 3 \times 0 = -10\,.$$

By this trick we have just got one term in the end rather than two.

2.1.2 Three-by-Three Determinants

All the above rules will be valid for a general $N \times N$ determinant. A 3×3 determinant can be expanded by the first row as

$$\Delta = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12} \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13} \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}$$

$$= a_{11}a_{22}a_{33} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31}\,. \qquad (2)$$

Thus we can express the 3×3 determinant as the sum of three 2×2 ones. **Note particularly the negative sign in front of the second 2×2 determinant.**

Alternatively, can expand the determinant by the second column say;

$$\Delta = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = -a_{12} \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{22} \begin{vmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{vmatrix} - a_{32} \begin{vmatrix} a_{11} & a_{13} \\ a_{21} & a_{23} \end{vmatrix},$$

and this gives exactly the same value as before. Pay special attention to the terms which pick up the minus sign. The pattern is:

$$\begin{vmatrix} + & - & + \\ - & + & - \\ + & - & + \end{vmatrix}.$$

The rule of Sarrus

One simple way of remembering how to expand a 3×3 determinant is the *rule of Sarrus*, which is **not** valid for a 4×4 or higher order determinant. In this prescription, write the first and second columns again at the end of the determinant as:

$$\begin{matrix} a_{11} & a_{12} & a_{13} & a_{11} & a_{12} \\ a_{21} & a_{22} & a_{23} & a_{21} & a_{22} \\ a_{31} & a_{32} & a_{33} & a_{31} & a_{32} \end{matrix} \qquad (3)$$

Now SIX diagonals that lead from the top row to the bottom. Ones pointing to the right get a plus sign, those to the left a minus sign. Thus $a_{12}a_{23}a_{31}$ is positive, whereas $a_{12}a_{21}a_{33}$ is negative. Agrees with the result given in Eq. (2).

Examples

1. Evaluate
$$\Delta = \begin{vmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{vmatrix} .$$
$$\Delta = 1 \begin{vmatrix} 5 & 6 \\ 8 & 9 \end{vmatrix} - 2 \begin{vmatrix} 4 & 6 \\ 7 & 9 \end{vmatrix} + 3 \begin{vmatrix} 4 & 5 \\ 7 & 8 \end{vmatrix} = (45 - 48) - 2\,(36 - 42) + 3\,(32 - 35) = 0 .$$
The answer is zero because the third row is twice the second minus the first.

2. Evaluate
$$\Delta = \begin{vmatrix} 1 & -3 & -3 \\ 2 & -1 & -11 \\ 3 & 1 & 5 \end{vmatrix} .$$
Add three times column 1 to both columns 2 and 3.
$$\Delta = \begin{vmatrix} 1 & 0 & 0 \\ 2 & 5 & -5 \\ 3 & 10 & 14 \end{vmatrix} = \begin{vmatrix} 5 & -5 \\ 10 & 14 \end{vmatrix} = \begin{vmatrix} 5 & 0 \\ 10 & 24 \end{vmatrix} = 120 .$$

On computers one generally involves subtracts linear combinations of rows (or columns) such that there is only one element at the top of the first column with zeros everywhere else. Reduces size of the determinant by one and can be applied systematically. With pencil and paper, this often involves keeping track of fractions. Different books call this technique by different names.

2.1.3 Higher order determinants

A 4×4 determinant can be reduced to four 3×3 determinants as
$$\begin{vmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{vmatrix}$$
$$= a_{11} \begin{vmatrix} a_{22} & a_{23} & a_{24} \\ a_{32} & a_{33} & a_{34} \\ a_{42} & a_{43} & a_{44} \end{vmatrix} - a_{12} \begin{vmatrix} a_{21} & a_{23} & a_{24} \\ a_{31} & a_{33} & a_{34} \\ a_{41} & a_{43} & a_{44} \end{vmatrix} + a_{13} \begin{vmatrix} a_{21} & a_{22} & a_{24} \\ a_{31} & a_{32} & a_{34} \\ a_{41} & a_{42} & a_{44} \end{vmatrix} - a_{14} \begin{vmatrix} a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \\ a_{41} & a_{42} & a_{43} \end{vmatrix} \tag{4}$$

Alternatively, can reduce the size of determinant by taking linear combinations of rows and/or columns. This can be generalised to higher dimensions.

2.1.4 Solving linear simultaneous equations: CRAMER's rule

Consider the simultaneous equations
$$\begin{aligned} a_{11}\,x_1 + a_{12}\,x_2 + a_{13}\,x_3 &= b_1 , \\ a_{21}\,x_1 + a_{22}\,x_2 + a_{23}\,x_3 &= b_2 , \\ a_{31}\,x_1 + a_{32}\,x_2 + a_{33}\,x_3 &= b_3 \end{aligned} \tag{5}$$

for the unknown x_i. The solution is

$$x_1 = \begin{vmatrix} b_1 & a_{12} & a_{13} \\ b_2 & a_{22} & a_{23} \\ b_3 & a_{32} & a_{33} \end{vmatrix} \Big/ \Delta\,, \quad x_2 = \begin{vmatrix} a_{11} & b_1 & a_{13} \\ a_{21} & b_2 & a_{23} \\ a_{31} & b_3 & a_{33} \end{vmatrix} \Big/ \Delta\,, \quad x_3 = \begin{vmatrix} a_{11} & a_{12} & b_1 \\ a_{21} & a_{22} & b_2 \\ a_{31} & a_{32} & b_3 \end{vmatrix} \Big/ \Delta\,, \quad (6)$$

where

$$\Delta = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}\,. \qquad (7)$$

Just replace the appropriate column with the column of numbers from the right hand side. This is called Cramer's rule and gives same results as matrix inversion — but rather quicker!

Example

Use Cramer's rule to solve the following simultaneous equations just for the variable x_1:

$$\begin{aligned} 3x_1 - 2x_2 - x_3 &= 4\,, \\ 2x_1 + x_2 + 2x_3 &= 10\,, \\ x_1 + 3x_2 - 4x_3 &= 5\,. \end{aligned}$$

We can expand the determinant appearing here by the first row as

$$\Delta = \begin{vmatrix} 3 & -2 & -1 \\ 2 & 1 & 2 \\ 1 & 3 & -4 \end{vmatrix} = 3(-4-6) + 2(-8-2) - 1(6-1) = -55\,.$$

Alternatively, adding simultaneously columns 2 and 3 to column 1 gives

$$\Delta = \begin{vmatrix} 0 & -2 & -1 \\ 5 & 1 & 2 \\ 0 & 3 & -4 \end{vmatrix}\,.$$

Expand now by the first column (not forgetting the minus sign)

$$\Delta = -5(8+3) = -55\,.$$

Now by Cramer's rule,

$$\Delta \times x_1 = \begin{vmatrix} 4 & -2 & -1 \\ 10 & 1 & 2 \\ 5 & 3 & -4 \end{vmatrix} = \begin{vmatrix} 4 & -2 & -1 \\ 0 & -5 & 10 \\ 5 & 3 & -4 \end{vmatrix} = \begin{vmatrix} 4 & -2 & -5 \\ 0 & -5 & 0 \\ 5 & 3 & 2 \end{vmatrix} = -5(8+25) = -165\,.$$

Hence $x_1 = 3$.

2.2 Three-dimensional Vectors

You have met vectors an ordinary (real) 3-dimensional Euclidean space. Now generalise definitions and results to **complex** spaces with n-dimensions. This will be of importance for the 2B22 Quantum Mechanics course.

Define a three-dimensional Euclidean space by introducing three mutually orthogonal basis vectors $\hat{\imath}$, $\hat{\jmath}$ and $\hat{k}$. However, cannot generalise this notation to arbitrary number of dimensions, so use $\underline{\hat{e}}_1 = \hat{\imath}$, $\underline{\hat{e}}_2 = \hat{\jmath}$, and $\underline{\hat{e}}_3 = \hat{k}$ instead. These basis vectors have unit length,

$$\underline{\hat{e}}_1 \cdot \underline{\hat{e}}_1 = \underline{\hat{e}}_2 \cdot \underline{\hat{e}}_2 = \underline{\hat{e}}_3 \cdot \underline{\hat{e}}_3 = 1\,, \tag{8}$$

and are perpendicular to each other;

$$\underline{\hat{e}}_1 \cdot \underline{\hat{e}}_2 = \underline{\hat{e}}_2 \cdot \underline{\hat{e}}_3 = \underline{\hat{e}}_3 \cdot \underline{\hat{e}}_1 = 0\,. \tag{9}$$

Summarised in one equation as

$$\underline{\hat{e}}_i \cdot \underline{\hat{e}}_j = \delta_{ij}\,, \tag{10}$$

where the Kronecker delta δ_{ij} is shorthand for

$$\delta_{ij} = \begin{cases} 1 & \text{if } \; i = j \\ 0 & \text{if } \; i \neq j \end{cases} \tag{11}$$

Any vector $\underline{v}$ in this three-dimensional space may be written down in terms of its components along the $\underline{\hat{e}}_i$. Switching notation here so that vectors are underline, rather than having arrows on top, in line with notation for matrices. Thus

$$\underline{v} = v_1\,\underline{\hat{e}}_1 + v_2\,\underline{\hat{e}}_2 + v_3\,\underline{\hat{e}}_3\,,$$

where the coefficients v_i may be obtained by taking the scalar product of $\underline{v}$ with the basis vector $\underline{\hat{e}}_i$;

$$v_i = \underline{\hat{e}}_i \cdot \underline{v}\,. \tag{12}$$

This follows because the $\underline{\hat{e}}_i$ are perpendicular and have length one.

If we know two vectors $\underline{v}$ and $\underline{u}$ in terms of their components, then their scalar product is

$$\underline{u} \cdot \underline{v} = (u_1\,\underline{\hat{e}}_1 + u_2\,\underline{\hat{e}}_2 + u_3\,\underline{\hat{e}}_3)\cdot(v_1\,\underline{\hat{e}}_1 + v_2\,\underline{\hat{e}}_2 + v_3\,\underline{\hat{e}}_3) = u_1\,v_1 + u_2\,v_2 + u_3\,v_3 = \sum_{i=1}^{3} u_i\,v_i\,. \tag{13}$$

A particularly important case is that of the scalar product of a vector with itself, which gives rise to Pythagoras's theorem

$$v^2 = \underline{v} \cdot \underline{v} = v_1^2 + v_2^2 + v_3^2\,. \tag{14}$$

The length of a vector $\underline{v}$ is

$$v = \mid \underline{v} \mid = \sqrt{v^2} = \sqrt{v_1^2 + v_2^2 + v_3^2}\,. \tag{15}$$

A unit vector has length one.

A vector is the zero vector if and only if all its components vanish. Thus

$$\underline{v} = \underline{0} \qquad \Longleftrightarrow \qquad (v_1\,, v_2\,, v_3) = (0,0,0)\,. \tag{16}$$

The vector $\underline{v}$ is a linear combination of the basis vectors $\underline{\hat{e}}_i$. Note that the basis vectors themselves are linearly independent, because there is no linear combination of the $\underline{\hat{e}}_i$ which vanishes – unless all the coefficients are zero. Putting it in other words,

$$\underline{\hat{e}}_3 \neq \alpha\,\underline{\hat{e}}_1 + \beta\,\underline{\hat{e}}_2\,, \tag{17}$$

where α and β are scalars. Clearly, something in the x-direction plus something else in the y-direction cannot give something lying in the z-direction.

On the other hand, for three vectors taken at random, one might well be able to express one of them in terms of the other two. For example, consider the three vectors given in component form by

$$\underline{u} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} : \quad \underline{v} = \begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix} : \quad \underline{w} = \begin{pmatrix} 9 \\ 12 \\ 15 \end{pmatrix}\,. \tag{18}$$

Then

$$\underline{w} = \underline{u} + 2\underline{v}\,. \tag{19}$$

We then say that $\underline{u}$, $\underline{v}$ and $\underline{w}$ are linearly dependent. This is an important concept.

The three-dimensional space S_3 is defined as one where there are three, **BUT NO MORE**, orthonormal linearly independent vectors $\underline{\hat{e}}_i$. Any vector lying in this three-dimensional space can be written as a linear combination of the basis vectors. All this is really saying is that we can always write $\underline{v}$ in the component form;

$$\underline{v} = v_1\,\underline{\hat{e}}_1 + v_2\,\underline{\hat{e}}_2 + v_3\,\underline{\hat{e}}_3\,.$$

Note the $\underline{\hat{e}}_i$ are not unique. Could, for example rotate, the system through 45° and use these new axes as basis vectors.

Now generalise this to an arbitrary number of dimensions and letting the components become complex.

2.3 Linear Vector Space

A linear vector space S is a set of abstract quantities $\underline{a}\,, \underline{b}\,, \underline{c}\,, \cdots$, called vectors, which have the following properties:

1. If $\underline{a} \in S$ and $\underline{b} \in S$, then

$$\begin{aligned} \underline{a} + \underline{b} &= \underline{c} \in S. \\ \underline{c} = \underline{a} + \underline{b} &= \underline{b} + \underline{a} \quad \text{(Commutative law)} \\ (\underline{a} + \underline{b}) + \underline{c} &= \underline{a} + (\underline{b} + \underline{c}) \quad \text{(Associative law)}\,. \end{aligned} \tag{20}$$

2. Multiplication by a scalar (possibly complex)

$$\begin{aligned} \underline{a} \in S \quad &\Longrightarrow \quad \lambda\,\underline{a} \in S \quad (\lambda \text{ a complex number})\,, \\ \lambda\,(\underline{a} + \underline{b}) \quad &= \quad \lambda\,\underline{a} + \lambda\,\underline{b}\,, \\ \lambda\,(\mu\,\underline{a}) \quad &= \quad (\lambda\,\mu)\,\underline{a} \quad (\mu \text{ another complex number})\,. \end{aligned} \tag{21}$$

3. There exists a null (zero) vector $\underline{0} \in S$ such that

$$\underline{a} + \underline{0} = \underline{a} \tag{22}$$

for all vectors $\underline{a}$.

4. For every vector $\underline{a}$ there exists a unique vector $-\underline{a}$ such that

$$\underline{a} + (-\underline{a}) = \underline{0}\,. \tag{23}$$

5. **Linear Independence**
A set of vectors $\underline{X}_1\,,\ \underline{X}_2\,,\ \cdots\ \underline{X}_n$ are linearly dependent when it is possible to find a set of scalar coefficients c_i (not all zero) such that

$$c_1\,\underline{X}_1 + c_2\,\underline{X}_2\ \cdots\ c_n\,\underline{X}_n = \underline{0}\,.$$

If no such constants c_i exist, then the $\underline{X}_i$ are linearly independent.

By definition, an n-dimensional complex vector space S_n contains just n linearly independent vectors. Hence any vector $\underline{X}$ can be written as a linear combination

$$\underline{X} = c_1\,\underline{X}_1 + c_2\,\underline{X}_2\ \cdots\ c_n\,\underline{X}_n\,. \tag{24}$$

6. **Basis vectors and components**
Any set of n linearly independent vectors can be used as a basis for an n-dimensional vector space, which means that the basis is not unique. Once the basis has been chosen, any vector can be written uniquely as a linear combination

$$\underline{v} = \sum_{i=1}^{n} v_i\,\underline{X}_i\,.$$

Have not assumed that the basis vectors are orthogonal. For certain physical problems, convenient to work with basis vectors which are not perpendicular — eg when dealing with crystals with hexagonal symmetry. Here we will only work with basis vectors $\underline{\hat{e}}_i$ which are orthogonal and of unit length.

7. **Definition of scalar product**
Let the coefficients c_i in Eq. (24) be complex. Such complex spaces are important for Quantum Mechanics.

Write vector $\underline{v}$ in terms of its components v_i along basis vectors $\underline{\hat{e}}_i$, and similarly for another vector $\underline{u}$. Then the scalar product of these two vectors will be defined by

$$(\underline{u}\,,\underline{v}) = \underline{u}\cdot\underline{v} = u_1^*\,v_1 + u_2^*\,v_2 + \cdots + u_n^*\,v_n\,. \tag{25}$$

Only difference to the usual form on the right hand side is the complex conjugation on all the components u_i since the vectors have to be allowed to be complex. This is the only essential difference with real vectors. To stress this difference though, sometimes use a different notation on the left hand side and denote the scalar product by $(\underline{u}\,,\,\underline{w})$ rather than $\underline{u}\cdot\underline{w}$.

Note that

$$(\underline{v}\,,\,\underline{u}) = v_1^*\,u_1 + v_2^*\,u_2 + \cdots + v_n^*\,u_n = (\underline{u}\,,\,\underline{v})^*\,. \tag{26}$$

Thus, in general, the scalar product is a complex scalar.

8. **Consequences of the definition**
   (a) If $\underline{y} = \alpha\,\underline{u} + \beta\,\underline{v}$ then $(\underline{w}\,,\,\underline{y}) = \alpha\,(\underline{w}\,,\,\underline{u}) + \beta(\underline{w}\,,\,\underline{v})$.
   (b) Putting $\underline{u} = \underline{v}$, we see that

   $$u^2 = (\underline{u}\,,\,\underline{u}) = u_1^*\,u_1 + u_2^*\,u_2 + \cdots + u_n^*\,u_n = \mid u_1 \mid^2 + \mid u_2 \mid^2 + \cdots + \mid u_n \mid^2\,. \tag{27}$$

   Generalisation of Pythagoras's theorem for complex numbers. Since the $\mid u_i \mid^2$ are real and cannot be negative, then $u^2 \geq 0$. Can talk about $u = \sqrt{u^2}$ as the real length of a complex vector. In particular, if $u = 1$, $\underline{u}$ is a unit vector.
   (c) Two vectors are orthogonal if $(\underline{u}\,,\,\underline{v}) = 0$.
   (d) Components of a vector are given by the scalar product $v_i = (\underline{\hat{e}}_i\,,\,\underline{v})$.

Representations

Given a set of basis vectors $\underline{\hat{e}}_i$, any vector $\underline{v}$ in an n-dimensional space can be written uniquely in the form $\underline{v} = \sum_{i=1}^{n} v_i\,\underline{\hat{e}}_i$. The set of numbers v_i, $i = 1, \cdots, n$ (the components) are said to represent the vector $\underline{v}$ in that basis. The concept of a vector is more general and abstract than that of the components. The components are somehow man-made. If we rotate the coordinate system then the vector stays in the same direction but the components change. This whole business of matrices (and much of third year Quantum Mechanics) is connected with what happens when we change basis vectors.

2.4 Linear Transformations

Perform some operation on vector $\underline{v}$ which changes it into another vector in the space S_n. For example, rotate the vector. Denote the operation by $\hat{A}$ and, instead of tediously saying that $\hat{A}$ acts on $\underline{v}$, write it symbolically as $\hat{A}\,\underline{v}$. By assumption, therefore, $\underline{u} = \hat{A}\,\underline{v}$ is another vector in the same space S_n. To agree with the notation of the 2B22 Quantum Mechanics course, put hats on all the operators.

Have seen that manipulation of vectors is simplified by working with components. To investigate this further, see how the operation $\hat{A}$ changes the basis vectors $\underline{\hat{e}}_1, \underline{\hat{e}}_2, \cdots, \underline{\hat{e}}_n$. For the sake of definiteness, let us look at $\underline{\hat{e}}_1$, which has a 1 in the first position and zeros everywhere else:

$$\underline{\hat{e}}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \qquad (n \text{ terms in the column}). \tag{28}$$

as $\hat{\underline{e}}_1$ with the operator $\hat{A}$. This gives rise to a vector which we shall denote by $\underline{a}_1$ because it started from $\hat{\underline{e}}_1$. Thus

$$\underline{a}_1 = \hat{A}\,\hat{\underline{e}}_1\,. \tag{29}$$

To write this in terms of components, must introduce a second index

$$\underline{a}_1 = \begin{pmatrix} a_{11} \\ a_{21} \\ a_{31} \\ \vdots \\ a_{n1} \end{pmatrix}. \tag{30}$$

To specify action of $\hat{A}$ completely, must define how it acts on <u>all</u> the basis vectors $\hat{\underline{e}}_i$;

$$\underline{a}_i = \hat{A}\,\hat{\underline{e}}_i = \begin{pmatrix} a_{1i} \\ a_{2i} \\ a_{3i} \\ \vdots \\ a_{ni} \end{pmatrix}. \tag{31}$$

This requires n^2 numbers a_{ji}, $(j = 1, 2, \cdots, n; i = 1, 2, \cdots, n)$.

Instead of writing $\underline{a}_i$ explicitly as a column vector, can use the basis vectors once again to show that

$$\underline{a}_i = a_{1i}\,\hat{\underline{e}}_1 + a_{2i}\,\hat{\underline{e}}_2 + a_{3i}\,\hat{\underline{e}}_3 + \cdots + a_{ni}\,\hat{\underline{e}}_n = \sum_{j=1}^{n} a_{ji}\,\hat{\underline{e}}_j\,. \tag{32}$$

as $\hat{\underline{e}}_i$ has 1 in the i'th position and 0's everywhere else.

Knowing the basis vectors transformation, it is (in principle) easy to evaluate the action of $\hat{A}$ on some vector $\underline{v} = \sum_i v_i\,\hat{\underline{e}}_i$. Then

$$\underline{u} = \hat{A}\,\underline{v} = \sum_i (\hat{A}\,\hat{\underline{e}}_i)\,v_i = \sum_{i,j} a_{ji}\,v_i\,\hat{\underline{e}}_j\,. \tag{33}$$

But, writing $\underline{u}$ in terms of components as well,

$$\underline{u} = \sum_j u_j\,\hat{\underline{e}}_j\,, \tag{34}$$

and comparing coefficients of $\hat{\underline{e}}_j$, we find

$$u_j = \sum_{i=1}^{n} a_{ji}\,v_i\,. \tag{35}$$

This is just the law for matrix multiplication. Many of you will have seen it for 2×2 matrices. For $n \times n$, the sums are just a bit bigger! Note that basis vectors transform with $\sum_j a_{ji}\,\hat{\underline{e}}_j$, whereas the components involve the other index $\sum_i a_{ji}\,v_i$.

The set of numbers a_{ij} represents the abstract operator $\hat{A}$ in the particular basis chosen; these n^2 numbers determine completely the effect of $\hat{A}$ on any arbitrary vector:

the vector undergoes a linear transformation. It is convenient to arrange all these numbers into a square array

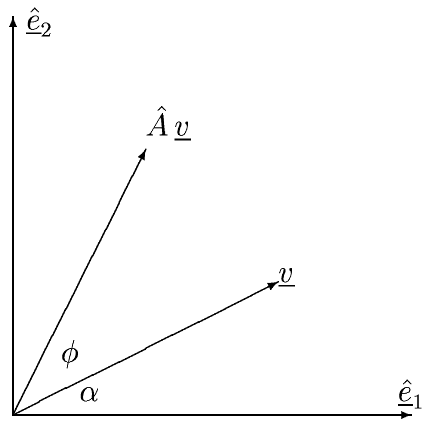
$$\underline{A} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}, \tag{36}$$

called matrix. This one is in fact a square matrix with n rows and n columns.

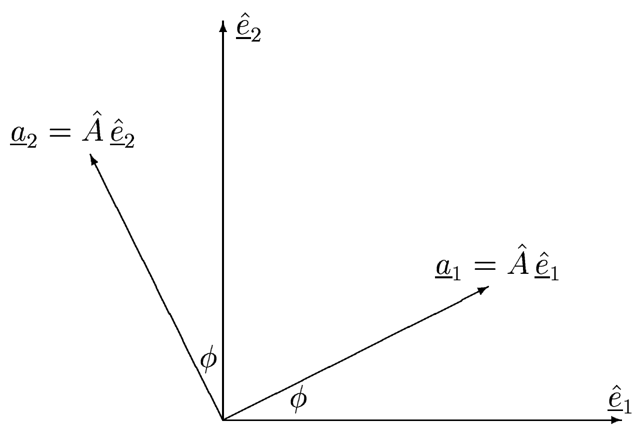
Signify a vector by putting an arrow on the top, underline it, or put a tilde under or over it or write it in bold in order to distinguish it from a scalar. Similarly must write something on the $\underline{A}$ in order to show that it is a matrix. The textbooks tend to use bold face — here we are going just to underline the symbol.

Example 1

Let $\hat{A}$ be the operator which rotates a vector in two dimensions through an angle ϕ anticlockwise.



Want to find the matrix representation of operator $\hat{A}$. Do this by looking at what happens to the basis vectors under the rotation.



Using simple trigonometry,

$$\underline{a}_1 = \hat{A}\,\underline{\hat{e}}_1 = \cos\phi\,\underline{\hat{e}}_1 + \sin\phi\,\underline{\hat{e}}_2$$

$$= a_{11}\,\underline{\hat{e}}_1 + a_{21}\,\underline{\hat{e}}_2\,.$$

Hence $a_{11} = \cos\phi$ and $a_{21} = \sin\phi$.

Similarly,

$$\begin{aligned} \underline{a}_2 = \hat{A}\,\underline{\hat{e}}_2 &= -\sin\phi\,\underline{\hat{e}}_1 + \cos\phi\,\underline{\hat{e}}_2 \\ &= a_{12}\,\underline{\hat{e}}_1 + a_{22}\,\underline{\hat{e}}_2\,, \end{aligned}$$

so that $a_{12} = -\sin\phi$ and $a_{22} = \cos\phi$.

The two-dimensional rotation matrix therefore takes the form

$$\underline{A} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = \begin{pmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{pmatrix}\,. \tag{37}$$

We now have to check whether this gives an answer which is consistent with the first picture. Here

$$\begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} v\cos\alpha \\ v\sin\alpha \end{pmatrix}$$

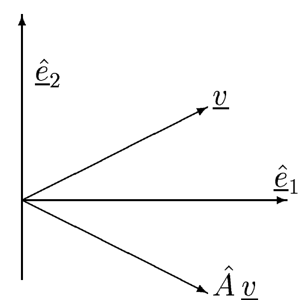
so that

$$\begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = \begin{pmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{pmatrix} \begin{pmatrix} v\cos\alpha \\ v\sin\alpha \end{pmatrix} = \begin{pmatrix} v\cos\alpha\cos\phi - v\sin\alpha\sin\phi \\ v\sin\alpha\cos\phi + v\cos\alpha\sin\phi \end{pmatrix} = \begin{pmatrix} v\cos(\alpha+\phi) \\ v\sin(\alpha+\phi) \end{pmatrix}.$$

Exactly what you get from applying trigonometry to the diagram.

Concrete example #2

Matrix representation for a reflection in the x-axis.



In this case

$$\begin{aligned} \hat{A}\,\underline{\hat{e}}_1 &= \underline{\hat{e}}_1 \\ \hat{A}\,\underline{\hat{e}}_2 &= -\underline{\hat{e}}_2\,. \end{aligned}$$

Hence $a_{11} = 1$, $a_{22} = -1$, $a_{21} = a_{12} = 0$ and

$$\underline{A} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\,.$$

As a test, see what happens to the vector in the picture:

$$\underline{w} = \underline{A}\,\underline{v} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} v_1 \\ -v_2 \end{pmatrix}\,,$$

as expected.

2.5 Multiple Transformations; Matrix Multiplication

Suppose that we know the action of some operator $\hat{A}$ on any vector and also the action of another operator $\hat{B}$. What is the action of the combined operation of $\hat{B}$ followed by $\hat{A}$? Consider

$$\begin{aligned} \underline{w} &= \hat{B}\,\underline{v} \\ \underline{u} &= \hat{A}\,\underline{w}\,. \\ \underline{u} &= \hat{A}\,\hat{B}\,\underline{v} = \hat{C}\underline{v}\,. \end{aligned} \tag{38}$$

To find the matrix representation of $\hat{C}$, write the above equations in component form:

$$\begin{aligned} w_i &= \sum_j b_{ij}\,v_j \\ u_k &= \sum_i a_{ki}\,w_i \\ &= \sum_{i,j} a_{ki}\,b_{ij} v_j \\ &= \sum_j c_{kj}\,v_j\,. \end{aligned} \tag{39}$$

Since this is supposed to hold for any vector $\underline{v}$, it requires that

$$c_{kj} = \sum_{i=1}^{n} a_{ki}\,b_{ij}\,. \tag{40}$$

This is the law for the multiplication of two matrices $\underline{A}$ and $\underline{B}$. The product matrix has the elements c_{kj}. For 2×2 matrices you had the rule at A-level or even at GCSE!

Matrices can be used to represent the action of linear operations, such as reflection and rotation, on vectors. Now that we know how to combine such operations through matrix multiplication, we can build up quite complicated operations. This leads us quite naturally to the study of the properties of matrices in general.

2.6 Properties of Matrices

In general a matrix is a set of elements, which can be either numbers or variables, set out in the form of an array. For example

$$\underset{\text{(rectangular)}}{\begin{pmatrix} 2 & 6 & 4 \\ -1 & i & 7 \end{pmatrix}} \quad \text{or} \quad \underset{\text{(square)}}{\begin{pmatrix} 0 & -i \\ 3+6i & x^2 \end{pmatrix}}$$

A matrix having n rows and m columns is called an $n \times m$ matrix. The above examples are 2×3 and 2×2. A square matrix clearly has $n = m$. The general matrix is written

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

There is often confusion between the matrices and determinants. The notational difference is that a matrix is an array surrounded by brackets whereas a determinant has vertical lines. They are, however, very different beasts. The determinant $\mid A \mid$ is a single number (or algebraic expression). A matrix $\underline{A}$ is a whole array of $n \times m$ numbers which represents a transformation.

A vector is a simple matrix which is $n \times 1$ (column vector) or $1 \times n$ (row vector), as in

$$\begin{pmatrix} v_1 \\ v_2 \\ v_3 \\ \cdots \\ v_n \end{pmatrix} \quad \text{or} \quad (v_1,\, v_2,\, v_3,\, \cdots, v_n)\ .$$

Rules

1. Two matrices $\underline{A}$ and $\underline{B}$ are equal if they have the same number n of rows and m of columns and if all of the corresponding elements are equal.

2. There exists an $n \times m$ zero-matrix where all the elements are zero.

3. There exists a unit matrix. This is an $n \times n$ square matrix with ones down the diagonal and zeros everywhere else.

$$\underline{I} = \underline{E} = \begin{pmatrix} 1 & 0 & 0 & \cdots \\ 0 & 1 & 0 & \cdots \\ 0 & 0 & 1 & \cdots \\ \cdots & \cdots & \cdots & \cdots \end{pmatrix}\ .$$

   Some books do use $\underline{E}$ for this. In component form

$$I_{ij} = \delta_{ij}\ ,$$

   where the Kronecker-delta has been employed.

4. Addition or Subtraction.
   The sum of two matrices $\underline{A}$ and $\underline{B}$ can only be defined if they have the same number of n rows and the same number m of columns. If this is the case, then the matrix $\underline{C}$ is also $n \times m$ and has elements

$$c_{ij} = a_{ij} + b_{ij}\ .$$

   It follows immediately that $\underline{A} + \underline{B} = \underline{B} + \underline{A}$ (commutative law of addition) and $(\underline{A} + \underline{B}) + \underline{C} = \underline{A} + (\underline{B} + \underline{C})$ (associative law).

5. Multiplication by a scalar.

$$\underline{B} = \lambda \underline{A} \quad \Longrightarrow \quad b_{ij} = \lambda a_{ij}\ .$$

6. Matrix multiplication:

$$\underline{C} = \underline{A}\,\underline{B} \quad \Longrightarrow \quad c_{ij} = \sum_{k=1}^{n} a_{ik}\, b_{kj}\ .$$

Note that matrix multiplication can only be defined if the number of columns in $\underline{A}$ is equal to the number of rows in $\underline{B}$. Then if $\underline{A}$ is $m \times n$ and $\underline{B}$ is $n \times p$, then $\underline{C}$ is $m \times p$.

Note that matrix multiplication is **NOT** commutative; $\underline{A}\,\underline{B} \neq \underline{B}\,\underline{A}$. One of the multiplications might not even be defined! If $\underline{A}$ is $m \times n$ and $\underline{B}$ is $n \times m$, then $\underline{A}\,\underline{B}$ is $m \times m$ and $\underline{B}\,\underline{A}$ is $n \times n$.

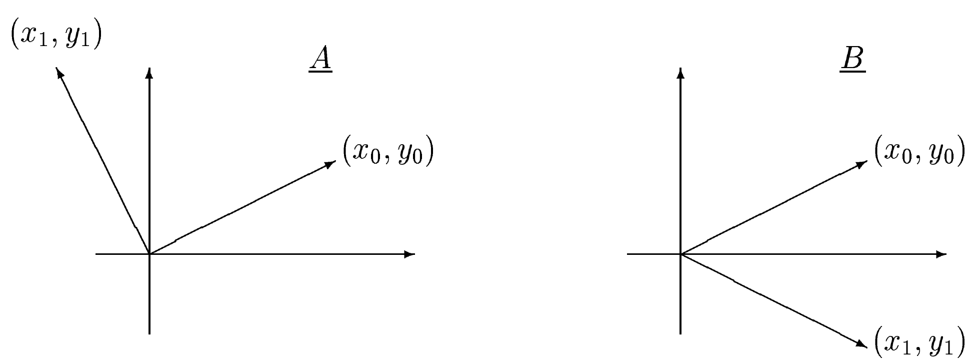
Matrices do not commute because they are constructed to represent linear operations and, in general, such operations do not commute. It can matter in which order you do certain operations.

On the other hand,

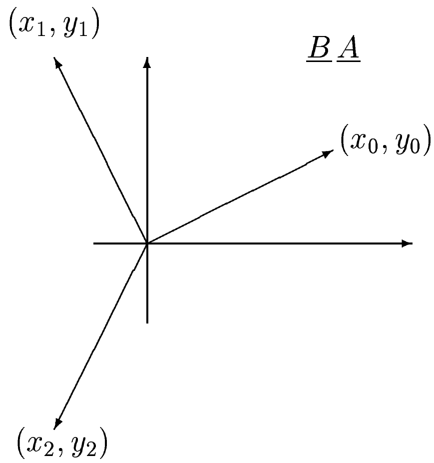
$$\begin{aligned} \underline{A}\,(\underline{B}\,\underline{C}) &= (\underline{A}\,\underline{B})\,\underline{C} \\ \underline{A}\,(\underline{B} + \underline{C}) &= \underline{A}\,\underline{B} + \underline{A}\,\underline{C}\,. \end{aligned}$$

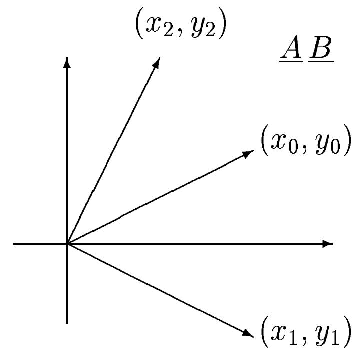
Will assume you are familiar with the actual multiplication process in practice. If not, you have been warned!

Example 1
Let $\underline{A}$ represent a rotation of 90° around the z-axis and $\underline{B}$ a reflection in the x-axis.



For the combination $\underline{B}\,\underline{A}$, we first act with $\underline{A}$ and then $\underline{B}$. In the case of $\underline{A}\,\underline{B}$ it is the other way around and this leads to a different result, as shown in the picture.





Clearly the end point (x_2, y_2) is different in the two cases so that the operations corresponding to $\underline{A}$ and $\underline{B}$ obviously don't commute. We now want to show exactly the same results using matrix manipulation, in order to illustrate the power of matrix multiplication.

The 2×2 matrix representing the two-dimensional rotation through angle ϕ.

$$\underline{A} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = \begin{pmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \quad \text{for } \phi = 90^\circ\,.$$

Similarly, for the reflection in the x-axis,

$$\underline{B} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\,.$$

Hence

$$\begin{aligned} \underline{A}\,\underline{B} &= \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \\ \underline{B}\,\underline{A} &= \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}\,, \end{aligned}$$

so that in the $\underline{A}\,\underline{B}$ case $x_2 = y_0$ and $y_2 = x_0$. The x and y coordinates are simply interchanged. In the other case both x_2 and y_2 get an extra minus sign. This is exactly what we see in the picture.

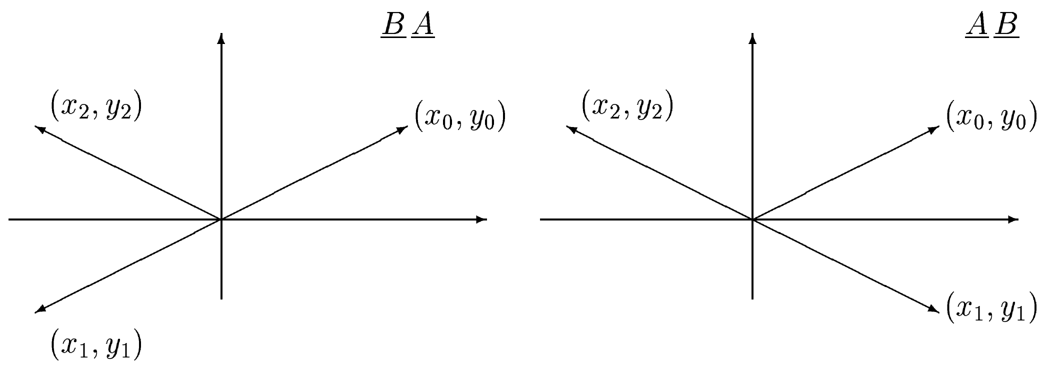
Example 2

It may of course happen that two operators commute, as for example when one represents a rotation through 180° and the other a reflection in the x-axis. Then

$$\underline{A} = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}\,, \quad \underline{B} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\,, \quad \underline{A}\,\underline{B} = \underline{B}\,\underline{A} = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}\,.$$

The combined operation describes a reflection in the y-axis.

Geometrically these operations correspond to



You should note that the final point $(x_2, y_2) = (-x_0, y_0)$ is the same in both diagrams, but the intermediate point is not.

Determinant of a Matrix Product

By writing out both sides explicitly, it is straightforward to show that for 2×2 or 3×3 square matrices the determinant of a product of two matrices is equal to the product of the determinants.

$$\mid \underline{A}\,\underline{B} \mid = \mid \underline{A} \mid \times \mid \underline{B} \mid \ . \tag{41}$$

However, this result is true in general for $n \times n$ square matrices of any size.

One consequence of this is that, although $\underline{A}\,\underline{B} \neq \underline{B}\,\underline{A}$, their determinants are equal. In the first example that I gave of matrix multiplication, we see that $\mid \underline{A}\,\underline{B} \mid = \mid \underline{B}\,\underline{A} \mid = -1$. This result for the determinant of products will prove very useful later.

2.7 Special Matrices

Multiplication by the unit matrix

Let $\underline{A}$ be an $n \times n$ matrix and $\underline{I}$ the $n \times n$ unit matrix. Then

$$(\underline{A}\,\underline{I})_{ij} = \sum_k a_{ik}\,\delta_{kj} = a_{ij}\ ,$$

since the Kronecker-delta δ_{ij} vanishes unless $i = j$. Thus

$$\underline{A}\,\underline{I} = \underline{A}\,. \tag{42}$$

Similarly

$$(\underline{I}\,\underline{A})_{ij} = \sum_k \delta_{ik}\,a_{kj} = a_{ij}\ ,$$

and

$$\underline{I}\,\underline{A} = \underline{A}\,. \tag{43}$$

The multiplication on the left or right by $\underline{I}$ does not change a matrix $\underline{A}$. In particular, the unit matrix $\underline{I}$ (or any multiple of it) commutes with any other matrix of the appropriate size.

Diagonal matrices

A diagonal matrix is a square matrix with elements only along the diagonal:

$$\underline{A} = \begin{pmatrix} a_1 & 0 & 0 & \cdots \\ 0 & a_2 & 0 & \cdots \\ 0 & 0 & a_3 & \cdots \\ \cdots & \cdots & \cdots & \cdots \end{pmatrix} .$$

Thus

$$(\underline{A})_{ij} = a_i \, \delta_{ij} \, .$$

Now consider two diagonal matrices $\underline{A}$ and $\underline{B}$ of the same size.

$$(\underline{A}\,\underline{B})_{ij} = \sum_k A_{ik} \, B_{kj} = \sum_k a_i \, \delta_{ik} \, \delta_{kj} \, b_k = (a_i \, b_i) \, \delta_{ij} \, .$$

Hence $\underline{A}\,\underline{B}$ is also a diagonal matrix with elements equal to the products of the corresponding individual elements. Note that for diagonal matrices, $\underline{A}\,\underline{B} = \underline{B}\,\underline{A}$, so that $\underline{A}$ and $\underline{B}$ commute.

Transposing matrices

The transposed matrix $\underline{A}^T$ is just the original matrix $\underline{A}$ with its rows and columns interchanged. Hence

$$(\underline{A}^T)_{ij} = (\underline{A})_{ji} \, . \tag{44}$$

The transpose of an $n \times m$ matrix is $m \times n$.

Consequences

a) Clearly $(\underline{A}^T)^T = \underline{A}$.

b) If $\underline{A}^T = \underline{A}$, $\underline{A}$ is symmetric.
If $\underline{A}^T = -\underline{A}$, $\underline{A}$ is anti-symmetric.

c) Transposing matrix products. Look at $\underline{C} = \underline{A}\,\underline{B}$, which has elements

$$c_{ij} = \sum_k a_{ik} \, b_{kj} \, .$$

Now

$$(\underline{C}^T)_{ji} = c_{ij} = \sum_k a_{ik} \, b_{kj} = \sum_k (\underline{A}^T)_{ki} \, (\underline{B}^T)_{jk} = \sum_k (\underline{B}^T)_{jk} \, (\underline{A}^T)_{ki} = (\underline{B}^T \, \underline{A}^T)_{ji} \, .$$

Hence

$$(\underline{A}\,\underline{B})^T = \underline{B}^T \, \underline{A}^T \, . \tag{45}$$

Transposing a product of matrices, reverses the order of multiplication. True no matter how many terms there are;

$$(\underline{A}\,\underline{B}\,\underline{C})^T = \underline{C}^T \, \underline{B}^T \, \underline{A}^T \, .$$

This rule, which is also true for operators, will be used by the Quantum Mechanics lecturers in the second and third years.

d) If $\underline{A}^T\,\underline{A} = \underline{I}$, $\underline{A}$ is an orthogonal matrix. Check that the two-dimensional rotation matrix

$$\underline{A} = \begin{pmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{pmatrix} .$$

is orthogonal. For this, need $\cos^2\phi + \sin^2\phi = 1$. Matrix $\underline{A}$ rotates the system through angle ϕ, while the transpose matrix $\underline{A}^T$ rotates it back through angle $-\phi$. Because of this, orthogonal matrices are of great practical use in different branches of Physics.

Taking the determinant of the defining equation, and using the determinant of a product rule gives

$$\mid \underline{A}^T \mid \mid \underline{A} \mid = \mid \underline{I} \mid = 1 \, .$$

But the determinant of a transpose of a matrix is the same as the determinant of the original matrix — it doesn't matter if you switch rows and columns in a determinant. Hence

$$\mid \underline{A} \mid \mid \underline{A} \mid = \mid \underline{A} \mid^2 = 1 \, ,$$

so $\mid \underline{A} \mid = \pm 1$.

e) Suppose $\underline{A}$ and $\underline{B}$ are orthogonal matrices. Their product $\underline{C} = \underline{A}\,\underline{B}$ is also orthogonal.

$$\underline{C}^T\,\underline{C} = (\underline{A}\,\underline{B})^T\,(\underline{A}\,\underline{B}) = \underline{B}^T\,\underline{A}^T\,\underline{A}\,\underline{B} = \underline{B}^T\,\underline{I}\,\underline{B} = \underline{B}^T\,\underline{B} = \underline{I} \, .$$

Physical meaning: since the matrix for rotation about the x-axis is orthogonal and so is rotation about the y-axis, then the matrix for rotation about the y-axis followed by one about the x-axis is also orthogonal.

Complex conjugation

To take the complex conjugate of a matrix, just complex-conjugate all its elements:

$$(\underline{A}^*)_{ij} = a^*_{ij} \, . \tag{46}$$

For example

$$\underline{A} = \begin{pmatrix} -i & 0 \\ 3-i & 6+i \end{pmatrix} \Longrightarrow \underline{A}^* = \begin{pmatrix} +i & 0 \\ 3+i & 6-i \end{pmatrix} .$$

If $\underline{A} = \underline{A}^*$, the matrix is real.

Hermitian conjugation

Combines complex conjugation and transposition; it is probably more important than either – especially in Quantum Mechanics. Sometimes called the **Hermitian adjoint** and denoted by a dagger (†).

$$\underline{A}^\dagger = (\underline{A}^T)^* = (\underline{A}^*)^T \, . \tag{47}$$

Thus $(\underline{A}^\dagger)^\dagger = \underline{A}$.

For example

$$\underline{A} = \begin{pmatrix} -i & 0 \\ 3-i & 6+i \end{pmatrix} \Longrightarrow \underline{A}^\dagger = \begin{pmatrix} +i & 3+i \\ 0 & 6-i \end{pmatrix} .$$

If $\underline{A}^\dagger = \underline{A}$, $\underline{A}$ is Hermitian.
If $\underline{A}^\dagger = -\underline{A}$, $\underline{A}$ is anti-Hermitian.

All real symmetric matrices are Hermitian, but also other possibilities. Eg

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

is Hermitian.

Rule for Hermitian conjugates of products is the same as for transpositions:

$$(\underline{A}\,\underline{B})^\dagger = \underline{B}^\dagger\,\underline{A}^\dagger\,. \tag{48}$$

Unitary Matrices

Matrix $\underline{U}$ is unitary if

$$\underline{U}^\dagger\,\underline{U} = \underline{I}\,. \tag{49}$$

Unitary matrices are very important in Quantum Mechanics!

Again consider the determinant product rule.

$$|\,\underline{U}^\dagger\,|\,|\,\underline{U}\,| = |\,\underline{I}\,| = 1\,.$$

Changing rows and columns in a determinant does nothing, but Hermitian conjugate also involves complex conjugation. Hence

$$|\,\underline{U}\,|^*\,|\,\underline{U}\,| = 1\,,$$

and so $|\,\underline{U}\,| = e^{i\phi}$, with ϕ being real.

2.8 Matrix Inversion

Explicit 2×2 evaluation

Define the inverse of a square matrix $\underline{A}$ and evaluate it. The inverse, $\underline{B} = \underline{A}^{-1}$, is defined to be that matrix which, when multiplied by $\underline{A}$, gives the unit matrix;

$$\underline{B}\,\underline{A} = \underline{I}\,.$$

Consider

$$\underline{A} = \begin{pmatrix} 1 & 2 \\ 4 & 3 \end{pmatrix} \quad \text{and} \quad \underline{B} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}\,.$$

Need to determine unknown numbers a, b, c, d from the condition

$$\underline{B}\,\underline{A} = \begin{pmatrix} a+4b & 2a+3b \\ c+4d & 2c+3d \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\,,$$

gives

$$\begin{aligned} a + \tfrac{3}{2}b &= 0 \qquad & c + 4d &= 0\,, \\ a + 4b &= 1 \qquad & c + \tfrac{3}{2}d &= \tfrac{1}{2}\,. \end{aligned}$$

These simultaneous equations have solutions $a = -\frac{3}{5}$, $b = \frac{2}{5}$, $c = \frac{4}{5}$, and $d = -\frac{1}{5}$. In matrix form

$$\underline{A}^{-1} = \tfrac{1}{5}\begin{pmatrix} -3 & 2 \\ 4 & -1 \end{pmatrix}.$$

Rule for 2×2 matrices

Need some automated way of evaluating inverse matrices. Motivate the result with this example, then generalise and only justify iafterwards.

$$\underline{A} = \begin{pmatrix} 1 & 2 \\ 4 & 3 \end{pmatrix} \quad \text{and} \quad (\underline{A}^{-1})^T = -\tfrac{1}{5}\begin{pmatrix} 3 & -4 \\ -2 & 1 \end{pmatrix}.$$

Notice that inside the bracket, all the coefficients are exchanged across the diagonal between $\underline{A}$ and $\underline{A}^{-1}$. There are a couple of minus signs, but these are coming in exactly the positions that one gets minus signs when expanding out a 2×2 determinant. The only remaining puzzle is the origin of the factor $-\frac{1}{5}$. Well this is precisely

$$\frac{1}{\mid A \mid} = \frac{1}{(1 \times 3 - 4 \times 2)} = -\frac{1}{5}.$$

The determinant $\mid A \mid$ has come in useful after all.

This simple observation is true for the inverse of any 2×2 matrix. Consider

$$\underline{A} = \begin{pmatrix} \alpha & \gamma \\ \beta & \delta \end{pmatrix}.$$

According to the hand-waving observation above, one would expect

$$(\underline{A}^{-1})^T = \frac{1}{(\alpha\delta - \beta\gamma)}\begin{pmatrix} \delta & -\beta \\ -\gamma & \alpha \end{pmatrix} \quad \text{and} \quad \underline{A}^{-1} = \frac{1}{(\alpha\delta - \beta\gamma)}\begin{pmatrix} \delta & -\gamma \\ -\beta & \alpha \end{pmatrix}.$$

Verify that the $\underline{A}^{-1}$ defined in this way does indeed satisfy $\underline{A}^{-1}\,\underline{A} = \underline{I}$.

IMPORTANT: Do not forget the minus signs and do not forget to transpose the matrix afterward.

Cofactors and minors

A 3×3 determinant can be expanded by the first row (Laplace's rule) as

$$\Delta = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11}\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12}\begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13}\begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}.$$

2×2 sub-determinants are obtained by striking out the rows and columns containing respectively a_{11}, a_{12} and a_{13}. These sub-determinants are the 2×2 **minors** of determinant Δ.

Define the 2×2 minor obtained by striking out the i'th row and j'th column to be M_{ij}. From the examples

$$\Delta = \sum_j a_{ij}\, M_{ij}\, (-1)^{i+j} = \sum_i a_{ij}\, M_{ij}\, (-1)^{i+j}. \tag{50}$$

The first form corresponds to expanding by row-i, the second to column-j. After summing over j, the answer does not depend upon the value of i, *i.e.* on which row has been used for the expansion.

One trouble about this formula is the $(-1)^{i+j}$ factor which always arises in expanding determinants. One can define the **cofactor** matrix $\underline{C} = [C_{ij}]$ with this explicit factor included:

$$C_{ij} = (-1)^{i+j}\, M_{ij}\,, \tag{51}$$

so that

$$\Delta = \sum_{i\;or\;j} a_{ij}\, C_{ij}\,. \tag{52}$$

This merely puts the minus sign problem somewhere else!

If $\underline{A}$ is a 3×3 matrix, then so is $\underline{C}$. We define the **adjoint** matrix to be the transpose of $\underline{C}$, which means that the indices i and j are switched around:

$$[\underline{A}^{\text{adj}}]_{ij} = C_{ji}\,. \tag{53}$$

Theorem

For any square matrix,

$$\underline{A}^{-1} = \underline{A}^{\text{adj}}/\mid A\mid\,. \tag{54}$$

This agrees with our experience in the case of a 2×2 matrix. For a 3×3 matrix one can write down the most general form, carry out the operations outlined above, to show explicitly that $\underline{A}^{-1}\,\underline{A} = \underline{I}$. Eq. (54) is valid for any size matrix, but here won't need to work out anything bigger than 3×3. Now show how to carry out these operations in practice.

Example Find the inverse of

$$\underline{A} = \begin{pmatrix} -1 & 2 & 3 \\ 2 & 0 & -4 \\ -1 & -1 & 1 \end{pmatrix}.$$

Matrix of minors is

$$\underline{M} = \begin{pmatrix} -4 & -2 & -2 \\ 5 & 2 & 3 \\ -8 & -2 & -4 \end{pmatrix}.$$

Cofactor matrix changes a few signs to give

$$\underline{C} = \begin{pmatrix} -4 & 2 & -2 \\ -5 & 2 & -3 \\ -8 & 2 & -4 \end{pmatrix}.$$

Adjoint matrix involves changing rows and columns:

$$\underline{A}^{\text{adj}} = \begin{pmatrix} -4 & -5 & -8 \\ 2 & 2 & 2 \\ -2 & -3 & -4 \end{pmatrix}.$$

Now

$$\mid A\mid = -1\times(-4) - 2\times(-2) + 3\times(-2) = 2\,.$$

Hence

$$\underline{A}^{-1} = \frac{1}{2} \begin{pmatrix} -4 & -5 & -8 \\ 2 & 2 & 2 \\ -2 & -3 & -4 \end{pmatrix} .$$

Can check that this is right by doing the explicit $\underline{A}^{-1}\underline{A}$ multiplication.

Note that if $\mid A \mid = 0$, twe say the determinant is singular; $\underline{A}^{-1}$ does not exist . [It has some infinite elements.]

Lots of other ways to do matrix inversion: Gaussian or Gauss-Jordan elimination, as described by Boas. These methods become more important as the size of the matrix goes up.

Properties of the inverse matrix

a) $\underline{A}\,\underline{A}^{-1} = \underline{A}^{-1}\,\underline{A} = \underline{I}$; a matrix commutes with its inverse.

b) $(\underline{A}^{-1})^T = (\underline{A}^T)^{-1}$; the operations of inversion and transposition commute.

c) If $\underline{C} = \underline{A}\,\underline{B}$, what is $\underline{C}^{-1}$? Consider

$$\underline{B}^{-1}\,\underline{A}^{-1}\,\underline{I} = \underline{B}^{-1}\,\underline{A}^{-1}\,\underline{C}\,\underline{C}^{-1} = \underline{B}^{-1}\,\underline{A}^{-1}\,\underline{A}\,\underline{B}\,\underline{C}^{-1} = \underline{B}^{-1}\,\underline{B}\,\underline{C}^{-1} = \underline{C}^{-1} = (\underline{A}\,\underline{B})^{-1} .$$

Hence

$$(\underline{A}\,\underline{B})^{-1} = \underline{B}^{-1}\,\underline{A}^{-1} . \tag{55}$$

reverse the order before inverting each matrix.

d) If $\underline{A}$ is orthogonal, *i.e.* $\underline{A}^T\,\underline{A} = \underline{I}$, then $\underline{A}^{-1} = \underline{A}^T$.

e) If $\underline{A}$ is unitary, *i.e.* $\underline{A}^\dagger\,\underline{A} = \underline{I}$, then $\underline{A}^{-1} = \underline{A}^\dagger$.

f) Using the determinant of a product rule, it follows immediately that
$\mid \underline{A}^{-1} \mid = 1/ \mid \underline{A} \mid$.

g) **Matrix division**
Division of matrices is not really defined, but one can multiply by the inverse matrix. Unfortunately, in general,

$$\underline{A}\,\underline{B}^{-1} \neq \underline{B}^{-1}\,\underline{A} .$$

2.9 Solution of Linear Simultaneous Equations

Know how to solve simultaneous equations of form

$$\begin{aligned} a_{11}\,x_1 + a_{12}\,x_2 + a_{13}\,x_3 &= b_1 , \\ a_{21}\,x_1 + a_{22}\,x_2 + a_{23}\,x_3 &= b_2 , \\ a_{31}\,x_1 + a_{32}\,x_2 + a_{33}\,x_3 &= b_3 \end{aligned}$$

for unknown x_i as the ratio of two determinants. Result proved in 2×2 case, here give indication of a more general proof.

Write eq. in matrix form

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} ,$$

that is

$$\underline{A}\,\underline{x} = \underline{b} \quad \text{or} \quad \sum_j a_{ij}\,x_j = b_i\,.$$

Can write formal solution immediately by multiplying both sides by $\underline{A}^{-1}$:

$$\underline{x} = \underline{A}^{-1}\,\underline{b}\,.$$

All that remains is to evaluate the result!

Using the previous expression for the inverse matrix,

$$x_j = \sum_i (\underline{A}^{\text{adj}})_{ji}\,b_i / \mid A \mid\ .$$

If the determinant does not vanish, this leads to Cramer's rule discussed in the first lecture. $\sum_i (\underline{A}^{\text{adj}})_{ji}\,b_i$ is the determinant obtained by replacing the j'th column of $\underline{A}$ by the column vector $\underline{b}$.

There are many special cases of this formula; consider only two:

a) If $\mid A \mid = 0$. Then matrix $\underline{A}$ is singular and inverse matrix cannot be defined. Provided that the equations are mutually consistent, this means that (at least) one of the equations is not linearly independent of the others. Do not have n equations for n unknowns but rather only $n-1$ equations. Can only try to solve the equations for $n-1$ of the x_i in terms of the b_i and one of the x_i. It might take some trial and error to find which of the equations to throw away.

b) If all $b_i = 0$, have to look for a solution of the homogeneous equation

$$\underline{A}\,\underline{x} = \underline{0}\,.$$

There is, of course, the uninteresting solution where all the $x_i = 0$. Can there be a more interesting solution? The answer is yes, provided that $\mid A \mid = 0$.

2.10 Eigenvalues and Eigenvectors

Let $\underline{A}$ be an $n \times n$ square matrix and $\underline{X}$ an $n \times 1$ column vector such that

$$\underline{A}\,\underline{X} = \lambda\,\underline{X} = \lambda\,\underline{I}\,\underline{X}\,, \tag{56}$$

where λ is some scalar number. λ is an **eigenvalue** of matrix $\underline{A}$ and $\underline{X}$ is the corresponding eigenvector. Half of Quantum Mechanics seems to be devoted to searching for eigenvalues!

To attack the problem, rearrange Eq. (56) as

$$(\underline{A} - \lambda\,\underline{I})\,\underline{X} = \underline{0}\,. \tag{57}$$

Set of n homogeneous linear equations which has interesting solutions if

$$\mid \underline{A} - \lambda\,\underline{I} \mid = 0\,. \tag{58}$$

Explicitly:

$$\begin{vmatrix} (a_{11} - \lambda) & a_{12} & a_{13} & \cdots \\ a_{21} & (a_{22} - \lambda) & a_{23} & \cdots \\ a_{31} & a_{32} & (a_{33} - \lambda) & \cdots \\ \cdots & \cdots & \cdots & \cdots \end{vmatrix} = 0\,. \tag{59}$$

Eq. for the required eigenvalues λ is a polynomial of degree n in λ and hence there are n solutions. These are:

1. not necessarily real (even if all the a_{ij} are real);
2. may be equal to others.

This is the characteristic equation of the eigenvalue problem.

Label roots as

$$\lambda_1,\ \lambda_2,\ \cdots,\lambda_n\,.$$

If two of the eigenvalues are equal, then the eigenvalue has a two-fold degeneracy, or that it is doubly-degenerate. Similarly, if there are r equal roots then this corresponds to an r-fold degeneracy.

Suppose we know eigenvalues λ_i. Have to solve

$$(\underline{A} - \lambda_i\,\underline{I})\,\underline{X}_i = \underline{0}$$

to find corresponding eigenvector $\underline{X}_i$. There are n eigenvectors $\underline{X}_i$ which can be written in terms of components as

$$\underline{X}_i = \begin{pmatrix} x_{1i} \\ x_{2i} \\ \vdots \\ \vdots \\ x_{ni} \end{pmatrix}.$$

Example Find the eigenvalues and eigenvectors of the matrix

$$\underline{A} = \begin{pmatrix} 3 & 2 \\ 1 & 4 \end{pmatrix}.$$

The characteristic equation is

$$| \underline{A} - \lambda \underline{I} | = \begin{vmatrix} (3-\lambda) & 2 \\ 1 & (4-\lambda) \end{vmatrix} = (3-\lambda)(4-\lambda) - 2 = 0\,,$$

giving solutions $\lambda_1 = 5$ and $\lambda_2 = 2$.

In the case of $\lambda_1 = 5$, we have

$$\begin{pmatrix} (3-\lambda) & 2 \\ 1 & (4-\lambda) \end{pmatrix} \begin{pmatrix} x_{11} \\ x_{21} \end{pmatrix} = \begin{pmatrix} -2 & 2 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} x_{11} \\ x_{21} \end{pmatrix} = \underline{0}\,.$$

This gives the two equations

$$\begin{aligned} -2x_{11} + 2x_{21} &= 0, \\ x_{11} - x_{21} &= 0. \end{aligned}$$

Equations not linearly independent so solution involves some arbitrary constant c_1;

$$x_{11} = x_{21} = c_1\,.$$

Similarly, for $\lambda_2 = 2$, we get

$$x_{12} = c_2, \quad x_{22} = -\tfrac{1}{2}c_2\,.$$

In summary

$$\begin{aligned} \lambda_1 = 5 \quad &\Longrightarrow \quad \underline{X}_1 = c_1 \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \\ \lambda_2 = 2 \quad &\Longrightarrow \quad \underline{X}_2 = c_2 \begin{pmatrix} 1 \\ -\frac{1}{2} \end{pmatrix}. \end{aligned}$$

c_i is an arbitrary constants but convenient to choose it so that $\underline{X}_i$ is a unit vector. This vector is then normalised. Scalar product of two (possibly complex) vectors was defined

$$(\underline{a}\,,\underline{b}) = a_1^*\,b_1 + a_2^*\,b_2 + \cdots + a_n^*\,b_n = \underline{a}^\dagger\,\underline{b}\,.$$

For lengths of eigenvectors to be unity, need

$$\underline{X}_1^\dagger\,\underline{X}_1 = \underline{X}_2^\dagger\,\underline{X}_2 = 1\,.$$

First eqs. gives

$$(c_1^* \;\; c_1^*) \begin{pmatrix} c_1 \\ c_1 \end{pmatrix} = 2\,|\,c_1\,|^2 = 1\,.$$

The phase of c_1 is completely arbitrary — equation only fixes the magnitude of the (complex) number c_1. Taking it to be real and positive, $c_1 = 1/\sqrt{2}$.

Second eqs. gives

$$(c_2^* \;\; -\tfrac{1}{2}c_2^*) \begin{pmatrix} c_2 \\ -\frac{1}{2}c_2 \end{pmatrix} = \tfrac{5}{4}\,|\,c_2\,|^2 = 1\,,$$

and so $c_2 = 2/\sqrt{5}$.

Final answer is

$$\begin{aligned} \lambda_1 = 5 \quad &\Longrightarrow \quad \underline{X}_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \\ \lambda_2 = 2 \quad &\Longrightarrow \quad \underline{X}_2 = \frac{2}{\sqrt{5}} \begin{pmatrix} 1 \\ -\frac{1}{2} \end{pmatrix}. \end{aligned}$$

2.11 Eigenvalues of Unitary Matrices

Any unitary matrix $\underline{U}$ satisfies

$$\underline{U}^\dagger \, \underline{U} = \underline{U} \, \underline{U}^\dagger = \underline{I} \, .$$

To find eigenvalues, solve

$$\underline{U} \, \underline{X} = \lambda \, \underline{I} \, \underline{X} \, . \tag{60}$$

Take the Hermitian conjugate of Eq. (60),

$$\begin{aligned} (\underline{U} \, \underline{X})^\dagger &= (\lambda \, \underline{I} \, \underline{X})^\dagger \\ \underline{X}^\dagger \, \underline{U}^\dagger &= \lambda^* \, \underline{X}^\dagger \, \underline{I} \, . \end{aligned} \tag{61}$$

Note that the Hermitian conjugate interchanges the order in a product.

Multiply the left hand sides of Eqs. (60, 61) together and also the right hand sides:

$$\underline{X}^\dagger \, \underline{U}^\dagger \, \underline{U} \, \underline{X} = \lambda^* \lambda \, \underline{X}^\dagger \, \underline{X} \, . \tag{62}$$

But $\underline{U}^\dagger \, \underline{U} = \underline{I}$, and $\underline{X}^\dagger \, \underline{X} = X^2$. Hence

$$X^2 = \mid \lambda \mid^2 \; X^2 \, . \tag{63}$$

Since $X^2 \neq 0$, can divide by this to get $\mid \lambda \mid = 1$, *i.e.* all the eigenvalues are (possibly complex) numbers of unit modulus;

$$\lambda = e^{i\phi} \quad \text{with } \phi \text{ real.} \tag{64}$$

2.12 Eigenvalues of Hermitian Matrices

A Hermitian matrix is one for which $\underline{H} = \underline{H}^\dagger$. Consider two eigenvector equations eigenvalues $\lambda_i \neq \lambda_j$;

$$\underline{H} \, \underline{X}_i = \lambda_i \, \underline{X}_i \, , \tag{65}$$

$$\underline{H} \, \underline{X}_j = \lambda_j \, \underline{X}_j \, . \tag{66}$$

Take the Hermitian conjugate of Eq. (65);

$$\begin{aligned} (\underline{H} \, \underline{X}_i)^\dagger &= (\lambda_i \, \underline{X}_i)^\dagger \, , \\ \underline{X}_i^\dagger \, \underline{H}^\dagger = \underline{X}_i^\dagger \, \underline{H} &= \lambda_i^* \, \underline{X}_i^\dagger \, . \end{aligned} \tag{67}$$

Now multiply Eq. (67) on the right by $\underline{X}_j$

$$\underline{X}_i^\dagger \, \underline{H} \, \underline{X}_j = \lambda_i^* \, \underline{X}_i^\dagger \, \underline{X}_j \, . \tag{68}$$

Go back to Eq.(66) and multiply it on the left by $\underline{X}_i^\dagger$;

$$\underline{X}_i^\dagger \, \underline{H} \, \underline{X}_j = \lambda_j \, \underline{X}_i^\dagger \, \underline{X}_j \, . \tag{69}$$

The left hand sides of Eqs. (68) and (69) are identical and so, for all i and j, the right hand sides have to be as well;

$$(\lambda_i^* - \lambda_j) \, \underline{X}_i^\dagger \, \underline{X}_j = 0 \, . \tag{70}$$

Take first $i = j$ **:**

$$(\lambda_i^* - \lambda_i)\, \underline{X}_i^\dagger\, \underline{X}_i = (\lambda_i^* - \lambda_i)\, X_i^2 = 0\,. \tag{71}$$

But since all X_i^2 are non-zero

$$\lambda_i^* - \lambda_i = 0\,, \tag{72}$$

which means that all the eigenvalues are <u>**real**</u>.

Now take $i \neq j$:

If the eigenvalues are non-degenerate, *i.e.* $i \neq j \implies \lambda_i \neq \lambda_j$, then

$$\underline{X}_i^\dagger\, \underline{X}_j = 0\,, \tag{73}$$

which means that the corresponding eigenvectors are orthogonal.

If two eigenvalues are the same, *i.e.* a particular root is doubly degenerate, then the proof fails because one can then have $\lambda_i - \lambda_j = 0$ for $i \neq j$. Nevertheless, can still <u>choose</u> linear combinations of corresponding eigenvectors to make all eigenvectors orthogonal.

Orthogonal basis set

Normalise the eigenvectors of a Hermitian matrix as in the 2×2 example. Then the $\underline{\hat{X}_i}$ are unit orthogonal vectors; can take as basis vectors for this n-dimensional space *i.e.* ny vector can be written as

$$\underline{V} = \sum_i V_i\, \underline{\hat{X}_i}\,.$$

This simple result will be used extensively in the second and third year Quantum Mechanics course. The Hamiltonian (Energy) operator is Hermitian and so its eigenfunctions are orthogonal. Any wavefunction can be expanded in terms of these eigenfunctions.

2.13 Useful Rules for Eigenvalues

1. If we group all the different $\underline{\hat{X}_i}$ column vectors together in a single $n \times n$ matrix $\underline{X}$, then the eigenvector equation can then be written in the form

$$\underline{A}\,\underline{X} = \underline{X}\,\underline{\Lambda}\,, \tag{74}$$

where $\underline{\Lambda}$ is the diagonal matrix of eigenvalues

$$\underline{\Lambda} = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}\,. \tag{75}$$

Take the determinant of Eq. (74), the determinant of a product rule shows that

$$|\, \underline{A} \,|\,|\, \underline{X} \,| = |\, \underline{\Lambda} \,|\,|\, \underline{X} \,|\,.$$

Hence

$$\mid \underline{\Lambda} \mid = \mid \underline{A} \mid \ .$$

Use this to check that we got the right answer for 2×2 matrix $\begin{pmatrix} 3 & 2 \\ 1 & 4 \end{pmatrix}$. Has determinant $\Delta = 10$ which is equal to product of eigenvalues 5 and 2.

2. The trace of a matrix is defined as the sum of diagonal elements;

$$tr\{\underline{A}\} = \sum_i a_{ii} \ . \tag{76}$$

For example, the 2×2 matrix above has $tr\{\underline{A}\} = 7$, which equals the sum of the eigenvalues 2 and 5. Is this just luck or is it much deeper?

Rewrite Eq. (74) by taking $\underline{X}$ over to the other side as an inverse matrix.

$$\underline{A} = \underline{X}\,\underline{\Lambda}\,\underline{X}^{-1} \ .$$

Take the trace. Writing it out explicitly

$$tr\{\underline{A}\} = \sum_i a_{ii} = \sum_{i,j,k} (\underline{X})_{ij}\,(\underline{\Lambda})_{jk}\,(\underline{X}^{-1})_{ki}$$

$$= \sum_{i,j,k} (\underline{\Lambda})_{jk}\,(\underline{X}^{-1})_{ki}\,(\underline{X})_{ij} = tr\{\underline{\Lambda}\,\underline{X}^{-1}\,\underline{X}\} = tr\{\underline{\Lambda}\} = \sum_i \lambda_i \ .$$

The trace of a matrix is equal to the sum of its eigenvalues.

3. If matrix $\underline{A}$ is Hermitian, then $\underline{X}$ is <u>unitary</u> because

$$\underline{X}_i^\dagger\,\underline{X}_j = \delta_{ij} \ .$$

2.14 Real Quadratic Forms

A general real quadratic form is written as

$$F = \underline{X}^T\,\underline{A}\,\underline{X} = \sum_{i,j} a_{ij}\,x_i\,x_j \ . \tag{77}$$

Simplify by assuming matrix $\underline{A}$ is symmetric, *i.e.* $a_{ij} = a_{ji}$. Coefficients can be read off by inspection. Eg if (Boas, p.422),

$$F = x^2 + 6xy - 2y^2 - 2yz + z^2 \ ,$$

then $a_{11} = 1$ is the coefficient of the x^2 term. Similarly, $a_{12} = a_{21} = 3$ is half the coefficient of the xy term. The coefficient is shared between two equal elements of the matrix. The equation can thus be re-written as

$$F = \begin{pmatrix} x \\ y \\ z \end{pmatrix} \begin{pmatrix} 1 & 3 & 0 \\ 3 & -2 & -1 \\ 0 & -1 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} \ .$$

Rotate the coordinate system

$$\underline{X} = \underline{R}\,\underline{Y} \tag{78}$$

such that the quadratic form has no cross terms of the kind $y_1 y_2$.

$$F = \underline{Y}^T\,\underline{R}^T\,\underline{A}\,\underline{R}\,\underline{Y} = \underline{Y}^T\,\underline{D}\,\underline{Y}\,, \tag{79}$$

where $\underline{D}$ is a diagonal matrix.

For rotating the axes the matrix $\underline{R}$ is orthogonal, $\underline{R}^T\,\underline{R} = \underline{I}$. From Eq. (79), Need to find $\underline{R}$ such that

$$\underline{R}^T\,\underline{A}\,\underline{R} = \underline{D}\,. \tag{80}$$

In principle have already solved this problem. $\underline{D}$ is the diagonal matrix of eigenvalues $\underline{\Lambda}$, and $\underline{R}$ is the matrix of eigenvectors.

Example

Diagonalise the quadratic form

$$F = 5x^2 - 4xy + 2y^2\,.$$

In terms of a matrix

$$F = (x\,,\,y)\begin{pmatrix} 5 & -2 \\ -2 & 2 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix},$$

which has eigenvalues

$$\begin{vmatrix} (5-\lambda) & -2 \\ -2 & (2-\lambda) \end{vmatrix} = \lambda^2 - 7\lambda + 6 = 0\,.$$

Two solutions, $\lambda_1 = 6$ and $\lambda_2 = 1$. [You could check these by showing that the trace of the matrix equals 7 and its determinant equals 6.]

For $\lambda_1 = 6$, the eigenvector equation is

$$\begin{pmatrix} -1 & -2 \\ -2 & -4 \end{pmatrix}\begin{pmatrix} r_{11} \\ r_{21} \end{pmatrix} = 0\,,$$

which gives $r_{11} = -2r_{21}$. Normalisation gives

$$\underline{r}_1 = \frac{1}{\sqrt{5}}\begin{pmatrix} -2 \\ 1 \end{pmatrix}.$$

For $\lambda_2 = 1$, the eigenvector equation is

$$\begin{pmatrix} 4 & -2 \\ -2 & 1 \end{pmatrix}\begin{pmatrix} r_{12} \\ r_{22} \end{pmatrix} = 0\,,$$

which gives $r_{22} = 2r_{12}$. The normalised eigenvector is

$$\underline{r}_2 = \frac{1}{\sqrt{5}}\begin{pmatrix} 1 \\ 2 \end{pmatrix},$$

and the rotation matrix

$$\underline{R} = \begin{pmatrix} -2/\sqrt{5} & 1/\sqrt{5} \\ 1/\sqrt{5} & 2/\sqrt{5} \end{pmatrix} .$$

Thus

$$F = 6x'^2 + y'^2 ,$$

where

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \underline{R}^T \begin{pmatrix} x \\ y \end{pmatrix} ,$$

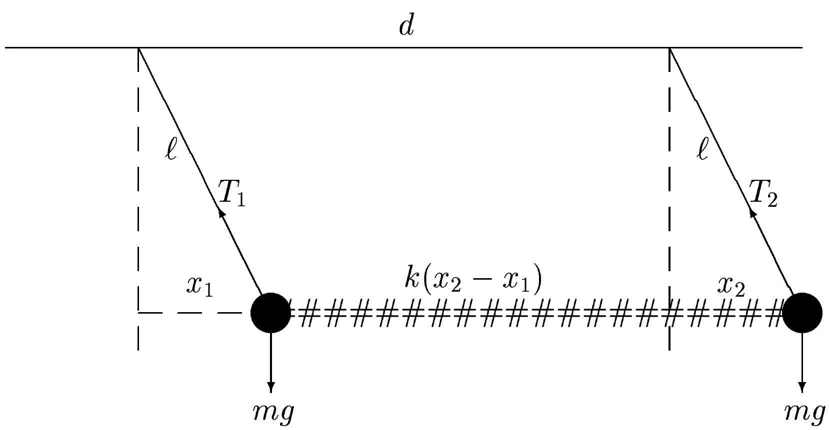
i.e.

$$\begin{aligned} x' &= \frac{1}{\sqrt{5}}(-2x+y) , \\ y' &= \frac{1}{\sqrt{5}}(x+2y) . \end{aligned}$$

You should check this by putting expressions for x' and y' into the new expression for F.

2.15 Normal Modes of Oscillation

Consider two point particles, each mass m, attached by light inextensible strings of length ℓ to a horizontal beam, the points of suspensions being a distance d apart. Connect the two masses by a light spring of natural length d and spring constant d. The force pulling the two masses together is $k(x_2 - x_1)$, where x_2 and x_1 are the instantaneous displacements of the masses from equilibrium. The tension T_i in the string produces a restoring horizontal force of mgx_i/ℓ (for small displacements).



The equations of motion of the system are

$$\begin{aligned} m\frac{d^2x_1}{dt^2} &= -\frac{mg}{\ell}x_1 + k(x_2 - x_1) , \\ m\frac{d^2x_2}{dt^2} &= -\frac{mg}{\ell}x_2 + k(x_1 - x_2) . \end{aligned}$$

In matrix form

$$\frac{d^2\underline{X}}{dt^2} = \underline{A}\,\underline{X} ,$$

where

$$\underline{A} = \begin{pmatrix} \alpha & \beta \\ \beta & \alpha \end{pmatrix} = \begin{pmatrix} -g/\ell - k/m & k/m \\ k/m & -g/\ell - k/m \end{pmatrix} .$$

These equations are coupled, in that $\ddot{x}_1$ depends also upon the value of x_2. Now find linear combinations of x_i such that equations become uncoupled. Let

$$\underline{X} = \underline{R}\,\underline{Y} ,$$

where $\underline{R}$ is an orthogonal matrix which does not depend upon time. Hence

$$\underline{R}\,\frac{d^2\underline{Y}}{dt^2} = \underline{A}\,\underline{R}\,\underline{Y} .$$

Multiply on the left by $\underline{R}^T$ and use $\underline{R}^T\,\underline{R} = \underline{I}$ to obtain

$$\frac{d^2\underline{Y}}{dt^2} = \underline{R}^T\,\underline{A}\,\underline{R}\,\underline{Y} .$$

For eqs to be uncoupled, need right-hand side to be a diagonal matrix which, as for the quadratic form problem, is the eigenvalue matrix, $\underline{\Lambda}$:

$$\underline{R}^T\,\underline{A}\,\underline{R} = \underline{\Lambda} ,$$

where $\underline{R}$ is the matrix of normalised eigenvectors. The new variables y_i satisfy the uncoupled equations

$$\ddot{y} = \lambda_i\, y .$$

First determine the eigenvalues from

$$\begin{vmatrix} -g/\ell - k/m - \lambda & k/m \\ k/m & -g/\ell - k/m - \lambda \end{vmatrix} = 0 .$$

Has the two solutions $\lambda_1 = -g/\ell$ and $\lambda_2 = -g/\ell - 2k/m$. Eqs of motion are

$$\begin{aligned} \ddot{y}_1 &= -\omega_1^2\, y_1 &&= -\frac{g}{\ell}\, y_1 , \\ \ddot{y}_2 &= -\omega_2^2\, y_2 &&= -\left(\frac{g}{\ell} + 2\frac{k}{m}\right)\, y_2 , \end{aligned}$$

which have general solution

$$\begin{aligned} y_1 &= \alpha_1 \sin\omega_1 t + \beta_1 \cos\omega_1 t , \\ y_2 &= \alpha_2 \sin\omega_2 t + \beta_2 \cos\omega_2 t \end{aligned}$$

The relation between x_i and y_i is given by rotation matrix $\underline{R}$, *i.e.* the eigenvectors of $\underline{A}$. For $\lambda_1 = -g/\ell$,

$$\begin{pmatrix} -k/m & k/m \\ k/m & -k/m \end{pmatrix} \begin{pmatrix} r_{11} \\ r_{21} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} .$$

gives $r_{11} = r_{21}$ and normalised eigenvector $\begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix}$.

For $\lambda_2 = -g/\ell - 2k/m$,

$$\begin{pmatrix} k/m & k/m \\ k/m & k/m \end{pmatrix} \begin{pmatrix} r_{12} \\ r_{22} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

gives $r_{12} = -r_{22}$, normalised eigenvector $\begin{pmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{pmatrix}$. The rotation matrix is then

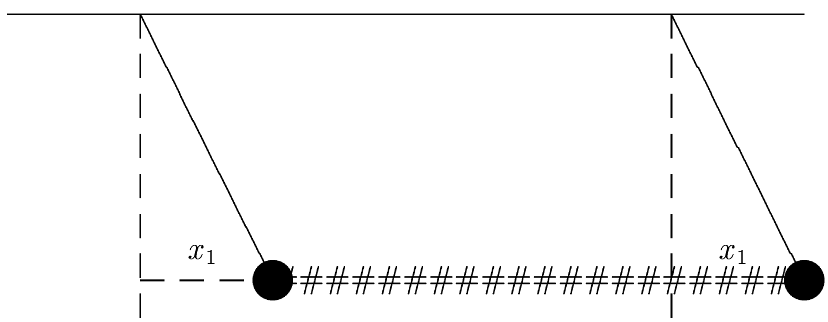
$$\underline{R} = \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{pmatrix}.$$

Old and new coordinates related by

$$\begin{aligned} x_1 &= \frac{1}{\sqrt{2}}(y_1 + y_2) \quad &:& \quad y_1 = \frac{1}{\sqrt{2}}(x_1 + x_2), \\ x_2 &= \frac{1}{\sqrt{2}}(y_1 - y_2) \quad &:& \quad y_2 = \frac{1}{\sqrt{2}}(x_1 - x_2). \end{aligned}$$

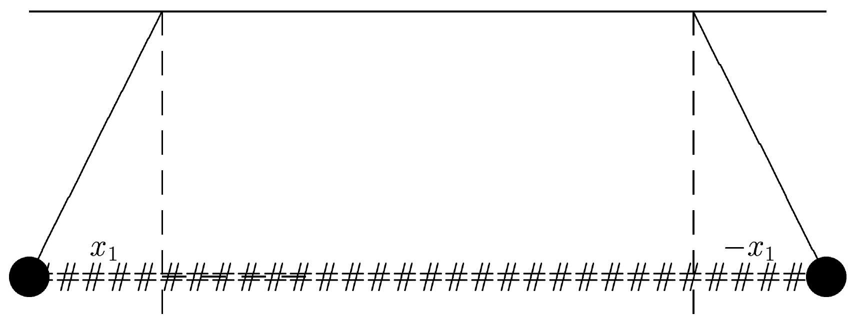
Call an uncoupled modes of oscillation a normal mode. Depending upon the boundary conditions, it is possible to excite one normal mode independently of the other. What do the normal modes look like in terms of the x_i.

Normal mode 1: $y_2 = 0$, and $x_1 = x_2 = y_1/\sqrt{2}$.



The two pendulums swing together in phase and of course, since the two pendulums are identical, the spring is neither stretched nor compressed. Effectively the spring doesn't influence this mode at all. Frequency $\omega_1 = \sqrt{g/\ell}$ is that for a free pendulum of the same length.

Normal mode 2: $y_1 = 0$, and $x_1 = -x_2 = y_2/\sqrt{2}$.



The two pendulums oscillate out of phase: the spring is alternately stretched and compressed. Compared to the first normal mode, the restoring forces are here <u>increased</u> because the spring is contributing something. Hence the frequency is higher:

$$\omega_2 = \sqrt{\frac{g}{\ell} + \frac{2k}{m}} \,.$$

A real problem has boundary conditions. Eg at time $t = 0$ take pendulum 1 to be at rest at equilibrium and pendulum 2 to be at rest at displacement $x_2 = a$. What is the subsequent motion? In terms of the y_i variables, at $t = 0$,

$$\begin{array}{rcl} y_1 = \dfrac{a}{\sqrt{2}} & : & y_2 = -\dfrac{a}{\sqrt{2}} \,, \\ \dot{y}_1 = 0 & : & \dot{y}_2 = 0 \,. \end{array}$$

Hence, at later times, the solutions are

$$\begin{aligned} y_1 &= \frac{a}{\sqrt{2}} \cos \omega_1 t \,, \\ y_2 &= -\frac{a}{\sqrt{2}} \cos \omega_2 t \,. \end{aligned}$$

In terms of the physical variables,

$$\begin{aligned} x_1 &= \frac{a}{2} \left(\cos \omega_1 t - \cos \omega_2 t \right) \,, \\ x_2 &= \frac{a}{2} \left(\cos \omega_1 t + \cos \omega_2 t \right) \,. \end{aligned}$$