CHEM2024 - Week 24 Lecture 2 - Algebraic foundations of quantum theory II

Chapter 1 of Atkins, "Molecular Quantum Mechanics", 5th edition.

1. Linear operators

A *linear operator* is defined as a map between two spaces that creates a correspondence between their elements in a way that preserves addition and multiplication by a scalar. If $\mathcal{V}$ and $\mathcal{W}$ are two spaces over the same field $\mathbb{F}$, then an operator $\mathbf{A}:\mathcal{V}\rightarrow\mathcal{W}$ is called linear if:

1. The sum of any two elements of $\mathcal{V}$ is mapped into the sum of their images in $\mathcal{W}$:

$$\forall\vec{a},\vec{b}\in\mathcal{V}\qquad\mathbf{A}\left(\vec{a}+\vec{b}\right)=\mathbf{A}\vec{a}+\mathbf{A}\vec{b}$$

2. Multiplying an element of $\mathcal{V}$ by a scalar also multiplies its image in $\mathcal{W}$ by the same scalar:

$$\forall\vec{a}\in\mathcal{V}\;\;\forall\alpha\in\mathbb{F}\qquad\mathbf{A}\left(\alpha\vec{a}\right)=\alpha\mathbf{A}\vec{a}$$

This definition is the same for vectors and functions.

Example 1: multiplication of a vector by a matrix is a linear operation. In particular, rotation matrices are linear operators because, for any rotation matrix $\mathbf{R}$ and any two vectors $\vec{u}$ and $\vec{v}$:

$$\mathbf{R}\left(\alpha\vec{u}+\beta\vec{v}\right)=\alpha\mathbf{R}\vec{u}+\beta\mathbf{R}\vec{v}$$

where α and β are arbitrary complex scalars.

Example 2: the second derivative operator is linear, because

$$\frac{\partial^2}{\partial x^2}\left[\alpha f\left(x\right)+\beta g\left(x\right)\right]=\alpha\frac{\partial^2}{\partial x^2}f\left(x\right)+\beta\frac{\partial^2}{\partial x^2}g\left(x\right)\qquad\forall f,g,\alpha,\beta$$

2. Matrix-vector representations

Because the expansion of $f\left(x\right)$ in an orthonormal basis $\left\{g_n\left(x\right)\right\}$ is unique, the function may be identified with the corresponding coefficients that may be written out as elements of a column vector:

$$f\left(x\right)=\sum_{n=1}^{\infty}a_n g_n\left(x\right)\qquad\Leftrightarrow\qquad\vec{f}=\begin{pmatrix}a_1\\a_2\\\vdots\end{pmatrix}\tag{1}$$

Such a vector is called a *representation* of the function $f\left(x\right)$ in the basis set $\left\{g_n\left(x\right)\right\}$.

Example 3: find a vector representation of $1+\cos x$ in the plane wave basis $g_n\left(x\right)=e^{-inx}/\sqrt{2\pi}$ where n is an integer and $x\in\left[0,2\pi\right]$.

Solution: it may be verified by direct inspection that the basis is orthonormal

$$\left\langle g_n\left(x\right)\middle|g_k\left(x\right)\right\rangle=\int_0^{2\pi}g_n^*\left(x\right)g_k\left(x\right)dx=\delta_{nk}$$

Expansion coefficients are obtained by computing scalar products:

$$a_n=\left\langle g_n\left(x\right)\middle|1+\cos x\right\rangle=\frac{1}{\sqrt{2\pi}}\int_0^{2\pi}e^{-inx}\left(1+\frac{e^{-ix}+e^{+ix}}{2}\right)dx=\begin{cases}\sqrt{\pi/2} & n=\pm1\\ \sqrt{2\pi} & n=0\end{cases}$$

All other scalar products are zero. Thus, the representation is an infinite vector with three non-zero elements: $a_0 = \sqrt{2\pi}$ and $a_{\pm 1} = \sqrt{\pi/2}$.

The same technique may be applied to operators. For example, consider the eigensystem problem for some linear differential operator $\hat{H}$:

$$\hat{H}\psi(x) = E\psi(x) \tag{2}$$

Depending on the exact form of $\hat{H}$, such equations may be exceedingly hard to solve analytically. However, if we pick an orthonormal basis set $\{g_n(x)\}$ and expand the function:

$$\psi(x) = \sum_k a_k g_k(x) \tag{3}$$

then Equation (2) acquires the following form:

$$\sum_k a_k \hat{H} g_k(x) = E\sum_k a_k g_k(x) \tag{4}$$

After taking the scalar product of both sides of this equation with $g_m(x)$, we obtain the following system:

$$\begin{gathered} {\scriptstyle m}\left\{\sum_k a_k \left\langle g_m(x)\middle|\hat{H}\middle|g_k(x)\right\rangle = E\sum_k a_k \left\langle g_m(x)\middle|g_k(x)\right\rangle\right. \\ \left\langle g_m(x)\middle|\hat{H}\middle|g_k(x)\right\rangle = \int g_m^*(x)\hat{H}g_k(x)\,dx \end{gathered} \tag{5}$$

The right hand side simplifies because the set $\{g_n(x)\}$ is orthonormal:

$$\sum_k a_k \left\langle g_m(x)\middle|g_k(x)\right\rangle = \sum_k a_k \delta_{mk} = a_m \tag{6}$$

and on the left hand side, each term $\left\langle g_m(x)\middle|\hat{H}\middle|g_k(x)\right\rangle$ is an integral, but ultimately just a number that we will call H_{mk}. This dramatically simplifies the system:

$${\scriptstyle m}\left\{\sum_k H_{mk} a_k = E a_m\right. \tag{7}$$

We can now recognise the matrix-vector product on the left hand side. The equation therefore simplifies further into a matrix eigensystem problem that a computer can easily solve:

$$\mathbf{H}\vec{a} = E\vec{a} \tag{8}$$

where $\vec{a}$ is a vector that represents the function $\psi(x)$ in the basis $\{g_n(x)\}$, and $\mathbf{H}$ is a matrix with elements $H_{mk} = \left\langle g_m(x)\middle|\hat{H}\middle|g_k(x)\right\rangle$. This *matrix representation* is equivalent to the differential equation formalism, but it is much easier to handle on digital computers.

4. Mathematical framework of non-relativistic quantum mechanics

The following statements may be viewed as a summary of one-dimensional non-relativistic quantum mechanics from the mathematical perspective:

1. Any quantum system can be described by a complex-valued function, called *wavefunction*, of coordinates and time. Such functions populate a metric space, called the *state space* of the system. The norm and the scalar product on this space are defined in the usual way:

$$\left\|\psi(x)\right\| = \sqrt{\int \psi^*(x)\psi(x)\,dx} \qquad \left\langle \psi(x)\middle|\varphi(x)\right\rangle = \int \psi^*(x)\varphi(x)\,dx \tag{9}$$

where the star denotes complex conjugation and the integration is carried out over all coordinates of the system. Wavefunctions of physical systems are continuous and have an infinite number of continuous derivatives.

2. The *expectation value* (defined as the average outcome of many repeated experiments) of an observable property $O(t)$ of a quantum system with a wavefunction $\psi(x,t)$ is related to the wavefunction in the following way:

$$O(t) = \langle \psi(x,t) | \hat{O} | \psi(x,t) \rangle = \int \psi^*(x,t) \hat{O} \psi(x,t) dx \tag{10}$$

where $\hat{O}$ is a linear multiplicative or differential operator. Simple commonly occurring operators are coordinate $\hat{x}$, momentum $\hat{p}$, and kinetic energy $\hat{T}$:

$$\hat{x} = x \qquad \hat{p} = -i\frac{\partial}{\partial x} \qquad \hat{T} = \frac{1}{2m}\frac{\partial^2}{\partial x^2} \tag{11}$$

It may be shown that eigenvalues of such operators are real and their eigenfunctions belonging to different eigenvalues are orthogonal.

3. The wavefunctions obey the *time-dependent Schrödinger equation*:

$$\frac{\partial}{\partial t}|\psi(x,t)\rangle = -i\hat{H}(x,t)|\psi(x,t)\rangle \tag{12}$$

where $\hat{H}(x,t)$ is the operator that corresponds to the energy of the system. By analogy with classical mechanics it is called *Hamiltonian*. Its eigenvalues are known in chemistry as *energy levels* and its eigenfunctions, in the case of molecular quantum systems, are called *orbitals*. A plausible derivation for the Schrödinger equation may be built on the philosophical principle of causality and the theory of Lie groups. A rigorous derivation does not appear to exist.

4. The absolute square of the wavefunction $|\psi(x,t)|^2$ is the *probability density* of the system being found at the coordinate x at time t. The probability theory requirement for the sum total of all probabilities to be equal to 1

$$\int |\psi(x,t)|^2 dx = 1$$

then yields the *normalization condition* for the wavefunction:

$$\|\psi(x,t)\| = 1$$

5. A measurement of an observable property O performed on a system in a state described by an eigenfunction $|\varphi_k\rangle$ of the corresponding operator $\hat{O}$

$$\hat{O}|\varphi_k\rangle = \lambda_k |\varphi_k\rangle$$

yields the corresponding eigenvalue λ_k with the probability of 1. A measurement of O performed on a system that is described by a linear combination of eigenfunctions of $\hat{O}$:

$$|\psi\rangle = a_1|\varphi_1\rangle + a_2|\varphi_2\rangle + a_3|\varphi_3\rangle + \ldots$$

yields λ_1 with probability $|a_1|^2$, λ_2 with probability $|a_2|^2$ and so on. The normalization condition for linear combinations of wavefunctions is therefore $|a_1|^2 + |a_2|^2 + |a_3|^2 + \ldots = 1$.

5. Interpretations of quantum mechanics

Almost anything that one could put in writing about what quantum mechanics "actually means" is arguable. The primary problem appears to be in the human concept of *meaning* – recognition, valuation, emotional association and incorporation into the knowledge framework that a person already holds. Quantum mechanics does not appear to "mean" anything recognisable in the terms that the human brain had evolved to use. Scientists themselves interpret quantum theory using its mathematical properties, which are well understood and currently used with great effect in research and engineering.

This does not prevent "philosophers of physics" from making attempts at interpreting quantum theory, but few of their propositions are not dismissed by the actual physics community as ludicrous. Some lingering misconceptions had been created by the physicists themselves: *e.g.* the "measurement paradox" is not a paradox at all, but a collision of approximations between quantum and classical physics. For the purposes of this course we shall subscribe to the *instrumentalist interpretation* of quantum mechanics that is based on David Mermin's famous *dictum*: "shut up and calculate".