Orthogonal Function Solution of Differential Equations

Introduction

A given ordinary differential equation will have solutions in terms of "its own" functions. Thus, for example, the solution of the SHO Schrödinger equation is expressed in terms of Hermite polynomials multiplying a gausian:

$$\text{SHO Schrödinger equation} \quad \rightarrow \quad H_n(x)e^{-x^2/2}.$$

We are talking here about the *eigenfunctions* of the equation and their *eigenvalues* (*n* in this case).

Now we are going to look into using the *wrong* functions to solve a differential equation. Why might we want to do this? There are a number of reasons.

- We could use the set of $H_n(x)e^{-x^2/2}$ to see how much the behaviour of the system was "like" a quantum SHO.

- Similarly, we could use the set $e^{i(kx-\omega t)}$ to see how much the behaviour was "like" that of free traveling waves.

These procedures could be important in their own right, or as part of a perturbation or approximation scheme.

- The practical consequence of this procedure is that differential equations are transformed into *matrix* equations. (And of course this is at the heart of the equivalence of the Schrödinger and the Heisenberg approaches to quantum mechanics.

The fundamental ideas

The functions which arise from the solution of a differential equation "often" have some important properties. The equations which arise in physical problems can be assumed to comprise a *complete orthogonal set* of functions. (Formally, this result follows from the Sturm-Liouville theorem.)

Orthogonal

We know what orthogonal means. For the set of functions $\{\varphi_n(x)\}$ there is an orthogonality integral

$$\int \varphi_n(x)\varphi_m(x)\mathrm{d}x$$

(it might also have a weight function.).

The integral must be evaluated over the correct interval and it is zero unless $n = m$. If the functions are *normalised* then the integral is unity when $n = m$.

Complete

If the set of functions is complete, this means that *any* function $f(x)$ can be expressed as a linear sum of the functions

$$f(x) = \sum_n \varphi_n(x) f_n .$$

This expression for $f(x)$ will be valid over the orthogonality interval.

You should remember how to find the coefficients f_n:

$$f_n = \frac{\int f(x)\varphi_n(x)\mathrm{d}x}{\int \varphi_n^2(x)\mathrm{d}x} .$$

The denominator is needed when the functions are not normalised.

Differentiation

An important and highly relevant example of the completeness property is the differentiation of a function.

Starting from a function $f(x)$, if we differentiate it we obtain a new function $\mathrm{d}f/\mathrm{d}x$. Now "completeness" tells us that any function $f(x)$ can be expressed as a linear sum of the basis functions $\varphi_n(x)$. But since $\mathrm{d}f/\mathrm{d}x$ is just another function, this too can be expressed as a linear sum of the basis functions.

We have

$$f(x) = \sum_n \varphi_n(x) f_n$$

and we similarly expect that

$$\frac{\mathrm{d}f}{\mathrm{d}x} = \sum_n \varphi_n(x) g_n .$$

The question, then, is how to find the coefficients g_n for the derivative.

Given the basis set $\{\varphi_n(x)\}$ the function $f(x)$ is specified by the coefficients f_n. We want to find the coefficients g_n which correspondingly specify the derivative $\mathrm{d}f/\mathrm{d}x$. Let's proceed in the following way.

Start from $f(x)$:

$$f(x) = \sum_n \varphi_n(x) f_n$$

and differentiate it

$$\frac{\mathrm{d}f}{\mathrm{d}x} = \sum_n \frac{\mathrm{d}\varphi_n(x)}{\mathrm{d}x} f_n \; .$$

Completeness assures us that each of the $\mathrm{d}\varphi_n(x)/\mathrm{d}x$ will be a linear function of the basis functions.

For example, the derivative of Hermite polynomials is expressed as

$$\frac{\mathrm{d}}{\mathrm{d}x} H_n(x) = 2nH_{n-1}(x) \, .$$

In general we will have

$$\frac{\mathrm{d}}{\mathrm{d}x} \varphi_n(x) = \sum_m \varphi_m(x) D_{mn} \; .$$

The coefficients D_{mn} need two indices, one to indicate which basis function we are differentiating (n), and the other to indicate how much of each of the basis functions we need in the derivative (m).

For the Hermite polynomial case we then have

$$D_{mn} = 2n\delta_{n+1,n} \, .$$

Check that you understand this.

Using the above result we can return to the task of differentiating our function $f(x)$.

$$\begin{aligned}
\frac{\mathrm{d}f}{\mathrm{d}x} &= \sum_n \frac{\mathrm{d}\varphi_n(x)}{\mathrm{d}x} f_n \\
&= \sum_n \sum_m \varphi_m(x) D_{mn} f_n \\
&= \sum_m \varphi_m(x) \left(\sum_n D_{mn} f_n \right).
\end{aligned}$$

We framed our question by saying that if we write the derivative as the expansion

$$\frac{\mathrm{d}f}{\mathrm{d}x} = \sum_m \varphi_m(x) g_m$$

then what are the coefficients g_m? We now have the answer:

$$g_m = \sum_n D_{mn} f_n \; .$$

You should observe that this has the form of a matrix product. If we represent g_m and f_n as column matrices then

$$\begin{pmatrix} g_1 \\ g_2 \\ g_3 \\ \vdots \end{pmatrix} = \begin{pmatrix} D_{11} & D_{12} & D_{13} & \vdots \\ D_{21} & D_{22} & \vdots & \vdots \\ D_{31} & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots \end{pmatrix} \times \begin{pmatrix} f_1 \\ f_2 \\ f_3 \\ \vdots \end{pmatrix}.$$

Notation and 3d vector analogy

With 3d vectors, once one has a given basis set $\{\hat{\mathbf{i}}, \hat{\mathbf{j}}, \hat{\mathbf{k}}\}$ then any vector can be expressed as a linear sum of these

$$\mathbf{v} = v_i \hat{\mathbf{i}} + v_j \hat{\mathbf{j}} + v_k \hat{\mathbf{k}} \, .$$

Once the basis set is decided upon then any physical vector **v** is specified by giving its coefficients v_i, v_j, and v_k and these coefficients can be assembled into a column matrix

$$\begin{pmatrix} v_i \\ v_j \\ v_k \end{pmatrix} .$$

Mathematicians might call such a column of numbers a vector. We may even occasionally succumb to this inexactitude. Strictly speaking, this column of numbers is a particular *representation* of the physical vector **v** – if we change the basis set $\{\hat{\mathbf{i}}, \hat{\mathbf{j}}, \hat{\mathbf{k}}\}$ then we need a different column of coefficients to represent the same old vector **v**.

Now consider *functions*. With a given basis set $\{\varphi_n(x)\}$ any function $f(x)$ can be expressed as a linear sum of these

$$f(x) = \sum_n \varphi_n(x) f_n \, .$$

Once the basis set is decided upon then any function $f(x)$ is specified by giving its coefficients $\{f_n\}$ and these numbers can be assembled into a column matrix

$$\begin{pmatrix} f_1 \\ f_2 \\ f_3 \\ \vdots \end{pmatrix} .$$

This column of coefficients (a mathematician's vector) is a particular representation of the function $f(x)$. If we changed the basis set $\{\varphi_n(x)\}$ we were using then we would need a different column of coefficients to represent the same old function.

Taking the notational analogy further we could denote the column of coefficients by a "vector" symbol **f**. That is, we could make the association

$$f(x) \sim \mathbf{f} \, .$$

As an example of using this notational analogy consider the differentiation result of before.

$$g_m = \sum_n D_{mn} f_n$$

Here the f_n are the coefficients of the function $f(x)$ and the g_m are the coefficients of its derivative. If we denote these columns of coefficients by the vectors $\mathbf{f}$ and $\mathbf{f}'$ then the differentiation result becomes the vector-matrix equation

$$\mathbf{f}' = \mathbf{D}\,\mathbf{f}$$

where **D** represents the matrix whose elements are D_{mn}.

This shows us that multiplying with the matrix **D** has the effect of differentiation.

Matrix representation of differential equations

The above analogies suggest that differential equations may be represented by matrix equations. This may be done formally in the following way.

Starting from the simple example

$$g(x) = \frac{\mathrm{d}}{\mathrm{d}x} f(x),$$

in terms of the basis set $\{\varphi_n(x)\}$ this may be expressed as

$$\sum_n \varphi_n(x) g_n = \sum_m \varphi_m(x) \left(\sum_n D_{mn} f_n \right).$$

(Remember that the summation indices are completely arbitrary.) The "trick" to turn this into a matrix equation is to multiply both sides by one of the basis functions and then perform an orthogonality integral. Let's multiply by $\varphi_p(x)$ and integrate. On the left hand side as *n* varies through its range of values the orthogonality integral of $\varphi_n(x)$ with $\varphi_p(x)$ will always be zero except when $n = p$. Thus we pick out just the *p* term from the sum. On the right hand side, as *m* varies through its range we will similarly pick out the *p* term. We then end up with the equation

$$g_p = \sum_n D_{pn} f_n ,$$

or

$$\mathbf{g} = \mathbf{D}\,\mathbf{f}$$

which is the matrix representation of the original equation.

Differential equations with constant coefficients

Homogeneous equations

Let's consider a second order (homogeneous) differential equation with constant coefficients

$$A \frac{\mathrm{d}^2 f(x)}{\mathrm{d}x^2} + B \frac{\mathrm{d}f(x)}{\mathrm{d}x} + Cf(x) = 0 .$$

This is transformed into a matrix equation in four steps:

1 Express *f* as a linear sum of the basis functions:

$$f(x) = \sum_n \varphi_n(x) f_n .$$

2 Substitute this into the differential equation:

$$A\sum_n \frac{\mathrm{d}^2\varphi_n(x)}{\mathrm{d}x^2} f_n + B\sum_n \frac{\mathrm{d}\varphi_n(x)}{\mathrm{d}x} f_n + C\sum_n \varphi_n(x) f_n = 0 .$$

3 Multiply by $\varphi_p(x)$ and do the orthogonality integral:

$$A\sum_n \left\{\int \varphi_p(x)\frac{\mathrm{d}^2\varphi_n(x)}{\mathrm{d}x^2}\mathrm{d}x\right\} f_n + B\sum_n \left\{\int \varphi_p(x)\frac{\mathrm{d}^2\varphi_n(x)}{\mathrm{d}x^2}\mathrm{d}x\right\} f_n + C\sum_n \left\{\int \varphi_p(x)\varphi_n(x)\mathrm{d}x\right\} f_n = 0$$

4 Identify the matrix form of each term. By comparison with the derivation of the differentiation matrix **D** above, the differential equation may be written in matrix form as

$$A\mathbf{D}^2\mathbf{f} + B\mathbf{D}\mathbf{f} + \mathbf{1}\mathbf{f} = 0$$

or

$$\left\{A\mathbf{D}^2 + B\mathbf{D} + \mathbf{1}\right\}\mathbf{f} = 0$$

where **1** is the unit matrix.

Now $A\mathbf{D}^2 + B\mathbf{D} + \mathbf{1}$ is just another matrix, which we can denote by **M**. Thus our differential equation has been transformed, with the basis function set $\{\varphi_n(x)\}$, into the matrix equation

$$\mathbf{M}\,\mathbf{f} = 0 .$$

The homogeneous differential equation has been transformed into a homogeneous matrix equation. This would then be solved by the usual procedures for solving homogeneous matrix equations.

Inhomogeneous equations

Let us now consider an inhomogeneous equation such as

$$A\frac{\mathrm{d}^2 f(x)}{\mathrm{d}x^2} + B\frac{\mathrm{d}f(x)}{\mathrm{d}x} + Cf(x) = s(x)$$

where $s(x)$ is the *source* term, making the equation inhomogeneous. Now both f and s must be expressed as linear sums of the basis functions:

$$f(x) = \sum_n \varphi_n(x) f_n$$

$$s(x) = \sum_n \varphi_n(x) s_n .$$

We use the vector notation analogy to denote the column of coefficients of $f(x)$ by the vector symbol **f** and similarly the coefficients of $s(x)$ by **s**.

$$f(x) \sim \mathbf{f}$$

$$s(x) \sim \mathbf{s}.$$

Then, using the same procedure as in the previous case, we arrive at the inhomogeneous matrix equation

$$\mathbf{M}\,\mathbf{f} = \mathbf{s}\,.$$

The inhomogeneous differential equation has been transformed into an inhomogeneous matrix equation. This would then be solved by the usual procedures for solving inhomogeneous matrix equations. The formal solution is particularly straightforward: we multiply both equations, from the left, by the inverse of **M**. Then we have

$$\mathbf{M}^{-1}\mathbf{M}\,\mathbf{f} = \mathbf{M}^{-1}\mathbf{s}$$

or

$$\mathbf{f} = \mathbf{M}^{-1}\mathbf{s}\,.$$

This solves the problem; $f(x)$ is given in terms of $s(x)$.

Differential equations with variable coefficients

When the coefficients of the differential equation are variable, that is when we have an equation of the form

$$A(x)\frac{\mathrm{d}^2 f(x)}{\mathrm{d}x^2} + B(x)\frac{\mathrm{d}f(x)}{\mathrm{d}x} + C(x)f(x) = 0$$

things are slightly more complicated.

We proceed to transform this into a matrix equation using the same four steps as before:

1 Express f as a linear sum of the basis functions:

$$f(x) = \sum_n \varphi_n(x) f_n\,.$$

2 Substitute this into the differential equation:

$$A(x)\sum_n \frac{\mathrm{d}^2\varphi_n(x)}{\mathrm{d}x^2} f_n + B(x)\sum_n \frac{\mathrm{d}\varphi_n(x)}{\mathrm{d}x} f_n + C(x)\sum_n \varphi_n(x) f_n = 0\,.$$

3 Multiply by $\varphi_p(x)$ and do the orthogonality integral:

$$\sum_n \left\{\int \varphi_p(x)A(x)\frac{\mathrm{d}^2\varphi_n(x)}{\mathrm{d}x^2}\mathrm{d}x\right\} f_n + \sum_n \left\{\int \varphi_p(x)B(x)\frac{\mathrm{d}^2\varphi_n(x)}{\mathrm{d}x^2}\mathrm{d}x\right\} f_n +$$
$$+\sum_n \left\{\int \varphi_p(x)C(x)\varphi_n(x)\mathrm{d}x\right\} f_n = 0$$

4 Identify the matrix form of each term. In this case the we don't have a direct connection with the differentiation matrix **D** we had before. Nevertheless this still has the form of a matrix equation

$$\mathbf{M}\,\mathbf{f} = 0$$

where the elements of the matrix **M** are now.

$$M_{pn} = \int \varphi_p(x)A(x)\frac{\mathrm{d}^2\varphi_n(x)}{\mathrm{d}x^2}\mathrm{d}x + \int \varphi_p(x)B(x)\frac{\mathrm{d}^2\varphi_n(x)}{\mathrm{d}x^2}\mathrm{d}x + \int \varphi_p(x)C(x)\varphi_n(x)\mathrm{d}x$$