

# Orthogonal Polynomials

## 1 Introduction

Mathematically *orthogonal* means perpendicular, that is, *at right angles*. For example, the set of vectors  $\{x, y, z\}$  in three-dimensional space are orthogonal. The concept of orthogonality can be formalised, with respect to vectors, as

$$x \cdot y = 0; \quad x \cdot z = 0; \quad y \cdot z = 0,$$

that is, the dot product of two orthogonal vectors is zero. Indeed the set of vectors  $\{x, y, z\}$  is an *orthogonal set*. If  $\{x, y, z\}$  are all unit vectors we say that  $\{x, y, z\}$  is an *orthonormal set* of vectors. Orthogonal vectors are linearly independent and  $\{x, y, z\}$  therefore spans three-dimensional space. Thus any point in space may be written in terms of its components in the  $x$ ,  $y$  and  $z$  directions (or indeed in terms of components in the directions of any three orthogonal vectors). In this way  $\{x, y, z\}$  forms a **basis** for three-dimensional space.

### 1.1 Orthogonal Functions

Two functions,  $f_i(x)$  and  $f_j(x)$ , are orthogonal on an interval  $[a, b]$  if

$$\int_a^b f_i(x)f_j(x)dx = 0 \tag{1}$$

Mathematically we say that the *inner product* of the functions  $f_i(x)$  and  $f_j(x)$  is zero. The functions are *orthonormal* if

$$\int_a^b f_i(x)f_j(x)dx = \delta_{ij} \tag{2}$$

where

$$\begin{aligned} \delta_{ij} &= 1 & i = j \\ &= 0 & i \neq j \end{aligned}$$

is the Kronecker delta function.

## 2 Orthogonal Polynomials

A set of orthogonal polynomials is an infinite sequence of polynomials,  $p_0(x), p_1(x), p_2(x), \dots$ , where  $p_n(x)$  has degree  $n$  and any two polynomials in the set are orthogonal to each other, that is,

$$\int_a^b f_i(x)f_j(x)dx = 0 \quad \text{for } i \neq j \tag{3}$$

The set of polynomials is orthonormal if

$$\int_a^b f_i(x)f_j(x)dx = \delta_{ij} \tag{4}$$

The interval  $[a, b]$  is called the interval of orthogonality and may be infinite at one or both ends.

Any infinite sequence of polynomials  $\{p_n\}$ , with  $p_n$  having degree  $n$  forms a basis for the infinite-dimensional vector space of all polynomials. Such a sequence can be turned into an orthogonal basis using the **Gram-Schmidt** orthogonalisation process, by projecting out the components of each polynomial that are orthogonal to the polynomials already chosen.

## 2.1 Hilbert Space

Hilbert (David Hilbert 1862-1943) space generalises the idea of Euclidean space (that is, three-dimensional vector space etc) to infinite-dimensional spaces. Mathematically, a Hilbert space is an *inner product* space that is *complete*.

Hilbert spaces typically arise as infinite-dimensional *function spaces*. An element of a Hilbert space can be expressed by its coordinates with respect to an orthonormal basis of the space (directly analogous to  $\{x, y, z\}$  coordinates in three-dimensional vector space).

*Note:*

*Complete* means that as a sequence of functions approaches some limit, then the limit is also part of the Hilbert space. For example, if a molecular wave function is expanded in terms of orthogonal basis functions of a Hilbert space, then the molecular wavefunction is also an element of the Hilbert space.

The concept of Hilbert space offers one of the best mathematical formulations of quantum mechanics. Quantum mechanical states (wavefunctions) are described by vectors in a Hilbert space, quantum mechanical observables can be expressed by linear operators and quantum measurement is related to orthogonal projection.

## 2.2 Properties of Orthogonal Polynomials

All sets of orthogonal polynomials have a number of fascinating properties.

- Any polynomial  $f(x)$  of degree  $n$  can be expanded in terms of  $p_0, p_1, \dots, p_n$ , that is, there exist coefficients  $a_i$  such that  $f(x) = \sum_{i=0}^n a_i p_i(x)$ .
- Given an orthogonal set of polynomials  $\{p_0(x), p_1(x) \dots\}$ , each polynomial,  $p_k(x)$  is orthogonal to any polynomial of degree  $< k$ .
- Any orthogonal set of polynomials  $\{p_0(x), p_1(x) \dots\}$  has a recurrence formula that relates any three consecutive polynomials in the sequence, that is, the relation  $p_{n+1} = (a_n x + b_n)p_n - c_n p_{n-1}$  exists, where the coefficients  $a$ ,  $b$  and  $c$  depend on  $n$ . Such a recurrence formula is often used to generate higher order members in the set.
- Each polynomial in  $\{p_0(x), p_1(x) \dots\}$  has all  $n$  of its roots **real**, distinct and strictly within the interval of orthogonality (ie not on its ends). This is an extremely unusual property! It is particularly important when considering the classes of polynomials that arise as quantum mechanical solutions to a given Hamiltonian (or other Hermitian) operator - see below.
- Furthermore the roots of the  $n^{th}$  degree polynomial,  $p_n$  lie strictly inside the roots of the  $(n + 1)^{th}$  degree polynomial  $p_{n+1}$ .

## 2.3 Examples of Orthogonal Polynomials

The field of orthogonal polynomials was developed in the late 19th century and many of the sets of orthogonal polynomials described arose from descriptions of specific physical problems.

### Legendre Polynomials

These polynomials are orthogonal on the interval  $[-1, 1]$ . They arise as solutions to Legendre's differential equation (Adrien-Marie Legendre 1752-1833),

$$(1 - x^2) \frac{\partial^2 P(x)}{\partial x^2} - 2x \frac{\partial P(x)}{\partial x} + \lambda P(x) = 0 \quad \text{with } \lambda = n(n+1) \quad (5)$$

This equation is frequently encountered in Physics, particularly when spherical polar coordinates are used and where the problem has cylindrical symmetry. Legendre polynomials form part of the quantum mechanical solution for rotational motion.

$$\begin{aligned} P_0(x) &= 1 \\ P_1(x) &= x \\ P_2(x) &= \frac{1}{2}(3x^2 - 1) \\ P_3(x) &= \frac{1}{2}(5x^3 - 3x) \\ P_4(x) &= \frac{1}{8}(35x^4 - 30x^2 + 3) \\ P_5(x) &= \frac{1}{8}(63x^5 - 70x^3 + 15x) \\ P_6(x) &= \frac{1}{16}(231x^6 - 315x^4 + 105x^2 - 5) \dots \end{aligned}$$

### Laguerre Polynomials

The Laguerre polynomials,  $L_n$  are solutions to Laguerre's differential equation (Edmond Laguerre 1834-1886),

$$x \frac{\partial^2 L_n}{\partial x^2} + (1 - x) \frac{\partial L_n}{\partial x} + \lambda L_n = 0 \quad \text{with } \lambda = n \quad (6)$$

This equation arises in the radial part of the Hamiltonian for one-electron atoms.

$$\begin{aligned} L_0(x) &= 1 \\ L_1(x) &= -x + 1 \\ L_2(x) &= \frac{1}{2}(x^2 - 4x + 2) \\ L_3(x) &= \frac{1}{6}(-x^3 + 9x^2 - 18x + 6) \\ L_4(x) &= \frac{1}{24}(x^4 - 16x^3 + 72x^2 - 96x + 24) \\ L_5(x) &= \frac{1}{120}(-x^5 + 25x^4 - 200x^3 + 600x^2 - 600x + 120x) \\ L_6(x) &= \frac{1}{720}(x^6 - 36x^5 + 450x^4 - 2400x^3 + 5400x^2 - 4320x + 720) \dots \end{aligned}$$

The Laguerre polynomials are orthogonal over the interval  $[0, +\infty)$  with respect to the weight function  $e^{-x}$ . Indeed the radial component of the wavefunction of the one-electron atom is obtained by multiplying the Laguerre polynomials by  $e^{-x}$

### Hermite Polynomials

Hermite polynomials,  $H_n(x)$ , arise as part of the solution to the quantum harmonic oscillator Hamiltonian. They are solutions to Hermite's equation (Charles Hermite 1822-1901):

$$\frac{\partial H_n^2}{\partial^2 H_n} - 2xH_n + \lambda H_n = 0, \quad \text{with } \lambda = n \quad (7)$$

The Hermite polynomials are orthogonal over the interval  $(-\infty, +\infty)$  with respect to the weight function  $e^{-x^2/2}$ . Indeed the quantum harmonic oscillator wavefunctions are obtained by multiplying the Hermite polynomials by  $e^{-x^2/2}$ .

$$\begin{aligned} H_0(x) &= 1 \\ H_1(x) &= 2x \\ H_2(x) &= 4x^2 - 2x + 2 \\ H_3(x) &= 8x^3 - 12x \\ H_4(x) &= 16x^4 - 48x^2 + 12 \\ H_5(x) &= 32x^5 - 160x^3 + 120x \\ H_6(x) &= 64x^6 + 480x^4 + 720x^2 - 120 \dots \end{aligned}$$

## 3 Operators

Every observable physical quantity can be characterised by an **operator**.

An operator is a mathematical device that converts one function into another. It is one step up from a function, which is a device that converts one number (or collection of numbers) into another.

In general a caret or hat is used to denote an operator.

The simplest form of operator arises for physical observables that are just functions of position coordinates. These just multiply the functions on which they operate.

**Example:** The classical potential energy of a harmonic oscillator is  $\frac{1}{2}kx^2$ . The associated quantum mechanical operator  $\hat{V}$  just multiplies any function by  $\frac{1}{2}kx^2$ , that is

$$\hat{V}\Psi(x) = \frac{1}{2}kx^2\Psi(x)$$

A more interesting type of operator is a **differential operator**.

For instance, the operator  $\frac{d}{dx}$  changes the function  $\sin x$  to  $\cos x$ :

$$\frac{d}{dx} \sin x = \cos x$$

If a particle's position is described by coordinates  $(x, y, z)$ , then its momentum in the x direction is given by the operator

$$\hat{p}_x = -i\hbar \frac{\partial}{\partial x}$$

Where partial derivatives have been used because there are often several coordinates to consider. The form of the quantum operator for momentum,  $\hat{p}_x$  is unusual, it does not appear to have any correspondance with

classical momentum,  $mv$ . This form represents another of the basic axioms on which quantum mechanics has been developed.

If we can write some physical quantity in terms of position and momentum variables, then the rules for constructing operators are:

- All position variables remain unchanged
- A momentum in the direction of a coordinate  $q$  is replaced by the operator  $-i\hbar\frac{\partial}{\partial q}$ .

*Notes:*

- We almost always work with momentum in quantum mechanics rather than velocity.
- This second rule applies to angular coordinates as well as Cartesian ones. For example, rotation about the  $z$  axis is often described using a polar coordinate  $\phi$ . Angular momentum about the  $z$  axis is then described by the operator  $-i\hbar\frac{\partial}{\partial\phi}$ .

### 3.1 Expectation Values

In quantum mechanics every physical observable,  $Q$ , can be characterised by an operator,  $\hat{Q}$ , and the expectation value of  $\hat{Q}$  is calculated as

$$\langle Q \rangle = \frac{\int \Psi^* \hat{Q} \Psi d\tau}{\int \Psi^* \Psi d\tau} \quad (8)$$

Here  $\int \dots d\tau$  is a conventional notation meaning that we integrate over **all** the variables and over **all** space.

The order of the factors in the numerator is important: operators operate to the **right** only, so in this case  $\hat{Q}$  operates only on  $\Psi$ , not on  $\Psi^*$ . This does not matter for position operators, which operate by multiplication, but is crucial for differential operators.

### 3.2 Dirac Notation

The notation  $|\Psi\rangle$  is often used for the wavefunction  $\Psi$ . Often we abbreviate, writing  $|\Psi_n\rangle$  as  $|n\rangle$ .  $|\Psi_n\rangle$  or  $|n\rangle$  is called a **ket** vector.

$\langle\Psi|$  is called a **bra** vector. The bra notation implies the complex conjugate  $\Psi^*$ .

a complete *bra-ket* expression like  $\langle\Psi|\Psi\rangle$  or  $\langle\Psi|\hat{Q}|\Psi\rangle$  implies integration over all space.

Thus

$$\begin{aligned} \langle\Psi|\Psi\rangle &= \int \Psi^* \Psi d\tau \\ &= \int |\Psi|^2 d\tau \\ \langle\Psi|\hat{Q}|\Psi\rangle &= \int \Psi^* \hat{Q} \Psi d\tau \end{aligned}$$

and

$$\langle Q \rangle = \frac{\langle\Psi|\hat{Q}|\Psi\rangle}{\langle\Psi|\Psi\rangle}$$

### 3.3 Eigenfunctions and Eigenvalues

Suppose that we have an operator  $\hat{Q}$  and a wavefunction  $\Psi_q$  that satisfy the equation:

$$\hat{Q}\Psi_q = q\Psi_q \quad (9)$$

where  $q$  is just a number (with the appropriate dimensions). That is, if  $\Psi_q$  depends on some set of variables  $(x_1, x_2, \dots)$  then  $q$  is **independent** of all these variables.

Equation (9) is an **eigenvalue equation**. The wavefunction  $\Psi_q$  is an **eigenfunction** and  $q$  is the corresponding **eigenvalue**.

Much of the effort in practical quantum chemistry involves finding eigenfunctions and eigenvalues of interesting operators.

The operator for energy is called the **Hamiltonian** and its symbol is  $\mathcal{H}$ . Its eigenvalue equation is Schrödinger's time-independent equation:

$$\mathcal{H}\Psi = E\Psi \quad (10)$$

In general this equation has many solutions

$$\mathcal{H}\Psi_n = E_n\Psi_n, \quad n = 1, 2, 3, \dots$$

and we have a set of eigenfunctions  $\psi_n$  for Hamiltonian. In such cases the symbol  $\psi$  does not carry any information, and the label  $n$  is important, so it is common to use the Dirac notation  $|n\rangle$ , which gives the label more prominence.

A general property of the set of eigenfunctions of an operator like the Hamiltonian is that they are **orthogonal**;

$$\langle m|n\rangle \equiv \int_0^a \psi_m^* \psi_n dx = 0 \quad \text{if } m \neq n$$

Another important property is that the eigenfunctions form a **complete set**, provided that  $E_n \rightarrow \infty$  as  $n \rightarrow \infty$ . That is, any function of the same variables with the same boundary conditions can be expressed as a linear combination of the  $\psi_n$ :

$$\Psi = \sum_n c_n \psi_n$$

### 3.4 Hermiticity

This is an important technicality. Hermiticity is a generalisation of complex conjugation: under complex conjugation an operator  $\hat{Q}$  becomes its Hermitian conjugate,  $\hat{Q}^\dagger$ .

Consider the complex conjugation:

$$\begin{aligned} \langle \psi_1 | \hat{Q} | \psi_2 \rangle^* &= \left\{ \int \psi_1^* (\hat{Q} \psi_2) d\tau \right\}^* \\ &= \int \psi_1 (\hat{Q} \psi_2)^* d\tau \\ &= \langle \hat{Q} \psi_2 | \psi_1 \rangle \end{aligned}$$

However, we can consider the bra vector,  $\langle \hat{Q} \psi_2 |$  in terms of a new operator,  $\hat{Q}^\dagger$ . Where we **define**

$$\{\hat{Q} | \psi_2 \rangle\}^* = |\hat{Q} \psi_2 \rangle^* = \langle \hat{Q} \psi_2 | = \langle \psi_2 | \hat{Q}^\dagger$$

Therefore we can write

$$\langle \psi_1 | \hat{Q} | \psi_2 \rangle^* = \langle \hat{Q} \psi_2 | \psi_1 \rangle = \langle \psi_2 | \hat{Q}^\dagger | \psi_1 \rangle \quad (11)$$

Thus every operator  $\hat{Q}$  has a **Hermitian conjugate**,  $\hat{Q}^\dagger$ , which can be defined using eq. (11), or its integral equivalent:

$$\int \psi_1^* \hat{Q} \psi_2 d\tau = \int (\hat{Q}^\dagger \psi_1)^* \psi_2 d\tau$$

An operator that is equal to its Hermitian conjugate is said to be **Hermitian**.

**Example:** Find the Hermitian conjugates of the operators  $\frac{d}{dx}$  and  $i\frac{d}{dx}$

$$\begin{aligned} \int_{-\infty}^{\infty} \psi_1^* \frac{d\psi_2}{dx} dx &= [\psi_1^* \psi_2]_{-\infty}^{\infty} - \int_{-\infty}^{\infty} \frac{d\psi_1^*}{dx} \psi_2 dx \\ &= \int_{-\infty}^{\infty} \left(-\frac{d\psi_1}{dx}\right)^* \psi_2 dx \end{aligned}$$

$$\text{therefore} \quad \left(\frac{d}{dx}\right)^\dagger = -\frac{d}{dx}$$

$$\begin{aligned} \int_{-\infty}^{\infty} \psi_1^* i \frac{d\psi_2}{dx} dx &= [\psi_1^* i \psi_2]_{-\infty}^{\infty} + \int_{-\infty}^{\infty} \left(-i \frac{d\psi_1^*}{dx}\right) \psi_2 dx \\ &= \int_{-\infty}^{\infty} \left(i \frac{d\psi_1}{dx}\right)^* \psi_2 dx \end{aligned}$$

$$\text{therefore} \quad \left(i \frac{d}{dx}\right)^\dagger = i \frac{d}{dx}$$

Thus  $i\frac{d}{dx}$  is Hermitian (cf the momentum operator) but  $\frac{d}{dx}$  is not.

### 3.5 Properties of Hermitian operators:

- Their eigenvalues are always real
- Eigenfunctions corresponding to different eigenvalues are orthogonal

**Proof:** Show the eigenfunctions of a Hermitian operator are real and that eigenfunctions with different eigenvalues are orthogonal

Suppose  $\psi_1$  and  $\psi_2$  are eigenfunctions of an Hermitian operator  $\hat{B}$  with eigenvalues  $b_1$  and  $b_2$ . Then, using Dirac notation:

$$\hat{B}|\psi_1\rangle = b_1|\psi_1\rangle \quad (a)$$

$$\hat{B}|\psi_2\rangle = b_2|\psi_2\rangle \quad (b)$$

Premultiply eq. (a) by  $\psi_2^*$  and integrate over all space (this is a standard *trick* and is used in most proofs in quantum mechanics):

$$\langle \psi_2 | \hat{B} | \psi_1 \rangle = b_1 \langle \psi_2 | \psi_1 \rangle$$

Now take the complex conjugate:

$$\langle \psi_2 | \hat{B} | \psi_1 \rangle^* = b_1^* \langle \psi_2 | \psi_1 \rangle^*$$

Now use Hermiticity:

$$\langle \psi_1 | \hat{B} | \psi_2 \rangle = b_1^* \langle \psi_1 | \psi_2 \rangle$$

Premultiply eq. (b) by  $\psi_1^*$  and integrate over all space:

$$\langle \psi_1 | \hat{B} | \psi_2 \rangle = b_2 \langle \psi_1 | \psi_2 \rangle$$

Combining the last two results:

$$(b_1^* - b_2) \langle \psi_1 | \psi_2 \rangle = 0$$

Now, if we had started with  $\psi_2 = \psi_1$  and  $b_2 = b_1$  then we would have

$$(b_1^* - b_1) \langle \psi_1 | \psi_1 \rangle = 0$$

But the integral of  $|\psi_1|^2$  over all space is non-zero, (the basic postulate of quantum mechanics that  $|\psi|^2$  is proportional to probability) so for the equation to be true we must have  $b_1^* = b_1$ , that is, the eigenvalues must be real.

Alternately, if we had started with  $b_1 \neq b_2$  then we must conclude that the integral  $\langle \psi_1 | \psi_2 \rangle = 0$ . Hence the eigenfunctions of a Hermitian operator with different eigenvalues are orthogonal. This is one side of an *if and only if* proof, it can also be shown that if an operator has all its eigenvalues real and its eigenfunctions orthogonal it must be Hermitian.

Other useful properties of Hermitian operators are, if  $\hat{A}$  and  $\hat{B}$  are Hermitian:

- If  $\hat{A}$  is invertible so is  $\hat{A}^\dagger$  and  $(\hat{A}^\dagger)^{-1} = (\hat{A}^{-1})^\dagger$
- $(\hat{A} + \hat{B})^\dagger = \hat{A}^\dagger + \hat{B}^\dagger$
- $(\lambda \hat{A})^\dagger = \lambda^* \hat{A}^\dagger$  where  $\lambda^*$  is the complex conjugate of the scalar  $\lambda$
- $(\hat{A}\hat{B})^\dagger = \hat{B}^\dagger \hat{A}^\dagger$

All observable (measurable) properties of a system must be real and hence the operators associated with observable properties, like the Hamiltonian, momentum, dipole moment etc, are Hermitian. Thus, when we deal with such operators, we can make use of all the useful properties of Hermitian operators.

Another property of operators is that, if two observables are to have simultaneously precisely defined values, their corresponding operators must commute. If you would like more details on this please let me know (if nothing else it will tell me if anyone has actually read this far!!)

This (non-examinable) description of some of the mathematics behind quantum mechanics touches on the 20th century advances in mathematical theory. (Most of the maths we use is 19th century or earlier). The most elegant derivation of quantum mechanics (to me anyway) comes from the recognition that the quantum mechanical *commutator*, identified as the *Lie bracket*, generates a *Lie Algebra*. The eigenvalues of the operators are the *roots* of the Lie Algebra. As such the recurrence relations between the various eigenfunctions (and step-up and step-down operators) arise as the operators that step from one root of the Lie algebra to the next.