PHAS2246E Mathematical Methods III

Intoduction

I am Jochen Weller
Room G10 in KLB
Phone 7679 7788 (internal 37788)
Email: jochen.weller@ucl.ac.uk
www: http://www.star.ucl.ac.uk/∼jweller/Teaching/mathmethods3.html
Books:

KF Riley, MP Hobson and SJ Bence, Mathematical Methods for Physics and Engineering (Cambridge, 2002).
For a simpler view see
KA Stroud, Engineering Mathematics (Palgrave, 2001).

Course aims:

- Last compulsory maths course
- Provide mathematical foundations for remaining physics
- Particularly geared to 2B22 Quantum Physics course

Course Syllabus:

- Linear Vector Spaces and Matrices
- Partial Differential Equations
- Series Solution of Differential Equations
- Legendre Functions
- Fourier Analysis
- Vector Operators

Coursework - 4 Problem sheets

- Coursework 1 - due 31/10/07
- Coursework 2 - due 21/11/07
- Coursework 3 - due 12/12/07
- Coursework 4 - due 9/1/08

75% of the exercises count (the ones where most points where achieved), i.e. last year there where 17 problems in all four courseworks, this results in 12.75 exercises. This is always rounded up, i.e. 13 in the example. This is then converted into a percentage. Note that a minimum of 15% in the coursework is required.

1 Linear Vector Spaces and Matrices

1.1 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,

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

and are perpendicular to each other;

$$\underline{\hat{e}}_1 \cdot \underline{\hat{e}}_2 = \underline{\hat{e}}_2 \cdot \underline{\hat{e}}_3 = \underline{\hat{e}}_3 \cdot \underline{\hat{e}}_1 = 0\,. \tag{2}$$

Summarised in one equation as

$$\underline{\hat{e}}_i \cdot \underline{\hat{e}}_j = \delta_{ij}\,, \tag{3}$$

where the Kronecker delta δ_{ij} is shorthand for

$$\delta_{ij} = \begin{cases} 1 & \text{if } \; i = j \\ 0 & \text{if } \; i \neq j \end{cases} \tag{4}$$

Any vector $\underline{v}$ in this three-dimensional space may be written down in terms of its <u>components</u> 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 $\underline{\hat{e}}_i$;

$$v_i = \underline{\hat{e}}_i \cdot \underline{v}\,. \tag{5}$$

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

$$\underline{u} \cdot \underline{v} = (u_1\,\underline{\hat{e}}_1 + u_2\,\underline{\hat{e}}_2 + u_3\,\underline{\hat{e}}_3) \cdot (v_1\,\underline{\hat{e}}_1 + v_2\,\underline{\hat{e}}_2 + v_3\,\underline{\hat{e}}_3) = u_1\,v_1 + u_2\,v_2 + u_3\,v_3 = \sum_{i=1}^{3} u_i\,v_i\,. \tag{6}$$

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

$$v^2 = \underline{v} \cdot \underline{v} = v_1^2 + v_2^2 + v_3^2\,. \tag{7}$$

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

$$v = \mid \underline{v} \mid = \sqrt{v^2} = \sqrt{v_1^2 + v_2^2 + v_3^2}\,. \tag{8}$$

A unit vector has length one.

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

$$\underline{v} = \underline{0} \qquad \Longleftrightarrow \qquad (v_1\,, v_2\,, v_3) = (0, 0, 0)\,. \tag{9}$$

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

$$\hat{\underline{e}}_3 \neq \alpha\,\hat{\underline{e}}_1 + \beta\,\hat{\underline{e}}_2\,, \tag{10}$$

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

$$\underline{u} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} : \quad \underline{v} = \begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix} : \quad \underline{w} = \begin{pmatrix} 7 \\ 8 \\ 9 \end{pmatrix} . \tag{11}$$

Then

$$\underline{w} = 2\underline{v} - \underline{u}\,. \tag{12}$$

We then say that $\underline{u}$, $\underline{v}$ and $\underline{w}$ are <u>linearly dependent</u>. 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 $\hat{\underline{e}}_i$. 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;

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

Note the $\hat{\underline{e}}_i$ 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.

1.2 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

$$\begin{aligned} \underline{a} + \underline{b} &= \underline{c} \in S. \\ \underline{c} = \underline{a} + \underline{b} &= \underline{b} + \underline{a} \ \ \text{(Commutative law)} \\ (\underline{a} + \underline{b}) + \underline{c} &= \underline{a} + (\underline{b} + \underline{c}) \ \ \text{(Associative law)}\,. \end{aligned} \tag{13}$$

2. Multiplication by a scalar (possibly complex)

$$\begin{array}{rcl} \underline{a} \in S & \Longrightarrow & \lambda\,\underline{a} \in S \quad (\lambda \text{ a complex number})\,, \\ \lambda\,(\underline{a} + \underline{b}) & = & \lambda\,\underline{a} + \lambda\,\underline{b}\,, \\ \lambda\,(\mu\,\underline{a}) & = & (\lambda\,\mu)\,\underline{a} \quad (\mu \text{ another complex number})\,. \end{array} \tag{14}$$

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

$$\underline{a} + \underline{0} = \underline{a} \tag{15}$$

for all vectors $\underline{a}$.

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

$$\underline{a} + (-\underline{a}) = \underline{0}\,. \tag{16}$$

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

$$\underline{X} = c_1\,\underline{X}_1 + c_2\,\underline{X}_2\ \cdots\ c_n\,\underline{X}_n\,. \tag{17}$$

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. (17) 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

$$(\underline{u}\,,\underline{v}) = \underline{u}\cdot\underline{v} = u_1^*\,v_1 + u_2^*\,v_2 + \cdots + u_n^*\,v_n\,. \tag{18}$$

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

$$(\underline{v}\,,\,\underline{u}) = v_1^*\,u_1 + v_2^*\,u_2 + \cdots + v_n^*\,u_n = (\underline{u}\,,\,\underline{v})^*\,. \tag{19}$$

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

   $$u^2 = (\underline{u}\,,\,\underline{u}) = u_1^*\,u_1 + u_2^*\,u_2 + \cdots + u_n^*\,u_n = \mid u_1\mid^2 + \mid u_2\mid^2 + \cdots + \mid u_n\mid^2\,. \tag{20}$$

   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 $v_i = (\hat{\underline{e}}_i\,,\,\underline{v})$.

Representations

Given a set of basis vectors $\hat{\underline{e}}_i$, any vector $\underline{v}$ in an n-dimensional space can be written uniquely in the form $\underline{v} = \sum_{i=1}^{n} v_i\,\hat{\underline{e}}_i$. 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.

1.3 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. A linear transformation $\hat{A}$ is defined by the following properties:

1. $\hat{A}(\underline{u} + \underline{u}) = \hat{A}\underline{u} + \hat{A}\underline{v}$

2. $\hat{A}(\alpha\underline{u}) = \alpha\hat{A}\underline{u}$,

with α a complex scalar.

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:

$$\underline{\hat{e}}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \quad (n \text{ terms in the column}). \tag{21}$$

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

$$\underline{a}_1 = \hat{A}\,\underline{\hat{e}}_1\,. \tag{22}$$

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

$$\underline{a}_1 = \begin{pmatrix} a_{11} \\ a_{21} \\ a_{31} \\ \vdots \\ a_{n1} \end{pmatrix}. \tag{23}$$

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

$$\underline{a}_i = \hat{A}\,\underline{\hat{e}}_i = \begin{pmatrix} a_{1i} \\ a_{2i} \\ a_{3i} \\ \vdots \\ a_{ni} \end{pmatrix}. \tag{24}$$

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

$$\underline{a}_i = a_{1i}\,\underline{\hat{e}}_1 + a_{2i}\,\underline{\hat{e}}_2 + a_{3i}\,\underline{\hat{e}}_3 + \cdots + a_{ni}\,\underline{\hat{e}}_n = \sum_{j=1}^{n} a_{ji}\,\underline{\hat{e}}_j\,. \tag{25}$$

as $\underline{\hat{e}}_i$ 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 $\underline{v} = \sum_i v_i\,\underline{\hat{e}}_i$. Then

$$\underline{u} = \hat{A}\,\underline{v} = \sum_i (\hat{A}\,\underline{\hat{e}}_i)\,v_i = \sum_{i,j} a_{ji}\,v_i\,\underline{\hat{e}}_j\,. \tag{26}$$

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

$$\underline{u} = \sum_j u_j\,\underline{\hat{e}}_j\,, \tag{27}$$

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

$$u_j = \sum_{i=1}^{n} a_{ji}\, v_i \,. \tag{28}$$

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 $\sum_j a_{ji}\, \underline{\hat{e}}_j$, 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

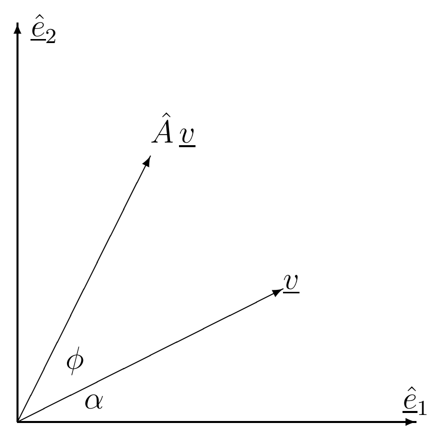
$$\underline{A} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} , \tag{29}$$

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

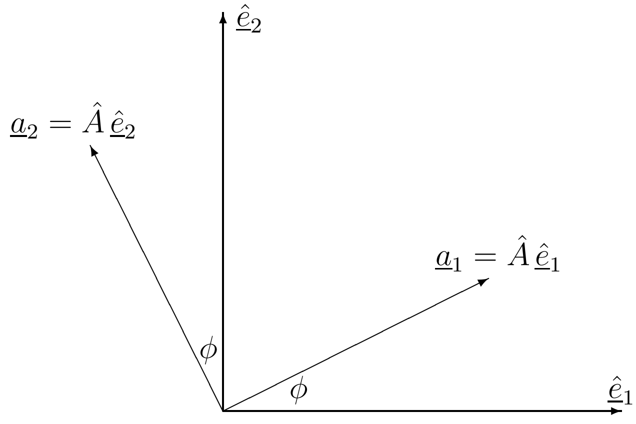
Signify a vector by putting an arrow on the top, underline it, or put a tilde under or over it or write it in bold in order to distinguish it from a scalar. Similarly must write something on the $\underline{A}$ in order to show that it is a matrix. The textbooks tend to use bold face — here we are going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

$$\begin{aligned} \underline{a}_1 = \hat{A}\,\underline{\hat{e}}_1 &= \cos\phi\,\underline{\hat{e}}_1 + \sin\phi\,\underline{\hat{e}}_2 \\ &= a_{11}\,\underline{\hat{e}}_1 + a_{21}\,\underline{\hat{e}}_2\,. \end{aligned}$$

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

$$\underline{A} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = \begin{pmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{pmatrix}\,. \tag{30}$$

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.

1.4 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

$$
\begin{aligned}
\underline{w} &= \hat{B}\,\underline{v} \\
\underline{u} &= \hat{A}\,\underline{w}\,. \\
\underline{u} &= \hat{A}\,\hat{B}\,\underline{v} = \hat{C}\underline{v}\,.
\end{aligned}
\tag{31}
$$

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

$$
\begin{aligned}
w_i &= \sum_j b_{ij}\,v_j \\
u_k &= \sum_i a_{ki}\,w_i \\
&= \sum_{i,j} a_{ki}\,b_{ij}v_j \\
&= \sum_j c_{kj}\,v_j\,.
\end{aligned}
\tag{32}
$$

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

$$
c_{kj} = \sum_{i=1}^{n} a_{ki}\,b_{ij}\,. \tag{33}
$$

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.

1.5 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

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

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 $|\,A\,|$ 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 $n \times n$.

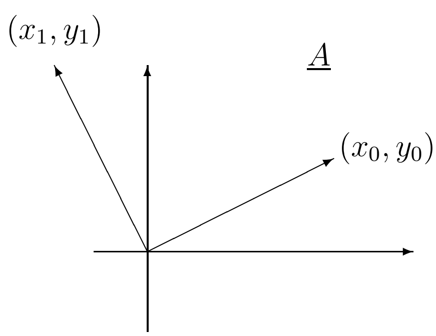
Matrices do not commute because they are constructed to represent linear operations and, in general, such operations do not commute. It can matter in which order you do certain operations.

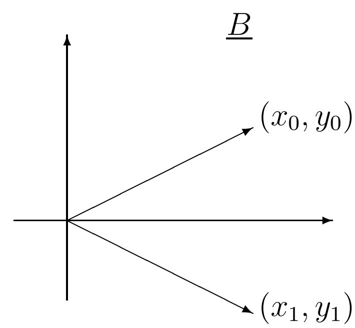
On the other hand,

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

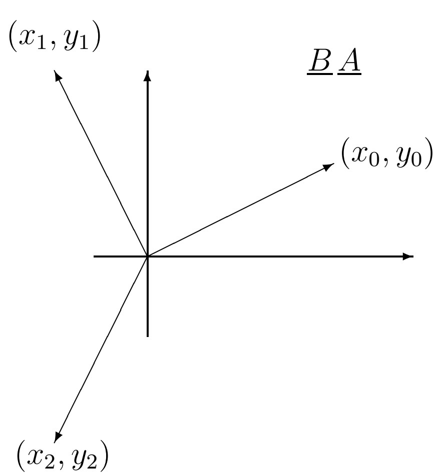
Will assume you are familiar with the actual multiplication process in practice. If not, you have been warned!

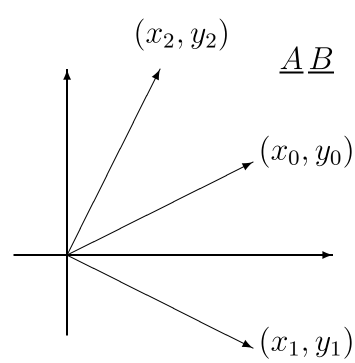
Example 1
Let $\underline{A}$ represent a rotation of 90° around the z-axis and $\underline{B}$ a reflection in the x-axis.





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.

Determinants

A 4×4 determinant can be reduced to four 3×3 determinants as

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{vmatrix}$$

$$= a_{11}\begin{vmatrix} a_{22} & a_{23} & a_{24} \\ a_{32} & a_{33} & a_{34} \\ a_{42} & a_{43} & a_{44} \end{vmatrix} - a_{12}\begin{vmatrix} a_{21} & a_{23} & a_{24} \\ a_{31} & a_{33} & a_{34} \\ a_{41} & a_{43} & a_{44} \end{vmatrix} + a_{13}\begin{vmatrix} a_{21} & a_{22} & a_{24} \\ a_{31} & a_{32} & a_{34} \\ a_{41} & a_{42} & a_{44} \end{vmatrix} - a_{14}\begin{vmatrix} a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \\ a_{41} & a_{42} & a_{43} \end{vmatrix} \tag{34}$$

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

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.

$$|\,\underline{A}\,\underline{B}\,| = |\,\underline{A}\,| \times |\,\underline{B}\,|\,. \tag{35}$$

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 $|\,\underline{A}\,\underline{B}\,| =$ $|\,\underline{B}\,\underline{A}\,| = -1$. This result for the determinant of products will prove very useful later.

1.6 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

$$\underline{A}\,\underline{I} = \underline{A}\,. \tag{36}$$

Similarly

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

and

$$\underline{I}\,\underline{A} = \underline{A}\,. \tag{37}$$

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

$$(\underline{A}^T)_{ij} = (\underline{A})_{ji}\;. \tag{38}$$

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

$$(\underline{A}\,\underline{B})^T = \underline{B}^T\, \underline{A}^T \; . \tag{39}$$

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 <u>orthogonal</u> 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:

$$(\underline{A}^*)_{ij} = a^*_{ij} \; . \tag{40}$$

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 (†).

$$\underline{A}^\dagger = (\underline{A}^T)^* = (\underline{A}^*)^T . \tag{41}$$

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:

$$(\underline{A}\,\underline{B})^\dagger = \underline{B}^\dagger\,\underline{A}^\dagger . \tag{42}$$

Unitary Matrices

Matrix $\underline{U}$ is unitary if

$$\underline{U}^\dagger\,\underline{U} = \underline{I} . \tag{43}$$

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.

1.7 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

$$\begin{array}{cc} a + \frac{3}{2}b = 0 & c + 4d = 0\,, \\ a + 4b = 1 & c + \frac{3}{2}d = \frac{1}{2}\,. \end{array}$$

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

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

The determinant $|\,A\,|$ 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

$$\Delta = \sum_j a_{ij}\, M_{ij}\, (-1)^{i+j} = \sum_i a_{ij}\, M_{ij}\, (-1)^{i+j} \,. \tag{44}$$

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:

$$C_{ij} = (-1)^{i+j}\, M_{ij} \,, \tag{45}$$

so that

$$\Delta = \sum_{i\ or\ j} a_{ij}\, C_{ij} \,. \tag{46}$$

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:

$$[\underline{A}^{\mathrm{adj}}]_{ij} = C_{ji} \,. \tag{47}$$

Theorem

For any square matrix,

$$\underline{A}^{-1} = \underline{A}^{\mathrm{adj}} / \mid A \mid \,. \tag{48}$$

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. (48) 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:

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

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 $\mid A \mid = 0$, we say the matrix 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

$$(\underline{A}\,\underline{B})^{-1} = \underline{B}^{-1}\,\underline{A}^{-1} . \tag{49}$$

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,

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

1.8 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,

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

If the determinant does not vanish, this leads to Cramer's rule discussed in the first lecture. $\sum\limits_i (\underline{A}^{\mathrm{adj}})_{ji}\,b_i$ 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 $\mid A \mid = 0$.

1.9 Eigenvalues and Eigenvectors

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

$$\underline{A}\,\underline{X} = \lambda\,\underline{X} = \lambda\,\underline{I}\,\underline{X}\,, \tag{50}$$

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. (50) as

$$(\underline{A} - \lambda\,\underline{I})\,\underline{X} = \underline{0}\,. \tag{51}$$

Set of n homogeneous linear equations which has interesting solutions if

$$|\,\underline{A} - \lambda\,\underline{I}\,| = 0\,. \tag{52}$$

Explicitly:

$$\begin{vmatrix} (a_{11}-\lambda) & a_{12} & a_{13} & \cdots \\ a_{21} & (a_{22}-\lambda) & a_{23} & \cdots \\ a_{31} & a_{32} & (a_{33}-\lambda) & \cdots \\ \cdots & \cdots & \cdots & \cdots \end{vmatrix} = 0\,. \tag{53}$$

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

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

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

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

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

Final answer is

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

Eigenvalues of Unitary Matrices

(Skipped)

Any unitary matrix $\underline{U}$ satisfies

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

To find eigenvalues, solve

$$\underline{U}\,\underline{X} = \lambda\,\underline{I}\,\underline{X}\,. \tag{54}$$

Take the Hermitian conjugate of Eq. (54),

$$\begin{aligned} (\underline{U}\,\underline{X})^\dagger &= (\lambda\,\underline{I}\,\underline{X})^\dagger \\ \underline{X}^\dagger\,\underline{U}^\dagger &= \lambda^*\,\underline{X}^\dagger\,\underline{I}\,. \end{aligned} \tag{55}$$

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

Multiply the left hand sides of Eqs. (54, 55) together and also the right hand sides:

$$\underline{X}^\dagger\,\underline{U}^\dagger\,\underline{U}\,\underline{X} = \lambda^*\lambda\,\underline{X}^\dagger\,\underline{X}\,. \tag{56}$$

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

$$X^2 = \mid\lambda\mid^2\ X^2\,. \tag{57}$$

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;

$$\lambda = e^{i\phi}\ \text{ with } \phi \text{ real.} \tag{58}$$

1.10 Eigenvalues of Hermitian Matrices

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

$$\underline{H}\,\underline{X}_i = \lambda_i\,\underline{X}_i\,, \tag{59}$$

$$\underline{H}\,\underline{X}_j = \lambda_j\,\underline{X}_j\,. \tag{60}$$

Take the Hermitian conjugate of Eq. (59);

$$\begin{aligned} (\underline{H}\,\underline{X}_i)^\dagger &= (\lambda_i\,\underline{X}_i)^\dagger\,, \\ \underline{X}_i^\dagger\,\underline{H}^\dagger = \underline{X}_i^\dagger\,\underline{H} &= \lambda_i^*\,\underline{X}_i^\dagger\,. \end{aligned} \tag{61}$$

Now multiply Eq. (61) on the right by $\underline{X}_j$

$$\underline{X}_i^\dagger\,\underline{H}\,\underline{X}_j = \lambda_i^*\,\underline{X}_i^\dagger\,\underline{X}_j\,. \tag{62}$$

Go back to Eq.(60) and multiply it on the left by $\underline{X}_i^\dagger$;

$$\underline{X}_i^\dagger\,\underline{H}\,\underline{X}_j = \lambda_j\,\underline{X}_i^\dagger\,\underline{X}_j\,. \tag{63}$$

The left hand sides of Eqs. (62) and (63) are identical and so, for all i and j, the right hand sides have to be as well;

$$(\lambda_i^* - \lambda_j)\, \underline{X}_i^\dagger\, \underline{X}_j = 0\,. \tag{64}$$

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

$$(\lambda_i^* - \lambda_i)\, \underline{X}_i^\dagger\, \underline{X}_i = (\lambda_i^* - \lambda_i)\, X_i^2 = 0\,. \tag{65}$$

But since all X_i^2 are non-zero

$$\lambda_i^* - \lambda_i = 0\,, \tag{66}$$

which means that all the eigenvalues are **real**.

Now take $i \neq j$:

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

$$\underline{X}_i^\dagger\, \underline{X}_j = 0\,, \tag{67}$$

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 choose 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.

1.11 Useful Rules for Eigenvalues

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

$$\underline{A}\,\underline{W} = \underline{W}\,\underline{\Lambda}\,, \tag{68}$$

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

$$\underline{\Lambda} = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ : & : & \cdots & : \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}\,. \tag{69}$$

Take the determinant of Eq. (68), the determinant of a product rule shows that

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

Hence

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

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;

$$tr\{\underline{A}\} = \sum_i a_{ii} \ . \tag{70}$$

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. (68) by taking $\underline{W}$ over to the other side as an inverse matrix.

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

Take the trace. Writing it out explicitly

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

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

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

3. If matrix $\underline{A}$ is Hermitian, then $\underline{W}$ is unitary because

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

1.12 Real Quadratic Forms

A general real quadratic form is written as

$$F = \underline{X}^T\,\underline{A}\,\underline{X} = \sum_{i,j} a_{ij}\,x_i\,x_j \ . \tag{71}$$

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

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

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

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

Rotate the coordinate system

$$\underline{X} = \underline{R}\,\underline{Y} \tag{72}$$

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

$$F = \underline{Y}^T\,\underline{R}^T\,\underline{A}\,\underline{R}\,\underline{Y} = \underline{Y}^T\,\underline{D}\,\underline{Y}\,, \tag{73}$$

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. (73), Need to find $\underline{R}$ such that

$$\underline{R}^T\,\underline{A}\,\underline{R} = \underline{D}\,. \tag{74}$$

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

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

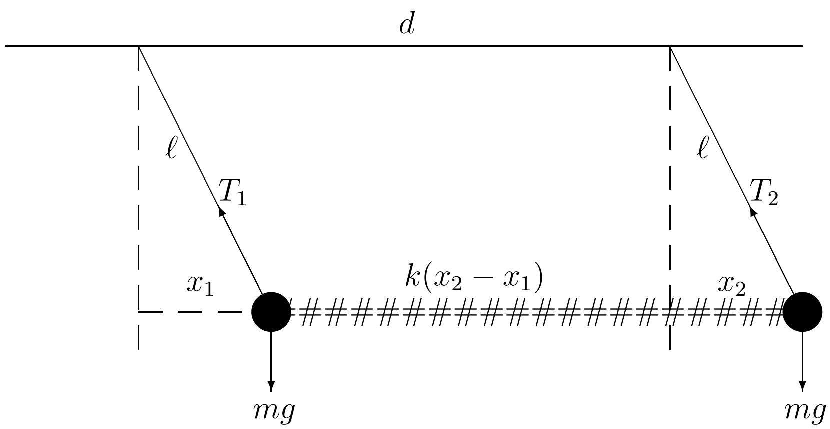
i.e.

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

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

1.13 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 k. 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

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

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

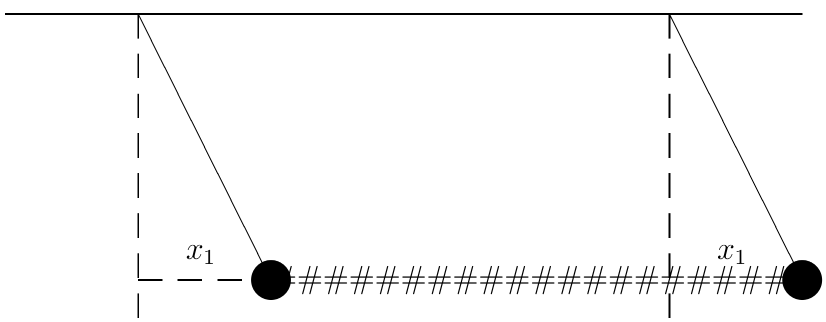
$$\underline{R} = \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{pmatrix}.$$

Old and new coordinates related by

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

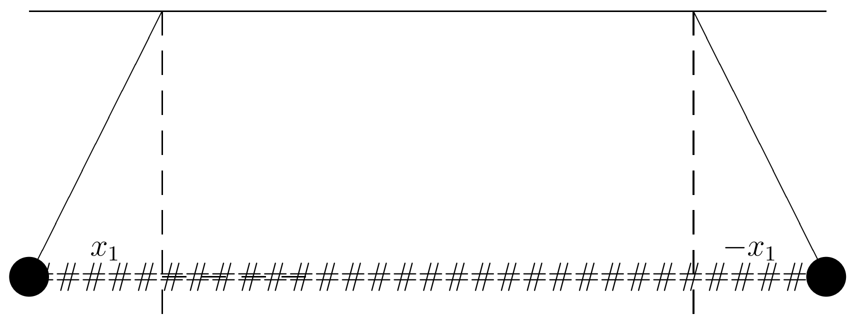
Call an uncoupled modes of oscillation a <u>normal mode</u>. Depending upon the boundary conditions, it is possible to excite one normal mode independently of the other. What do the normal modes look like in terms of the x_i.

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 $t = 0$,

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

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,

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

2 Partial Differential Equations

2.1 Introduction

Have solved ordinary differential equations, *i.e.* ones where there is one independent and one dependent variable. Only ordinary differentiation is therefore involved. As the world is three-dimensional, most differential equations are functions of three spatial variables, eg $(x\,,\,y\,,\,z)$, and maybe time t also. Typical example is Laplace equation

$$\nabla^2\, V(\underline{r}) = 0\,,$$

where $V(\underline{r})$ is the electrostatic potential in region where there is no charge. The operator ∇^2, called the Laplacian, was introduced last year. In Cartesian coordinates

$$\nabla^2\, V = \frac{\partial^2 V}{\partial x^2} + \frac{\partial^2 V}{\partial y^2} + \frac{\partial^2 V}{\partial z^2}\,. \tag{75}$$

Another important example is the time-independent Schrödinger equation for 1 particle

$$-\frac{\hbar^2}{2m}\,\nabla^2\,\Psi(\underline{r}) + V(\underline{r})\,\Psi(\underline{r}) = E\,\Psi(\underline{r})\,, \tag{76}$$

for the quantum-mechanical motion of a particle of mass m in a potential $V(\underline{r})$. $\Psi(\underline{r})$ is the particle's wave function and $\hbar = h/2\pi$, where h is Planck's constant. There are many more examples that you will come across later in your degree programme.

2.2 Classification of Differential Equations

Before considering various differential equations (DE) in detail it is worth defining some of the terms used to classify these equations into different types. The following terms are used: **Order**. The order of a DE is the order of its <u>highest</u> derivative, so

$$a_n(x)\frac{d^n y}{dx^n} + a_{n-1}(x)\frac{d^{n-1}y}{dx^{n-1}} + \ldots a_0(x)y = 0 \tag{77}$$

is a DE of order n. This definition holds even if there are several variables, so

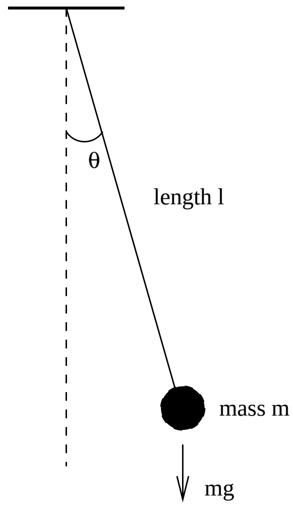
$$\frac{\partial^3 y}{\partial x^3} + \frac{\partial^2 y}{\partial t^2} = 0\,, \tag{78}$$

is a third-order.

Linearity. A linear DE can be written entirely as a linear function. *i.e.* no powers above the first power, of the unknown function and its derivative. So

$$a_n(x)\frac{d^n y}{dx^n} + a_{n-1}(x)\frac{d^{n-1}y}{dx^{n-1}} + \ldots a_0(x)y = b(x) \tag{79}$$

is linear if the a_i's and b are functions of x only. It is non-linear if any of the a_i's depend on y. Example: a pendulum



$$\frac{d^2\theta}{dt^2} + \frac{g}{l}\sin\theta = 0 \tag{80}$$

is non-linear in θ. However if θ is small then $\sin\theta \approx \theta$ and the DE

$$\frac{d^2\theta}{dt^2} + \frac{g}{l}\theta = 0 \tag{81}$$

is linear. Linear DE's are important because they are easier to solve.

Ordinary/Partial. If an unknown function, eg y, is a function of only one variable, eg x, then one gets ordinary DEs such as

$$\frac{dy}{dx} = c\ . \tag{82}$$

If y is function of more than one variable, eg x and t then one gets a partial DE eg

$$\frac{\partial y(x,t)}{\partial x} + \frac{\partial y(x,t)}{\partial t} = c \tag{83}$$

provided the variables, x and t, are independent. If the variables are dependent, eg $x = f(s,t)$, then it is necessary to specify which are held constant

$$\left.\frac{\partial y(x,t)}{\partial x}\right|_t = c(s,t) \tag{84}$$

Such constructions are familiar from thermodynamics where P (pressure), V (volume) and T (temperature) are all inter-related eg by the ideal gas equation $PV = nRT$ and many functions, such as entropy S, have to written as partial DEs. This means that

$$\left.\frac{\partial S}{\partial T}\right|_P \neq \left.\frac{\partial S}{\partial T}\right|_V \tag{85}$$

Homogeneous. Means slightly different things for linear and non-linear DEs. Will only consider the linear DE case.

A matrix equation such as $\underline{Ax} = \underline{b}$ is homogeneous if $\underline{b} = 0$. Similarly, a (second-order) DE

$$P(x)\frac{d^2y}{dx^2} + Q(x)\frac{dy}{dx} + R(x)y = G(x) \tag{86}$$

is homogeneous if $G(x) = 0$ and is inhomogenous if $G(x) \neq 0$. Solving the homogeneous DE is usually the first step in solving an inhomogenous DE. We will restrict ourselves to homogenous DEs.

Solutions. By a solution of an ordinary DE

$$F(x, y(x), \frac{dy}{dx}, \frac{d^2y}{dx^2}, \ldots) = 0$$

we mean some function $y = u(x)$ in the range $a < x < b$ for which the problem is defined. This solution can always be verfied by direct substitution. Does

$$F(x, u(x), \frac{du}{dx}, \frac{d^2u}{dx^2}, \ldots) = 0 \ ?$$

Uniqueness. A DE in general will have more than one solution because:

1. There are unknown constants which can only be determined by the <u>boundary conditions</u>. Boundary conditions give information about the unknown function (or its derivatives) at some point. Eg $y = 0$ at $x = 1$ is a boundary condition. n boundary conditions are required to determine constants for an n^{th}-order equation. So a second-order DE requires 2 boundary conditions.

2. For an n^{th}-order DE there are usually n independent functions, $u(x)$, satisfying the DE. So a second-order DE has 2 solutions. Which solution is correct is often determined by the physics of the problem.

Existence. There is no guarantee that a DE will have a solution of the form $u(x)$.

Superposition Principle

If V_1 and V_2 are two solutions of any linear, homogeneous DE such as $\nabla^2 V(\underline{r}) = 0$, then $V = c_1 V_1 + c_2 V_2$, where c_1 and c_2 are arbitrary constants, is another solution. Used extensively for ordinary DEs, eg simple harmonic motion problem; is equally valid for partial DEs. This ability to add solutions is called the Superposition Principle. Of fundamental importance in Quantum Mechanics. Will exploit the superposition principle extensively when solving partial DEs.

2.3 Separation of variables

Most DEs that characterise physical problems depend on many variables and cannot be directly solved.Sometimes can solve these multi-dimensional problems by separation of variables which turns a partial DE in n variables into n ordinary DEs each in one variable. Take an $n = 2$ example

$$a(x,y)\frac{\partial^2 u}{\partial x^2} + b(x,y)\frac{\partial^2 u}{\partial y^2} = 0\,. \tag{87}$$

If this is separable we can write $u(x,y) = X(x)Y(y)$ which gives

$$a(x,y)Y(y)\frac{d^2 X}{dx^2} + b(x,y)X(x)\frac{d^2 Y}{dy^2} = 0\,, \tag{88}$$

or, dividing through by XY and re-arranging:

$$\frac{a(x,y)}{X(x)}\frac{d^2 X}{dx^2} = -\frac{b(x,y)}{Y(y)}\frac{d^2 Y}{dy^2}\,. \tag{89}$$

This equation is separable *provided* that the left-hand side can be written totally in terms of x and the right-hand side totally in terms of y. This may require some re-arrangement between $a(x,y)$ and $b(x,y)$ to give $A(x)$ and $B(y)$, respectively functions of x and y only.

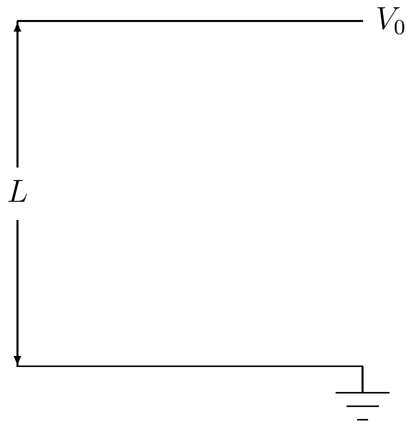
If eq. is separable, then have relationship of form $f(x) = g(y)$. Since relationship holds for all values of x and y, must mean that $f(x) = c = g(y)$, where c is some constant, often for convenience written as a square eg l^2. Can solve separately two equations

$$\frac{A(x)}{X(x)}\frac{d^2 X}{dx^2} = c, \qquad \frac{B(y)}{Y(y)}\frac{d^2 Y}{dy^2} = -c\,. \tag{90}$$

Note that separability depends on the coordinates chosen, it may be necessary to change coordinates.

Laplace's equation in Cartesian coordinates

Let us illustrate this with a physical example. Consider two infinitely large conducting plates. The one at $z = 0$ is earthed while that at $z = L$ is kept at a constant voltage V_0.



What is the potential between the two plates? You all know that the answer must be $V = V_0 z/L$ but we are going to derive this by solving the partial differential equation. This will demonstrate the techniques to be used in more complex cases.

Between the two plates, there is no charge and so the potential in this region satisfies Laplace's equation

$$\nabla^2 V(\underline{r}) = 0 \,.$$

The boundary conditions to be applied are that, independent of the values of x and y, on the plates

$$\begin{aligned} V = 0 \qquad &\text{at} \qquad z = 0\,, \\ V = V_0 \qquad &\text{at} \qquad z = L\,. \end{aligned} \tag{91}$$

Since the boundary conditions are expressed easily in terms of Cartesian coordinates, it makes obvious sense to attack the problem in this coordinate system. [Could also use cylindrical polar coordinates.] In this system, Laplace's equation becomes

$$\nabla^2 V = \frac{\partial^2 V}{\partial x^2} + \frac{\partial^2 V}{\partial y^2} + \frac{\partial^2 V}{\partial z^2} = 0\,.$$

Let us try for a solution of the form

$$V(x, y, z) = (\text{function of } x) \times (\text{function of } y) \times (\text{function of } z),$$

$$V(x, y, z) = X(x)\, Y(y)\, Z(z)\,. \tag{92}$$

At the moment we are just trying to get a single solution of the equation. If there is no solution of this kind then we will have to try something else — but of course there will be! Substituting the product form of Eq. (92) into Laplace's equation, we get

$$Y\,Z\,\frac{d^2X}{dx^2} + X\,Z\,\frac{d^2Y}{dy^2} + X\,Y\,\frac{d^2Z}{dz^2} = 0\,. \tag{93}$$

Note that we now have complete differentials (straight d's) because X is a function of only one variable (x), and similarly for Y and Z. Now divide through the equation by the product $V = X\,Y\,Z$ to get

$$\frac{1}{X}\left(\frac{d^2X}{dx^2}\right) + \frac{1}{Y}\left(\frac{d^2Y}{dy^2}\right) + \frac{1}{Z}\left(\frac{d^2Z}{dz^2}\right) = 0\,. \tag{94}$$

Now the first term in Eq. (94) is a function only of x, the second only of y, and the third only of z. **BUT** x, y, and z are independent variables. This means that we could keep y and z fixed and vary just x. In so doing, the second and third terms remain fixed because they only depend upon y and z respectively. Hence the first term must also remain fixed even if x changes. That is, the first term is a constant, as are the second and third. Thus

$$\begin{aligned}\frac{1}{X}\left(\frac{d^2X}{dx^2}\right) &= -\ell^2\,,\\ \frac{1}{Y}\left(\frac{d^2Y}{dy^2}\right) &= -m^2\,,\\ \frac{1}{Z}\left(\frac{d^2Z}{dz^2}\right) &= +n^2\,.\end{aligned} \tag{95}$$

with

$$n^2 = \ell^2 + m^2\,. \tag{96}$$

Note that n^2, ℓ^2 and m^2 are as yet arbitrary constants and could be negative. ℓ, m, n are not necessarily integers.

Have to solve

$$\frac{d^2X}{dx^2} = -\ell^2\,X\,. \tag{97}$$

For real $\ell \neq 0$, this is the simple harmonic oscillator equation

$$X = a_\ell \cos \ell x + b_\ell \sin \ell x\,, \tag{98}$$

where a_ℓ and b_ℓ are arbitrary constants which must be fixed by the boundary conditions. For special case $\ell = 0$, solution simplifies to

$$X = a_0 + b_0\,x\,. \tag{99}$$

If ℓ^2 is negative, put $\ell = il$; the $\cos lx$ and $sinlx$ become $\cosh \ell x$ and $i \sinh \ell x$. Have seen such changes before when studying the damped oscillator in 1B27.

Solutions for Y are similar to those for X, but with m replacing l. For Z have

$$\left(\frac{d^2 Z}{dz^2}\right) = +n^2\, Z\,. \tag{100}$$

Has solutions

$$\begin{aligned} Z = &\; e_n \cosh nz + f_n \sinh nz \qquad (n \neq 0)\,, \\ = &\; e_0 + f_0\, z \qquad\qquad\qquad (n = 0)\,. \end{aligned} \tag{101}$$

As a consequence, solutions of the separable form <u>do</u> exist. For example, one solution would be with $\ell = 3$, $m = 4$, and $n = 5$.

$$V(x, y, z) = \text{Constant} \times (\sin 3x) \times (\cos 4y) \times (\sinh 5z)$$

is a solution of Laplace's equation, but many more with different values of $(\ell,\, m,\, n)$ exist.

Most general solution is

$$V(x, y, z) = \text{Constant} \times \left\{ \begin{array}{c} \sin \ell x \\ \cos \ell x \end{array} \right\} \times \left\{ \begin{array}{c} \sin my \\ \cos my \end{array} \right\} \times \left\{ \begin{array}{c} \sinh nz \\ \cosh nz \end{array} \right\}$$

with constraint $n^2 = \ell^2 + m^2$.

By the superposition principle, any linear combination of such solutions is also a solution. The most general superposition is

$$\begin{aligned} V(x, y, z) \;\; = \;\; & \sum_{\ell, m} \{a_{\ell m} \cos \ell x + b_{\ell m} \sin \ell x\} \times \{c_{\ell m} \cos my + d_{\ell m} \sin my\} \\ & \times \{e_{\ell m} \cosh nz + f_{\ell m} \sinh nz\}\;. \end{aligned} \tag{102}$$

Here the solutions from $l, m, n = 0$ have to be added. For any choice of ℓ and m, with $n = \sqrt{\ell^2 + m^2}$, the above product is a solution. Hence the sum is also a solution. Note ℓ and m do not have to be integers and so the above need not be a discrete sum. Also note that if $\ell \to 0$, cosine is replaced by 1 and sine by x.

Imposing boundary conditions

Solution Eq. (102) is quite general, need to relate it potential problem of two parallel plates: have to impose the boundary conditions.

At $z = 0$,

$$V(z=0) = \sum_{\ell m} e_{\ell m} \, \{a_{\ell m} \, \cos \ell x + b_{\ell m} \, \sin \ell x\} \times \{c_{\ell m} \, \cos my + d_{\ell m} \, \sin my\} = 0$$

for all values of x and y. Hence $e_{\ell m} = 0$ for all ℓ and m. Most general solution simplifies to

$$V(x,y,z) = \sum_{\ell m} \sinh nz \times \{a_{\ell m} \, \cos \ell x + b_{\ell m} \, \sin \ell x\} \times \{c_{\ell m} \, \cos my + d_{\ell m} \, \sin my\} \; , \quad (103)$$

where coefficient $f_{\ell m}$ has been absorbed into redefined $a_{\ell m}$ and $b_{\ell m}$.
At $z = L$,

$$V(z=L) = \sum_{\ell m} \sinh nL \times \{a_{\ell m} \, \cos \ell x + b_{\ell m} \, \sin \ell x\} \times \{c_{\ell m} \, \cos my + d_{\ell m} \, \sin my\} = V_0 \; ,$$

for all x and y. Clearly, only solution which gives something independent of x and y is the special case of $\ell = m = n = 0$. Write this explicitly as

$$V(x,y,z) = z \, \{a + bx\}\{c + dy\} \; . \quad (104)$$

At $z = L$,

$$V_0 = L \, \{a + bx\}\{c + dy\}$$

for all (x, y) so that $b = d = 0$ and $ac = V_0/L$. The final solution is, from Eq. (104), the expected

$$V = \frac{V_0 z}{L} \; .$$

Comments

1. Method of solution is *Separation of Variables*: look for a solution which is a product of a function of x times a function of y times a function of z. Reduces problem to that of solving three ordinary differential equations in x, y and z.

2. Have found an infinite number of solutions of the Laplace equation, but have **not** shown that we have found them all.

3. In more complicated examples the ordinary differential equations may be very much harder to solve than the simple oscillator equations here.

4. Unlike the present case, in general you cannot guess the final answer at the start!

2.4 One-dimensional Wave Equation

Seen the wave equation in the first year 1B24 Waves and Optics course. In one dimension, for example a guitar string clamped at $x = 0$ and $x = L$, the displacement $y(x,t)$ obeys

$$\frac{\partial^2 y}{\partial x^2} - \frac{1}{c^2}\frac{\partial^2 y}{\partial t^2} = 0\,, \tag{105}$$

where t is the time variable and c the (constant) speed of wave propagation.

Looking for a solution in the form of a product

$$y(x,t) = X(x)\,T(t) \tag{106}$$

leads to

$$T\,\frac{d^2X}{dx^2} - \frac{1}{c^2}\,X\,\frac{d^2T}{dt^2} = 0\,. \tag{107}$$

After dividing out by $y = X\,T$ and taking one term over to the right hand side, we are left with

$$\frac{1}{X}\left(\frac{d^2X}{dx^2}\right) = \frac{1}{c^2T}\left(\frac{d^2T}{dt^2}\right)\,. \tag{108}$$

The left hand side is a function only of x and the right hand side purely of t. Since x and t are independent variables, this means that both sides are equal to a constant, which we shall call $-\omega^2$.

Reduced to solution of two ordinary differential equations

$$\begin{aligned}\left(\frac{d^2X}{dx^2}\right) + \omega^2\,X &= 0\,,\\ \left(\frac{d^2T}{dt^2}\right) + \omega^2c^2\,T &= 0\,.\end{aligned} \tag{109}$$

Solution of the x equation is

$$X(x) = C\cos\omega x + D\sin\omega x\,,$$

where C and D are arbitrary constants.

Since the boundary conditions are true for all time, we can impose them directly onto $X(x)$. At $x = 0$,

$$X(x=0) = 0 = C\,, \quad\Longrightarrow\quad C = 0\,,$$

whereas at $x = L$,

$$X(x=L) = 0 = D\,sin(\omega L) \quad\Longrightarrow\quad \omega = n\pi/L\,,$$

where $n = 1, 2, 3, \cdots$.

Solving the corresponding "t" equation,

$$\left(\frac{d^2T}{dt^2}\right) + (n\pi c/L)^2\, T = 0\ ,$$

gives

$$T = A\,\cos(n\pi ct/L) + B\,\sin(n\pi ct/L)\ ,$$

and a total solution of

$$y(x,t) = D\sin(n\pi x/L) \times \{A\,\cos(n\pi ct/L) + B\,\sin(n\pi ct/L)\}\ .$$

This is but one solution and, to get more, we use the superposition principle to find

$$y(x,t) = \sum_{n=1}^{\infty} \sin(n\pi x/L) \times \{A_n\,\cos(n\pi ct/L) + B_n\,\sin(n\pi ct/L)\}\ . \tag{110}$$

Constant D has been absorbed into constants A_n and B_n.

To go further need to impose extra boundary conditions eg shape of string at time $t = 0$. Will look at such problems under Fourier series.

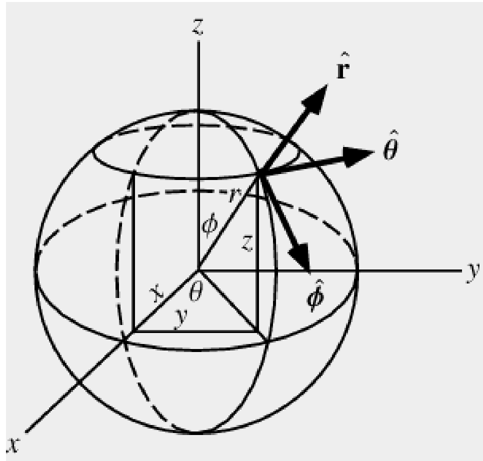
2.5 Laplace's Equation in Spherical Polar Coordinates

Switch to problems with spherical symmetry, important for Quantum Mechanics and atomic physics. If one needs to know the potential due to a charged sphere, it would be perverse to work in Cartesian coordinates. Choose a coordinate system which is appropriate to the boundary conditions to be imposed and, in this case, one should write things down in the spherical polar variables. Last year wrote ∇^2 in plane polar coordinates and it was messy. Unfortunately, in spherical polar coordinates, (r,θ,ϕ), it is even worse! Come to a simpler derivation later in the course. Now

$$\begin{aligned} x =&\ \ r\sin\theta\cos\phi\ ,\\ y =&\ \ r\sin\theta\sin\phi\ ,\\ z =&\ \ r\cos\theta\ . \end{aligned} \tag{111}$$

The partial derivatives of the Cartesian variables with respect to the polar coordinates are

$$\frac{\partial x}{\partial r} = \sin\theta\cos\phi\ ,\quad \frac{\partial y}{\partial r} = \sin\theta\sin\phi\ ,\quad \frac{\partial z}{\partial r} = \cos\theta\ ,$$



$$\frac{\partial x}{\partial \theta} = r\cos\theta\cos\phi\,, \quad \frac{\partial y}{\partial \theta} = r\cos\theta\sin\phi\,, \quad \frac{\partial z}{\partial \theta} = -r\sin\theta\,.$$
$$\frac{\partial x}{\partial \phi} = -r\sin\theta\sin\phi\,, \quad \frac{\partial y}{\partial \phi} = r\sin\theta\cos\phi\,, \quad \frac{\partial z}{\partial \phi} = 0\,. \tag{112}$$

Using the chain rule for partial differentiation, we get

$$\begin{aligned} \frac{\partial}{\partial r} &= \sin\theta\cos\phi\,\frac{\partial}{\partial x} + \sin\theta\sin\phi\,\frac{\partial}{\partial y} + \cos\theta\,\frac{\partial}{\partial z}\,, \\ \frac{\partial}{\partial \theta} &= r\cos\theta\cos\phi\,\frac{\partial}{\partial x} + r\cos\theta\sin\phi\,\frac{\partial}{\partial y} - r\sin\theta\,\frac{\partial}{\partial z}\,, \\ \frac{\partial}{\partial \phi} &= -r\sin\theta\sin\phi\,\frac{\partial}{\partial x} + r\sin\theta\cos\phi\,\frac{\partial}{\partial y}\,. \end{aligned} \tag{113}$$

These equations can be inverted to find the differentials with respect to Cartesians in terms of those with respect to polar coordinates:

$$\begin{aligned} \frac{\partial}{\partial x} &= \sin\theta\cos\phi\,\frac{\partial}{\partial r} + \frac{\cos\theta\cos\phi}{r}\,\frac{\partial}{\partial \theta} - \frac{\sin\phi}{r\sin\theta}\,\frac{\partial}{\partial \phi}\,, \\ \frac{\partial}{\partial y} &= \sin\theta\sin\phi\,\frac{\partial}{\partial r} + \frac{\cos\theta\sin\phi}{r}\,\frac{\partial}{\partial \theta} + \frac{\cos\phi}{r\sin\theta}\,\frac{\partial}{\partial \phi}\,, \\ \frac{\partial}{\partial z} &= \cos\theta\,\frac{\partial}{\partial r} - \frac{\sin\theta}{r}\,\frac{\partial}{\partial \theta}\,. \end{aligned} \tag{114}$$

The Laplacian operator is the sum of the squares of these three operators,

$$\nabla^2 = \left(\frac{\partial}{\partial x}\right)^2 + \left(\frac{\partial}{\partial y}\right)^2 + \left(\frac{\partial}{\partial z}\right)^2 = \left(\sin\theta\cos\phi\,\frac{\partial}{\partial r} + \frac{\cos\theta\cos\phi}{r}\,\frac{\partial}{\partial \theta} - \frac{\sin\phi}{r\sin\theta}\,\frac{\partial}{\partial \phi}\right)^2$$
$$+\left(\sin\theta\sin\phi\,\frac{\partial}{\partial r} + \frac{\cos\theta\sin\phi}{r}\,\frac{\partial}{\partial \theta} + \frac{\cos\phi}{r\sin\theta}\,\frac{\partial}{\partial \phi}\right)^2 + \left(\cos\theta\,\frac{\partial}{\partial r} - \frac{\sin\theta}{r}\,\frac{\partial}{\partial \theta}\right)^2\,. \tag{115}$$

Remember that the partial derivative with respect to θ acts for example on the $\sin\theta$ as well. Finally end up with

$$\nabla^2 V = \frac{\partial^2 V}{\partial r^2} + \frac{2}{r}\frac{\partial V}{\partial r} + \frac{1}{r^2}\frac{\partial^2 V}{\partial \theta^2} + \frac{1}{r^2}\cot\theta\frac{\partial V}{\partial \theta} + \frac{1}{r^2\sin^2\theta}\frac{\partial^2 V}{\partial \phi^2}\,. \tag{116}$$

This is expression for the Laplacian operator in spherical polar coordinates. Can be written in the slightly more compact form

$$\nabla^2 V = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial V}{\partial r}\right) + \frac{1}{r^2\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial V}{\partial\theta}\right) + \frac{1}{r^2\sin^2\theta}\left(\frac{\partial^2 V}{\partial\phi^2}\right)\,. \tag{117}$$

As a check on the form of the operator, consider

$$V = 2x^2 - y^2 - z^2 = r^2\left(2\sin^2\theta\,\cos^2\phi - \sin^2\theta\,\sin^2\phi - \cos^2\theta\right)\,.$$

In Cartesian coordinates, it follows immediately that $\nabla^2 V = 0$. In spherical polar coordinates,

$$\begin{aligned}
\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial V}{\partial r}\right) &= 6(2\sin^2\theta\,\cos^2\phi - \sin^2\theta\,\sin^2\phi - \cos^2\theta)\,.\\
\frac{\partial V}{\partial\theta} &= r^2(4\sin\theta\cos\theta\cos^2\phi - 2\sin\theta\cos\theta\sin^2\phi + 2\cos\theta\sin\theta)\,.\\
\sin\theta\frac{\partial V}{\partial\theta} &= r^2(4\sin^2\theta\cos\theta\cos^2\phi - 2\sin^2\theta\cos\theta\sin^2\phi + 2\cos\theta\sin^2\theta)\,.\\
\frac{1}{r^2\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial V}{\partial\theta}\right) &= 8\cos^2\theta\cos^2\phi - 4\sin^2\theta\cos^2\phi - 4\cos^2\theta\sin^2\phi\\
&\quad +2\sin^2\theta\sin^2\phi - 2\sin^2\theta + 4\cos^2\theta\,.\\
\frac{1}{r^2\sin^2\theta}\frac{\partial^2 V}{\partial\phi^2} &= 12\sin^2\phi - 6\,.
\end{aligned}$$

Remarkably enough, the sum of these three terms does in fact vanish!

2.6 Separation of Laplace's equation in Spherical Polar Coordinates

Look for a solution of the equation

$$\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial V}{\partial r}\right) + \frac{1}{r^2\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial V}{\partial\theta}\right) + \frac{1}{r^2\sin^2\theta}\left(\frac{\partial^2 V}{\partial\phi^2}\right) = 0 \tag{118}$$

in the form

$$V(r,\,\theta,\,\phi) = R(r)\times\Theta(\theta)\times\Phi(\phi)\,. \tag{119}$$

Involves functions which depend purely upon one variable each, *viz* r, θ and ϕ. Inserting this into Laplace's equation

$$\Theta\,\Phi\,\frac{1}{r^2}\frac{d}{dr}\left(r^2\,\frac{dR}{dr}\right)+R\,\Phi\,\frac{1}{r^2\sin\theta\,d\theta}\frac{d}{d\theta}\left(\sin\theta\,\frac{d\Theta}{d\theta}\right)+R\,\Theta\,\frac{1}{r^2\sin^2\theta}\left(\frac{d^2\Phi}{d\phi^2}\right)=0\,.$$

After dividing by $R\,\Theta\,\Phi$ and multiplying by $r^2\sin^2\theta$, find

$$\frac{\sin^2\theta}{R}\,\frac{d}{dr}\left(r^2\,\frac{dR}{dr}\right)+\frac{1}{\Theta}\,\sin\theta\,\frac{d}{d\theta}\left(\sin\theta\,\frac{d\Theta}{d\theta}\right)+\frac{1}{\Phi}\,\left(\frac{d^2\Phi}{d\phi^2}\right)=0\,.$$

First two terms here depend upon r and θ but third is function purely of azimuthal angle ϕ. Since r, θ and ϕ are independent variables, means that third term must be some constant, denote by $-m^2$. Hence

$$\frac{\partial^2\Phi}{\partial\phi^2}=-m^2\,\Phi\,, \tag{120}$$

which has solutions $e^{\pm im\phi}$ or, alternatively, $\cos m\phi$ and $\sin m\phi$.

As far as DE concerned, m could have any value, even complex. However Physics imposes a fairly general boundary condition. When ϕ increases by 2π, the vector position returns to the same point; expect same physical solution. Thus $\Phi(\phi+2\pi)=\Phi(\phi)$. Can only be accomplished if m is a real integer. Then $\Phi(\phi)$ is clearly a periodic function.

The remainder of the equation can be manipulated into

$$\frac{1}{R}\frac{d}{dr}\left(r^2\,\frac{dR}{dr}\right)=\frac{m^2}{\sin^2\theta}-\frac{1}{\Theta}\,\frac{1}{\sin\theta}\,\frac{d}{d\theta}\left(\sin\theta\,\frac{d\Theta}{d\theta}\right)\,.$$

Left hand side is function only of r, while right hand side depends only on θ. Means that both sides must be equal to some constant, denote by λ. Results in two ordinary DEs:

$$\frac{d}{dr}\left(r^2\,\frac{dR}{dr}\right)\;=\;\lambda\,R\,, \tag{121}$$

$$\frac{d}{d\theta}\left(\sin\theta\,\frac{d\Theta}{d\theta}\right)+\left(\lambda\,\sin\theta-\frac{m^2}{\sin\theta}\right)\,\Theta\;=\;0\,. \tag{122}$$

Now look at the radial equation of Eq. (121), rewritten as

$$r^2\,\left(\frac{d^2R}{dr^2}\right)+2r\,\left(\frac{dR}{dr}\right)-\lambda\,R=0\,. \tag{123}$$

This is a special kind of homogeneous equation which is unchanged if the r-variable is scaled as $r\to\alpha\,r$, where α is some constant. Try for a solution of the form $R(r)\sim r^\beta$, since this also stays in same form under the $r\to\alpha r$ scaling. Hence

$$\beta(\beta-1)\,r^\beta+2\beta\,r^\beta-\lambda\,r^\beta=0$$

Cancelling out the r^β factor, which cannot vanish, gives $\beta^2 + \beta = \lambda$, has solutions

$$\beta = \left(-1 \pm \sqrt{1+4\lambda}\right)/2\,.$$

Get exactly the same result by trying for the more general series solution. Standard manipulation leads to

$$\sum_{n=0}^{\infty} a_n \left\{(n+k)(n+k+1) - \lambda\right\} r^{n+k} = 0$$

The indicial equation leads to exactly the same result with β replaced by k. For higher values of n have

$$a_n \left\{(n+k)(n+k+1) - \lambda\right\} = a_n\, n(2k+1) = 0\,.$$

But $2k+1 = 2\beta+1 = \pm\sqrt{1+4\lambda}$ doesn't vanish. Hence $a_n = 0$ for $n \geq 1$ and get back to the single-term solution derived above.

To make things look a bit simpler, define eparation constant to be $\lambda \equiv \ell(\ell+1)$, where ℓ is not necessarily an integer. Then

$$\begin{aligned} \beta &= \left(-1 \pm \sqrt{1+4\ell(\ell+1)}\right)/2 \\ &= \ell \;\;\text{or}\;\; -\ell - 1\,. \end{aligned}$$

Most general form of the radial solution is

$$R(r) = A\,r^\ell + \frac{B}{r^{\ell+1}}\,. \tag{124}$$

In order not to interchange the two solutions, adopt the convention $\ell \geq -\frac{1}{2}$.

Left only with the θ equation which, with new separation constant $\ell(\ell+1)$, becomes

$$\frac{d}{d\theta}\left(\sin\theta\,\frac{d\Theta}{d\theta}\right) + \left(\ell(\ell+1)\,\sin\theta - \frac{m^2}{\sin\theta}\right)\,\Theta = 0\,, \tag{125}$$

which does not look very attractive. A little more tractable with the variable $\mu = \cos\theta$ rather than θ. Then $d\mu/d\theta = -\sin\theta$ and

$$\frac{d}{d\theta} = -\sin\theta\,\frac{d}{d\mu} = -\sqrt{1-\mu^2}\,\frac{d}{d\mu}\,.$$

Hence

$$\frac{d}{d\mu}\left[(1-\mu^2)\frac{d\Theta}{d\mu}\right] + \left[\ell(\ell+1) - \frac{m^2}{1-\mu^2}\right]\,\Theta = 0\,. \tag{126}$$

This is the famous Legendre differential equation important for quantum mechanics. Legendre discovered his equation when trying to interpret planetary gravitational fields, "Recherches sur la figure des planètes" (1784). This is about 150 years before the discovery of the Schrödinger equation and so you shouldn't blame quantum mechanics for the introduction of Legendre polynomials.

3 Series Solution of Differential Equations: Frobenius's Method

Have shown how and when differential equations (DEs) can be separated but need a general strategy for solving the resulting DEs.

Series solutions

Want to solve general, linear, homogenous, ordinary, second-order DE:

$$P(x)\frac{d^2y}{dx^2} + Q(x)\frac{dy}{dx} + R(x)y = 0\,. \tag{127}$$

One general method is expand y as a series:

$$y(x) = \sum_{n=0}^{\infty} a_n (x - x_0)^n \tag{128}$$

about some point, x_0. For this method we need to:

1. What point, x_0, to use for the expansion;
2. Determine what values of a_n satisfy the DE;
3. Decide for what, if any, values of x the series converges;

3.1 Simple Series Solution of Second Order Equations

Use classical and quantal simple harmonic oscillator (HO) as an example.

Classical HO

Particle mass m; restoring force constant K; equation

$$m\frac{d^2y}{dt^2} = -Ky \tag{129}$$

or

$$\frac{d^2y}{dt^2} + \omega^2 y = 0; \qquad \omega = \left(\frac{K}{m}\right)^{\frac{1}{2}} \tag{130}$$

Most general solution is

$$y = A\,\cos\omega t + B\,\sin\omega t\,, \tag{131}$$

where A and B are arbitrary constants fixed by boundary conditions. A second order linear equation has two arbitrary constants.

The two individual solutions of this equation, *viz* $\cos\omega t$ and $\sin\omega t$, are respectively even and odd functions of the independent variable t. Why? Write Eq. (130) for a function $y = f(t)$ and then let $t \to -t$. Have the two equations

$$\begin{aligned} \frac{d^2 f(t)}{dt^2} + \omega^2 f(t) &= 0\,, \\ \frac{d^2 f(-t)}{dt^2} + \omega^2 f(-t) &= 0\,. \end{aligned} \tag{132}$$

Thus $f(-t)$ satisfies the same equation as $f(t)$ because all operators in Eq. (130) are even; $\dfrac{d^2}{dt^2}$ doesn't change when $t \to -t$. Any linear combinations of $f(t)$ and $f(-t)$ also satisfy the equations. In particular, the even and odd combinations

$$f_e(t) = \tfrac{1}{2}[f(t) + f(-t)]\,, \tag{133}$$

$$f_o(t) = \tfrac{1}{2}[f(t) - f(-t)] \tag{134}$$

also satisfy the equation. This is the real reason why $\cos\omega t$ and $\sin\omega t$ are solutions to the oscillator equation. This argument doesn't show that the basic solutions <u>have</u> to be either even or odd, but one can always choose them so to be. Will use this argument when we discuss other DEs.

Now try for a series solution of the HO equation;

$$\begin{aligned} y &= \sum_{n=0}^{\infty} a_n\, x^n\,, \\ \frac{dy}{dx} &= \sum_{n=0}^{\infty} (n)\, a_n\, x^{n-1}\,, \\ \frac{d^2y}{dx^2} &= \sum_{n=0}^{\infty} (n)(n-1)\, a_n\, x^{n-2}\,. \end{aligned} \tag{135}$$

Inserting these into Eq. (130), we find that

$$\sum_{n=0}^{\infty} (n)(n-1)\, a_n\, x^{n-2} + \omega^2 \sum_{n=0}^{\infty} a_n\, x^n = 0\,. \tag{136}$$

Changing the dummy index in the first sum by $n \to n+2$,

$$\sum_{n=-2}^{\infty} (n+2)(n+1)\, a_{n+2}\, x^n + \omega^2 \sum_{n=0}^{\infty} a_n\, x^n = 0\,, \tag{137}$$

all the powers now look the same; compare coefficients to obtain the recurrence relation

$$(n+2)(n+1)\, a_{n+2} + \omega^2\, a_n = 0\,, \tag{138}$$

$$a_{n+2} = -\frac{\omega^2}{(n+2)(n+1)}\, a_n \,. \tag{139}$$

Given the value of a_0, this allows us to evaluate a_2, and then a_4 *etc.* The odd a_n are completely independent and, as far as getting a solution is concerned, we can put them all to zero. This independence of the odd and even a_n is a consequence of the fact that odd and even solutions of the differential equation are possible. It therefore follows from the fact that the differential operator is even in x, as shown by Eq. (132). In order to generate these purely odd/even solutions, it is easiest to put $a_1 = 0$. Does not create extra solutions, merely mixes some of the odd solution with even ones.

The recurrence relation

$$a_{n+2} = -\frac{\omega^2}{(n+2)(n+1)}\, a_n \,. \tag{140}$$

has the solution for n even

$$\begin{aligned} a_n = & \quad (-\omega^2)^{n/2}\, a_0/n! \qquad (n \text{ even})\,, \\ = & \qquad\qquad 0 \qquad\qquad (n \text{ odd})\,. \end{aligned} \tag{141}$$

The total solution is then

$$y = a_0 \sum_{n\ even} (-1)^{n/2} (\omega x)^n \,\frac{1}{n!} = a_0\, \cos \omega x \,. \tag{142}$$

To generate the odd solutions, set $a_0 = 0$ and start from a_1. To do this set $b_m = a_{n+1}$ and $m = n - 1$.

The recurrence relation is

$$b_{m+2} = -\frac{\omega^2}{(m+3)(m+2)}\, b_m \,, \tag{143}$$

so that

$$\begin{aligned} b_m = & \quad (-\omega^2)^{m/2}\, b_0/(m+1)! \qquad (m \text{ even})\,, \\ = & \qquad\qquad 0 \qquad\qquad (m \text{ odd})\,, \end{aligned} \tag{144}$$

and

$$y = b_0\, x \sum_{m\ even} (-1)^{m/2} (\omega x)^m \,\frac{1}{(m+1)!} = \frac{b_0}{\omega}\, \sin \omega x \,. \tag{145}$$

Quantum HO

Potential corresponding to force $-Kx$ is

$$V(x) = \frac{1}{2} K x^2 \tag{146}$$

Schrödinger equation:

$$-\frac{\hbar^2}{2m}\frac{d^2\psi}{dx^2} + \frac{1}{2}Kx^2\psi(x) = E\psi(x) \tag{147}$$

Change to dimenionless variables

$$y = \left(\frac{mK}{\hbar^2}\right)^{\frac{1}{4}} x = \alpha x; \qquad \epsilon = \frac{2E}{\hbar}\left(\frac{m}{K}\right)^{\frac{1}{2}} = \frac{2E}{\hbar\omega} = \frac{2E}{h\nu} \tag{148}$$

giving

$$\frac{d^2\psi}{dy^2} - y^2\psi(y) = -\epsilon\psi(y) \tag{149}$$

Complimentary solution

First solve simpler equation

$$\frac{d^2\psi}{dy^2} - y^2\psi(y) = 0 \tag{150}$$

(can think of this as equation as $|y| \to \infty$). Gives

$$\psi(y) = A\exp\left(-\frac{1}{2}y^2\right) + B\exp\left(\frac{1}{2}y^2\right) \tag{151}$$

Boundary conditions for a localised problem give $B = 0$ so that $\psi \to 0$ as $|y| \to \infty$. Assume full solution of form

$$\psi(y) = H(y)\exp\left(-\frac{1}{2}y^2\right)$$

$$\frac{d\psi}{dy} = \frac{dH}{dy}\exp\left(-\frac{1}{2}y^2\right) - y\psi$$

$$\frac{d^2\psi}{dy^2} = \frac{d^2H}{dy^2}\exp\left(-\frac{1}{2}y^2\right) - y\frac{dH}{dy}\exp\left(-\frac{1}{2}y^2\right) - \psi - y\frac{dH}{dy}\exp\left(-\frac{1}{2}y^2\right) + y^2\psi \tag{152}$$

which gives

$$\frac{d^2\psi}{dy^2} - y^2\psi(y) = -\epsilon\psi(y) = \exp\left(-\frac{1}{2}y^2\right)\left[\frac{d^2H}{dy^2} - 2y\frac{dH}{dy} - H\right] \tag{153}$$

so the equation to solve is

$$\frac{d^2H}{dy^2} - 2y\frac{dH}{dy} + (\epsilon - 1)H = 0 \tag{154}$$

This equation has no singular points. So can obtain two simple series solution about $y = 0$, these will have radius of convegence, $\rho = \infty$ (see 4.2). Also note that the equation is **even** so expect separate even and odd solutions

$$H(y) = \sum_{n=0}^{\infty} a_n y^n;$$

$$\frac{dH}{dy} = \sum_{n=0}^{\infty} n a_n y^{n-1};$$

$$\frac{d^2H}{dy^2} = \sum_{n=0}^{\infty} n(n-1) a_n y^{n-2} \tag{155}$$

so

$$\sum_{n=0}^{\infty} n(n-1)a_n y^{n-2} - 2\sum_{n=0}^{\infty} n a_n y^n + (\epsilon - 1)\sum_{n=0}^{\infty} a_n y^n = 0 \tag{156}$$

tidying this up and changing the dummy variable on the first sum by $n \to n+2$ gives

$$\sum_{n=-2}^{\infty} (n+1)(n+2)a_n y^n + \sum_{n=0}^{\infty} (\epsilon - 1 - 2n) a_n y^n = 0 \tag{157}$$

For this equation to be true for **all** values of y, the coefficient of each power of y must be **separately** equated to zero. This gives

$$2a_2 + (\epsilon - 1)a_0 = 0 \qquad \text{coef. of } y^0;$$

$$a_{j+2}(j+2)(j+1) - [\epsilon - 1 - 2j]a_j = 0 \qquad \text{coef. of } y^j. \tag{158}$$

giving a recurrence relation

$$a_{j+2} = \frac{2j - \epsilon + 1}{(j+1)(j+2)} a_j \qquad j = 0, 1, 2, \ldots \tag{159}$$

The series must **terminate** otherwise $H(y)$ and hence $\psi(x)$ go as $\exp(y^2/2)$, ie as the solution already rejected. If highest power of y in a solution is y^n, then a_{n+1} and a_{n+2} **must be zero**. This means

$$a_{n+2} = 0 = \frac{2n - \epsilon + 1}{(n+1)(n+2)} a_n \tag{160}$$

which gives

$$2n - \epsilon + 1 = 0 \tag{161}$$

or $\epsilon = 2n + 1$ as the physically allowed levels of the HO, which are

$$E = (n + \frac{1}{2})h\nu = (n + \frac{1}{2})\hbar\omega \qquad n = 0, 1, 2, \ldots \tag{162}$$

The polynomials $H(y)$ are called **Hermite Polynomials**, generally written $H_n(y)$. By convention they are written so that $a_n = 2^n$. They have recurrence relation

$$a_{j+2} = \frac{2(j-n)}{(j+1)(j+2)} a_j. \tag{163}$$

First few Hermite polynomials

$$H_0(y) = 1,$$

$$H_1(y) = 2y,$$

$$H_2(y) = 4y^2 - 2, \tag{164}$$

Normalisation constant:

$$N_n = \left(\frac{\alpha}{\pi^{\frac{1}{2}} 2^n n!}\right)^{\frac{1}{2}} \tag{165}$$

3.2 Special points

To know about the solubility of DE (133) it is necessary to analyse its structure. To do this it is convenient to re-write the equation by dividing through by $P(x)$ to give

$$\frac{d^2y}{dx^2} + p(x)\frac{dy}{dx} + q(x)y = 0\,, \tag{166}$$

where

$$p(x) = \frac{Q(x)}{P(x)}, \quad q(x) = \frac{R(x)}{P(x)}\,.$$

A DE of this form can have three types of points:

1. **Ordinary points**. x_o is an ordinary point if

$$\lim_{x\to x_0}[p(x)] \quad \text{and} \quad \lim_{x\to x_0}[q(x)] \tag{167}$$

are both finite. Most points are ordinary, eg all x in both classical and quantal HO problems. Legendre's equation (126) can be arrange into the form

$$\frac{d^2y}{dx^2} - \frac{2x}{1-x^2}\frac{dy}{dx} + \frac{\ell(\ell+1)}{1-x^2}y = 0\,. \tag{168}$$

It is ordinarly at $x_0 = 0$.

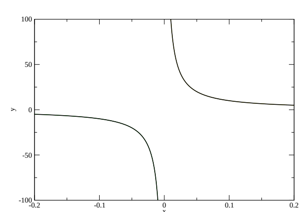


Figure 1: Plot of $y = \frac{1}{x}$ showing the singularity at $x = 0$.

Singular points best understood by considering $y = x^{-1}$ at $x = 0$, see figure 1. y is singular at $x = 0$.

Singularities in p and q determine if and how the DE can be solved. If $p(x_0)$ and/or $q(x_0)$ are not ordinary, they are singular. For example $x_0 = \pm 1$ above or $\tan x$ is singular at $(2n+1)\frac{\pi}{2}$ for integer n.

Simple series solutions of the type (128) can be used to find both solutions about an ordinary point.

2. **Regular singular points** are singularities such that

$$p_0 = \lim_{x \to x_0} [(x - x_0)p(x)] \quad \text{and} \quad q_0 = \lim_{x \to x_0} [(x - x_0)^2 q(x)] \tag{169}$$

are finite. For example $x_0 = 1$ is a regular singular point in Legendre's equation (168).

Frobenius' method works about regular singular points. Furthermore, it can be shown (Fuch's theorem) that there exists at least one series solution about any regular singular point.

3. **Essential singular points** are such that p_0 and/or q_0 are singular.

Series solution methods cannot be used for essentual singular points.

From now on will always assume that $x_0 = 0$. If this is not the case it is easier to make a change of variable using $t = x - x_0$ than to work with an expansion about $x_0 \neq 0$.

For series solutions the **radius of convergence**, which is the largest value of x for which the series can be used, is is given by the singularity nearest to x_0 in the *complex plane*. Examples will be given below.

3.3 Indicial equations

Frobenius' method is based on using a generalisation of the series solution (128) used to solve for series about regular points. Assuming one is expanding about $x_0 = 0$, this takes the form:

$$y(x) = \sum_{n=0}^{\infty} a_n x^{n+k} \tag{170}$$

with constraint that $a_0 \neq 0$. This equation has an extra x^k compared to expansion (128). k can take any value: it does not need to be integer, positive or even real. It is determined from something called the indicial equation.

To derive the indicial equation insert the expansion

$$\begin{aligned} y &= \sum_{n=0}^{\infty} a_n\, x^{n+k} \,, \\ \frac{dy}{dx} &= \sum_{n=0}^{\infty} (n+k)\, a_n\, x^{n+k-1} \,, \\ \frac{d^2y}{dx^2} &= \sum_{n=0}^{\infty} (n+k)(n+k-1)\, a_n\, x^{n+k-2} \,. \end{aligned} \tag{171}$$

into DE (166) and obtain

$$\sum_{n=0}^{\infty} (n+k)(n+k-1)\, a_n\, x^{n+k-2} + p(x) \sum_{n=0}^{\infty} (n+k)\, a_n\, x^{n+k-1} + q(x) \sum_{n=0}^{\infty} a_n\, x^{n+k} = 0 \,. \tag{172}$$

Multiply this equation through by x^{2-k}

$$\sum_{n=0}^{\infty} (n+k)(n+k-1)\, a_n\, x^{n} + p(x) \sum_{n=0}^{\infty} (n+k)\, a_n\, x^{n+1} + q(x) \sum_{n=0}^{\infty} a_n\, x^{n+2} = 0 \,, \tag{173}$$

then consider its value at the expansion point $x = 0$. At this point it is only necessary to consider terms with $n = 0$ since those with $n > 0$ are all zero at $x = 0$. This gives

$$k(k-1)\, a_0 + p(x)\, k\, a_0\, x + q(x)\, a_0\, x^2 = 0 \,. \tag{174}$$

Dividing through by a_0, since $a_0 \neq 0$, and remembering the definitions of p_0 and q_0 given by (169) one obtains

$$k(k-1) + p_0 k + q_0 = 0 \tag{175}$$

which is a quadratic equation for k and is known as the indicial equation. I find this general form easiest to use but, as shown below, the indicial equation can be derived for each case.

If k_1 and k_2 are the two solutions of the indicial equation, then there are two possibilities:

1. If $k_1 - k_2 \neq$ an integer, then both solutions can be obtained in the form:

$$\begin{aligned} y_1(x) &= x^{k_1}[a_0 + \sum_{n=1}^{\infty} a_n\, x^n] \,, \\ y_2(x) &= x^{k_2}[b_0 + \sum_{n=1}^{\infty} b_n\, x^n] \,, \end{aligned} \tag{176}$$

2. Otherwise, assuming $k_1 \leq k_2$, one solution has form of $y_2(x)$. Other may look like $y_1(x)$, but Fuch's theorem only guarantees that solution with k_2 will exist in this in form.

   One can find the second ("irregular") solution by letting $y(x) = y_2(x)\, v(x)$ and getting a simpler equation for $v(x)$. Often $v(x)$ has a nasty $\ell n(x)$ term in it. This is always the case if the indicial equation has equal roots, *i.e.* $k_1 = k_2$. This happens for Bessel's equation which one often comes across in problems with cylindrical symmetry.

Roots differing by an integer

Consider the equation

$$x(x-1)\frac{d^2y}{dx^2} + 3x\frac{dy}{dx} + y = 0\,.$$

Comparing this with the standard form gives $p(x) = \dfrac{3}{(x-1)}$ and $q(x) = \dfrac{1}{x(x-1)}$. Thus $x = 0$ and $x = 1$ are regular points of the differential equation and so we can expect to get at least one power series solution in x. At $x = 0$, get $k = 0$ and 1.

Inserting

$$\begin{aligned} y &= \sum_{n=0}^{\infty} a_n\, x^{n+k}\,, \\ \frac{dy}{dx} &= \sum_{n=0}^{\infty} (n+k)\, a_n\, x^{n+k-1}\,, \\ \frac{d^2y}{dx^2} &= \sum_{n=0}^{\infty} (n+k)(n+k-1)\, a_n\, x^{n+k-2} \end{aligned}$$

into the differential equation,

$$\sum_{n=0}^{\infty} (n+k)(n+k-1)\, a_n\, (x^{n+k} - x^{n+k-1}) + \sum_{n=0}^{\infty} 3(n+k)\, a_n\, x^{n+k} + \sum_{n=0}^{\infty} a_n\, x^{n+k} = 0\,.$$

Hence

$$\sum_{n=0}^{\infty} a_n\, x^{n+k}\, [(n+k)(n+k-1) + 3(n+k) + 1] = \sum_{n=-1}^{\infty} (n+k+1)(n+k)\, a_{n+1}\, x^{n+k}\,.$$

Indicial equation comes from looking at the lowest power of x, given by $n = -1$ on the right hand side. Gives $k(k-1) = 0$, *i.e.* $k = 1$ or $k = 0$ as above.

Equating higher powers of x gives <u>recurrence relation</u>:

$$(n+k+1)^2\, a_n = (n+k+1)(n+k)\, a_{n+1}\,,$$

$$a_{n+1} = \left(\frac{n+k+1}{n+k}\right) a_n \, .$$

Recurrence relations allow us to evaluate all the higher coefficients from the first one. To fix the first term one has to remember that $a_0 \neq 0$. This term then acts as a scale constant for the whole solution, which must be determined from the boundary conditions.

Taking the index $k = 1$ and putting $a_0 = 1$, we get $a_1 = 2$, $a_2 = 3$ *etc.* The full solution is

$$y_1(x) = x(1 + 2x + 3x^2 + 4x^3 + \cdots) = \frac{x}{(1-x)^2} \, .$$

Note that this series converges for $\mid x \mid < 1$; the divergence at $x \geq 1$ is due to the singular point there.

On the other hand, when the index $k = 0$, we are in trouble because the recurrence relation is

$$a_{n+1} = \left(\frac{n+1}{n}\right) a_n \, .$$

If you try to calculate a_1 by putting $n = 0$ you see that the whole thing blows up. Hence there is not a second *series* solution at $x = 0$. Fuch's theorem only guaranteed that there would be one solution of this kind; the other solution is going to be nasty at $x = 0$.

3.4 Applying Frobenius's method

Consider the equation

$$2x^2 \frac{d^2y}{dx^2} - x\frac{dy}{dx} + (1+x)y = 0 \, . \tag{177}$$

For this $p(x) = \dfrac{-1}{2x}$ and $q(x) = \dfrac{1+x}{2x^2}$. This eq. has a regular singular point at $x = 0$ with $p_0 = \dfrac{-1}{2}$ and $q_0 = \dfrac{1}{2}$, giving an indicial eq.

$$k(k-1) - \frac{1}{2}k + \frac{1}{2} = 0 \, . \tag{178}$$

$$2k^2 - 3k + 1 = (2k-1)(k-1) = 0 \, . \tag{179}$$

Hence $k = \frac{1}{2}$ and 1.

Expanding as a series

$$\begin{aligned} y &= \sum_{n=0}^{\infty} a_n \, x^{n+k} \, , \\ \frac{dy}{dx} &= \sum_{n=0}^{\infty} (n+k) \, a_n \, x^{n+k-1} \, , \\ \frac{d^2y}{dx^2} &= \sum_{n=0}^{\infty} (n+k)(n+k-1) \, a_n \, x^{n+k-2} \, . \end{aligned} \tag{180}$$

and substituting into DE (177) gives

$$\sum_{n=0}^{\infty} 2(n+k)(n+k-1)\, a_n\, x^{n+k} - \sum_{n=0}^{\infty} (n+k)\, a_n\, x^{n+k} + \sum_{n=0}^{\infty} a_n\, x^{n+k} + \sum_{n=0}^{\infty} a_n\, x^{n+k+1} = 0. \quad (181)$$

For (181) to be satisfied for all values of x then the coefficient of each power of x must be zero. For x^k this gives

$$2k(k-1) - k + 1 = 2k^2 - 3k + 1 = 0\,, \quad (182)$$

assuming $a_0 \neq 0$. This is the indical eq. derived above. Equating powers of x^{n+k} get

$$[2(k+n)(k+n-1) - (k+n) + 1]a_n + a_{n-1} = 0$$

$$a_n = \frac{-1}{2(k+n)^2 - 3(k+n) + 1} a_{n-1} = \frac{-1}{[2(k+n)-1][(k+n)-1]} a_{n-1}\,, \qquad n \geq 1\,. \quad (183)$$

Consider $k = 1$ and $k = \frac{1}{2}$ in turn.

$k = 1$

$$a_n = \frac{-1}{[2n+1]n} a_{n-1}\,, \qquad n \geq 1\,. \quad (184)$$

So

$$a_1 = -\frac{a_0}{3.1}\,,$$
$$a_2 = -\frac{a_1}{5.2} = \frac{a_0}{3.5(1.2)}\,,$$
$$a_3 = -\frac{a_2}{7.3} = \frac{a_0}{(3.5.7)(1.2.3)}\,,$$

etc. In general

$$a_n = \frac{(-1)^n}{[3.5.7\ldots(2n+1)]n!} a_0\,, \qquad n \geq 1\,. \quad (185)$$

Gives first general solution of (177), omiting constant multiplier a_0,

$$y_1(x) = x\left[1 + \sum_{n=1}^{\infty} \frac{(-1)^n x^n}{[3.5.7\ldots(2n+1)]n!}\right]. \quad (186)$$

Use the ratio test to determine the radius of convergence

$$\lim_{n\to\infty}\left|\frac{a_{n+1}x^{n+1}}{a_n x^n}\right| = \lim_{n\to\infty}\frac{|x|}{(2n+3)(n+1)} = 0\,, \quad (187)$$

so series converges for all x. Can be seen as no poles in $p(x)$ and $q(x)$ except at expansion point $x = 0$.

$k = \frac{1}{2}$

Same method gives independent solution

$$y_2(x) = x^{\frac{1}{2}} \left[1 + \sum_{n=1}^{\infty} \frac{(-1)^n x^n}{[3.5.7\ldots(2n-1)]n!} \right] . \tag{188}$$

which is slso convergent for all x.

Now apply this method to commonly occuring physical problems, particularly those that arise from trying to solve Schrödinger's equation.

Laguerre's equation

In PHAS2222 the radial eq. of the H atom is written:

$$\frac{d^2F}{dr^2} - 2\kappa\frac{dF}{dr} + \left[\frac{2Z}{r} - \frac{\ell(\ell+1)}{r^2}\right] F = 0$$

where Z is the charge on the atom, $\kappa = \sqrt{-2E}$, ℓ the orbital angular momentum of the electron is an integer ≥ 0, κ its energy, and the range is $0 \leq r \leq \infty$. This equation is also known as Laguerres equation.

Analysing this eq. for $r \to 0$ shows:

$$p(r) = -2\kappa, \qquad p_0 = 0;$$

$$q(r) = \left[\frac{2Z}{r} - \frac{\ell(\ell+1)}{r^2}\right], \qquad q_0 = -\ell(\ell+1). \tag{189}$$

There are no other singular points so expanding about this point will give solutions with radius of convergence, $\rho = \infty$.

Indicial equation:

$$k(k-1) + kp_0 + q_0 = 0$$

$$k(k-1) - \ell(\ell+1) = 0$$

$$k^2 - k - \ell(\ell+1) = 0$$

$$(k+\ell)(k-(\ell+1)) = 0$$

$$k = -\ell, \ell+1$$

Solution with $k = -\ell$

$$F(r) = \sum_{n=0} a_n r^{n-\ell}, \qquad a_0 \neq 0$$

is unbounded (ie ∞) at $r = 0$, so unphysical.

Solution $k = \ell + 1$ gives

$$F(r) = \sum_{n=0} a_n r^{n+\ell+1}, \qquad a_0 \neq 0$$

rewrite as

$$F(r) = \sum_{j=\ell+1} a_j r^j, \qquad a_{\ell+1} \neq 0$$

$$\frac{dF}{dr} = \sum_j j a_j r^{j-1}$$

$$\frac{d^2F}{dr^2} = \sum_j j(j-1) a_j r^{j-2}$$

Substituting in

$$\sum_{j=\ell+1} j(j-1)a_j r^{j-2} - 2\kappa \sum_{j=\ell+1} j a_j r^{j-1} + 2Z \sum_{j=\ell+1} a_j r^{j-1} - \ell(\ell+1) \sum_{j=\ell+1} a_j r^{j-2} = 0$$

$$\sum_{j=\ell+1} [j(j-1) - \ell(\ell+1)]a_j r^{j-2} = \sum_{j=\ell+1} [2\kappa j - 2Z]a_j r^{j-1}$$

$$\sum_{j=\ell} [j(j+1) - \ell(\ell+1)]a_{j+1} r^{j-1} = \sum_{j=\ell+1} [2\kappa j - 2Z]a_j r^{j-1}.$$

Equating powers of r^{j-1} gives

$$[j(j+1) - \ell(\ell+1)]a_{j+1} = [2\kappa j - 2z]a_j$$

and the recurrence relation

$$a_{j+1} = \frac{(2\kappa j - 2Z)}{j(j+1) - \ell(\ell+1)} a_j$$

with $j > \ell$.

As $j \to \infty$

$$\frac{a_{j+1}}{a_j} \to \frac{2\kappa}{j+1}$$

which means that for large r, $F(r)$ behaves as $\exp(2\kappa r)$, remember

$$\exp(2\kappa r) = 1 + 2\kappa r + \frac{(2\kappa r)^2}{2!} + \ldots \frac{(2\kappa r)^n}{n!} \ldots$$

$\exp(2\kappa r)$ is not bounded so series must terminate.

Let n be highest term allowed, then $a_{n+1} = 0$

$$(2\kappa n - 2Z) = 0$$

$$\kappa = \frac{Z}{n} = (-2E)^{\frac{1}{2}}$$

which leads directly to the energy levels of the H atom

$$E = -\frac{Z^2}{2n^2}, \qquad\qquad n = 1, 2, 3, \ldots$$

with the condition $n > \ell$.

These solutions are known as Laguerre polynomials (terms can be written down using recurrence relation). These are orthogonal polynomials (like Hermite and Legendre). In fact there are a set for each value of ℓ.

4 Legendre Functions

Solve Legendre's differential equation

$$\frac{d}{d\mu}\left[(1-\mu^2)\frac{d\Theta}{d\mu}\right] + \left[\ell(\ell+1) - \frac{m^2}{1-\mu^2}\right]\Theta = 0\,. \tag{190}$$

For $m = 0$ there is no azimuthal dependence on the angle ϕ:

$$\frac{d}{d\mu}\left[(1-\mu^2)\frac{d\Theta}{d\mu}\right] + \ell(\ell+1)\,\Theta = 0\,. \tag{191}$$

Special case of $\ell = 0$

Start with the even simpler case that we can treat by A-level methods. For $\ell = 0$,

$$\frac{d}{d\mu}\left[(1-\mu^2)\frac{d\Theta}{d\mu}\right] = 0\,.$$

This means that the quantity inside the square bracket must be some constant C;

$$(1-\mu^2)\frac{d\Theta}{d\mu} = C\,.$$

This equation separates as

$$\int d\Theta = \int \frac{C}{(1-\mu^2)}\,d\mu\,,$$

giving the solution

$$\Theta = C\,\tfrac{1}{2}\ell n\left(\frac{1+\mu}{1-\mu}\right) + D\,. \tag{192}$$

Legendre equation is ordinary second-order DE. so solution contains two arbitrary integration constants, written here as C and D. There are two independent solutions of the equation

$$P_0(\mu) = 1\,, \tag{193}$$

$$Q_0(\mu) = \tfrac{1}{2}\ell n\left(\frac{1+\mu}{1-\mu}\right)\,, \tag{194}$$

where subscript denotes the value of ℓ.

Since Legendre equation is homogeneous, most general solution is a linear superposition of P_0 and Q_0,

$$\Theta(\mu) = C\,Q_0(\mu) + D\,P_0(\mu)\,.$$

Note that $Q_0(\mu)$ diverges at $\theta = 0$, *i.e.* $\mu = \cos\theta = +1$.

Away from $\ell = m = 0$, the solutions are more complicated. In general, one solutions is finite at $\mu = \pm 1$, whereas the other one blows up there. To find such solutions, we must apply series methods.

4.1 Series solution

First note the eq. remains unchanged if we let $\mu \to -\mu$. As before, this means can write independent solutions as either odd or even functions of μ. This condition is satisfied by solutions obtained for $\ell = m = 0$, (194).

Carrying out a differentiation, Legendre's eq. becomes

$$\frac{d^2\Theta}{d\mu^2} - \frac{2\mu}{(1-\mu^2)}\frac{d\Theta}{d\mu} + \frac{\ell(\ell+1)}{(1-\mu^2)}\Theta = 0\,, \tag{195}$$

which is ordinary at $\mu = 0$ but has regular singularities at $\mu = \pm 1$, expect series solutions about $\mu = 0$ will converge for $|\mu| < 1$.

Look for solutions in the series form

$$\begin{aligned}
\Theta &= \sum_{n=0}^{\infty} a_n\, \mu^n\,, \\
\Theta' &= \sum_{n=0}^{\infty} n\, a_n\, \mu^{n-1}\,, \\
\Theta'' &= \sum_{n=0}^{\infty} n(n-1)\, a_n\, \mu^{n-2}\,.
\end{aligned} \tag{196}$$

Inserting the expressions for Θ, Θ' and Θ'' into Eq. (195), we find

$$\sum_{n=0}^{\infty} n(n-1)\, a_n \left[\mu^{n-2} - \mu^n\right] - 2\sum_{n=0}^{\infty} n\, a_n\, \mu^n + \ell(\ell+1)\sum_{n=0}^{\infty} a_n\, \mu^n = 0\,. \tag{197}$$

Grouping together all similar powers of μ simplifies things a bit:

$$\sum_{n=0}^{\infty} n(n-1)\, a_n\, \mu^{n-2} = \sum_{n=0}^{\infty} \{n(n+1) - \ell(\ell+1)\}\; a_n\, \mu^n\,. \tag{198}$$

To get recurrence relation, change dummy variable $n \to n+2$ on left of Eq. (198):

$$\sum_{n=-2}^{\infty} (n+1)(n+2)\, a_{n+2}\, \mu^n = \sum_{n=0}^{\infty} \{n(n+1) - \ell(\ell+1)\}\; a_n\, \mu^n\,. \tag{199}$$

Comparing coefficients of powers of μ then gives

$$a_{n+2} = \frac{(n+1)n - \ell(\ell+1)}{(n+1)(n+2)}\, a_n\,. \tag{200}$$

Recurrence relation links terms differing by two units in n. As for harmonic oscillator equation, is a direct consequence of the DE being even under $\mu \to -\mu$, means that there are odd and even solutions of Legendre's equation.

Even solutions

$$a_{n+2} = \frac{n(n+1) - \ell(\ell+1)}{(n+1)(n+2)}\, a_n = \frac{(n-\ell)(n+\ell+1)}{(n+1)(n+2)}\, a_n \,. \tag{201}$$

Solution is therefore

$$p_\ell(\mu) = a_0 \left[1 - \frac{\ell(\ell+1)}{2!}\, \mu^2 + \frac{(\ell-2)(\ell)(\ell+1)(\ell+3)}{4!}\, \mu^4 + \cdots \right] . \tag{202}$$

Odd solutions

Starts from a_1 giving

$$q_\ell(\mu) = a_1 \left[\mu - \frac{(\ell-1)(\ell+2)}{3!}\, \mu^3 + \frac{(\ell-3)(\ell-1)(\ell+2)(\ell+4)}{5!}\, \mu^5 + \cdots \right] . \tag{203}$$

Clear that $p_\ell(\mu)$ is even function of μ, as a_1 etc $= 0$, whereas $q_\ell(\mu)$ is odd function. Most general solution is

$$f_\ell(\mu) = A\, p_\ell(\mu) + B\, q_\ell(\mu) \,. \tag{204}$$

Earlier found the explicit forms of the solutions for $\ell = 0$, *viz*

$$P_0(\mu) = 1 \qquad \text{and} \qquad Q_0 = \tfrac{1}{2} \ell n \left(\frac{1+\mu}{1-\mu} \right) .$$

Since $P_0(\mu)$ is even and $Q_0(\mu)$ is odd, want to identify P_0 with p_0 and Q_0 with q_0.

Putting $\ell = 0$ in Eq. (202),

$$p_0(\mu) = a_0 \left[1 - \frac{0(1)}{2!}\, \mu^2 + \frac{(-2)(0)(1)(3)}{4!}\, \mu^4 + \cdots \right] = a_0\, P_0(\mu) \,. \tag{205}$$

Every term (except first) has an ℓ factor which kills it.

Odd solution of Eq. (203) is a bit more complicated:

$$q_0(\mu) = a_1 \left[\mu - \frac{(-1)(2)}{3!}\, \mu^3 + \frac{(-3)(-1)(2)(4)}{5!}\, \mu^5 + \cdots \right] = a_1 \left[\mu + \frac{1}{3}\, \mu^3 + \frac{1}{5}\, \mu^5 + \cdots \right] , \tag{206}$$

which is the series expansion of $\frac{1}{2} \ell n \left(\frac{1+\mu}{1-\mu} \right)$.

Shows all is OK for $\ell = 0$. Get an infinite series for $q_0(\mu)$ but one which <u>terminates</u> for $p_0(\mu)$.

Does the infinite series converge? Apply D'Alembert ratio test to investigate.

4.2 Range of Convergence

Series goes up by steps of two in n. For D'Alembert ratio test have then to compare $(n+2)$'nd term with n'th

$$R = \left| \frac{a_{n+2}\, \mu^{n+2}}{a_n\, \mu^n} \right| = \left| \frac{(n+1)n - \ell(\ell+1)}{(n+1)(n+2)}\, \mu^2 \right| . \tag{207}$$

Convergence depends upon large n, where

$$R \longrightarrow \left[1 - \frac{2}{n}\right] \mu^2 \longrightarrow \mu^2 \, . \tag{208}$$

$R < 1$ guarantees convergence and this is true if $|\mu| < 1$, exactly as expected because of the regular singular points of Legendre equation at $\mu = \pm 1$.

To see that $Q_0(\mu) = \frac{1}{2}\ell n\left(\dfrac{1+\mu}{1-\mu}\right)$ does indeed blow up at $\mu = \pm 1$; just put it into your calculator and see the error message flashing! Important point to note that $\mu = \cos\theta = \pm 1$ corresponds to $\theta = 0°$ and $180°$, need finite answers at these two points. Why should the electrostatic potential be infinite at the top and bottom of a sphere?

Avoided this problem with $P_0(\mu) = 1$ because series *terminates* with finite number of terms (for $\ell = 0$ just a single one). No issue with convergence. Only way out; to get finite answers at $\mu = \pm 1$ series must terminate. End up with a polynomial rather than an infinite series.

Going back to recurrence relation of Eq. (200),

$$a_{n+2} = \frac{(n+1)n - \ell(\ell+1)}{(n+1)(n+2)}\, a_n \, ,$$

series terminates if numerator on right hand side vanishes for some value of n;

$$(n+1)n - \ell(\ell+1) = (n-\ell)(n+1+\ell) = 0 \, . \tag{209}$$

The convention that $Re\{\ell\} \geq -\frac{1}{2}$ means that we need

$$\ell = n \, . \tag{210}$$

For even solution need ℓ to be any positive <u>even</u> integer,
For odd solution need ℓ to be any positive <u>odd</u> integer.

2B22 Quantum Mechanics course shows that condition ℓ be an integer corresponds to the quantisation of orbital angular momentum in integral units of $\hbar$. Result obtained here is fundamental to this, Atomic and other branches of Physics.

For any (non-negative) integer N, $p_{2N}(\mu)$ and $q_{2N+1}(\mu)$ are polynomials in μ, but that $p_{2N+1}(\mu)$ and $q_{2N}(\mu)$ are infinite series which diverge at $\mu = 1$. Clearly interested in solutions which are finite at $\theta = 0°$; group them with a common notation. Let

$$P_\ell(\mu) = \begin{cases} p_\ell(\mu) & \ell \text{ even}\, , \\ q_\ell(\mu) & \ell \text{ odd}\, , \end{cases}$$
$$Q_\ell(\mu) = \begin{cases} p_\ell(\mu) & \ell \text{ odd}\, , \\ q_\ell(\mu) & \ell \text{ even}\, . \end{cases} \tag{211}$$

$P_\ell(\mu)$ is a polynomial but $Q_\ell(\mu)$ is an infinite series which blows up at $\mu = 1$. Furthermore

$$
\begin{aligned}
P_\ell(-\mu) &= (-1)^\ell\, P_\ell(\mu)\,, \\
Q_\ell(-\mu) &= (-1)^{\ell+1}\, Q_\ell(\mu)\,.
\end{aligned}
\tag{212}
$$

Summary

Only for non-negative integers ℓ do we have solutions of Legendre's equation which are finite at $\mu = \pm 1$. These are the Legendre polynomials $P_\ell(\mu)$. There are also Legendre functions of the second kind, $Q_\ell(\mu)$, but these blow up at $\mu = \pm 1$. The Q_ℓ are far less important in Physics and will be largely neglected.

Standard to "normalise" Legendre polynomials such that

$$
P_\ell(1) = 1\,. \tag{213}
$$

From the series representation of Eqs. (202) and (203), we then see that

$$
\begin{aligned}
P_0(\mu) &= 1\,, \\
P_1(\mu) &= \mu\,, \\
P_2(\mu) &= \tfrac{1}{2}(3\mu^2 - 1)\,.
\end{aligned}
\tag{214}
$$

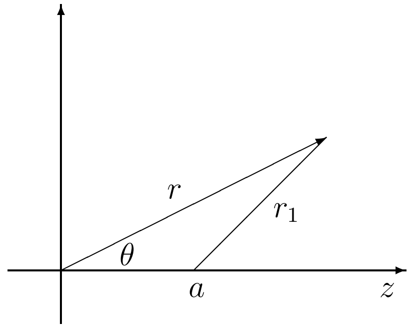
Going back to Eq. (118), original Laplace equation in spherical coordinates, most general solution which has <u>no</u> ϕ dependence is

$$
V(r\,,\,\theta) = \sum_{\ell=0}^{\infty} \left[\alpha_\ell\, r^\ell + \beta_\ell\, r^{-\ell-1}\right]\; P_\ell(\cos\theta)\,, \tag{215}
$$

where sum is over discrete integers ℓ, and α_ℓ and β_ℓ are constants fixed by boundary conditions.

4.3 Generating Function for Legendre Polynomials

Working out electrostatic potential due to point charge q at $z = a$.



In terms of the distance from the point $z = a$, potential is simply

$$V = \frac{q}{4\pi\varepsilon_0}\,\frac{1}{r_1}\,. \tag{216}$$

To evaluate it in terms of r and θ, use the cosine rule to obtain

$$r_1^2 = r^2 + a^2 - 2r\,a\cos\theta\,, \tag{217}$$

which leads to a potential of

$$V(r\,,\,\theta) = \frac{q}{4\pi\varepsilon_0}\,[r^2 + a^2 - 2r\,a\cos\theta]^{-\frac{1}{2}}\,. \tag{218}$$

There is no ϕ dependence because the charge was placed on the z-axis.

If we are interested in the potential in the region $r > a$, then we can expand Eq. (218) in powers of a/r,

$$\begin{aligned} V(r\,,\,\theta) &= \frac{q}{4\pi\varepsilon_0\,r}\left[1 + \left(\frac{a}{r}\right)^2 - 2\left(\frac{a}{r}\right)\cos\theta\right]^{-\frac{1}{2}}\,. \\ &\approx \frac{q}{4\pi\varepsilon_0\,r}\left[1 - \frac{a^2}{2r^2} + \frac{a}{r}\cos\theta + \frac{3a^2}{2r^2}\cos^2\theta + \cdots\right] \\ &= \frac{q}{4\pi\varepsilon_0\,r}\left[1 + \frac{a}{r}\cos\theta + \frac{a^2}{r^2}\,\tfrac{1}{2}\,(3\cos^2\theta - 1) + \cdots\right] \end{aligned} \tag{219}$$

Already know the general solution for Laplace's equation in any region where there is no charge. If the potential is to remain finite at large r, all the α_ℓ coefficients in Eq. (215) must vanish and so

$$V(r\,,\,\theta) = \frac{1}{r}\sum_{\ell=0}^{\infty}\frac{\beta_\ell}{r^\ell}\,P_\ell(\cos\theta)\,. \tag{220}$$

To determine the values of the β_ℓ, look what happens at $\theta = 0$ where, by definition, $P_\ell(\cos\theta) = 1$. In the forward direction $r = r_1 + a$, and so

$$V(r\,,\,\theta) = \frac{1}{r}\sum_{\ell=0}^{\infty}\frac{\beta_\ell}{r^\ell} = \frac{q}{4\pi\varepsilon_0\, r}\,\frac{1}{(1 - a/r)} = \frac{q}{4\pi\varepsilon_0\, r}\sum_{\ell=0}^{\infty}\frac{a^\ell}{r^\ell}\,. \tag{221}$$

Comparing different powers of r in the two sums, can read off

$$\beta_\ell = \frac{q}{4\pi\varepsilon_0}\,a^\ell\,. \tag{222}$$

Final solution at all angles

$$V(r\,,\,\theta) = \frac{q}{4\pi\varepsilon_0}\,\frac{1}{\sqrt{r^2 + a^2 - 2r\,a\cos\theta}} = \frac{q}{4\pi\varepsilon_0\, r}\sum_{\ell=0}^{\infty}\frac{a^\ell}{r^\ell}\,P_\ell(\cos\theta)\,. \tag{223}$$

Have solved a problem in electrostatics but result gives general method to derive Legendre polynomials. Comparing the two expressions for the potential in Eq. (223), and dropping elecrostatics factor outside gives

$$\frac{1}{r}\,\frac{1}{\sqrt{1 + a^2/r^2 - 2\,(a/r)\,\cos\theta}} = \frac{1}{r}\sum_{\ell=0}^{\infty}\frac{a^\ell}{r^\ell}\,P_\ell(\cos\theta)\,. \tag{224}$$

Change to notation where $t = a/r$ and $x = \cos\theta$,

$$g(x,t) \equiv \frac{1}{\sqrt{1 - 2x\,t + t^2}} = \sum_{\ell=0}^{\infty} P_\ell(x)\,t^\ell\,. \tag{225}$$

This is the generating function for the Legendre polynomials. Only valid if $\mid t \mid < 1$, which corresponds to $r > a$, otherwise there are convergence problems.

If you expand the square root using the binomial expansion, and compare powers of t, then you get the same answers as we got before, *viz*

$$\begin{aligned} P_0(x) &= 1\,, \\ P_1(x) &= x\,, \\ P_2(x) &= \tfrac{1}{2}(3x^2 - 1)\,. \end{aligned} \tag{226}$$

etc.

4.4 Recurrence Relations

Apart from physical interpretation, generating function helps derive recurrence relations between Legendre polynomials. In practice most efficient way of deriving polynomials.

Differentiate the generating function of Eq. (225) partially with respect to t

$$\frac{\partial g(x,t)}{\partial t} = \frac{x-t}{(1-2x\,t+t^2)^{3/2}} = \sum_{n=0}^{\infty} n\,P_n(x)\,t^{n-1}\,. \tag{227}$$

Multiply both sides by $1-2x\,t+t^2$ to give

$$(x-t)\,\frac{1}{(1-2x\,t+t^2)^{\frac{1}{2}}} = (1-2x\,t+t^2)\sum_{n=0}^{\infty} n\,P_n(x)\,t^{n-1}\,. \tag{228}$$

On the left-hand side see once generating function for Legendre polynomials, which means

$$(x-t)\sum_{n=0}^{\infty} P_n(x)\,t^n = (1-2x\,t+t^2)\sum_{n=0}^{\infty} n\,P_n(x)\,t^{n-1}\,. \tag{229}$$

This equation is a power series in t which is supposed to be valid for a range of values of t. Consequently it must be valid separately for each power of t. Exactly same argument used in the series solution of DEs. Writing out explicitly all different powers,

$$x\sum_{n=0}^{\infty} P_n(x)\,t^n - \sum_{m=0}^{\infty} P_m(x)\,t^{m+1} = \sum_{\ell=0}^{\infty} \ell\,P_\ell(x)\,t^{\ell-1} - 2x\sum_{n=0}^{\infty} n\,P_n(x)\,t^n + \sum_{m=0}^{\infty} m\,P_m(x)\,t^{m+1}. \tag{230}$$

Formula written with different dummy indices ℓ, m, and n so can change a couple of them easily. Let

$$\begin{aligned} m &\longrightarrow n-1\,,\\ \ell &\longrightarrow n+1\,. \end{aligned}$$

Then all terms in the sums contain same t^n factor. Reading off coefficient get

$$x\,P_n(x) - P_{n-1}(x) = (n+1)\,P_{n+1}(x) - 2n\,x\,P_n(x) + (n-1)\,P_{n-1}(x)\,.$$

Grouping like terms together gives the recurrence relation

$$(2n+1)\,x\,P_n(x) = (n+1)\,P_{n+1}(x) + n\,P_{n-1}(x)\,. \tag{231}$$

Thus, if you know $P_n(x)$ and $P_{n-1}(x)$, the recurrence relation allows you to obtain the formula for $P_{n+1}(x)$.

For example, putting $n=1$ in Eq. (231), we get

$$3x\,P_1(x) = 2P_2(x) + P_0(x)\,. \tag{232}$$

Since $P_0(x)=1$ and $P_1(x)=x$, this then immediately gives $P_2(x)=\frac{1}{2}\,(3x^2-1)$.

Using instead $n = 2$, we obtain

$$5x\, P_2(x) = 3P_3(x) + 2P_1(x)\,, \tag{233}$$

which means that

$$P_3(x) = \tfrac{1}{2}\,(5x^3 - 3x)\,.$$

4.5 Orthogonality of Legendre Polynomials

The Legendre differential equations for $P_n(x)$ and $P_m(x)$ are

$$\begin{aligned} \frac{d}{dx}\left[(1-x^2)\,P_n'(x)\right] + n(n+1)\,P_n(x) &= 0\,, \\ \frac{d}{dx}\left[(1-x^2)\,P_m'(x)\right] + m(m+1)\,P_m(x) &= 0\,. \end{aligned} \tag{234}$$

Multiply the first of Eqs. (234) by $P_m(x)$ and the second by $P_n(x)$ and subtract one from the other to find:

$$P_m(x)\,\frac{d}{dx}\left[(1-x^2)\,P_n'(x)\right] - P_n(x)\,\frac{d}{dx}\left[(1-x^2)\,P_m'(x)\right] = \left[m(m+1) - n(n+1)\right]\,P_m(x)P_n(x).$$

Now integrate both sides of this expression over x from -1 to $+1$:

$$\begin{aligned} &\int_{-1}^{+1}\left\{P_m(x)\,\frac{d}{dx}\left[(1-x^2)\,P_n'(x)\right] - P_n(x)\,\frac{d}{dx}\left[(1-x^2)\,P_m'(x)\right]\right\}\,dx \\ &\qquad = \left[m(m+1) - n(n+1)\right]\int_{-1}^{+1} P_m(x)P_n(x)\,dx\,. \end{aligned} \tag{235}$$

What we have to do now is show that the left hand side of Eq. (235) vanishes. This we do through integrating by parts.

$$\int_{-1}^{+1} P_m(x)\,\frac{d}{dx}\left[(1-x^2)\,P_n'(x)\right]\,dx = \left[P_m(x)\,(1-x^2)\,P_n'(x)\right]_{-1}^{+1} - \int_{-1}^{+1}(1-x^2)\,P_n'(x)P_m'(x)\,dx. \tag{236}$$

Now the first term on the RHS of Eq. (236) equals zero because $(1-x^2) = 0$ at both the limits. On the other hand, the second term is cancelled by an identical one coming from the second term in Eq.(235) where $m \Leftrightarrow n$. Hence

$$\left[m(m+1) - n(n+1)\right]\int_{-1}^{+1} P_m(x)P_n(x)\,dx = 0\,, \tag{237}$$

which means that, if $n \neq m$,

$$\int_{-1}^{+1} P_m(x)P_n(x)\,dx = 0\,. \tag{238}$$

This the *Orthogonality relation*. It is analogous to orthogonality of two vectors except that have integral over a continuous variable rather than sum over components.

To construct equivalent of a unit vector, evaluate integral when $m = n$:

$$I_n = \int_{-1}^{+1} \left[P_n(x)\right]^2 \, dx \, . \tag{239}$$

Many ways of working this out, but easiest uses generating function of Eq. (225). Writing this twice gives

$$\frac{1}{\sqrt{1-2x\,t+t^2}} = \sum_{n=0}^{\infty} P_n(x)\, t^n \, ,$$

$$\frac{1}{\sqrt{1-2x\,t+t^2}} = \sum_{n=0}^{\infty} P_m(x)\, t^m \, . \tag{240}$$

Multiply these two expressions together to give a double summation over n and m.

$$\frac{1}{1-2x\,t+t^2} = \sum_{n=0}^{\infty}\sum_{m=0}^{\infty} P_n(x)\, P_m(x)\, t^{n+m} \, . \tag{241}$$

Now integrate over x from -1 to $+1$:

$$\int_{-1}^{+1} \frac{dx}{1-2x\,t+t^2} = \sum_{n=0}^{\infty}\sum_{m=0}^{\infty} \int_{-1}^{+1} P_n(x)\, P_m(x)\, t^{n+m} dx \, . \tag{242}$$

The integral on the left gives

$$\frac{1}{2t}\,\ell n\left(\frac{(1+t)^2}{(1-t)^2}\right) = \frac{1}{t}\,\ell n\left(\frac{1+t}{1-t}\right) = 2\sum_{n=0}^{\infty} \frac{t^{2n}}{2n+1} \, . \tag{243}$$

Have already shown that integral on RHS vanishes unless $n = m$ and so only have a single sum:

$$\mathrm{RHS} = \sum_{n=0}^{\infty} I_n\, t^{2n} \, . \tag{244}$$

Comparing coefficients of t^{2n} on left and right hand sides gives

$$I_n = \int_{-1}^{+1} \left[P_n(x)\right]^2 \, dx = \frac{2}{2n+1} \, . \tag{245}$$

The orthogonality and normalisation of Legendre polynomials can be written as

$$\int_{-1}^{+1} P_m(x)\, P_n(x)\, dx = \frac{2}{2m+1}\, \delta_{m\,n} \, , \tag{246}$$

where the Kronecker delta symbol for two integers m and n is defined by

$$\delta_{m\,n} = \begin{cases} 1 & \text{for } m = n \\ 0 & \text{for } m \neq n \, . \end{cases} \tag{247}$$

4.6 Expansion in series of Legendre polynomials

Last year learned how to expand a function $f(x)$ in a Maclaurin series:

$$f(x) = \sum_{n=0}^{\infty} \alpha_n \, x^n$$

for $-1 < x < 1$. Also, in the first year Waves and Optics course, learned about expanding functions in series of sine and cosine functions. Such Fourier expansions will be the topic of the next section. Here expand $f(x)$ in an infinite series of Legendre polynomials:

$$f(x) = \sum_{n=0}^{\infty} a_n \, P_n(x) \; . \tag{248}$$

Start with a simple example:

$$f(x) = \frac{15}{2} x^2 - \frac{3}{2} = \frac{15}{2} \cdot \frac{1}{3} \left(2P_2(x) + P_0(x) \right) - \frac{3}{2} \, P_0(x) = P_0(x) + 5P_2(x) \; .$$

Whenever the power series for $f(x)$ only has a finite number of terms, *i.e.* is a polynomial, can calculate coefficients by solving system of algebraic equations. Example above is of this kind. If $f(x)$ is not a polynomial then can still calculate the coefficients using the orthonormality integral of Eq. (247). Multiplying Eq. (248) by $P_m(x)$ and integrating from -1 to $+1$ gives

$$\int_{-1}^{+1} f(x) \, P_m(x) \, dx = \sum_{n=0}^{\infty} a_n \int_{-1}^{+1} P_n(x) \, P_m(x) \, dx = \sum_{n=0}^{\infty} a_n \, \frac{2}{2n+1} \, \delta_{m\,n} = \frac{2}{2m+1} \, a_m \; . \tag{249}$$

Hence

$$a_m = \frac{2m+1}{2} \int_{-1}^{+1} f(x) \, P_m(x) \, dx \; . \tag{250}$$

Example 1. Calculate the Legendre coefficients for $f(x) = \frac{15}{2}x^2 - \frac{3}{2}$. Putting in the explicit forms for the Legendre polynomials, we have

$$a_0 = \frac{1}{2} \int_{-1}^{+1} \left(\frac{15}{2} x^2 - \frac{3}{2} \right) \, dx = \frac{1}{2} \left(\frac{15}{3} - 3 \right) = 1 \; ,$$

$$a_1 = \frac{3}{2} \int_{-1}^{+1} \left(\frac{15}{2} x^2 - \frac{3}{2} \right) \, x \, dx = 0 \quad \text{(integrand odd)},$$

$$a_2 = \frac{5}{2} \int_{-1}^{+1} \left(\frac{15}{2} x^2 - \frac{3}{2} \right) \left(\frac{3x^2}{2} - \frac{1}{2} \right) \, dx = \frac{5}{4} \left(\frac{45}{5} - \frac{24}{3} + 3 \right) = 5 \; .$$

These agree with what we found using direct algebra.

Example 2. Obtain the first two the Legendre coefficients for $f(x) = e^{\alpha x}$.

$$a_0 = \frac{1}{2}\int_{-1}^{+1} e^{\alpha x}\, dx = \frac{1}{2\alpha}\left(e^{\alpha} - e^{-\alpha}\right) = \frac{\sinh\alpha}{\alpha}\,,$$

$$a_1 = \frac{3}{2}\int_{-1}^{+1} x\, e^{\alpha x}\, dx = 3\left(\frac{\cosh\alpha}{\alpha} - \frac{\sinh\alpha}{\alpha^2}\right)\,,$$

where we had to do some integration by parts.

4.7 Return to the Potential Problem

Laplace equation in spherical polar coordinates and with axial symmetry (no ϕ dependence) has general solution for electrostatic potential in charge-free space

$$V(r, \cos\theta) = \sum_{\ell=0}^{\infty}\left(A_\ell r^\ell + \frac{B_\ell}{r^{\ell+1}}\right) P_\ell(\cos\theta)\,. \tag{251}$$

Suppose that $V(r) \to 0$ as $r \to \infty$. This means that all the $A_\ell = 0$ and

$$V(r, \cos\theta) = \sum_{\ell=0}^{\infty}\frac{B_\ell}{r^{\ell+1}} P_\ell(\cos\theta)\,. \tag{252}$$

To fix the values of B_ℓ coefficients, need another boundary condition:

Example

Suppose that on an isolated sphere of radius a the electrostatic potential varies like $V(r = a, \theta) = V_0\, e^{\alpha\cos\theta}$. How does the potential behave for large distances?

Using the Legendre series example already worked out,

$$B_0 = V_0 a\,\frac{\sinh\alpha}{\alpha}\,,$$

$$B_1 = 3V_0 a^2\left(\frac{\cosh\alpha}{\alpha} - \frac{\sinh\alpha}{\alpha^2}\right)\,,$$

and

$$V(r, \cos\theta) = V_0\left[\frac{a}{r}\,\frac{\sinh\alpha}{\alpha} + 3\frac{a^2}{r^2}\left(\frac{\cosh\alpha}{\alpha} - \frac{\sinh\alpha}{\alpha^2}\right)\cos\theta + 0\left(\frac{1}{r^3}\right)\right]\,.$$

For those of you familiar with electrostatics, the $\cos\theta$ term corresponds to the electric dipole moment and the discarded next term the quadrupole moment *etc.*

4.8 Associated Legendre Functions

In general ϕ dependence of solutions to Laplace's equation is of form $e^{im\phi}$, where m is an integer. To get the θ dependence, have to solve

$$\frac{d}{dx}\left[(1-x^2)\frac{dP_\ell^m}{dx}\right] + \left[\ell(\ell+1) - \frac{m^2}{1-x^2}\right] P_\ell^m = 0\,. \tag{253}$$

Have here replaced $\mu \to x$ and called $\Theta(\mu) \to P_\ell^m(x)$.] Only for $m = 0$ do we get the Legendre polynomials $P_\ell(x)$. To solve the equation for $m \neq 0$ is even more tedious than for $m = 0$. But results are important for Quantum Mechanics, where ℓ is known as the angular momentum quantum number and m the magnetic quantum number.

Well behaved solutions of Legendre's equation are only possible if

- ℓ is a non-negative integer.
- m is an integer with $-\ell \leq m \leq \ell$.

Solutions for m and $-m$ are the same since only m^2 occurs in Legendre's equation. For $m > 0$ the associated Legendre functions, P_ℓ^m, can be derived from the Legendre polynomials using

$$P_\ell^m(x) = (1-x^2)^{m/2}\left(\frac{d}{dx}\right)^m P_\ell(x)\,. \tag{254}$$

The orthogonality relation is also a bit more complicated than that of Eq. (246):

$$\int_{-1}^{+1} P_\ell^m(x)\, P_n^m(x)\, dx = \frac{2}{2\ell+1}\,\frac{(\ell+m)!}{(\ell-m)!}\,\delta_{\ell n}\,. \tag{255}$$

Specific cases

$\ell = 1$, $m = 1$:

$$P_1^1(x) = (1-x^2)^{1/2}\,\frac{d}{dx}\,x = (1-x^2)^{1/2} = \sin\theta\,.$$

$\ell = 2$, $m = 1$:

$$P_2^1(x) = (1-x^2)^{1/2}\,\frac{d}{dx}\,\tfrac{1}{2}(3x^2-1) = 3x\,(1-x^2)^{1/2} = 3\sin\theta\cos\theta\,.$$

$\ell = 2$, $m = 2$:

$$P_2^2(x) = (1-x^2)\,\frac{d^2}{dx^2}\,\tfrac{1}{2}(3x^2-1) = 3(1-x^2) = 3\sin^2\theta\,.$$

As a couple of examples to check the orthogonality relations, consider

$$\int_{-1}^{+1} P_2^1(x)\, P_1^1(x)\, dx = \int_{-1}^{+1} 3x(1-x^2)\, dx = 0\,.$$

$$\int_{-1}^{+1}\left[P_2^1(x)\right]^2\, dx = \int_{-1}^{+1} 9x^2(1-x^2)\, dx = 2\,\frac{9}{3} - 2\,\frac{9}{5} = \frac{12}{5}\,.$$

The last one agrees with the $\frac{2}{5}\,3!$ of Eq. (255).

4.9 Spherical Harmonics

Often (eg in Quantum Mechanics) gets the θ and ϕ dependence packaged together as one function called a spherical harmonic $Y_\ell^m(\theta, \phi)$. Thus

$$Y_\ell^m(\theta, \phi) = c_{\ell,m}\, P_\ell^m(\cos\theta)\, e^{im\phi} \tag{256}$$

is a solution of Legendre equation. Here the constants $c_{\ell,m}$ could be chosen many ways.By convention Y_ℓ^m os normalised so that

$$\int_{-1}^{+1} d(\cos\theta) \int_0^{2\pi} d\phi\, Y_\ell^{m\,*}(\theta, \phi)\, Y_{\ell'}^{m'}(\theta, \phi) = \delta_{\ell,\ell'}\, \delta_{m,m'}\,. \tag{257}$$

Using Eq. (255), this is achieved with

$$c_{\ell,m} = (-1)^m \sqrt{\frac{(\ell - m)!\,(2\ell + 1)}{(\ell + m)!\, 4\pi}}\,. \tag{258}$$

In Quantum Mechanics you will at some stage need to remember the orthogonality/normalisation relation of Eq. (257) but you will <u>NOT</u> be required to remember the actual algebraic form of Eq. (258).

The first few spherical harmonics are

$$\begin{aligned}
Y_0^0(\theta, \phi) &= \frac{1}{\sqrt{4\pi}}\,, \\
Y_1^1(\theta, \phi) &= -\sqrt{\frac{3}{8\pi}}\, \sin\theta\, e^{i\phi}\,, \\
Y_1^0(\theta, \phi) &= \sqrt{\frac{3}{4\pi}}\, \cos\theta\,, \\
Y_1^{-1}(\theta, \phi) &= \sqrt{\frac{3}{8\pi}}\, \sin\theta\, e^{-i\phi}\,.
\end{aligned} \tag{259}$$

5 Fourier Analysis

Have expanded arbitrary vector $\underline{A}$ in terms of basis vectors $\hat{\underline{e}}_i$ which are orthogonal (and normalised);

$$\hat{\underline{e}}_m \cdot \hat{\underline{e}}_n = \delta_{m\,n}\,. \tag{260}$$

Have also seen that an arbitrary function of $\cos\theta$ can be expanded as a series of Legendre polynomials, which are orthogonal (and normalised);

$$\int_{-1}^{+1} P_m(x)\, P_n(x)\, dx = \frac{2}{2m+1}\,\delta_{m\,n}\,. \tag{261}$$

The two formulae look very similar. The crucial difference between the two is that the scalar product is defined differently in the two cases. For the Legendre polynomials, the scalar product of two of them is given by an integral from -1 to $+1$. The factor of $2/(2m+1)$ on the right hand side is of no real importance — one doesn't <u>have</u> to work with basis vectors of length one.

5.1 Fourier Series

The other place where you have met the expansion of functions in terms of orthogonal functions is in the Fourier series that you saw in the Waves and Optics course in the first year. A pure note on a violin corresponds to a sinusoidal variation in both position x and time t. However, when the violinist bows the instrument, she or he excites a whole range of notes. To find the notes, the initial signal must be expanded in a Fourier series. We want here to justify and extend some of these first year techniques.

The orthogonality integral is most elegant in terms of complex exponentials:

$$\int_{-\pi}^{+\pi} e^{-imx}\, e^{inx}\, dx = 2\pi\delta_{m\,n}\,. \tag{262}$$

Using Euler's relation, $e^{imx} = \cos mx + i\sin mx$, we can then convert Eq. (262) into integrals for sines and cosines to give

$$\int_{-\pi}^{+\pi} \cos mx\, \sin nx\, dx = 0 \quad \text{for all } m, n. \tag{263}$$

Provided that $m, n \neq 0$,

$$\int_{-\pi}^{+\pi} \cos mx\, \cos nx\, dx = \int_{-\pi}^{+\pi} \sin mx\, \sin nx\, dx = \pi\,\delta_{m\,n}\,. \tag{264}$$

If $m = n = 0$, the the sine integral vanishes and we are just left with the cosine integral

$$\int_{-\pi}^{+\pi} 1\, dx = 2\pi \; . \tag{265}$$

Thus the functions

$$\frac{1}{\sqrt{2\pi}} \; , \quad \frac{1}{\sqrt{\pi}} \cos mx \; , \quad \frac{1}{\sqrt{\pi}} \sin mx \; , \tag{266}$$

with m a positive integer, form an orthonormal set with respect to integration over the interval $-\pi \le x \le +\pi$.

An arbitrary function $f(x)$ in the interval $-\pi \le x \le +\pi$ may be written in the form

$$f(x) = \tfrac{1}{2}a_0 + \sum_{n=1}^{\infty} a_n \cos nx + \sum_{n=1}^{\infty} b_n \sin nx \; , \tag{267}$$

where the Fourier coefficients are given by

$$\begin{aligned} a_m &= \frac{1}{\pi}\int_{-\pi}^{+\pi} f(x)\, \cos mx\, dx \; , \\ b_m &= \frac{1}{\pi}\int_{-\pi}^{+\pi} f(x)\, \sin mx\, dx \; . \end{aligned} \tag{268}$$

Proof: Multiply both sides of Eq. (267) by $\sin mx$ and integrate from $-\pi$ to $+\pi$.

$$\begin{aligned} \int_{-\pi}^{+\pi} f(x)\, \sin mx\, dx = \tfrac{1}{2}a_0 \int_{-\pi}^{+\pi} \sin mx\, dx + \sum_{n=1}^{\infty} a_n \int_{-\pi}^{+\pi} \sin mx\, \cos nx\, dx \\ + \sum_{n=1}^{\infty} b_n \int_{-\pi}^{+\pi} \sin mx\, \sin nx\, dx \; . \end{aligned} \tag{269}$$

The first two terms on the right hand side clearly vanish because the integrands are odd. The third term is only non-zero if $m = n$, which means that there is only one term in the sum. Using Eq. (264), this gives

$$\int_{-\pi}^{+\pi} f(x)\, \sin mx\, dx = \sum_{n=0}^{\infty} b_n\, \pi\, \delta_{m\,n} = \pi\, b_m \; , \tag{270}$$

as required.

Alternatively, multiplying by the cosine,

$$\begin{aligned} \int_{-\pi}^{+\pi} f(x)\, \cos mx\, dx = \tfrac{1}{2}a_0 \int_{-\pi}^{+\pi} \cos mx\, dx + \sum_{n=1}^{\infty} a_n \int_{-\pi}^{+\pi} \cos mx\, \cos nx\, dx \\ + \sum_{n=1}^{\infty} b_n \int_{-\pi}^{+\pi} \cos mx\, \sin nx\, dx \; . \end{aligned} \tag{271}$$

It is the third term which now vanishes for all m and n. If $m = 0$, only the first term survives and

$$\int_{-\pi}^{+\pi} f(x)\,\cos mx\,dx = \tfrac{1}{2}a_0\,2\pi = \pi\,a_0\,, \quad \text{for } m = 0\,. \tag{272}$$

On the other hand, if $m \neq 0$, it is the second term which is non-vanishing and

$$\int_{-\pi}^{+\pi} f(x)\,\cos mx\,dx = \sum_{n=1}^{\infty} a_n\,\pi\,\delta_{m\,n} = \pi\,a_m\,. \tag{273}$$

The $\frac{1}{2}$ factor in front of a_0 in Eq. (268) gives a consistent formula for all a_m.

Example 1: Rectangular wave. Consider the function

$$f_1(x) = \begin{cases} +1 & \text{for } \ 0 < x < \pi \\ -1 & \text{for } \ -\pi < x < 0 \end{cases}$$

Note that this is an <u>odd</u> function, *i.e.* $f(x) = -f(-x)$. Using Eq. (268), this means that all the even coefficients $a_n = 0$. [The $\cos nx$ are even functions and, when multiplied by $f(x)$, give odd integrands.] On the other hand

$$b_n = \frac{1}{\pi}\int_0^{\pi} \sin nx\,dx - \frac{1}{\pi}\int_{-\pi}^{0} \sin nx\,dx = \frac{1}{n\pi}\Big[-\cos nx\Big]_0^{\pi} - \frac{1}{n\pi}\Big[-\cos nx\Big]_{-\pi}^{0}$$

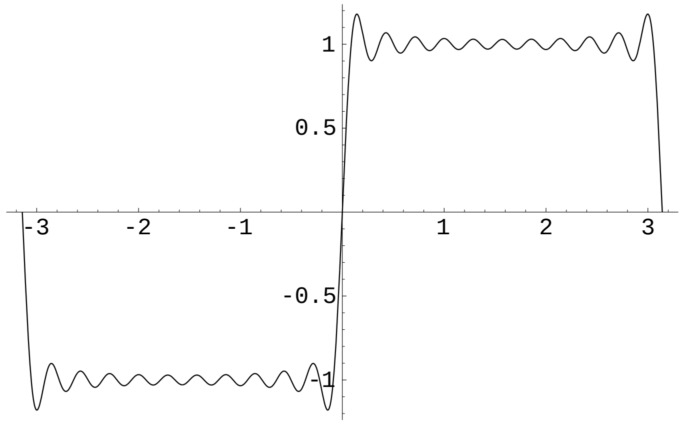
$$= \frac{1}{n\pi}[1 - \cos n\pi + 1 - \cos(-n\pi)] = \frac{2}{n\pi}[1 - \cos n\pi] = \frac{4}{n\pi} \times \begin{cases} 1 & \text{for } n \text{ odd} \\ 0 & \text{for } n \text{ even.} \end{cases}$$

Thus the Fourier series becomes

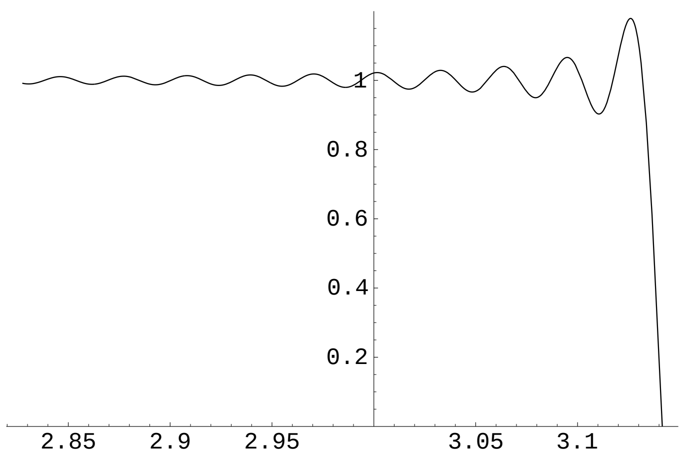
$$f_1(x) = \frac{4}{\pi}\sum_{n\ \mathrm{odd}} \frac{1}{n}\,\sin nx\,.$$

As an example of how things look if we just take a finite number of terms, the picture shows what happens when the series is truncated at $n = 21$. You see there are lots of little oscillations (typically 21) and the sizes of these oscillations get smaller as the number of terms increases. However, with a finite number of terms like this, the representation of a function which changes so sharply near $x = 0$ is not perfect! Note that the original function was not defined at $x = 0$ but the Fourier series has resulted in a representation with $f(0) = 0$. This is typical of a case where the function is discontinuous and the Fourier series will then converge to the mean of the results to the left and right of the discontinuity at $x = x_0$:

$$\longrightarrow \lim_{\epsilon\to 0} \tfrac{1}{2}\left\{f(x_0+\epsilon) + f(x_0-\epsilon)\right\}\,. \tag{274}$$



No matter how many terms you add, you never get it quite right at a discontinuity. Thus in the next picture there are are 100 terms and there is still an overshoot of about 18%. With more terms, the overshoot always stays the same size but it just gets squeezed into a smaller region in x. This is known as the Gibbs phenomenon. On the other hand, one could not get anything like as good a description of a discontinuous function using a power series.



Periodic Functions

So far we have looked at functions defined in the region $-\pi \leq x \leq +\pi$. What happens outside this region? Going back to Eq. (267), we see that the function is periodic with period 2π since

$$f(x+2\pi) = \tfrac{1}{2}a_0 + \sum_{n=1}^{\infty} a_n \cos n(x+2\pi) + \sum_{n=1}^{\infty} b_n \sin n(x+2\pi)$$

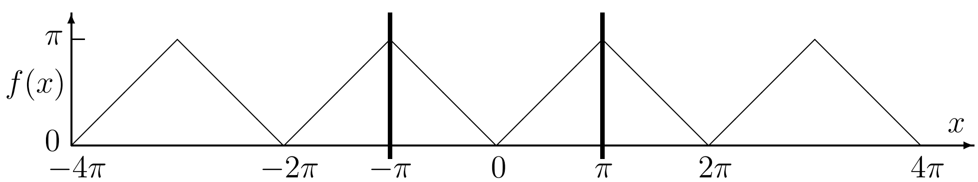
$$= \tfrac{1}{2}a_0 + \sum_{n=1}^{\infty} a_n \cos nx + \sum_{n=1}^{\infty} b_n \sin nx = f(x)\,. \tag{275}$$

Therefore we can use Fourier series either on a function which is defined just over an interval $-\pi \leq x \leq +\pi$, or one which is periodic with period 2π.

Example 2: A function $f_2(x)$, which is periodic with period 2π, is defined by

$$\begin{aligned} f_2(x) &= x\,, & 0 \leq x \leq \pi\,, \\ &= -x\,, & -\pi \leq x \leq 0\,. \end{aligned}$$

$f_2(x)$ is an even function, as shown in the picture.



Due to the eveness, the b_n all vanish since the integrands are odd. This is a very general trick — an odd function only has sines in its expansion whereas an even function has only cosines. The even coefficients are:

$$a_0 = \frac{2}{\pi}\int_0^{\pi} x\,dx = \left[\frac{x^2}{\pi}\right]_0^{\pi} = \pi\,,$$

where use has been made of the even character to integrate over half the interval.

$$a_n = \frac{2}{\pi}\int_0^{\pi} x\cos nx\,dx = \frac{2}{\pi}\left[\frac{x}{n}\sin nx + \frac{1}{n^2}\cos nx\right]_0^{\pi}\,.$$

The sine function vanishes at both limits, so that

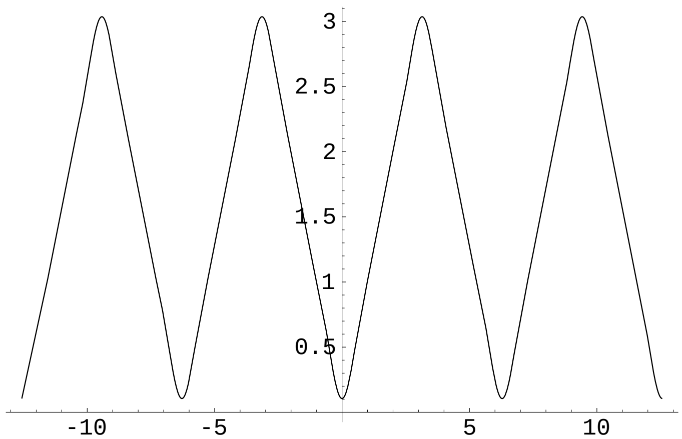
$$a_n = -\frac{2}{\pi n^2}\left[1-(-1)^n\right]\,.$$

This means that $a_n = -4/\pi n^2$ if n is odd but that $a_n = 0$ for even n.

Hence

$$f_2(x) = \frac{\pi}{2} - \frac{4}{\pi} \sum_{n \text{ odd}} \frac{1}{n^2} \cos nx \, .$$

This series converges rather faster than the first example since, keeping terms only up to $n = 5$, we get the picture:



The reason for this better behaviour is that, unlike the case in example 1, the original function $f_2(x)$ has no sudden jumps, although it has sudden changes in slope. Note, however, that with a finite number of terms the Fourier series never quite gets to zero.

Since $f_2(x)$ should vanish at $x = 0$, rearranging the Fourier series at this point gives

$$\sum_{n \text{ odd}} \frac{1}{n^2} = \frac{\pi^2}{8} \, .$$

The right hand side ≈ 1.234. Keeping only three terms on the left hand side gives $1 + 1/9 + 1/25 \approx 1.151$, which differs from the true answer by about 7%. This is another manifestation of the convergence of the Fourier series.

General interval

So far only looked at x between $\pm\pi$ but the same formulae are valid for any range of the same size, such as $0 \leq x \leq 2\pi$. If, on the other hand, the fundamental interval is $-L \leq x \leq +L$, just change variables to $y = \pi x/L$. Then

$$f(x) = \tfrac{1}{2} a_0 + \sum_{n=1}^{\infty} a_n \cos \tfrac{n\pi x}{L} + \sum_{n=1}^{\infty} b_n \sin \tfrac{n\pi x}{L} \, , \tag{276}$$

where the Fourier coefficients are given by

$$
\begin{aligned}
a_m &= \frac{1}{L}\int_{-L}^{+L} f(x)\,\cos\tfrac{m\pi x}{L}\,dx\,, \\
b_m &= \frac{1}{L}\int_{-L}^{+L} f(x)\,\sin\tfrac{m\pi x}{L}\,dx\,. \qquad (277)
\end{aligned}
$$

The conditions under which a Fourier expansion is valid go by the name of Dirichlet conditions, which are discussed by Boas. Roughly speaking, if $f(x)$ is periodic with period $2L$ with a finite number of discontinuities, then the expansion is valid provided that

$$\int_{-L}^{+L} |\,f(x)\,|\;dx \quad \text{is finite.} \qquad (278)$$

Differentiation of Fourier series

General motto: *be careful*! In the examples that we have looked at,

$$f_1(x) = \frac{d}{dx}\,f_2(x)\,.$$

Does this hold for their Fourier series? Just try and see!

$$\frac{d}{dx}\,f_2(x) = \frac{d}{dx}\left(\frac{\pi}{2} - \frac{4}{\pi}\sum_{n\ \mathrm{odd}}\frac{1}{n^2}\cos nx\right) = \frac{4}{\pi}\sum_{n\ \mathrm{odd}}\frac{1}{n}\sin nx = f_1(x)\,,$$

which works fine.

Now go one step further and look at the Fourier series for the next derivative

$$\frac{d}{dx}\,f_1(x) = \frac{4}{\pi}\sum_{n\ \mathrm{odd}}\cos nx\,.$$

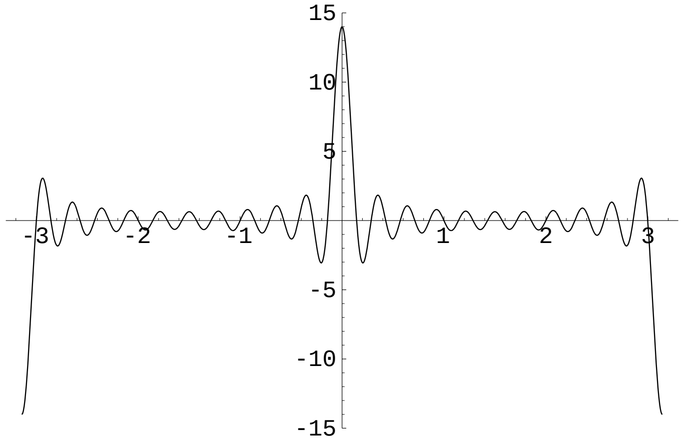
The series clearly must blow up at $x = 0$ or $x = \pm\pi$ because we then have an infinite number of terms equal to $+1$ or -1. Away from these points, the Fourier series oscillates around zero as in the picture, which was calculated with terms up to $n = 21$. If one takes more terms the peaks at $x = m\pi$ get higher but narrower.

To repeat, if the function is smooth then we can differentiate its Fourier series term by term. At any discontinuities we have to be careful — sometimes very careful!

Parseval's Identity

Suppose that $f(x)$ is periodic with period 2π such that it has the Fourier series

$$f(x) = \tfrac{1}{2}a_0 + \sum_{n=1}^{\infty} a_n\,\cos nx + \sum_{n=1}^{\infty} b_n\,\sin nx\,. \qquad (279)$$



Parseval's theorem is that the average value of f^2 is given by

$$< f^2(x) >= \frac{1}{2\pi} \int_{-\pi}^{+\pi} [f(x)]^2 \, dx = \left(\frac{a_0}{2} \right)^2 + \frac{1}{2} \sum_{n=1}^{\infty} (a_n^2 + b_n^2) \, . \tag{280}$$

We can insert the representation of Eq. (279) into the left hand side of Eq. (280) and simply carry out all the integrations. Now, by Eqs. (263, 264), there are no cross terms which survive the integration, so that

$$\frac{1}{2\pi} \int_{-\pi}^{+\pi} [f(x)]^2 \, dx = \frac{1}{2\pi} \int_{-\pi}^{+\pi} dx \left\{ \left(\tfrac{1}{2} a_0 \right)^2 + \sum_{n=1}^{\infty} a_n^2 \cos^2 nx + \sum_{n=1}^{\infty} b_n^2 \sin^2 nx \right\} \, . \tag{281}$$

Now the average value $< \cos^2 nx >=< \sin^2 nx >= \frac{1}{2}$, so that Parseval's identity follows immediately.

Let us now see what this gives us for the two examples that we have worked out. In the first case, $f_1^2(x) = 1$, and hence

$$1 = \frac{1}{2} \frac{16}{\pi^2} \sum_{n \text{ odd}} \frac{1}{n^2} \, ,$$

that is

$$\sum_{n \text{ odd}} \frac{1}{n^2} = \frac{\pi^2}{8} \, .$$

But this is already the result that we got from looking at the sum of the Fourier series in the second example at the position $x = 0$. The two answers are the same!

In the second case of the saw-tooth wave,

$$< f_2^2(x) >= \frac{1}{2\pi} \int_{-\pi}^{+\pi} x^2 \, dx = \left[\frac{x^3}{6\pi} \right]_{-\pi}^{+\pi} = \frac{\pi^2}{3} \, .$$

Hence

$$\frac{\pi^2}{3} = \left(\frac{\pi}{2} \right)^2 + \frac{8}{\pi^2} \sum_{n \text{ odd}} \frac{1}{n^4} \, .$$

Thus

$$\sum_{n \text{ odd}} \frac{1}{n^4} = \frac{\pi^4}{96} \, .$$

This agrees with the output from Mathematica!

Complex Fourier series

For those of you who are happy with complex numbers, the complex Fourier series are easier to handle than the real ones. If $f(x)$ is periodic, with period 2π, then

$$f(x) = \sum_{n=-\infty}^{+\infty} c_n \, e^{inx} \, , \tag{282}$$

where the *complex* coefficients are given by

$$c_n = \frac{1}{2\pi} \int_{-\pi}^{+\pi} e^{-inx} \, f(x) \, dx \, . \tag{283}$$

The proof of this follows immediately using the orthonormality condition of Eq. (262):

$$\int_{-\pi}^{+\pi} e^{-imx} \, e^{inx} \, dx = 2\pi \delta_{m\,n} \, .$$

Example: Let us do example 1 again and show that we get the same answer.

$$c_n = \frac{1}{2\pi} \int_{-\pi}^{0} e^{-inx} \, (-1) \, dx + \frac{1}{2\pi} \int_{0}^{+\pi} e^{-inx} \, (+1) \, dx = \frac{1}{2\pi} \left[\frac{-1}{-in} \, e^{-inx} \right]_{-\pi}^{0} + \frac{1}{2\pi} \left[\frac{1}{-in} \, e^{-inx} \right]_{0}^{\pi}$$

$$= \frac{1}{n\pi i} \left(1 - (-1)^n \right) \, .$$

This means that all the even coefficients vanish and the odd ones are $c_n = 2/(n\pi i)$. The complex Fourier series becomes

$$f(x) = \frac{2}{\pi i} \sum_{n=-\infty}^{+\infty} \frac{1}{n} \, e^{inx} \quad (n \text{ odd}) = \frac{2}{\pi i} \sum_{n=-\infty}^{-1} \frac{1}{n} \, e^{inx} + \frac{2}{\pi i} \sum_{n=1}^{+\infty} \frac{1}{n} \, e^{inx}$$

$$= -\frac{2}{\pi i} \sum_{n=1}^{\infty} \frac{1}{n} \, e^{-inx} + \frac{2}{\pi i} \sum_{n=1}^{+\infty} \frac{1}{n} \, e^{inx} = \frac{4}{\pi} \sum_{n=1}^{+\infty} \frac{1}{n} \, \frac{1}{2i} \left(e^{inx} - e^{-inx} \right) \, ,$$

where all the sums are over the odd values of n. You recognise the expression for $\sin nx$ in the last bracket and so we have obtained the same result as before.

There is a form of Parseval's identity which is valid for complex Fourier series:

$$\frac{1}{2\pi} \int_{-\pi}^{+\pi} | \, f(x) \, |^2 \, dx = \frac{1}{2\pi} \int_{-\pi}^{+\pi} f^*(x) \, f(x) \, dx = \sum_{n=-\infty}^{+\infty} | \, c_n \, |^2 \, . \tag{284}$$

5.2 Fourier Transforms

Go back to the expressions of eqs. (276, 277) for the Fourier series and Fourier coefficients in the case of an arbitrary interval of length $2L$:

$$f(x) = \tfrac{1}{2}a_0 + \sum_{n=1}^{\infty} a_n \cos\tfrac{n\pi x}{L} + \sum_{n=1}^{\infty} b_n \sin\tfrac{n\pi x}{L} ,$$

where

$$\begin{aligned} a_n &= \frac{1}{L}\int_{-L}^{+L} f(t) \cos\tfrac{n\pi t}{L}\, dt , \\ b_n &= \frac{1}{L}\int_{-L}^{+L} f(t) \sin\tfrac{n\pi t}{L}\, dt. \end{aligned}$$

Putting these together as one equation,

$$f(x) = \frac{1}{2L}\int_{-L}^{+L} f(t)\, dt + \frac{1}{L}\sum_{n=1}^{\infty} \cos\tfrac{n\pi x}{L} \int_{-L}^{+L} f(t) \cos\tfrac{n\pi t}{L} dt + \frac{1}{L}\sum_{n=1}^{\infty} \sin\tfrac{n\pi x}{L} \int_{-L}^{+L} f(t) \sin\tfrac{n\pi t}{L} dt. \tag{285}$$

Using the trigonometric addition formula

$$\cos\tfrac{n\pi x}{L}\cos\tfrac{n\pi t}{L} + \sin\tfrac{n\pi x}{L}\sin\tfrac{n\pi t}{L} = \cos\tfrac{n\pi}{L}(t-x) , \tag{286}$$

the Fourier series result can be written in the more compact form

$$f(x) = \frac{1}{2L}\int_{-L}^{+L} f(t)\, dt + \frac{1}{L}\sum_{n=1}^{\infty}\int_{-L}^{+L} f(t) \cos\tfrac{n\pi}{L}(t-x)\, dt . \tag{287}$$

Fourier transforms are what happens to Fourier series when the interval length $2L$ tends to infinity. To bring this about, define

$$\omega = \frac{n\pi}{L} \quad \text{and} \quad \Delta\omega = \frac{\pi}{L} . \tag{288}$$

In the limit that $L \to \infty$, the first term in Eq. (287) goes to zero provided the infinite integral converges. Hence

$$f(x) = \frac{1}{\pi}\sum_{n=1}^{\infty} \Delta\omega \int_{-\infty}^{+\infty} f(t) \cos\omega(t-x)\, dt . \tag{289}$$

Now since $\Delta\omega \to 0$ as $L \to \infty$, the sum in Eq. (289) can be replaced by an integral to reveal the fundamental expression of Fourier transforms:

$$f(x) = \frac{1}{\pi}\int_{0}^{\infty} d\omega \int_{-\infty}^{+\infty} f(t) \cos\omega(t-x)\, dt . \tag{290}$$

Since $\cos\omega(t-x)$ is an *even* function of ω, we can extend the integration limit to $-\infty$ provided that we divide by a factor of 2:

$$f(x) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} d\omega \int_{-\infty}^{+\infty} f(t)\,\cos\omega(t-x)\,dt\,. \tag{291}$$

On the other hand, $\sin\omega(t-x)$ is an *odd* function of ω, which means that

$$0 = \frac{1}{2\pi}\int_{-\infty}^{+\infty} d\omega \int_{-\infty}^{+\infty} f(t)\,\sin\omega(t-x)\,dt\,. \tag{292}$$

Adding i times Eq. (292) to Eq. (291), we get

$$f(x) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} d\omega \int_{-\infty}^{+\infty} f(t)\,[\cos\omega(t-x)+i\sin\omega(t-x)]\,dt = \frac{1}{2\pi}\int_{-\infty}^{+\infty} d\omega \int_{-\infty}^{+\infty} f(t)\,e^{i\omega(t-x)}\,dt. \tag{293}$$

Splitting up the exponential, we get to the final result that

$$f(x) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} d\omega\, e^{-i\omega x} \int_{-\infty}^{+\infty} f(t)\,e^{i\omega t}\,dt\,. \tag{294}$$

Now introduce the Fourier transform

$$g(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} f(x)\,e^{i\omega x}\,dx\,, \tag{295}$$

and its inverse

$$f(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} g(\omega)\,e^{-i\omega x}\,d\omega\,. \tag{296}$$

The two equations look the same except that in one there is a $+i$ in the exponent, whereas in the other there is a $-i$. Eqs.(295, 296) have been defined symmetrically, each with a $1\sqrt{2\pi}$ factor. Many books put a $1/2\pi$ factor in front of one integral and unity in front of the other.

The variable ω introduced here is an arbitrary mathematical variable but in most physical problems it corresponds to the angular frequency ω. The Fourier transform represents $f(x)$ in terms of a distribution of infinitely long sinusoidal wave trains where the frequency is a continuous variable. You will come across this in Quantum Mechanics, where such waves are eigenfunctions of the momentum operator $\hat{p}$. Then $g(\omega) = g(p)$ is the momentum-space representation of the function $f(x)$.

Example

Consider

$$E(t) = E_0\,e^{-\gamma t/2}\,e^{-i\omega_0 t} = E_0\,e^{-(i\omega_0+\frac{1}{2}\gamma)t} \quad \text{for } t \geq 0$$

and which vanishes for negative values of t. This could represent a damped oscillating electric field which was switched on at time $t = 0$. The Fourier transform is

$$g(\omega) = \frac{1}{\sqrt{2\pi}} E_0 \int_0^{+\infty} e^{i\omega t}\, e^{-i(\omega_0 - \frac{1}{2}i\gamma)t}\, dt = \frac{1}{\sqrt{2\pi}} E_0 \frac{1}{i\omega - i\omega_0 - \frac{1}{2}\gamma} \left[e^{i(\omega-\omega_0+\frac{1}{2}i\gamma)t} \right]_0^{\infty}$$

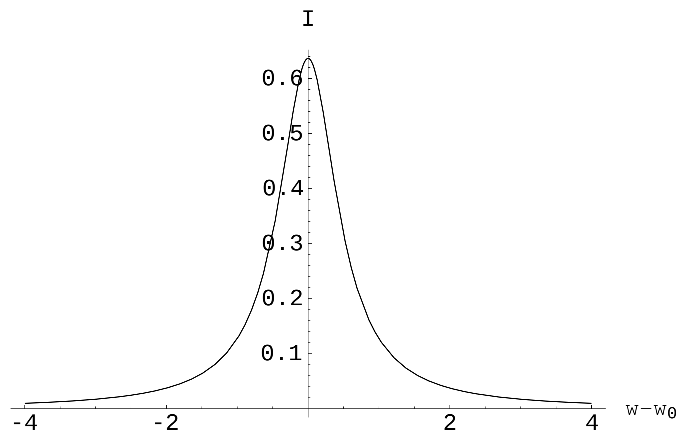
Because of the damping, the integated term vanishes at the upper limit and so we are left with

$$g(\omega) = \frac{1}{\sqrt{2\pi}} \frac{iE_0}{\omega - \omega_0 + \frac{1}{2}i\gamma}\,.$$

The intensity spectrum

$$I(\omega) = |\, g(\omega)\,|^2 = \frac{E_0^2}{2\pi} \frac{1}{(\omega-\omega_0)^2 + \gamma^2/4}$$

is peaked at $\omega = \omega_0$ with a width of γ. In plotting the figure we have taken $\gamma = 1$.



Useful results

1) If $f(x)$ is an *even* function of x, then

$$g(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} f(x)\, e^{i\omega x}\, dx = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} f(x)\, \cos\omega x\, dx$$

is an *even* function of ω.

2) Similarly, if $f(x)$ is an *odd* function of x, then $g(\omega)$ is an *odd* function of ω:

$$g(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} f(x)\, e^{i\omega x}\, dx = \frac{i}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} f(x)\, \sin\omega x\, dx\,.$$

3) Differentiating Eq. (296),

$$f'(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} [-i\omega\, g(\omega)]\, e^{-i\omega x}\, d\omega\,, \tag{297}$$

so that $-i\omega\, g(\omega)$ is the Fourier transform of $f'(x)$. By extension, $(-i\omega)^n\, g(\omega)$ is the Fourier transform of $d^n f/dx^n$.

4) From Eq. (296),

$$f(x+a) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} [e^{-i\omega a}\, g(\omega)]\, e^{-i\omega x}\, d\omega\ , \tag{298}$$

so that $e^{-i\omega a}\, g(\omega)$ is the Fourier transform of $f(x+a)$.

The Dirac delta function

The Kronecker delta symbol $\delta_{i\,j}$ has the property that

$$a_i = \sum_j \delta_{i\,j}\, a_j \tag{299}$$

for any vector a_j, provided that the sum includes the term where $i = j$. The Dirac delta function is the generalisation of this to the case where we have an integral rather than a sum, *i.e.* we want a function $\delta(x-t)$, such that

$$f(x) = \int_{-\infty}^{+\infty} \delta(x-t)\, f(t)\, dt\ . \tag{300}$$

This means that $\delta(t-x)$ is zero everywhere except the point $t = x$ but there it is so big that the integral is unity. This is rather like having a point charge in electrostatics — it is an idealisation. $\delta(t-x)$ is not a function in the normal sense; it is just too badly behaved.

Going back to Eq. (294),

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} d\omega\, e^{-i\omega x} \int_{-\infty}^{+\infty} f(t)\, e^{i\omega t}\, dt\ ,$$

we can rearrange it as

$$f(x) = \int_{-\infty}^{+\infty} f(t)\, dt \left\{ \frac{1}{2\pi} \int_{-\infty}^{+\infty} d\omega\, e^{i\omega(t-x)} \right\}\ . \tag{301}$$

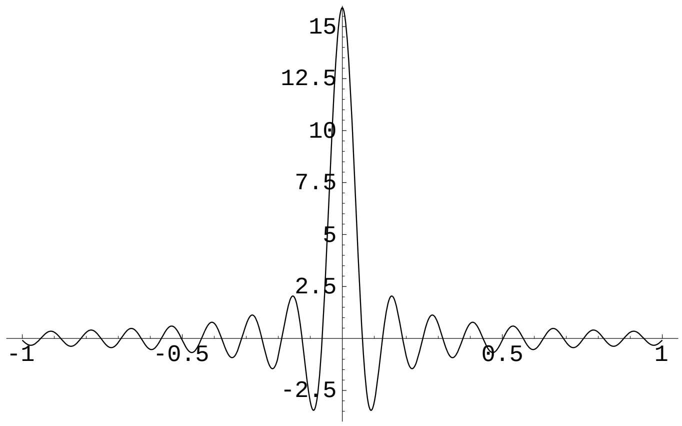
Thus we can identify

$$\delta(t-x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} d\omega\, e^{i\omega(t-x)}\ . \tag{302}$$

We have to be a bit careful about the convergence of the integral at large values of ω. Let us cut off the integration at $\omega = \pm N$ and then study what happens as N gets large.

$$\delta_N(t-x) = \frac{1}{2\pi} \int_{-N}^{+N} d\omega\, e^{i\omega(t-x)} = \frac{\sin N(t-x)}{\pi(t-x)}\ . \tag{303}$$

With $N = 50$, the figure shows a strong spike at $t = x$, but with lots of oscillations. The spike gets sharper as N gets larger, but the lobes at the bottom remain a constant fraction $2/3\pi$ of the central value. Note that $\delta(x) = \delta(-x)$; the Dirac delta is an even function.



Parseval's theorem

The equivalent of Parseval's theorem for Fourier transforms is easily proved using the Dirac delta-function. From Eq.(295),

$$g(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} f(x)\, e^{i\omega x}\, dx\ ,$$

$$g^*(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} f^*(y)\, e^{-i\omega y}\, dy\ , \tag{304}$$

where x is replaced by y in the second integral for clarity. Multiply the two expressions together and integrate over ω.

$$\int_{-\infty}^{\infty} g^*(\omega)\, g(\omega)\, d\omega = \frac{1}{2\pi} \int_{-\infty}^{+\infty} f^*(y)\, dy \int_{-\infty}^{+\infty} f(x)\, dx \int_{-\infty}^{+\infty} e^{i\omega(x-y)}\, d\omega\ . \tag{305}$$

But the last integral is just $2\pi\, \delta(y-x)$. The delta function removes the y integration and puts $y = x$ everywhere. The 2π factor knocks out the $1/2\pi$ factor outside and we are left with

$$\int_{-\infty}^{\infty} g^*(\omega)\, g(\omega)\, d\omega = \int_{-\infty}^{\infty} f^*(x)\, f(x)\, dx\ . \tag{306}$$

In words, the total intensity of a signal is equal to the total intensity of its Fourier transform. You should check this on the example given in class where

$$\mid E(t) \mid^2 = \mid E_0 \mid^2\ e^{-\gamma t} \quad \text{and} \quad \mid g(\omega) \mid^2 = \frac{E_0^2}{2\pi}\, \frac{1}{(\omega-\omega_0)^2 + \gamma^2/4}\ .$$