

1 Introduction

1.1 What this course is about

In brief, this course is a survey of the mathematical apparatus of Quantum Mechanics. More specifically, a survey of its general mathematical underpinning (the mathematical theory of Hilbert spaces) and of its formalism (bra and ket vectors, operators, matrix representation, unitary transformations, the Schrödinger and the Heisenberg pictures, wave mechanics, position and momentum spaces, creation and annihilation operators, with a special emphasis on the theory of angular momentum). The course also examines how this mathematical apparatus relates to what can be observed (i.e., how quantum states and physical observables are represented in the theory, and how predictions on the outcomes of experiments can be extracted from the formalism).

1.2 Revision: The hydrogen atom and the Stern-Gerlach experiment

As an introduction to the theoretical concepts addressed in this course, we start by revisiting two examples of quantum systems you have studied in Term 1.

Example 1: The hydrogen atom

Hydrogen molecules may be dissociated into individual hydrogen atoms by an electric discharge or some other interaction. The resulting atoms may be left in an excited state by this process, in which case they will decay to a state of lower energy sooner or later, usually by emitting a photon. Analysing the light emitted by an ensemble of such atoms reveals that the energy of these photons is narrowly distributed about discrete values. (An example of the resulting line spectrum was shown in the first lecture of the course and can be found in the Lecture Outlines folder.)

This experimental fact can be largely explained by calculations based on the following Hamiltonian:

$$H = -\frac{\hbar^2}{2\mu} \nabla^2 - \frac{e^2}{4\pi\epsilon_0} \frac{1}{r}, \quad (1.1)$$

where μ is the reduced mass of the electron - nucleus system, e is the charge of the electron in absolute value ($e > 0$), r is its distance to the nucleus, ϵ_0 is the vacuum permittivity, \hbar is the reduced Planck constant, and ∇^2 is the square of the gradient operator with respect to the coordinates of the electron. The latter

can be taken to be x , y and z , in which case

$$\nabla^2 \equiv \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}. \quad (1.2)$$

Alternatively, one can refer the position of the electron by the spherical polar coordinates r , θ and ϕ , in which case ∇^2 is a more complicated combination of first and second order partial derivatives.

As defined, H is an operator which acts on functions and transforms a function of x , y and z (or a function of r , θ , ϕ) into another function of the same variables. H is actually a combination of several operators: ∇^2 is an operator which, when acting on a function $\phi(x, y, z)$, transforms this function into the sum of its second order derivatives with respect to x , y and z . The $1/r$ term in the Hamiltonian is also an operator, which simply transforms $\phi(x, y, z)$ into the product of $\phi(x, y, z)$ with the potential energy. In particular, H acts on wave functions $\psi(r, \theta, \phi)$ representing bound states of the electron – nucleus system, i.e., states in which there is a vanishingly small probability that the electron is arbitrarily far from the nucleus. These wave functions can be normalized in such a way that

$$\int_0^\infty dr r^2 \int_0^\pi d\theta \sin \theta \int_0^{2\pi} d\phi |\psi(r, \theta, \phi)|^2 = 1. \quad (1.3)$$

As you have seen in the Term 1 Quantum Mechanics course, this Hamiltonian has discrete eigenvalues E_n ($n = 1, 2, \dots$) corresponding to bound energy eigenstates (the eigenvalues of the Hamiltonian are often called eigenenergies). I.e., in the notation used in that course, there exist bound state wave functions $\psi_{nlm}(r, \theta, \phi)$ such that the function obtained by letting H acts on $\psi_{nlm}(r, \theta, \phi)$ is simply $\psi_{nlm}(r, \theta, \phi)$ multiplied by a constant E_n :

$$H\psi_{nlm}(r, \theta, \phi) = E_n\psi_{nlm}(r, \theta, \phi). \quad (1.4)$$

As their name of eigenenergies indicates, the eigenvalues E_n have the physical dimensions of an energy. Calculations show that $E_n \propto -1/n^2$, $n = 1, 2, \dots$ for the above Hamiltonian. These eigenenergies thus form a discrete distribution of infinitely many energy levels (discrete meaning that these energy levels are separated by a gap from each other).¹

The eigenvalues of the Hamiltonian of Eq. (1.1) (the eigenenergies E_n) are related in a very simple way to the energy of the photons emitted by excited

¹Besides being eigenfunctions of H , the energy eigenfunctions $\psi_{nlm}(r, \theta, \phi)$, as defined, are also eigenfunctions of the angular momentum operators \mathbf{L}^2 and L_z . The quantum numbers l and m identify the corresponding eigenvalues. In the case of the Hamiltonian defined by Eq. (1.1), energy eigenfunctions with different values of l or m but the same value of the principal quantum number n correspond to the same eigenenergy E_n .

hydrogen atoms: each photon is emitted in a transition from an energy eigenstates to another one, and the photon energy is almost equal to the difference between the respective eigenenergies. The photon energy is not exactly equal to this difference for a variety of reasons — e.g., because describing the atom by way of this Hamiltonian amounts to neglecting spin-orbit coupling and other relativistic effects. However, what is important here is that there is a very close relation between the eigenvalues of the Hamiltonian and the results which would be found in an actual measurement of the energy of the photons.

Example 2: The Stern-Gerlach experiment

We refer here to an experiment of historical importance done by the two physicists whose names have remained associated with this type of measurements ever since, Otto Stern and Walther Gerlach, in the early 1920s. The aim of Stern and Gerlach was to test the predictions of the Bohr model of atomic structure in regards to the magnetic moment of atoms (modern Quantum Mechanics had not yet been developed). The principle of the experiment was simple: since particles with a magnetic moment $\boldsymbol{\mu}$ in a magnetic field \mathbf{B} experience a force \mathbf{F} equal to the gradient of $\boldsymbol{\mu} \cdot \mathbf{B}$, the magnetic moment of an atom can be inferred from how its trajectory is deflected when it passes through an inhomogeneous magnetic field. Stern and Gerlach directed a beam of silver atoms through a specially designed magnet producing a magnetic field $\mathbf{B} = B_z \hat{\mathbf{z}}$ such that $\nabla B_z \neq 0$. ($\hat{\mathbf{z}}$ is a unit vector in the z -direction.) Each of these atoms was therefore submitted to a force equal to $\mu_z \nabla B_z$, with μ_z the z -component of its magnetic moment. They observed that this beam was split into two when passing through this magnet, from which they concluded that only two values of μ_z were possible for these atoms. (The splitting is very visible in the image of the spatial distribution of the atoms shown in the first lecture of the course. A copy of this image can be found in the Lecture Outlines folder.)

At the time of this experiment Stern and Gerlach would not have been able to formulate these results in terms of modern concepts of Quantum Mechanics. However, we now know (1) that the magnetic moment of a silver atom originates almost entirely from the magnetic moment of the electrons it contains (the magnetic moment of the nucleus is comparatively much smaller and can be ignored in good approximation); (2) that for silver atoms in their ground state the contribution of the electrons to the total magnetic moment can be written in terms of a spin operator; and (3) that the two values of μ_z found in the experiment correspond to two different eigenvalues of this spin operator.

More specifically, one can say that observing whether an atom is deflected in one direction or the other amounts to a measurement of its spin in the z -direction. We can represent the spin state of the atoms deflected in one direction by a column vector χ_+ and the spin state of the atoms deflected in the other direction

by a column vector χ_- , and set

$$\chi_+ = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \text{and} \quad \chi_- = \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \quad (1.5)$$

As can be easily checked, these two column vectors are eigenvectors of the matrix

$$S_z = \begin{pmatrix} \hbar/2 & 0 \\ 0 & -\hbar/2 \end{pmatrix}. \quad (1.6)$$

In this formulation of the problem, this matrix is the spin operator mentioned in the previous paragraph. (A 2×2 matrix is an operator which transforms 2-component column vectors into 2-component column vectors since multiplying a 2-component column vector by a 2×2 matrix gives a 2-component column vector as a result. See Appendix A of these course notes, in a separate document, for a reminder of the rules of matrix multiplication.) The matrix S_z has two eigenvalues, $\hbar/2$ and $-\hbar/2$. Accordingly, in the experiment of Stern and Gerlach, the only possible values of μ_z were $\gamma\hbar/2$ and $-\gamma\hbar/2$, where γ is a certain constant whose value is not important for this discussion.

Before it enters the magnet, an individual atom could be in the spin state represented by the column vector χ_+ or in the spin state represented by the column vector χ_- . More generally, it could also be in a superposition state represented by a column vector of the form $c_+\chi_+ + c_-\chi_-$, where c_+ and c_- are two complex numbers such that

$$|c_+|^2 + |c_-|^2 = 1. \quad (1.7)$$

In the latter case, there would be a probability $|c_+|^2$ that it would be found to have $\mu_z = \gamma\hbar/2$ and a probability $|c_-|^2$ that it would be found to have $\mu_z = -\gamma\hbar/2$: only $\pm\gamma\hbar/2$ could be found for μ_z , even if the atom is initially in a superposition state. Supposing that neither $|c_+|^2$ nor $|c_-|^2$ is zero, then one cannot predict the value of μ_z which would be found for that individual atom. Only the probability of each of the two possible values of μ_z can be predicted.

Important observations

The quantum systems considered in these two examples are obviously quite different, both physically and mathematically — e.g., the mathematical objects representing the state of the system are functions of position co-ordinates in the first example and column vectors of numbers in the second example. Nonetheless, they are similar in regards to key aspects of their theoretical description:

- In both cases, the physical quantities measured in an experiment correspond to mathematical operators [the Hamiltonian of Eq. (1.1) in the first example, the spin matrix of Eq. (1.6) in the second example].

- The values these physical quantities could be found to have are given by the eigenvalues of the corresponding operator.
- Each of these values has a certain probability to be found, and these probabilities can be calculated from the wave function of column vector representing the state of the system.
- Whether a wave function or a column vector, the mathematical object representing the state of the system encapsulates all what can be known about the results of future measurements.

This relationship between measurements, operators and mathematical objects used to represent quantum states is actually very general and applies to any quantum system.

☞ Quantum systems can often be mathematically described in several different but equivalent ways. For instance, when discussing the Stern-Gerlach experiment, we have set the column vectors χ_+ and χ_- and the matrix S_z to be as given by Eqs. (1.5) and (1.6). However, we could equally well have decided to represent the two spin states by the column vectors

$$\chi'_+ = \begin{pmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix} \quad \text{and} \quad \chi'_- = \begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix} \quad (1.8)$$

and the relevant spin operator by the matrix

$$S'_z = \begin{pmatrix} 0 & -\hbar/2 \\ -\hbar/2 & 0 \end{pmatrix}. \quad (1.9)$$

As can be checked without difficulty, χ'_+ and χ'_- are eigenvectors of S'_z corresponding, respectively, to the eigenvalues $\hbar/2$ and $-\hbar/2$, exactly as χ_+ and χ_- are eigenvectors of S_z corresponding to the same eigenvalues. Formulating the problem as per Eqs. (1.8) and (1.9) instead of (1.5) and (1.6) changes the details of the mathematics involved in the calculations. However, these two formulations are equivalent from a Physics point of view.

As another example, consider the ground state wave function of a linear harmonic oscillator, $\psi_0(x)$. You have seen in the Term 1 Quantum Mechanics course that

$$\psi_0(x) = \left(\frac{m\omega}{\pi\hbar}\right)^{1/4} \exp(-m\omega x^2/2\hbar), \quad (1.10)$$

where m is the mass of the oscillator and ω its angular frequency. Clearly, this function is continuous and is such that the integral

$$\int_{-\infty}^{\infty} |\psi_0(x)| dx$$

exists. You may remember from one of your maths courses that these two mathematical facts guarantee that $\psi_0(x)$ can be written in the form of a Fourier integral. I.e., there exists a function $\phi_0(p)$ such that

$$\psi_0(x) = \frac{1}{(2\pi\hbar)^{1/2}} \int_{-\infty}^{\infty} \phi_0(p) \exp(ipx/\hbar) dp. \quad (1.11)$$

In fact, the function $\phi_0(p)$ can be obtained by taking the inverse transformation:

$$\phi_0(p) = \frac{1}{(2\pi\hbar)^{1/2}} \int_{-\infty}^{\infty} \psi_0(x) \exp(-ipx/\hbar) dx. \quad (1.12)$$

Thus knowing $\phi_0(p)$ is knowing $\psi_0(x)$ and knowing $\psi_0(x)$ is knowing $\phi_0(p)$. In other words, the ground state is represented by the function $\phi_0(p)$ as well as by the function $\psi_0(x)$: it is possible to use $\phi_0(p)$ rather than $\psi_0(x)$ if this would be convenient in some calculations, and the two formulations are completely equivalent from a Physics point of view. This topic will be discussed further in the course. [We just note at this stage that $\psi_0(x)$ is the “ground state wave function in position space” whereas $\phi_0(p)$ is the “ground state wave function in momentum space”.]

1.3 Quantum states and vector spaces

Of central importance in Quantum Mechanics is the mathematical theory of vector spaces, particularly of a kind of vector spaces called Hilbert spaces. The reason is that states of quantum systems correspond to vectors belonging to certain Hilbert spaces (e.g., wave functions, column vectors of numbers, or abstract vectors called ket vectors), and that physical quantities like the energy or the angular momentum correspond to operators acting on these vectors.