

Quantum Mechanics

L. Del Debbio, University of Edinburgh
Version 0.9

November 26, 2010

Contents

1	Quantum states	1
1.1	Introduction	2
1.1.1	Experiment with classical particles	2
1.1.2	Experiment with waves	3
1.1.3	Experiment with electrons	4
1.2	One-dimensional systems	5
1.2.1	Quantum states	5
1.2.2	Superposition principle	8
1.2.3	The uncertainty principle	9
1.2.4	Dynamical variables	10
1.3	Summary	11
2	Observables	13
2.1	Observing observables	14
2.2	Hermitean operators	16
2.2.1	Hermitean conjugate	16
2.2.2	Hermitean operators	17
2.3	Properties of Hermitean operators	18
2.4	Commutators	19
2.5	Compatible Observables	21
2.5.1	The Compatibility Theorem	21
2.6	Summary	22
3	Dynamics-1	23
3.1	Schrödinger equation	24
3.1.1	Equation of motion	24
3.1.2	Eigenstates of the Hamiltonian	24
3.1.3	General restrictions on $\Psi(x, t)$	26
3.1.4	Example: the quantum harmonic oscillator	26

3.1.5	Expansion in stationary states	30
3.2	Summary	30
4	Dynamics-2	33
4.1	Continuous spectrum	34
4.1.1	Eigenvalue equation	34
4.1.2	Orthonormality	34
4.1.3	Example: the free case	35
4.1.4	Spectral decomposition	39
4.2	General properties	39
4.2.1	Energy expectation value	39
4.2.2	Ehrenfest's theorem	40
4.2.3	Degeneracy and oscillations	40
4.3	Heisenberg representation	41
4.4	Summary	42
A	Postulates	43
5	Potentials	47
5.1	Potentials	48
5.1.1	Potential step	48
5.1.2	Tunneling	51
5.2	Summary	53
6	More potentials	55
6.1	Introduction	56
6.2	Infinite potential well	56
6.3	Parity	59
6.4	Finite potential well	59
6.5	Summary	61
7	More dimensions	63
7.1	Introduction	64
7.2	Quantum states	64
7.3	Observables	65
7.4	Dynamics	67
7.5	Degeneracy	68
7.6	Summary	69

8 Angular momentum	71
8.1 Introduction	72
8.2 Angular momentum operator	72
8.3 Square of the angular momentum	75
8.4 Eigenfunctions	76
8.5 Physical interpretation	77
9 Angular momentum - 2	79
9.1 Spectroscopic Notation	80
9.2 Dirac's notation	80
9.3 Experimental Evidence for Quantisation	81
9.4 Summary	83
10 Central potential	85
10.1 Introduction	86
10.2 Stationary states.	86
10.3 Physical interpretation.	88
10.4 Quantum rotator	88
10.5 Summary	91
11 The Hydrogen atom	93
11.1 Introduction	94
11.2 Stationary states	94
11.3 Solution of the radial equation	96
11.4 Physical interpretation	98
11.5 Summary	100
12 The harmonic oscillator	103
12.1 Introduction	104
12.2 Factorizing the Hamiltonian	104
12.3 Creation and annihilation	106
12.4 Eigensystem	107
12.4.1 Eigenvalues	107
12.4.2 Normalisation of Eigenstates	108
12.4.3 Wave functions	108
12.5 Summary	109
13 Angular momentum - Reloaded	111
13.1 Introduction	112
13.2 The Eigenvalue Spectra of \hat{J}^2 and \hat{J}_z	112

13.3	Nomenclature	116
13.4	Normalization	116
13.5	Matrix Representations	117
13.6	Summary	118
14	Spin	119
14.1	Introduction	120
14.2	Matrix representation	120
14.3	Eigenvectors	122
14.4	Scalar Products	123
14.5	Eigenvectors of \hat{J}_x	123
14.6	The Stern-Gerlach Experiment	126
14.7	Summary	128
15	Addition of angular momenta	129
15.1	Introduction	130
15.2	Total angular momentum operator	130
15.3	Addition Theorem	131
15.4	Summary	133
16	Identical Particles	135
16.1	Introduction	136
16.2	A first look at Helium	137
16.3	Two-electron wave function	138
16.4	More on the He atom	138
16.5	Pauli exclusion principle	140
16.6	Summary	141
17	Perturbation Theory	143
17.1	Introduction	144
17.2	Perturbative solution	144
17.3	Degenerate levels	147
17.4	Applications	148
17.4.1	Ground state of Helium	149
17.4.2	Spin-orbit effects in hydrogenic atoms	150
17.5	Summary	152

Lecture 1

Quantum states

1.1 Introduction

Quantum mechanics describes the behaviour of matter and light at the atomic scale ($d \sim 10^{-10}m$), where physical objects behave very differently from what we experience in our everyday's life. Because the atomic behaviour is so unusual, we need to develop abstract tools that allow us to *compute* the expected values of the observables. The theory can be tested by comparing our predictions to experimental results.

Experimental results at the beginning of the 20th century first highlighted behaviours at such small scales that were inconsistent with classical mechanics. It took a lot of effort until Planck, Bohr, Schrödinger, Heisenberg, and Born (who has been a professor in Edinburgh) obtained a consistent picture of the new dynamics.

Despite its unintuitive aspects, quantum mechanics describes very concrete features of the world as we know it, like e.g. the stability of the hydrogen atom. Some of its predictions have now been tested to great accuracy.

In fact, there are numerous experimental results that provide evidence in favour of quantum mechanics, e.g.

- double-slit experiments,
- photoelectric effect,
- stability of the H atom,
- black-body radiation.

We shall introduce the basic ideas of quantum mechanics by discussing briefly the **double-slit** experiment, which was first performed by Young in 1801, in order to resolve the question of the corpuscular nature of light ¹. The quantum behaviour of electrons will be compared with the more familiar behaviours of classical particles and waves.

1.1.1 Experiment with classical particles

Let us consider first a source of *classical* particles, emitting projectiles (bullets) in random directions. In front of the source we have a wall with two holes that are denoted **1** and **2**. A detector is placed behind the wall, counting the bullets that pass through the holes. The apparatus is schematically represented in Fig. 1.1.

When both holes are open the distribution of bullets on the detector is given by the curve P_{12} in (A). However when **2** is closed, and therefore the bullets can only pass through **1**, we obtain the distribution labelled P_1 in (B). Similarly, when the bullets can only go through **2**, the distribution is given by P_2 . The important feature, which is typical of *classical* mechanics,

¹An interesting account of Young's original experiment has been published in *The Physics Teacher*, **24** 217, 1986, and can be found at <http://cavendishscience.org/phys/tyoung/tyoung.htm>.

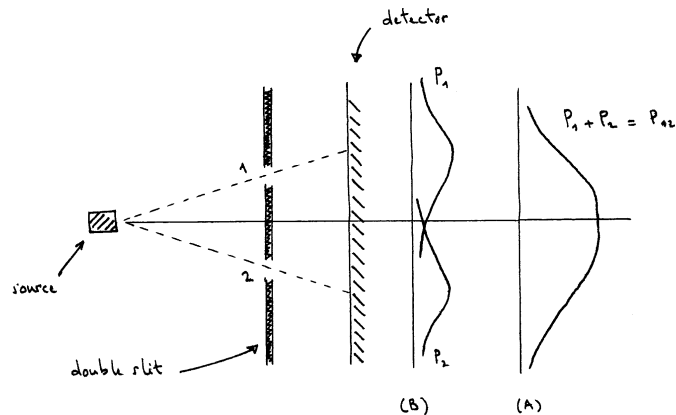


Figure 1.1: Double-slit experiment with classical particles.

is that $P_{12} = P_1 + P_2$. We call this result an observation of *no interference*, particles that go through **1** do not interfere with those that pass through **2**, and the probability add in an intuitive way.

1.1.2 Experiment with waves

We now consider the same experiment, but with a source emitting waves. The detector measures the *intensity* of the waves, i.e. the square of the height of the wave.

The original wave emitted by the source is diffracted at the holes, and two new circular waves spread from each hole. If we cover each hole at a time, we obtain respectively the intensity profiles I_1 and I_2 sketched in (B) in Fig. 1.2. However when both holes are open, we find the pattern I_{12} in (A).

Clearly in the case of diffracted waves $I_{12} \neq I_1 + I_2$. The two waves interfere. The maxima of I_{12} correspond to constructive interference, and viceversa for the minima. The height of the waves can be represented as (the real part of) a complex amplitude $h_1 e^{i\omega t}$, where h_1 is a complex number, and the intensity is given by $I_1 = |h_1|^2$. In the case of waves interference, the amplitudes add up, so that $I_{12} = |h_1 + h_2|^2$.

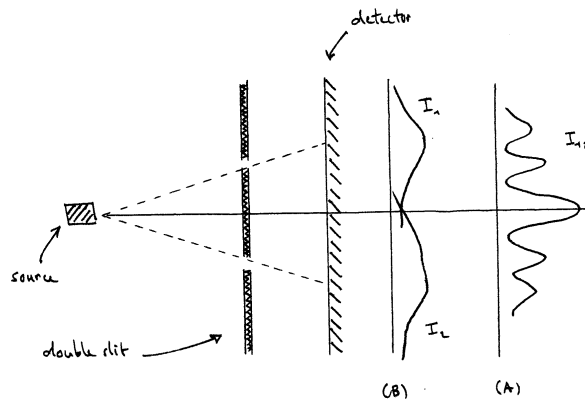


Figure 1.2: Double-slit experiment with waves.

1.1.3 Experiment with electrons

A conclusive double-slit experiment with electron beams was performed by Tonomura et al. at Hitachi (Japan)². This experiment confirmed the observation of an interference pattern similar to the one observed for waves. This experiment shows that the dynamics of the electrons, that can be considered as single indivisible particles, is influenced by the presence of both slits; a single particle at the atomic scale “behaves like a wave”. This property is known as *wave-particle duality*. Wave-particle duality was first formulated by De Broglie in 1925, who stated that a particle with momentum p behaves like a wave with wavelength $\lambda = h/p$, where h is Planck’s constant:

$$h = 6.62606896(33) \times 10^{-34} \text{ J} \cdot \text{s}. \quad (1.1)$$

Note that Planck’s constant has units $[h] = \text{Energy} \times \text{Time}$, it plays a fundamental role by defining the scale where quantum phenomena become relevant. We will often encounter the constant $\hbar = h/(2\pi)$.

²A. Tonomura et al., *Am. J. Phys.* **57**, 117, 1989.

1.2 One-dimensional systems

We shall now proceed to summarize the laws of quantum mechanics for one-dimensional systems. Before entering into the details, it is worthwhile to quote Feynman's lectures³:

"In this subject we have, of course, the difficulty that the quantum mechanical behavior of things is quite strange. Nobody has an everyday experience to lean on to get a rough, intuitive idea of what will happen. So there are two ways of presenting the subject: We could either describe what can happen in a rather rough physical way, telling you more or less what happens without giving the precise laws of everything; or we could, on the other hand, give the precise laws in their abstract form. But, then because of the abstractions, you wouldn't know what they were all about, physically. The latter method is unsatisfactory because it is completely abstract, and the first way leaves an uncomfortable feeling because one doesn't know exactly what is true and what is false. [...] Here, we will try to find a happy medium between the two extremes".

1.2.1 Quantum states

Let us begin with

the fundamental law of quantum mechanics

which summarizes the idea of wave-particle duality. The quantum state of a system is described by a complex function Ψ , which depends on the coordinate x and on time:

$$\boxed{\text{quantum state} \sim \Psi(x, t)} \quad (1.2)$$

The wave function does not depend on the momentum of the particle. Compared to classical mechanics, we seem to have lost the symmetry between coordinates and momenta. We shall revisit this issue later. The wave function encodes, in a probabilistic sense, all the information about the system. The probability of obtaining a given result can be computed from the wave function.

For instance,

$$\boxed{|\Psi(x, t)|^2 dx \text{ is the probability that a measurement of the position of the particle yields a result in the interval } x \rightarrow x + dx.}$$

Thus $|\Psi(x, t)|^2$ is a *probability per unit length* or *probability density*. The total probability of finding the particle somewhere along the real axis must be unity, thus:

$$\|\Psi\|^2 = \int dx |\Psi(x, t)|^2 = 1. \quad (1.3)$$

³R.P. Feynman, R.B. Leighton, M. Sands, *The Feynman lectures on physics - Quantum Mechanics*, Addison-Wesley, 1965.

Any function such that its integral along the real axis is finite can be normalized by multiplying by an appropriate constant. In practice two wave functions that differ by an arbitrary factor $c \in \mathbb{C}$ describe the same physical system.

Mathematical aside

Let us see an example of a normalizable function. The function $\psi(x) = e^{-x^2/2}$ is clearly normalizable. Its norm is

$$\|\psi\|^2 = \int dx e^{-x^2} = \sqrt{\pi}, \quad (1.4)$$

and therefore the *normalized wave function* is:

$$\psi(x) = \frac{1}{\pi^{1/4}} \exp[-x^2/2]. \quad (1.5)$$

On the other hand, the function $e^{x^2/2}$ is non-normalizable, and therefore does not represent a physical state.

In general, if $\int dx |\psi(x)|^2 = c$, then the normalized wave function is $\frac{1}{\sqrt{c}}\psi(x)$.

Example In order to understand better how the information about the system is encoded in the wave function, we shall start with a simpler version of our one-dimensional system. Let us consider a particle in a discretized space. The particle can only be in a *finite* number of positions along the real axis, as shown in Fig. 1.3. In this particular example, the particle can be in one of *six*⁴ points along the real axis labelled $0, \dots, 5$. The lattice spacing (i.e. the distance between two points) is denoted ϵ .

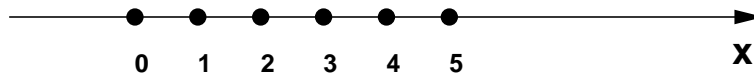


Figure 1.3: Discretized one-dimensional system with six sites. The particle can only be in one of the six points denoted by $0, \dots, 5$. The distance between successive points is ϵ .

According to the probabilistic interpretation of the wave function, the probability for the particle to be at x_i at time t is given by $|\Psi(x_i, t)|^2 \epsilon$. Note the factor ϵ that multiplies the modulo square of the wave function. This is needed because the modulo square is the probability density per unit length. Hence to find the probability we need to multiply by the distance between two points. If we redefine

$$\psi_i = \sqrt{\epsilon} \Psi(x_i, t), i = 0, \dots, 5, \quad (1.6)$$

⁴There is no particular significance in the fact that we have chosen here 6 points. The same example could have been worked out with two points, or any finite number of points.

then the whole information about this system is encoded in a six-dimensional *complex vector*:

$$|\Psi\rangle = (\psi_0, \dots, \psi_5). \quad (1.7)$$

In Eq. (1.7) we have used the *ket* notation introduced by Dirac to denote a state vector. We shall keep using this notation throughout these notes.

In this (simple) discretized example, we can see explicitly that the state of the system is represented by a vector. This is an important concept to remember.

The normalization condition, which amounts to require that the total probability of finding the particle is one, can be written:

$$\sum_{k=0}^5 |\psi_k|^2 = 1, \quad (1.8)$$

i.e. $|\Psi\rangle$ is a *complex* vector of unit norm. We shall frequently use the following convention:

$$\langle\Psi|\Psi\rangle = \sum_{k=0}^5 |\psi_k|^2, \quad (1.9)$$

where we have used Dirac's notation to indicate the scalar product of two vectors:

$$\langle\Phi|\Psi\rangle = \sum_k \phi_k^* \psi_k. \quad (1.10)$$

The wave function for a continuous system can be seen as the limit of the discretized case where the number of points goes to infinity, while the distance ϵ becomes infinitesimally small. In this limit, instead of a finite-dimensional vector, we obtain an infinite number of coordinates, encoded in a continuous function $\psi(x)$. We shall still refer to the wave function as the state vector, bearing in mind that in this case the vector is infinite-dimensional. There is one coordinate $\Psi(x, t)$ for each point x on the real axis, and there is an infinity of points along the real axis. We denote the state vector using a ket $|\Psi(t)\rangle$. Note that the time dependence of the state vector is explicit in Dirac's notation. The spatial coordinate x labels the components of the state vector.

The norm of the state vector can be written as the limit of the norm of the finite-dimensional vector for $\epsilon \rightarrow 0$. Starting from Eq. (1.9), and taking the limit:

$$\langle\Psi|\Psi\rangle = \lim_{\epsilon \rightarrow 0} \sum_k \epsilon |\Psi(x_k, t)|^2 = \int dx |\Psi(x, t)|^2, \quad (1.11)$$

we recover Eq. (1.3). Similarly the scalar product of two wave functions can be defined as the limit of the discrete case:

$$\langle \Phi | \Psi \rangle = \int dx \Phi(x, t)^* \Psi(x, t). \quad (1.12)$$

1.2.2 Superposition principle

State vectors can be combined linearly to obtain new admissible quantum states. If Ψ_1 and Ψ_2 are quantum states, then

$$\Psi(x, t) = c_1 \Psi_1(x, t) + c_2 \Psi_2(x, t), \quad c_1, c_2 \in \mathbb{C}, \quad (1.13)$$

is also a possible state of the system. In mathematical terms, the space of possible quantum states is called a *vector space*⁵.

Mathematical aside

As a consequence the time evolution of a quantum mechanical system must be determined by a *linear* equation:

$$\mathcal{L}\Psi = 0, \quad (1.14)$$

where \mathcal{L} is a linear operator, i.e. an operator such that:

$$\mathcal{L}(c_1 \Psi_1 + c_2 \Psi_2) = c_1 \mathcal{L}\Psi_1 + c_2 \mathcal{L}\Psi_2. \quad (1.15)$$

Example 1 Note that the concept of superposition of states is very different from anything we have encountered in classical mechanics. Consider two quantum states $|A\rangle$ and $|B\rangle$, such that the measurement of an observable O yields the result a when the system is in the state $|A\rangle$, and the result b when the system is in the state $|B\rangle$. The superposition principle states that the state vector:

$$|C\rangle = c_A |A\rangle + c_B |B\rangle, \quad (1.16)$$

where c_A and c_B are complex numbers such that $|c_A|^2 + |c_B|^2 \neq 0$, describes another possible physical state of the system. A measurement of the observable O in the state $|C\rangle$ can only yield the value a or b , with respective probabilities:

$$p_a = \frac{|c_A|^2}{|c_A|^2 + |c_B|^2}, \quad p_b = \frac{|c_B|^2}{|c_A|^2 + |c_B|^2}. \quad (1.17)$$

No other results are possible for the measured value of O in the state $|C\rangle$.

⁵You can find a summary on the properties of vector space at http://en.wikipedia.org/wiki/Vector_space, or in your lecture notes from Year 2.

Example 2 Let us consider again the discretized one-dimensional system in Fig. 1.3; we can have a state $|1\rangle$ where the particle is localized e.g. at site 1, and a state $|2\rangle$ where the particle is localized at site 2. Measuring the position of the particle in state $|1\rangle$ yields $x = 1$ with probability 1. Likewise we obtain $x = 2$ with probability 1 for a particle described by the state vector $|2\rangle$.

The state $\frac{1}{\sqrt{2}}|1\rangle + \frac{1}{\sqrt{2}}|2\rangle$ is an admissible quantum state. Measuring the position of the particle in this latter state, the outcome will be $x = 1$ or $x = 2$ with 50% probability. No other value is allowed in this state.

1.2.3 The uncertainty principle

In classical mechanics the state of a particle in a one-dimensional world is completely determined by the value of its position $x(t)$ and momentum $p(t)$, i.e. by its *trajectory*.

The situation is radically different in quantum mechanics. The probabilistic interpretation of the wave function implies that we can at best obtain the probability density for a particle to be at a given position x at time t . As a consequence the concept of *classical trajectory* used in Newtonian mechanics does not make sense in quantum mechanics. The position and momentum of the particle can be defined, but their values cannot be measured simultaneously.

On the other hand, when the scales in the problem are much larger than the Planck constant h , we expect to recover the classical results.

These two features are summarized in the so-called *uncertainty relations*, first derived by Heisenberg. The uncertainty relations state that, if the position and momentum are measured simultaneously, with respective precisions Δx and Δp , then:

$$\Delta x \cdot \Delta p \geq \frac{\hbar}{2}. \quad (1.18)$$

It is clear from Eq. (1.18) that if the position of the particle is known exactly, then the knowledge of its momentum is completely lost. In general the product of the two uncertainties has to be greater than $\hbar/2$.

It is important to appreciate that Heisenberg's inequalities reflect a physical limitation. A better experimental apparatus would **not** allow a higher precision to be obtained. The uncertainty is a property of the dynamics of the system.

They also encode the idea that in quantum mechanics the measurement of a quantity interferes with the dynamics. If we measure exactly the position of the particle, then we lose all knowledge of its momentum.

Hence the concept of determinism is lost during the measurement process. As we will discuss below, the evolution of a quantum system left unperturbed is completely determined by the Schrödinger equation, and therefore is completely deterministic. You should contrast this with the situation in classical mechanics, where we can assume that measurements do not perturb the state of the system.

1.2.4 Dynamical variables

We have seen above that the modulo square of the wave function $|\Psi(x, t)|^2$ yields the probability density of finding the particle at position. We can therefore compute the *mean value* of the particle position:

$$\langle x \rangle = \int dx x |\Psi(x, t)|^2. \quad (1.19)$$

If we define an operator \hat{X} acting on the wave function:

$$\hat{X}\Psi(x, t) = x\Psi(x, t); \quad (1.20)$$

the expectation value can then be written as:

$$\langle \hat{X} \rangle = \int dx \Psi(x, t)^* x \Psi(x, t) = \langle x \rangle. \quad (1.21)$$

More generally in quantum mechanics each observable O is associated to a *linear operator* \hat{O} acting on the wave function. A linear operator satisfies:

$$\hat{O}(c_1\Psi_1 + c_2\Psi_2) = c_1\hat{O}\Psi_1 + c_2\hat{O}\Psi_2, \quad (1.22)$$

where c_1 and c_2 are complex numbers. Note that the result of acting on a state vector with an operator produces *new* state vector.

$$\hat{O}\Psi(x, t) = \Psi'(x, t), \text{ or equivalently } \hat{O}|\Psi\rangle = |\Psi'\rangle. \quad (1.23)$$

Example For instance the Hamiltonian of the system may be written as $\hat{H} = \hat{T} + \hat{V}$, where the kinetic and potential energy operators are defined by:

$$\hat{T} = \frac{\hat{P}^2}{2m}, \quad \hat{V} = V(\hat{X}). \quad (1.24)$$

The operator $V(\hat{X})$, which is a function of the position operator \hat{X} , acts as:

$$V(\hat{X})\Psi(x, t) = V(x)\Psi(x, t). \quad (1.25)$$

Any other operator that is a function of \hat{X} acts in the same way. We shall see later how to define the operator \hat{p} , and more complicated operators.

1.3 Summary

Let us conclude this chapter by summarizing the main concepts introduced in this lecture.

- Quantum states are represented by a wave function $\Psi(x, t)$, or equivalently by a vector in an infinite-dimensional vector space.
- In a one-dimensional system, $|\Psi(x, t)|^2 dx$ yields the probability of finding the particle in an interval $[x, x + dx]$. Hence the normalization: $\int dx |\Psi(x, t)|^2 = 1$.
- Superposition principle: a linear combination of wave functions with arbitrary complex coefficients yields a possible quantum state.
- Uncertainty principle.
- Dynamical variables are associated to linear operators acting on the wave functions.

Lecture 2

Observables

2.1 Observing observables

We have seen at the end of the previous lecture that each dynamical variable is associated to a linear operator \hat{O} , and its expectation value in a quantum state can be computed:

$$\langle O \rangle = \int dx \Psi(x, t)^* \hat{O} \Psi(x, t). \quad (2.1)$$

Let us now clarify how these expectation values are related to what can be observed in experiments. If the observable O is measured several times under identical conditions, the results will be a set of values $\{O_1, O_2, \dots, O_n\}$.

The expectation value defined in Eq. (2.1) is equal to the average $\sum_{k=1}^n O_k/n$ in the limit where the number of measurements $n \rightarrow \infty$.

The possible outcomes of experiments O_k are the *eigenvalues* of the operator \hat{O} , i.e. the solutions of the eigenvalue equation:

$$\hat{O}\psi_k = O_k\psi_k, \quad (2.2)$$

where ψ_k is the *eigenfunction* corresponding to the eigenvalue O_k . The eigenfunction represents the wave function of a state in which the measurement of O yields the value O_k with probability 1. To check this statement, we can compute the variance of O in the state ψ_k :

$$\begin{aligned} \text{Var}_k[O] &= \langle \psi_k | \hat{O}^2 | \psi_k \rangle - \langle \psi_k | \hat{O} | \psi_k \rangle^2 = \\ &= O_k^2 \langle \psi_k | \psi_k \rangle - (O_k \langle \psi_k | \psi_k \rangle)^2 = 0, \end{aligned} \quad (2.3)$$

where we have used the fact that the eigenfunctions are normalized to one.

A generic state can be expressed as a superposition of eigenstates:

$$\psi(x) = \sum_k c_k \psi_k(x), \quad (2.4)$$

which can also be rewritten using Dirac's notation as:

$$|\psi\rangle = \sum_k c_k |\psi_k\rangle. \quad (2.5)$$

The set of eigenfunctions is called a *complete set* of states, or a *basis*. You can easily prove that a generic linear combination of eigenstates is not an eigenstate.

Using the fact that eigenfunctions are orthogonal (see below), you can readily check that:

$$c_m = \langle \psi_m | \psi \rangle = \int dx \psi_m(x)^* \psi(x), \quad (2.6)$$

and therefore given a state ψ , you can compute c_m if the eigenfunctions are known.

Given the decomposition in Eq. (2.4), the probability of finding the result O_k when measuring O in the state $\psi(x)$ is given by:

$$P_k = |c_k|^2. \quad (2.7)$$

Clearly the sum of probabilities should be properly normalized and therefore:

$$\sum_k P_k = \sum_k |c_k|^2 = 1. \quad (2.8)$$

Collapse of the wave function Another important feature of quantum mechanics is the following:

Immediately after a measurement that gave the result f_k , the system is in the state ψ_k . The state vector has been *projected* onto the eigenstate by the process of performing the measurement.

If we want to express the same concept using equations, we can say that immediately after a measurement yielding the value f_k :

$$\psi(x) \mapsto \psi_k(x), \quad (2.9)$$

$$|\psi\rangle \mapsto |\psi_k\rangle. \quad (2.10)$$

This is sometimes referred to as the *collapse of the wave function*. The operator that performs the state reduction is called a *projection operator*:

$$|\psi\rangle \mapsto \mathcal{P}_k |\psi\rangle. \quad (2.11)$$

After a measurement that yielded the value f_k , the wave function of the system coincides with the eigenfunction ψ_k . Then, as discussed below Eq. (2.2), if we perform immediately another measurement of f we will find the *same* value f_k with probability 1.

Conversely, if the wave function does not coincide with one of the eigenfunctions, then the observable f does not have a given value in the state Ψ . We can only compute the probability for each eigenvalue to be the outcome of the experiment.

Mathematical aside

A projection operator satisfies:

$$\mathcal{P}_k^2 = \mathcal{P}_k, \quad \mathcal{P}_k \mathcal{P}_l = 0, \quad l \neq k. \quad (2.12)$$

Verify that you recognize these properties by considering the more familiar case of projectors in three-dimensional Euclidean space.

Clearly these phenomena do not have a classical analogue. The description of a physical system in quantum mechanics is radically different from the classical one. You need to practice in order to get familiar with the quantum mechanical framework.

2.2 Hermitean operators

We stated in the first lecture that every observable is represented by an operator; the correspondence is

OBSERVABLE	\iff	OPERATOR
total energy	\iff	\hat{H}
position	\iff	\hat{X}
momentum	\iff	\hat{P}
\vdots	\iff	\vdots

Observables take *real* values only. Therefore we must require that the operators that represent observables have only *real* eigenvalues, since we want to identify the eigenvalues with the possible results of measurements. We can guarantee this if we only use *Hermitean* operators to represent observables.

2.2.1 Hermitean conjugate

We first define the so-called *Hermitean conjugate* \hat{O}^\dagger of an operator \hat{O} . Let $\psi(x)$ and $\phi(x)$ be arbitrary spatial wave functions (e.g. $\psi(x) \equiv \Psi(x, 0)$ and $\phi(x) \equiv \Phi(x, 0)$), then

$$\int_{-\infty}^{\infty} \phi^*(x) \hat{O}^\dagger \psi(x) dx \equiv \left(\int_{-\infty}^{\infty} \psi^*(x) \hat{O} \phi(x) dx \right)^* .$$

We can rewrite this relation using Dirac's notation as:

$$\boxed{\langle \phi | \hat{O}^\dagger | \psi \rangle = \left(\langle \psi | \hat{O} | \phi \rangle \right)^*} \quad (2.13)$$

Mathematical aside

Compare Eq. (2.13) with the more familiar expression from linear algebra:

$$O_{ij}^\dagger = O_{ji}^* . \quad (2.14)$$

The expressions in Eq. (2.13) are the *matrix elements* of the operator \hat{O} , just like O_{ij} are the matrix elements of a matrix O . The quantum states $|\phi\rangle$ and $|\psi\rangle$ are the “indices” that label the matrix elements in quantum mechanics.

Using this identification, many equations that we will encounter in quantum mechanics will look rather familiar.

Example Let $\hat{O} = \frac{d}{dx}$ then we can integrate by parts to obtain

$$\int_{-\infty}^{\infty} \psi^*(x) \frac{d}{dx} \phi(x) dx = [\phi(x)\psi^*(x)]_{-\infty}^{\infty} - \int_{-\infty}^{\infty} \phi(x) \frac{d}{dx} \psi^*(x) dx$$

We can discard the constant term on the right hand side, since physically acceptable wave functions vanish at $x = \pm\infty$, and if we then take the complex conjugate of the resulting equation we obtain

$$\begin{aligned} \left(\int_{-\infty}^{\infty} \psi^*(x) \frac{d}{dx} \phi(x) dx \right)^* &= - \int_{-\infty}^{\infty} \phi^*(x) \frac{d}{dx} \psi(x) dx \\ &\equiv \int_{-\infty}^{\infty} \phi^*(x) \left(\frac{d}{dx} \right)^\dagger \psi(x) dx \end{aligned}$$

from the definition of Hermitean conjugate. Thus we can make the identification

$$\left(\frac{d}{dx} \right)^\dagger = -\frac{d}{dx}.$$

2.2.2 Hermitean operators

We can now define a *Hermitean operator*; it is an operator for which

$$\boxed{\hat{O}^\dagger \equiv \hat{O}} \tag{2.15}$$

Example This is clearly not true of all operators; $\frac{d}{dx}$ is NOT Hermitean since we have just shown that

$$\left(\frac{d}{dx} \right)^\dagger = -\frac{d}{dx},$$

whereas the operator $-i\hbar \frac{d}{dx}$ IS Hermitean; the proof is straightforward and is left as an exercise.

2.3 Properties of Hermitean operators

1. Hermitean operators have real eigenvalues.

Proof:

Suppose \hat{A} is a Hermitean operator so that $\hat{A}^\dagger = \hat{A}$, and let \hat{A} have an eigenvalue a , with corresponding eigenfunction $\psi_a(x)$:

$$\hat{A} \psi_a(x) = a \psi_a(x)$$

Then

$$\int_{-\infty}^{\infty} \psi_a^*(x) \hat{A} \psi_a(x) dx = a \int_{-\infty}^{\infty} \psi_a^*(x) \psi_a(x) dx = a \int_{-\infty}^{\infty} |\psi_a(x)|^2 dx = a$$

If we take the complex conjugate of this equation, we obtain

$$\left(\int_{-\infty}^{\infty} \psi_a^*(x) \hat{A} \psi_a(x) dx \right)^* = a^*$$

but if we make use of the definition of the Hermitean conjugate, we can rewrite the left-hand side of this equation in terms of \hat{A}^\dagger and use the fact that $\hat{A}^\dagger = \hat{A}$ by hypothesis:

$$\left(\int_{-\infty}^{\infty} \psi_a^*(x) \hat{A} \psi_a(x) dx \right)^* = \int_{-\infty}^{\infty} \psi_a^*(x) \hat{A}^\dagger \psi_a(x) dx = \int_{-\infty}^{\infty} \psi_a^*(x) \hat{A} \psi_a(x) dx.$$

The right-hand side is now just the integral which appears in the first equation and is equal to a , so we have proved that

$$a^* = a$$

thus showing that the eigenvalue a is real as stated. **QED.**

2. The eigenfunctions of a Hermitean operator which belong to different eigenvalues are orthogonal.

Proof:

Suppose that

$$\hat{A} \psi_1(x) = a_1 \psi_1(x) \quad \text{and} \quad (2.16)$$

$$\hat{A} \psi_2(x) = a_2 \psi_2(x) \quad \text{with } a_1 \neq a_2 \quad (2.17)$$

From Eq. (2.16) we have

$$\int_{-\infty}^{\infty} \psi_2^*(x) \hat{A} \psi_1(x) dx = a_1 \int_{-\infty}^{\infty} \psi_2^*(x) \psi_1(x) dx \quad (2.18)$$

whereas from Eq. (2.17)

$$\int_{-\infty}^{\infty} \psi_1^*(x) \hat{A} \psi_2(x) dx = a_2 \int_{-\infty}^{\infty} \psi_1^*(x) \psi_2(x) dx \quad (2.19)$$

Taking the complex conjugate of Eq. (2.19) yields on the left hand side

$$\left(\int_{-\infty}^{\infty} \psi_1^*(x) \hat{A} \psi_2(x) dx \right)^* \equiv \int_{-\infty}^{\infty} \psi_2^*(x) \hat{A}^\dagger \psi_1(x) dx = \int_{-\infty}^{\infty} \psi_2^*(x) \hat{A} \psi_1(x) dx, \quad (2.20)$$

whereas the right hand side gives

$$a_2^* \int_{-\infty}^{\infty} \psi_1(x) \psi_2^*(x) dx = a_2 \int_{-\infty}^{\infty} \psi_2^*(x) \psi_1(x) dx, \quad (2.21)$$

using the fact that $a_2 = a_2^*$.

Comparing with Eq. (2.18) we see that

$$a_2 \int_{-\infty}^{\infty} \psi_2^*(x) \psi_1(x) dx = a_1 \int_{-\infty}^{\infty} \psi_2^*(x) \psi_1(x) dx, \quad (2.22)$$

which we can rearrange to yield the result

$$(a_2 - a_1) \int_{-\infty}^{\infty} \psi_2^*(x) \psi_1(x) dx = 0. \quad (2.23)$$

Given that $a_2 \neq a_1$ by hypothesis, this implies that

$$\int_{-\infty}^{\infty} \psi_2^*(x) \psi_1(x) dx \equiv \quad (2.24)$$

$$\equiv \langle \psi_2 | \psi_1 \rangle = 0, \quad (2.25)$$

which is the desired result.

QED.

2.4 Commutators

The product of two operators is defined as you would expect:

$$\hat{O}_1 \hat{O}_2 |\psi\rangle = \hat{O}_1 \left(\hat{O}_2 |\psi\rangle \right). \quad (2.26)$$

Note that the order in which the operators are applied to the state is important! The *commutator* of two operators is:

$$\left[\hat{O}_1, \hat{O}_2 \right] = \hat{O}_1 \hat{O}_2 - \hat{O}_2 \hat{O}_1. \quad (2.27)$$

In general the commutator does NOT vanish, and defines a third operator, acting on quantum states.

Example We have seen previously that the action of the position operator \hat{X} is:

$$\hat{X}\psi(x) = x\psi(x), \quad (2.28)$$

i.e. the wave function is simply multiplied by the value of x . The *momentum operator* is given by a differential operator:

$$\hat{P}\psi(x) = -i\hbar \frac{d}{dx}. \quad (2.29)$$

This is a simple realization of de Broglie's duality hypothesis. Remember that according to the wave-particle duality to each particle with momentum p we can associate a wave with wavelength h/p . A wave with a fixed wavelength is a plane wave, described by the function:

$$\psi_p(x) = \exp[ipx/\hbar]. \quad (2.30)$$

When we act with the operator \hat{P} defined in Eq. (2.29), we see that $\psi_p(x)$ is an eigenstate of \hat{P} with eigenvalue p . So the plane wave corresponds to a state with given momentum p . This justifies the definition of \hat{P} as a momentum operator.

Consider the case $\hat{O}_1 = \hat{X}$, $\hat{O}_2 = \frac{d}{dx}$. Then:

$$\hat{O}_1\hat{O}_2\psi(x) = \hat{X} \left(\frac{d}{dx}\psi(x) \right) \quad (2.31)$$

$$= x \frac{d}{dx}\psi(x), \quad (2.32)$$

while

$$\hat{O}_2\hat{O}_1\psi(x) = \hat{O}_2 \left(\hat{X}\psi(x) \right) \quad (2.33)$$

$$= \hat{O}_2(x\psi(x)) \quad (2.34)$$

$$= \frac{d}{dx}(x\psi(x)) \quad (2.35)$$

$$= \psi(x) + x \frac{d}{dx}\psi(x). \quad (2.36)$$

Putting the two results together, we obtain for this particular choice of \hat{O}_1 and \hat{O}_2 :

$$\left[\hat{O}_1, \hat{O}_2 \right] \psi(x) = -\psi(x), \quad (2.37)$$

i.e.

$$\boxed{\left[\hat{O}_1, \hat{O}_2 \right] = -1.} \quad (2.38)$$

From the example above we deduce the fundamental *canonical commutation relation*:

$$\boxed{\left[\hat{X}, \hat{P} \right] = i\hbar} \quad (2.39)$$

2.5 Compatible Observables

Suppose A and B are observables and we perform the following sequence of measurements in rapid succession on a single system:

1. measure A
2. measure B
3. remeasure A

Then *if and only if* the result of 3 is *certain* to be the same as the result of 1, we say that A and B are *compatible observables*. In general, this will not be the case: the measurement of B will “spoil” the result of 1.

In a little more detail, suppose that A and B are represented by operators \hat{A} and \hat{B} respectively, with

$$\begin{aligned}\hat{A} u_i(x) &= A_i u_i(x) \\ \hat{B} v_i(x) &= B_i v_i(x)\end{aligned}$$

Then measurement 1 yields some eigenvalue, A_j say, of \hat{A} , forcing the system into the state $u_j(x)$. Measurement 2 yields B_k say, forcing the system into the state $v_k(x)$, so that measurement 3 is made with the system in the state $v_k(x)$. The only way that 3 is *certain* to yield the result A_j as obtained in 1 is if $v_k(x) \equiv u_j(x)$. For this to be true in all circumstances it must be the case that each eigenfunction $v_k(x)$ of \hat{B} is identical with some eigenfunction $u_j(x)$ of \hat{A} . If there is no degeneracy this implies a 1-1 correspondence between the eigenfunctions of \hat{A} and the eigenfunctions of \hat{B} . We say that \hat{A} and \hat{B} have a *common eigenbasis*.

2.5.1 The Compatibility Theorem

Given two observables, A and B , represented by Hermitean operators \hat{A} and \hat{B} , then any one of the following three statements implies the other two:

1. A and B are compatible observables;
2. \hat{A} and \hat{B} have a common eigenbasis;
3. the operators \hat{A} and \hat{B} commute: $[\hat{A}, \hat{B}] = 0$

Example proof:

Let us show, for instance, that $3 \Rightarrow 2$. We have

$$\begin{aligned}\hat{A} u_i(x) &= A_i u_i(x) \\ \hat{B} v_i(x) &= B_i v_i(x)\end{aligned}$$

so that for *any* eigenfunction of \hat{A}

$$\begin{aligned}\hat{A}\hat{B}u_i(x) &= \hat{B}\hat{A}u_i(x) \text{ by virtue of 3} \\ &= \hat{B}A_i u_i(x) \\ &= A_i \hat{B}u_i(x)\end{aligned}$$

Thus $\hat{B}u_i(x)$ is an eigenfunction of \hat{A} belonging to the eigenvalue A_i . If we assume that the eigenvalues are non-degenerate, then $\hat{B}u_i(x)$ must be some multiple of $u_i(x)$:

$$\hat{B}u_i(x) = \rho u_i(x) \text{ say}$$

This just says that $u_i(x)$ is an eigenfunction of \hat{B} belonging to the eigenvalue ρ , and we must have that, for some j ,

$$\rho = B_j \quad \text{and} \quad u_i(x) = v_j(x)$$

Thus any eigenfunction of the set $\{u_i(x)\}$ coincides with some member of the set $\{v_j(x)\}$. The correspondence has to be 1-1 because both sets are orthonormal; if we assume that two functions in one set coincide with a single function in the other set, we are led to a contradiction that two orthogonal functions are identical to the same function. By simply relabelling all the functions in one set we can always ensure that

$$u_1(x) = v_1(x), u_2(x) = v_2(x), u_3(x) = v_3(x), \dots \text{etc.}$$

and this is the common eigenbasis. A more general proof, in the case where the eigenvalues are degenerate is left as an exercise in problem sheet 1.

2.6 Summary

Let us conclude this chapter by summarizing the main concepts introduced in this lecture.

- The result of a measurement of an observable O is one of the eigenvalues of the linear operator \hat{O} .
- Properties of Hermitean operators.
- A measurement that yields a result f_k , collapses the wave function into the eigenfunction ψ_k .
- Commutators and compatible observables.

Lecture 3

Dynamics-1

3.1 Schrödinger equation

3.1.1 Equation of motion

The time evolution of a quantum state is dictated by Schrödinger's equation:

$$\boxed{i\hbar \frac{\partial}{\partial t} \Psi(x, t) = \hat{H} \Psi(x, t)}, \quad (3.1)$$

where $\hat{H} = \hat{T} + \hat{V}$ is the Hamiltonian operator. Eq. (3.1) replaces Newton's equation of classical mechanics.

The operator \hat{H} is obtained from the classical Hamiltonian, by replacing the position and momentum with the corresponding operators \hat{X} and \hat{P} in Cartesian coordinates.

Schrödinger's equation states that the Hamiltonian operator determines the time shifts of the wave function, i.e. $\hat{H} = i\hbar \frac{\partial}{\partial t}$. Similarly, the momentum operator determines the space translations - remember that $\hat{P} = -i\hbar \frac{\partial}{\partial x}$. In a four-dimensional Minkowski space-time this correspondence can be written as a four-vector relation:

$$p_\mu = i\hbar \frac{\partial}{\partial x^\mu}. \quad (3.2)$$

It is interesting to note that special relativity can be easily "embedded" in the rules of quantum mechanics! The Standard Model of particle physics, which describes the interactions between the fundamental constituents of matter very successfully, is based on this fact. This will become more clear when you study quantum field theory.

We can now discuss briefly a heuristic explanation of the identification $\hat{H} = i\hbar \frac{\partial}{\partial t}$ by considering a one-dimensional plane wave

$$\Psi(x, t) = \text{const} \times e^{-i(\omega t - kx)}. \quad (3.3)$$

The plane wave describes e.g. the propagation of monochromatic light. Applying the operator $i\hbar \frac{\partial}{\partial t}$ we obtain:

$$i\hbar \frac{\partial}{\partial t} \Psi(x, t) = \hbar\omega \Psi(x, t) = h\nu \Psi(x, t) \quad (\omega = 2\pi\nu), \quad (3.4)$$

which is precisely the energy of the photon.

3.1.2 Eigenstates of the Hamiltonian

Let us consider a time-independent Hamiltonian:

$$\hat{H} = \hat{H}(\hat{X}, \hat{P}), \quad (3.5)$$

where there is no explicit dependence on time.

The time evolution for an energy eigenstate can be worked out easily. Schrödinger's equation becomes simply:

$$i\hbar \frac{\partial}{\partial t} \Psi(x, t) = E\Psi(x, t). \quad (3.6)$$

Integrating Eq. (3.6) with respect to time:

$$\Psi(x, t) = \psi_E(x) \exp \left[-\frac{i}{\hbar} Et \right], \quad (3.7)$$

where $\psi_E(x)$ is the spatial wavefunction corresponding to the eigenvalue E .

Such states are known as *stationary states* because the associated probability density is independent of time:

$$\begin{aligned} |\Psi_E(x, t)|^2 \equiv \Psi_E^*(x, t)\Psi_E(x, t) &= \psi_E^*(x) \exp\{iEt/\hbar\} \psi_E(x) \exp\{-iEt/\hbar\} \\ &= |\psi_E(x)|^2. \end{aligned}$$

Likewise, the expectation value of any operator is also independent of time:

$$\langle \Psi_E(t) | \hat{O} | \Psi_E(t) \rangle = \int dx \Psi_E(x, t)^* \hat{O} \Psi_E(x, t) \quad (3.8)$$

$$= \int dx \psi_E(x)^* e^{iEt/\hbar} \hat{O} e^{-iEt/\hbar} \psi_E(x) \quad (3.9)$$

$$= \int dx \psi_E(x)^* e^{iEt/\hbar} e^{-iEt/\hbar} \hat{O} \psi_E(x) \quad (3.10)$$

$$= \int dx \psi_E(x)^* \hat{O} \psi_E(x) \quad (3.11)$$

$$= \langle \Psi_E(0) | \hat{O} | \Psi_E(0) \rangle. \quad (3.12)$$

Note that *formally* we can write the solution of the Schrödinger equation as:

$$|\Psi(t)\rangle = \exp \left[-\frac{i}{\hbar} \hat{H}t \right] |\Psi(0)\rangle, \quad (3.13)$$

even for a state that is not an eigenstate of the energy. However in this case $\exp \left[-\frac{i}{\hbar} \hat{H}t \right]$ is a horribly complicated operator. It can be written explicitly using a Taylor expansion:

$$\exp \left[-\frac{i}{\hbar} \hat{H}t \right] = \sum_n \frac{1}{n!} \left(\frac{-i}{\hbar} \right)^n \hat{H}^n t^n. \quad (3.14)$$

When applied to an eigenstate of the energy it turns out to be a simple complex phase!! We shall discuss this in more detail in problem sheet 2.

The eigenvalue equation

$$\boxed{\hat{H}\psi_n(x) = E_n\psi_n(x)} \quad (3.15)$$

yields the eigenvalues E_n , i.e. all the possible values of the energy of the system, and the stationary states $\psi_n(x)$. Eq. (3.15) is often referred to as *time-independent Schrödinger equation*.

Using the explicit definition of \hat{H} in Eq. (1.24), we can write the time-independent Schrödinger equation as a second order differential equation:

$$\left[-\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} + V(x) \right] u_n(x) = E_n u_n(x). \quad (3.16)$$

Eigenvalues and eigenstates are found by looking for normalizable solutions of the differential equation, taking into account the appropriate boundary conditions. Note that most authors use the notation $\{u_n(x)\}$ for the eigenfunctions of the energy, which we adopt from now on.

3.1.3 General restrictions on $\Psi(x, t)$

We see from the form of Eq. (3.16) that its solutions have to satisfy a number of properties:

1. $\Psi(x, t)$ must be a single-valued function of x and t .
2. $\Psi(x, t)$ must be a continuous function of x and t .
3. $\frac{\partial \Psi}{\partial x}(x, t)$ must be a continuous function of x .

There is an exception to this set of rules.

1. Restriction 3 does not apply when the potential energy function $V(x)$ has infinite discontinuities e.g. the ‘particle-in-a-box’ problem met in Physics 2, for which

$$V(x) = \begin{cases} \infty & \text{if } x > a \text{ or } x < 0 \\ 0 & \text{otherwise} \end{cases}$$

We shall see later a detailed description of this system.

3.1.4 Example: the quantum harmonic oscillator

In order to give a concrete example of solutions of the Schrödinger equation, we shall briefly review the quantum harmonic oscillator.

Remember that the harmonic oscillator is of paramount importance in physics, because every system close to the minimum of the potential is described by a quadratic potential, i.e. it behaves like a harmonic oscillator.

Classically, the equation of motion for a particle of mass m , subject to a restoring force proportional to the displacement, x , is

$$m\ddot{x} = -kx \quad (3.17)$$

where the constant of proportionality, k , is usually called the spring constant. This has oscillatory solutions of angular frequency $\omega = \sqrt{k/m}$. We can obtain the energy equation by writing

$$\ddot{x} = v \frac{dv}{dx}, \quad (3.18)$$

and integrating to give

$$\frac{1}{2}mv^2 + \frac{1}{2}kx^2 = E, \quad (3.19)$$

so that the potential energy function, V , is given by

$$V(x) = \frac{1}{2}kx^2 = \frac{1}{2}m\omega^2x^2. \quad (3.20)$$

The evolution of the position of the oscillator as a function of time $x(t)$ is shown in Fig. 3.1 for two different values of ω . The total energy of the oscillator is $E = \frac{1}{2}m\omega^2L^2$, where L is the maximum amplitude of the oscillations. The larger the amplitude of the oscillations, the larger the energy of the system. In the case depicted in Fig. 3.1 the amplitude $L = 1$ for both curves.

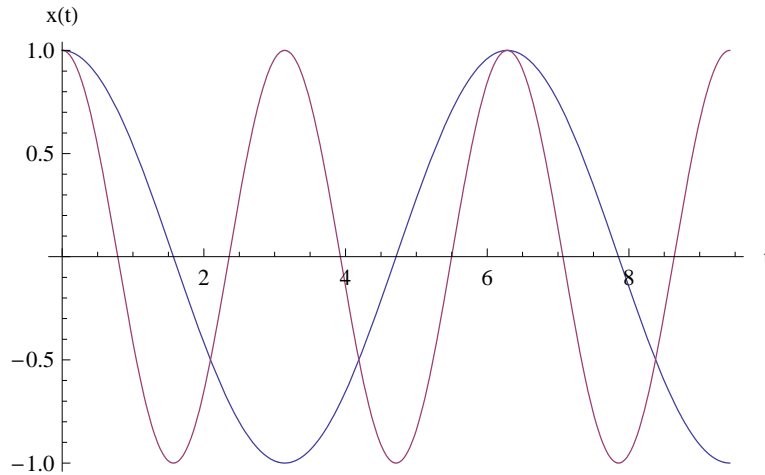


Figure 3.1: Classical trajectory of the harmonic oscillator for $\omega = 1$ (blue curve), and $\omega = 2$ (red curve).

For the quantum harmonic oscillator, the time-independent Schrödinger equation is thus

$$\left[-\frac{\hbar^2}{2m} \frac{d^2}{dx^2} + \frac{1}{2}m\omega^2x^2 \right] u_n(x) = E_n u_n(x). \quad (3.21)$$

We will not discuss the details of the solution of this differential equation, although you may have already seen it in a mathematical methods course and/or book. In any case we will develop a much more elegant method in a later lecture. It suffices to say that acceptable solutions only arise for certain values of the total energy i.e. the energy is *quantised*.

The energy eigenvalues are as follows:

$$E_n = (n + \frac{1}{2})\hbar\omega \quad n = 0, 1, 2, 3, \dots \infty \quad (3.22)$$

The state of lowest energy, or *ground state*, is labelled by $n = 0$ and has an energy $\frac{1}{2}\hbar\omega$ rather than zero.

The corresponding eigenfunctions may be written as

$$u_n(x) = C_n \exp\{-\alpha^2 x^2/2\} H_n(\alpha x) \quad (3.23)$$

where C_n is a normalisation constant, $\alpha^2 \equiv m\omega/\hbar$ and $H_n(\alpha x)$ are polynomials of degree n , known as *Hermite polynomials*. They satisfy the orthogonality relation

$$\int_{-\infty}^{\infty} \exp\{-s^2\} H_m(s) H_n(s) ds = 2^n \sqrt{\pi} n! \delta_{mn}, \quad (3.24)$$

implying that the energy eigenfunctions are orthogonal (orthonormal with an appropriate choice of the constant, C_n).

The first few Hermite polynomials are

$$\begin{aligned} H_0(s) &= 1 & H_2(s) &= 4s^2 - 2 \\ H_1(s) &= 2s & H_3(s) &= 8s^3 - 12s \end{aligned} \quad (3.25)$$

Note that the polynomials, and hence the energy eigenfunctions, are alternately even and odd functions, i.e. are of either even or odd parity. If we introduce a *parity operator*, $\hat{\mathcal{P}}$, whose action on a function, $f(x)$, is defined by

$$\hat{\mathcal{P}} f(x) = f(-x), \quad (3.26)$$

so that the eigenvalues of $\hat{\mathcal{P}}$ are ± 1 only, corresponding to the even and odd functions

$$f_+(x) \equiv \frac{1}{2}[f(x) + f(-x)] \quad \text{and} \quad f_-(x) \equiv \frac{1}{2}[f(x) - f(-x)]. \quad (3.27)$$

You can readily check that:

$$\hat{\mathcal{P}} u_n(x) = (-1)^n u_n(x), \quad (3.28)$$

and we say that the parity is $(-1)^n$.

The probability density to find the oscillator at a given position x is $|u_n(x)|^2$. The probability densities for the first five levels are shown in Fig. 3.2. Note the higher levels have a larger probability to find the oscillator at larger values of x .

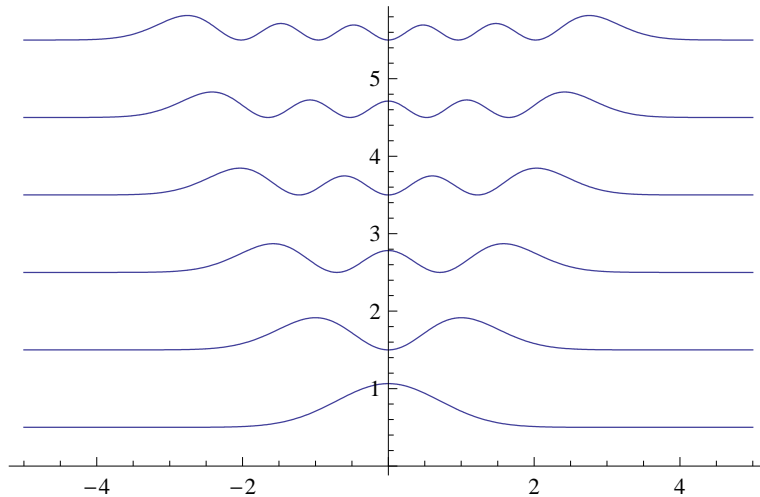


Figure 3.2: Probability to find the oscillator at x for the first five levels of the harmonic oscillator for $\alpha = 1$. Each curve is shifted upwards by $n + \frac{1}{2}$ to improve the readability of the plot.

In order to quantify the last sentence we can compute the average position, and the fluctuations of the position. The expectation value of the position of the quantum harmonic oscillator is given by:

$$\langle x \rangle_n = \int_{-\infty}^{+\infty} dx x |C_n \exp\{-\alpha^2 x^2/2\} H_n(\alpha x)|^2. \quad (3.29)$$

You can easily convince yourself that $\langle x \rangle = 0$, because the integrand is always an *odd* function of x . This is what we would have naively expected, since the system is symmetric under parity. The quantum oscillator spends an equal amount of time to the left and to the right of the origin at $x = 0$.

It is more interesting to compute the mean square deviation $\langle x^2 \rangle$, which measures the amplitude of the oscillations. You can compute the integrals explicitly for the first few Hermite polynomials, or simply look up the answer in a table of integrals:

$$\langle x^2 \rangle_n = \left(n + \frac{1}{2}\right). \quad (3.30)$$

Once again this is the quantitative expression of our intuitive understanding from classical mechanics: the larger the energy, the larger the oscillations!!

3.1.5 Expansion in stationary states

Given the expansion of the wave function at $t = 0$ in the $\{u_n(x)\}$ basis:

$$\Psi(x, 0) = \sum_n a_n u_n(x), \quad a_n = \langle u_n | \psi(0) \rangle, \quad (3.31)$$

we can express the evolved wave function at time t as:

$$\Psi(x, t) = \sum_n a_n e^{-iE_n t/\hbar} u_n(x). \quad (3.32)$$

If the Hamiltonian does not depend explicitly on time, then the energy levels E_n are constant in time:

$$\begin{aligned} \frac{d}{dt} E_n &= \frac{d}{dt} \langle \Psi_n(x, t) | \hat{H} | \Psi_n(x, t) \rangle = \frac{d}{dt} \int dx \Psi_n(x, t)^* \hat{H} \Psi_n(x, t) \\ &= \frac{1}{i\hbar} \int dx \Psi_n(x, t)^* [\hat{H}, \hat{H}] \Psi_n(x, t) = 0 \end{aligned} \quad (3.33)$$

An observable such that:

$$\frac{d}{dt} \langle \Psi(t) | \hat{O} | \Psi(t) \rangle = 0, \quad (3.34)$$

for *any* state $\Psi(x, t)$ is called a *conserved observable*.

A time-independent operator is conserved if and only if $[\hat{H}, \hat{O}] = 0$.

This is a consequence of the equation that determines the time-dependence of a matrix element:

$$i\hbar \frac{d}{dt} \langle \Psi(t) | \hat{O} | \Psi(t) \rangle = \langle \Psi(t) | [\hat{O}, \hat{H}] | \Psi(t) \rangle. \quad (3.35)$$

The proof of this result is left as an exercise in problem sheet 2.

3.2 Summary

Let us conclude this chapter by summarizing the main concepts introduced in this lecture.

- Time-dependent Schrödinger equation.
- Eigenstates of the Hamiltonian. Time-independent Schrödinger equation. Physical characterization of the stationary states.
- First properties of the solutions of the Schrödinger equation.
- The harmonic oscillator. Analytic solution, comparison between the quantum, and the classical behaviours.
- Completeness of the energy eigenstates. Expansion in stationary states.

Lecture 4

Dynamics-2

4.1 Continuous spectrum

Until now we have discussed a number of examples where operators have a discrete spectrum, i.e. where the eigenvalues are numbered by some integer index k . However there are operators that have a continuous spectrum, like e.g. the energy and the momentum of an unbound state, or the position operator \hat{X} . This requires us to generalize the statements we have made so far.

4.1.1 Eigenvalue equation

Let us denote by \hat{f} the operator associated to an observable with a continuum spectrum, the eigenvalue equation takes the form:

$$\hat{f}\psi_f(x) = f\psi_f(x), . \quad (4.1)$$

Once again, a generic state can be expanded as a superposition of eigenstates:

$$\psi(x) = \int df c(f)\psi_f(x); \quad (4.2)$$

you should compare this expression with its analogue Eq. (2.4) in the case of discrete eigenvalues.

The coefficient of the expansion Eq. (4.2) is obtained by taking the scalar product:

$$c(f) = \langle f|\psi\rangle = \int dx \psi_f(x)^*\psi(x). \quad (4.3)$$

The probabilistic interpretation of the wave function can be generalized:

The probability of finding a result between f and $f+df$ for the observable \hat{f} is given by:

$$|c(f)|^2 df .$$

Thus we derive the normalization condition:

$$\int df |c(f)|^2 = 1. \quad (4.4)$$

4.1.2 Orthonormality

Following the steps performed to obtain Eq. (2.24), we have:

$$\int dx \psi_{f'}(x)^* \hat{f}\psi_f(x) = f \int dx \psi_{f'}(x)^* \psi_f(x), \quad (4.5)$$

$$\int dx \psi_f(x)^* \hat{f}\psi_{f'}(x) = f' \int dx \psi_f(x)^* \psi_{f'}(x). \quad (4.6)$$

Taking the complex conjugate of Eq. (4.5), and using the fact that \hat{f} is Hermitean, yields:

$$\left(\int dx \psi_f(x)^* \hat{f} \psi_{f'}(x) \right)^* = \int dx \psi_{f'}(x)^* \hat{f}^\dagger \psi_f(x), \quad (4.7)$$

$$= \int dx \psi_{f'}(x)^* \hat{f} \psi_f(x). \quad (4.8)$$

Combining the results above:

$$(f - f') \int dx \psi_{f'}(x)^* \hat{f} \psi_f(x) = 0, \quad (4.9)$$

and therefore, if $f \neq f'$,

$$\langle f' | f \rangle = 0. \quad (4.10)$$

Now comes a subtle point. The norm of the state $|\psi\rangle$ is given by:

$$\langle \psi | \psi \rangle = \int df df' c(f)^* c(f') \langle f | f' \rangle \quad (4.11)$$

$$= \int df |c(f)|^2 \int df' \langle f | f' \rangle, \quad (4.12)$$

where we used the orthogonality result Eq. (4.10). Now the integral over df' in Eq. (4.12) vanishes for any finite value of $\langle f | f' \rangle$. Therefore we need to impose that $\langle f | f' \rangle$ is infinite, and normalized so that

$$\int df' \langle f' | f \rangle = 1. \quad (4.13)$$

The *delta function* introduced by Dirac satisfies precisely this condition, so we can write:

$$\langle f' | f \rangle = \int dx \psi_{f'}(x)^* \psi_f(x), \quad (4.14)$$

$$= \delta(f - f'). \quad (4.15)$$

Properties of the Dirac delta are summarized below.

4.1.3 Example: the free case

As an example of a system with a continuous spectrum we shall briefly recall the properties of the free quantum particle ($V = 0$). The Schrödinger equation reads:

$$-\frac{\hbar^2}{2m} \frac{d^2}{dx^2} \psi(x) = E \psi(x). \quad (4.16)$$

For $E > 0$ this equation has got two solutions:

$$u_{k,1}(x) = e^{ikx}, \quad u_{k,2}(x) = e^{-ikx}, \quad k = \sqrt{\frac{2mE}{\hbar^2}}. \quad (4.17)$$

Each level (except for $E = 0$) is doubly degenerate, with the two solutions representing respectively particles travelling to the right and to the left. Note that we have used to indices to identify the eigenfunctions: k is related to the value of the energy, and we can see that it spans a continuum spectrum. The second index is discrete, and runs from 1 to 2, simply to distinguish the two degenerate functions.

The time-dependent solution of Schrödinger equation reads:

$$\Psi_k(x, t) \equiv A \exp\{i(kx - \omega t)\} = A \exp\{i(px/\hbar - Et/\hbar)\}$$

which is an eigenfunction of both energy:

$$i\hbar \frac{\partial}{\partial t} \Psi_k(x, t) = \hbar\omega \Psi_k(x, t) \equiv E \Psi_k(x, t)$$

and momentum:

$$\hat{P} \Psi_k(x, t) = -i\hbar \frac{\partial}{\partial x} \Psi_k(x, t) = \hbar k \Psi_k(x, t) \equiv p \Psi_k(x, t)$$

and yet is not normalisable because

$$|\Psi_k(x, t)|^2 = |A|^2 \exp\{i(kx - \omega t)\} \exp\{-i(kx - \omega t)\} = |A|^2$$

so that

$$\int_{-\infty}^{\infty} |\Psi_k(x, t)|^2 dx = |A|^2 \int_{-\infty}^{\infty} dx = \infty$$

However, the physical interpretation is obvious; we are equally likely to find the particle *anywhere* in the interval $-\infty < x < \infty$, so that the probability density is uniform. This can be regarded as a direct consequence of the Heisenberg Uncertainty Principle: the uncertainty in momentum is zero and therefore the uncertainty in position must be infinite.

We can circumvent this technical difficulty by, for example, pretending that we are dealing with a finite system and only taking the limit of an infinite system at the end of any computation.

Note that solutions with $E < 0$ yield a purely imaginary k , and therefore the eigenfunctions are not even bounded. Hence they do not represent physical states. The only allowed states for the free particle are the ones with positive (kinetic) energy.

We can now readily check that the eigenfunctions of the continuous spectrum found above satisfy the orthogonality and completeness relations discussed in Sects. 4.1.1, and 4.1.2:

$$\psi(x) = \int_0^\infty dk \left[\tilde{f}(k)e^{ikx} + \tilde{f}(-k)e^{-ikx} \right], \quad (4.18)$$

$$\int dx e^{ix(k-k')} = 2\pi\delta(k-k'). \quad (4.19)$$

Eq. (4.18) is nothing but the well-known fact that normalizable functions admit a Fourier transform. On the other hand Eq. (4.19) is a property of the Dirac delta, that you can find in the box below ¹.

Note that Eq. (4.19) implies that the properly normalized eigenfunctions are:

$$u_{k,1}(x) = \frac{1}{\sqrt{2\pi}} e^{ikx}, \quad u_{k,2}(x) = \frac{1}{\sqrt{2\pi}} e^{-ikx}, \quad k = \sqrt{\frac{2mE}{\hbar^2}}. \quad (4.20)$$

¹You will see again both these results in one of the Mathematics courses

Mathematical aside

The delta function is the natural extension of the familiar Kronecker delta δ_{ij} to the case of continuous variables; it is defined as:

$$\delta(x) = 0, \text{ if } x \neq 0, \quad \delta(0) = \infty, \quad (4.21)$$

and

$$\int_a^b dx \delta(x)g(x) = \begin{cases} g(0), & \text{if } a < 0 < b, \\ 0, & \text{otherwise,} \end{cases} \quad (4.22)$$

for any continuous function $g(x)$.

The following properties are very useful when dealing with delta functions.

$$\int_a^b dx \delta(x - c)g(x) = \begin{cases} g(c), & \text{if } a < c < b, \\ 0, & \text{otherwise;} \end{cases} \quad (4.23)$$

$$\delta(-x) = \delta(x); \quad (4.24)$$

$$\delta(ax) = \frac{1}{|a|}\delta(x); \quad (4.25)$$

$$g(x)\delta(x - y) = g(y)\delta(x - y); \quad (4.26)$$

$$x\delta(x) = 0; \quad (4.27)$$

$$\frac{d}{dx}\theta(x) = \delta(x), \quad \text{where } \theta(x) = \begin{cases} 1, & \text{if } x \geq 0, \\ 0, & \text{if } x < 0; \end{cases} \quad (4.28)$$

$$\delta(g(x)) = \sum_{i=1}^r \frac{1}{|g'(x_i)|} \delta(x - x_i), \quad \text{where } g(x_i) = 0; \quad (4.29)$$

$$\int_0^\infty dx g(x)\delta(x) = \frac{1}{2}g(0). \quad (4.30)$$

The delta function can be defined as the limit of ordinary functions, here are some examples, that you are encouraged to sketch to visualize what happens.

$$\delta(x) = \lim_{\epsilon \rightarrow 0} \frac{1}{\sqrt{\pi\epsilon}} e^{-x^2/\epsilon^2}, \quad (4.31)$$

$$= \lim_{L \rightarrow \infty} \frac{\sin(Lx)}{\pi x}, \quad (4.32)$$

$$= \lim_{L \rightarrow \infty} \frac{\sin^2(Lx)}{\pi Lx^2}, \quad (4.33)$$

$$= \lim_{\epsilon \rightarrow 0} \frac{1}{\pi} \frac{\epsilon}{\epsilon^2 + x^2}. \quad (4.34)$$

The following representation for the delta function is very useful:

$$\int_{-\infty}^{\infty} dx e^{ix(k-k')} = 2\pi\delta(k-k'). \quad (4.35)$$

4.1.4 Spectral decomposition

For a generic operator \hat{O} , the eigenvalue spectrum can be made of discrete eigenvalues O_n and continuous values f . In this case we need to consider both the eigenfunctions of the discrete spectrum and those of the continuous spectrum to have a basis to expand quantum states in.

$$\psi(x) = \sum_n c_n \psi_n(x) + \int df c(f) \psi_f(x) \quad (4.36)$$

$$|\psi\rangle = \sum_n \langle n|\psi\rangle |n\rangle + \int df \langle f|\psi\rangle |f\rangle. \quad (4.37)$$

Eq. (4.36) summarizes the completeness relation for both discrete and continuous spectra.

4.2 General properties

There are some nice theorems that can be *derived* from the postulates of the theory. They are logical consequences of the principles that we have set up, and do not require further assumptions. They yield further information on the dynamics of the systems that we wish to study.

4.2.1 Energy expectation value

The expectation value of the energy in a quantum state is given by:

$$\langle \Psi | \hat{H} | \Psi \rangle = \int dx \Psi(x, t)^* \left(\frac{\hat{P}^2}{2m} + \hat{V} \right) \Psi(x, t) \quad (4.38)$$

$$= -\frac{\hbar^2}{2m} \int dx \Psi(x, t) \frac{d^2}{dx^2} \Psi(x, t) + \int dx \Psi(x, t) V(x) \Psi(x, t) \quad (4.39)$$

$$= \frac{\hbar^2}{2m} \int dx \left| \frac{d}{dx} \Psi(x, t) \right|^2 + \int dx V(x) |\Psi(x, t)|^2 \quad (4.40)$$

$$\geq V_{\min} \int dx |\Psi(x, t)|^2 = V_{\min}, \quad (4.41)$$

i.e. the expectation value of the energy is always larger than the minimum of the potential energy:

$$\boxed{\langle E \rangle \geq V_{\min}}. \quad (4.42)$$

4.2.2 Ehrenfest's theorem

Using Eq. (3.35) the following equations can be proved:

$$\frac{d}{dt} \langle m\hat{X} \rangle = \langle \hat{P} \rangle, \quad (4.43)$$

$$\frac{d}{dt} \langle \hat{P} \rangle = - \left\langle \frac{d}{dX} V(\hat{X}) \right\rangle. \quad (4.44)$$

This result is known as Ehrenfest's theorem. It is interesting to note that the expectation values of the position and momentum operators satisfy the same equations of motion that we find in Newtonian mechanics. Eq. (4.43) is easily proven using the fact that $[m\hat{X}, \hat{P}^2/(2m)] = i\hbar\hat{P}$. In order to obtain Eq. (4.44) you need to use the following relation:

$$[\hat{P}, V(\hat{X})] = i\hbar \frac{d}{dX} V(\hat{X}), \quad (4.45)$$

where $(d/dx)V$ denotes the derivative of V with respect to its argument.

4.2.3 Degeneracy and oscillations

The results in Secs. 4.2.1, and 4.2.2 apply to systems in any number of dimensions. The following results instead are only valid for one-dimensional systems.

The discrete energy levels in a one-dimensional potential well are not degenerate.

Proof. Let us assume the contrary, namely that there is an energy value E for which we have two degenerate eigenstates ψ_1 and ψ_2 . The two eigenfunctions satisfy the same equation:

$$\psi_1'' = -\frac{2m}{\hbar^2} (E - V(x)) \psi_1, \quad (4.46)$$

$$\psi_2'' = -\frac{2m}{\hbar^2} (E - V(x)) \psi_2. \quad (4.47)$$

Multiplying the first equation by ψ_2 and the second one by ψ_1 , and subtracting them, we obtain:

$$\psi_2 \psi_1'' - \psi_1 \psi_2'' = 0. \quad (4.48)$$

Integrating with respect to x yields:

$$\psi_1'(x)\psi_2(x) - \psi_2'(x)\psi_1(x) = \text{const} . \quad (4.49)$$

Since the eigenstates must be normalizable, we deduce that $\psi_1 = \psi_2 = 0$ at $x = +\infty$, and hence $\text{const} = 0$,

$$\psi_1'(x)\psi_2(x) - \psi_2'(x)\psi_1(x) = 0 . \quad (4.50)$$

Integrating again with respect to x :

$$\log \psi_1 = \log \psi_2 + \text{const} , \quad (4.51)$$

i.e. ψ_1 and ψ_2 are proportional and therefore describe the *same* eigenstate, which contradicts the initial statement.

The second theorem states that

The wave function of the n -th discrete energy level has $n - 1$ zeroes.

In particular, the wave function of the ground state does not vanish, for any value of x .

We shall not prove this theorem explicitly.

4.3 Heisenberg representation

Let us conclude this lecture by discussing a different way to characterize the time evolution of a quantum system.

In the Schrödinger picture the state of the system evolves in time, while the operators are usually time-independent. The expectation value of an observable O at time t is given by:

$$\langle \Psi(t) | \hat{O} | \Psi(t) \rangle = \langle \Psi(0) | e^{i\hat{H}t/\hbar} \hat{O} e^{-i\hat{H}t/\hbar} | \Psi(0) \rangle , \quad (4.52)$$

where we have used Eq. (3.13) to express $|\Psi(t)\rangle$ as a function of $|\Psi(0)\rangle$.

We can now interpret Eq. (4.52) as the expectation value of a time-dependent observable:

$$\hat{O}_H(t) = e^{i\hat{H}t/\hbar} \hat{O} e^{-i\hat{H}t/\hbar} , \quad (4.53)$$

whose expectation is computed between time-independent states. This is called the *Heisenberg representation*.

The dynamics in the Heisenberg representation is dictated by an equation describing the time evolution of the operators:

$$i\hbar \frac{d}{dt} \hat{O}_H(t) = \left[\hat{O}, \hat{H} \right] . \quad (4.54)$$

Depending on the problem that you are trying to solve, one representation may be more effective than the other. The physical predictions however are clearly identical!

4.4 Summary

As usual, we summarize the main concepts introduced in this lecture.

- Completeness and orthogonality relations for the continuous spectrum.
- Free particle as an example of the above concepts.
- More properties of the solutions of Schrödinger equation.
- The Heisenberg representation.

Interlude A

Postulates

The Postulates of Quantum Mechanics

We can now try to summarize the minimal set of assumptions that we have discussed so far to set up quantum mechanics. There is no unique choice and you will find a variety of formulations in the various recommended texts. This is a quick review of the concepts discussed in lectures 1-4; it will not be presented in a lecture, but should be used as a reference for the basic concepts. The rest of the course will present further developments of quantum mechanics that rely on these postulates.

Postulate 1:

Every possible physical state of a given system corresponds to some wavefunction $\Psi(x, t)$ that is a single-valued function of the parameters of the system and of time, and from which all possible predictions of the physical properties of the system can be obtained.

Note: the parameters could be, for example, coordinates but may also refer to internal variables such as ‘spin’.

Postulate 2:

Every observable may be represented by a Hermitean operator. To each such observable, \mathcal{A} , there corresponds an operator, \hat{A} , with a complete orthonormal set of eigenfunctions, $\{u_i(x)\}$, and a corresponding set of real eigenvalues, $\{A_i\}$:

$$\hat{A} u_i(x) = A_i u_i(x)$$

The only possible values which any measurement of \mathcal{A} can yield are the eigenvalues A_1, A_2, A_3, \dots

Notes:

- Orthonormality means as usual that

$$\int_{-\infty}^{\infty} u_i^*(x) u_j(x) dx = \delta_{ij}$$

- completeness means that an arbitrary wavefunction $\Psi(x, t)$ can be expanded as:

$$\Psi(x, t) = \sum_i c_i(t) u_i(x)$$

with coefficients c_i given by orthogonal projection:

$$c_i(t) = \int_{-\infty}^{\infty} u_i^*(x) \Psi(x, t) dx$$

The set of functions $\{u_i(x)\}$ is referred to as *the eigenbasis of \hat{A}* .

- the eigenvalues of \hat{A} may be discrete or continuous.

Postulate 3:

If the observable \mathcal{A} is measured on a system which, immediately prior to the measurement, is in the state $\Psi(x, t)$ then the strongest predictive statement that can be made about the result is

$$P(A_j), \text{ the probability of getting } A_j = \left| \int_{-\infty}^{\infty} u_j^*(x) \Psi(x, t) dx \right|^2 = |c_j(t)|^2$$

Notes:

- measurements are assumed to be *ideal*, i.e. to yield a single, errorless real number;
- the integral $\int_{-\infty}^{\infty} u_j^*(x) \Psi(x, t) dx$ is sometimes called an *overlap integral*;
- in general, *we cannot predict with certainty the outcome of a measurement*; only in the special case where $\Psi(x, t)$ coincides with an eigenfunction of \hat{A} , for example, $u_k(x)$ at the instant t , in which case

$$c_j(t) = \int_{-\infty}^{\infty} u_j^*(x) u_k(x) dx = \delta_{jk}$$

so that A_k will be obtained with probability 1;

- a measurement of \mathcal{A} on each of two identically prepared systems, both in the same quantum state Ψ , will not necessarily yield the same result.

Successive Measurements

What can we say about the state of a system *after* making a measurement of \mathcal{A} on it? Suppose that the result of our measurement was A_k . Then it is plausible that were we to immediately remeasure \mathcal{A} , we should get the same result A_k . Postulate 3 asserts that we can only be *certain* to get the result A_k if the system is described by the eigenfunction u_k corresponding to the eigenvalue A_k .

Postulate 4:

A measurement of an observable \mathcal{A} generally causes a drastic, uncontrollable change in the state of the system. Regardless of the form of $\Psi(x, t)$ just before the measurement, immediately after the measurement the wavefunction will coincide with the eigenfunction of \hat{A} corresponding to the eigenvalue obtained in the measurement of \mathcal{A} .

Notes:

- this is sometimes referred to as the *collapse of the wavefunction*; we also speak of *forcing the system into an eigenstate*;
- we have assumed that the eigenvalues and eigenfunctions are in 1-1 correspondence i.e. that there is no degeneracy;
- Postulate 3 guarantees that if, after measurement of \mathcal{A} , the wavefunction coincides with $u_k(x)$, then the probability of getting A_k is unity if we immediately remeasure \mathcal{A} ;
- if the wavefunction, $\Psi(x, t)$ before the measurement does not coincide with an eigenfunction of \hat{A} , then the observable \mathcal{A} cannot be said to have a value in the state $\Psi(x, t)$;
- more generally, we speak of a series of *successive measurements* being made, if the state of the system immediately prior to the $(n + 1)$ th measurement (of the same, or some other, observable) is that which resulted from the n th measurement, in contrast to the case of *repeated measurements* which are always made with the system in the same state immediately prior to each measurement.

Postulate 5:

The time development of a quantum system is determined by the Time-Dependent Schrödinger Equation:

$$\hat{H} \Psi(x, t) = i\hbar \frac{\partial}{\partial t} \Psi(x, t)$$

where the Hamiltonian operator, \hat{H} , is formed from the corresponding classical Hamiltonian function by operator substitution, and represents the total energy of the system.

Notes:

- \hat{H} possesses a complete orthonormal set of eigenfunctions $\{u_n(x)\}$ and a corresponding set of real eigenvalues $\{E_n\}$;
- if $\Psi(x, 0)$ is normalised to 1 then $\Psi(x, t)$ is also normalised to 1 for all t .

Lecture 5

Potentials

5.1 Potentials

In this lecture we will solve Schrödinger's equation for some simple one-dimensional potentials, and discuss the physical interpretation of the results.

The Schrödinger equation for these problems is:

$$i\hbar \frac{\partial}{\partial t} \Psi(x, t) = -\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} \Psi(x, t) + V(x) \Psi(x, t). \quad (5.1)$$

Thus the time-independent Schrödinger equation is:

$$-\frac{\hbar^2}{2m} \frac{d^2}{dx^2} \psi(x) + V(x) \psi(x) = E \psi(x). \quad (5.2)$$

This is a second order differential equation for the function ψ , which has two linearly independent solutions. Remember that ψ must be single-valued and continuous, and its first derivative with respect to x must also be continuous if the potential is not singular - see Sect. 3.1.3.

5.1.1 Potential step

The simplest case we can think of is a *potential step*:

$$V(x) = \begin{cases} 0, & \text{for } x < 0, \\ V > 0, & \text{for } x > 0, \end{cases} \quad (5.3)$$

as shown in Fig. 5.1.

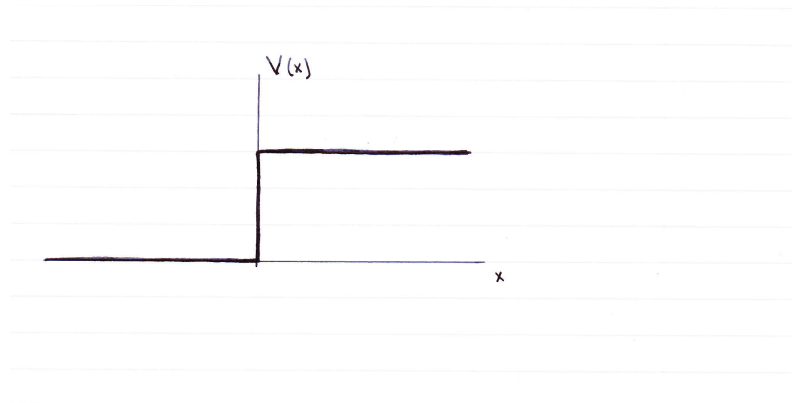


Figure 5.1: Potential step.

Note that the energy eigenvalues are always larger than the minimum of the potential, and hence in this case $E > 0$. There are two distinct cases that need to be analyzed.

Case 1. We shall start from the case where the total energy E is greater than V . We can then rewrite Eq. (5.2) as:

$$\begin{cases} \psi'' = -(p/\hbar)^2\psi(x), & \text{for } x < 0, \\ \psi'' = -(\bar{p}/\hbar)^2\psi(x), & \text{for } x > 0, \end{cases} \quad (5.4)$$

where we have introduced the real numbers:

$$p = \sqrt{2mE}, \quad \bar{p} = \sqrt{2m(E - V)}, \quad (5.5)$$

and the double prime symbol denotes two derivatives with respect to the position x .

In classical mechanics when a particle moving from left to right arrives at $x = 0$ its potential energy increases. In order for energy to be conserved, the kinetic energy of the particle has to decrease, i.e. the particle continues to move to the right, at a smaller velocity. The classical momenta of the particle to the left and to the right of the well are given respectively by p and \bar{p} . There is clearly no reflection when $E > V$ in classical mechanics.

The solutions of the eigenvalue equation for the energy in the quantum system are :

$$\psi(x) = \begin{cases} Ae^{ipx/\hbar} + Be^{-ipx/\hbar}, & \text{for } x < 0, \\ Ce^{i\bar{p}x/\hbar} + De^{-i\bar{p}x/\hbar}, & \text{for } x > 0. \end{cases} \quad (5.6)$$

They represent a superposition of plane waves travelling to the right and to the left.

We want to consider the case of an incident particle from the left that hits the barrier. Hence for $x > 0$ we do not expect to have a wave travelling to the left, and we can set $D = 0$.

We are now left with three unknown coefficients in Eq. (5.6), namely A, B , and C . One of them determines the overall normalization of the wave function, while the other two can be determined imposing the continuity of the wave function and its derivative at the origin:

$$\lim_{x \rightarrow 0^-} \psi(x) = \lim_{x \rightarrow 0^+} \psi(x), \quad (5.7)$$

$$\lim_{x \rightarrow 0^-} \psi'(x) = \lim_{x \rightarrow 0^+} \psi'(x). \quad (5.8)$$

You can readily check that:

$$\text{Eq. (5.7)} \Rightarrow A + B = C \quad (5.9)$$

$$\text{Eq. (5.8)} \Rightarrow p(A - B) = \bar{p}C. \quad (5.10)$$

The system can be easily solved:

$$B = \frac{p - \bar{p}}{p + \bar{p}}A, \quad (5.11)$$

$$C = \frac{2p}{p + \bar{p}}A. \quad (5.12)$$

Let us now discuss the physical content of this solution. To the right of the step, $x > 0$, we have a wave propagating to the right. C is the relative amplitude for the particle to be in the transmitted wave. To the left of the step, $x < 0$, we find a superposition of the incident wave and a *reflected* wave. The relative amplitudes for the particle to be in each of these states are respectively A and B . We have found a phenomenon that has no classical analogue: there is a possibility for the particle to be reflected by the potential step even if its energy is greater than the height of the barrier!!

Case 2. Let us now discuss the case where $E < V$. In classical mechanics this situation corresponds to an incident particle that does not have enough kinetic energy to get past the potential barrier, and therefore is reflected at $x = 0$. Formally the eigenvalue equations are the same as Eq. (5.4), but now \bar{p} is a purely imaginary number. In order to have a probability density that is bounded for $x > 0$, we need to choose the solution that yields a decreasing exponential:

$$\bar{p} = i\sqrt{2m(V - E)} \equiv i\tilde{p}, \quad (5.13)$$

$$\psi(x) = Ce^{-\tilde{p}x/\hbar}. \quad (5.14)$$

Note that in this case:

$$B = \frac{p - i\tilde{p}}{p + i\tilde{p}}A, \quad (5.15)$$

and therefore $|A|^2 = |B|^2$. The magnitudes of the incident and the reflected waves are the same. The two amplitudes are related by a simple phase shift:

$$\frac{p - i\tilde{p}}{p + i\tilde{p}} = -e^{2i\delta(E)}. \quad (5.16)$$

The phase shift can be expressed as a function of p and \tilde{p} :

$$\tilde{p} = p \cot \delta, \quad (5.17)$$

which shows that $\delta \rightarrow 0$ as $E \rightarrow 0$.

It is crucial to remark here that there is a non-vanishing probability to find the particle in the classically forbidden region $x < 0$. However this probability decays exponentially as x is increased. The typical range in which we can expect to find the particle is of the order of \hbar/p .

In the region where $E > V(x)$, the wave function has an oscillatory behaviour, and therefore a constant amplitude. On the other hand, in the classically forbidden region the amplitude of the wave function falls off exponentially.

5.1.2 Tunneling

We have seen in the previous Section that a quantum particle can access regions that are classically forbidden. This fact leads to a very important phenomenon called *tunneling*, which is characteristic of quantum mechanics.

A step potential of height V and size a as shown in Fig. 5.2 yields a simple example of tunneling that we can solve analytically.

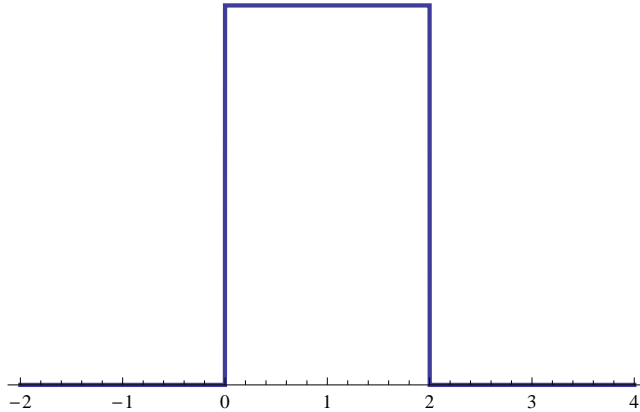


Figure 5.2: Potential step of height V and size a ($a = 2$ in this example) that allows the system to tunnel from one side to the other.

We are looking for solutions that would not penetrate the barrier classically, hence $0 < E < V$. In this case, following the method used in the previous Section, we can write the most general function for the wave function representing a wave incident from the left:

$$\psi(x) = \begin{cases} Ae^{ipx/\hbar} + Be^{-ipx/\hbar}, & \text{for } x < 0, \\ Ce^{-\tilde{p}x/\hbar} + De^{\tilde{p}x/\hbar}, & \text{for } 0 < x < a, \\ AS(E)e^{ip(x-a)/\hbar}, & \text{for } x > a. \end{cases} \quad (5.18)$$

Note that in this case we can keep the exponentially growing solution in the interval $0 < x < a$. Since this interval is finite, the wave function remains bounded. The fact that we have an incident wave from the left is encoded in the fact that there is no left-propagating wave for $x > a$, i.e. no term of the form $e^{-ipx/\hbar}$.

In this case we have a system with 5 unknown variables A , B , C , D , and $S(E)$. We have to impose the continuity equations at $x = 0$, and $x = a$. The latter are four equations, which enable us to determine the unknown coefficients up to an overall normalization. In particular here we are interested in computing the transition amplitude $S(E)$, i.e. the ratio of the probability amplitude for the outgoing wave at $x > a$ to the probability amplitude for

the incoming wave at $x < 0$. The computation is not difficult, but rather lengthy, and will be discussed in problem sheet 3.

The solution for the transmission amplitude is:

$$S(E) = \frac{2ip\tilde{p}}{(p^2 - \tilde{p}^2) \sinh(\frac{\tilde{p}a}{\hbar}) + 2ip\tilde{p} \cosh(\frac{\tilde{p}a}{\hbar})}. \quad (5.19)$$

The transmissivity T is defined as the probability for the particle to tunnel through the barrier:

$$T = |S(E)|^2 = \left[1 + \frac{\sinh^2(\tilde{p}a/\hbar)}{4(E/V)(1 - E/V)} \right]^{-1}. \quad (5.20)$$

This is a monotonically increasing function of E , as long as $E < V$. On the other hand, if the energy of the state is fixed, we see from Eq. (5.20) that T decreases exponentially with the size of the barrier a :

$$T \propto \exp \left[-2\sqrt{2m(V - E)}a/\hbar \right]. \quad (5.21)$$

Tunneling is a phenomenon that has no classical analogue. Its experimental observation is one of the many strong confirmations that our framework yields correct predictions. Tunneling has numerous applications in physics at the atomic and subatomic scale.

1. The α decay of nuclei can be modeled as a tunneling process. The α particle inside the nucleus can be described as quantum mechanical system with the potential depicted in Fig. 5.3. A particle with energy E would be confined inside the nucleus - i.e. inside the dip in the potential - according to classical mechanics. However, in quantum mechanics, the particle has a finite probability of tunneling to the right.

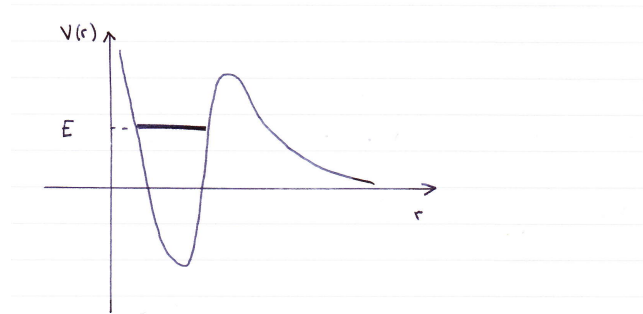


Figure 5.3: Potential felt by an α particle in the nucleus.

2. In solid state physics, if we separate two metals with a thin insulator layer, and apply a voltage across the metals, then the insulator will act as a potential barrier. Any current observed in such an experiment is due to the tunneling of the electrons across the potential barrier.

3. The scanning tunneling microscope allows individual atoms at the surface of a metal to be imaged. The layer of air between the surface of the metal and the STM's needle can be seen as a potential barrier that the electrons need to tunnel through. Since the tunneling probability depends exponentially on the distance between the STM's needle and the surface, one can determine such distance very precisely by measuring the current of electrons that are actually tunneling.

5.2 Summary

As usual, we summarize the main concepts introduced in this lecture.

- Solution of the Schrödinger equation for simple potentials.
- Potential step and continuity equations.
- Reflected and transmitted waves.
- Tunneling.
- Experimental observation of tunneling phenomena.

Lecture 6

More potentials

6.1 Introduction

Finally, in this lecture we are going to discuss potential wells. We are going to review briefly the infinite potential well, and then the finite one.

In the examples that we have seen in the previous lecture, the energy could take any value from zero to infinity - these are examples of *continuous spectra*. In the case of potential wells, we will find, in addition, bound states with only discrete values of the energy - i.e. states that are labelled by an integer.

6.2 Infinite potential well

This is a simple problem, that we have already seen in problem sheet 2. Let us review its solution, breaking it down into logical steps.

Physical setting The system we consider here is a single particle that evolves in the potential

$$V(x) = \begin{cases} 0, & \text{for } -\frac{a}{2} < x < \frac{a}{2}, \\ \infty, & \text{otherwise.} \end{cases} \quad (6.1)$$

The potential is sketched in Fig. 6.1. Note that the infinite potential well is completely specified by its width a . Therefore we expect the physics of this system to depend on a , and on the particle mass m .

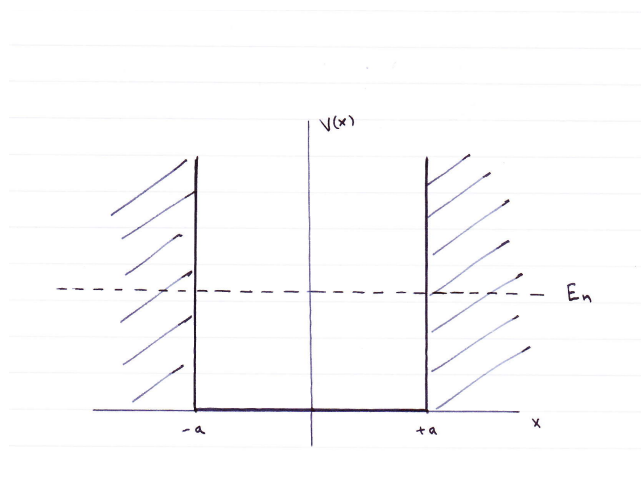


Figure 6.1: Infinite well potential.

As discussed in the previous lecture, if the walls were of finite height $V > E$, the wave function outside the wall would be of the form $e^{-\tilde{p}|x|/\hbar}$, and we would need to impose continuity at the boundary. We will consider this case in the next section. As the height of the potential barrier tends to infinity, we have that $\tilde{p} \rightarrow \infty$, and hence the wave function outside the potential $e^{-\tilde{p}|x|/\hbar} \rightarrow 0$. In other words we expect the wave function to vanish outside the potential well, thereby implying that the probability of finding the particle outside the well is exactly zero: the particle is confined inside the container.

We want to find the stationary states of this system, i.e. we want to solve the time-independent Schrödinger equation.

Eigenvalue equation In order to find the stationary states we need to solve the eigenvalue problem for the Hamiltonian, which turns the physical problem above into a well-defined mathematical problem. We need to find the solutions of the free particle equation:

$$-\frac{\hbar^2}{2m} \frac{d^2}{dx^2} \psi(x) = E\psi(x), \quad (6.2)$$

with the boundary conditions:

$$\psi(a/2) = \psi(-a/2) = 0. \quad (6.3)$$

All the information about the system is encoded in Eqs. (6.2), (6.3). In order to extract this information we need to solve the equation.

Solution of the eigenvalue problem Setting $k = \sqrt{\frac{2mE}{\hbar^2}}$, Eq. (6.2) becomes the familiar equation:

$$\psi''(x) = -k^2\psi(x). \quad (6.4)$$

The generic solution is a linear superposition of plane waves:

$$\psi(x) = Ae^{ikx} + Be^{-ikx}. \quad (6.5)$$

So far, the solution of the eigenvalue problem looks very similar to the solutions we discussed for the potential barriers in the previous lecture. The peculiar features of the infinite well are encoded in the boundary conditions specified in Eq. (6.3). The latter can be written as:

$$\psi(a/2) = Ae^{ika/2} + Be^{-ika/2} = 0, \quad (6.6)$$

$$\psi(-a/2) = Ae^{-ika/2} + Be^{ika/2} = 0, \quad (6.7)$$

Eq. (6.6) yields

$$A = -Be^{-ika}.$$

Replacing this result in Eq. (6.7) we obtain:

$$\sin(ka) = 0, \quad (6.8)$$

where we see that the boundary conditions actually restrict the possible values of k , and hence of the possible energy values. In order to satisfy Eq. (6.8), we need:

$$ka = n\pi, \quad n \text{ integer}. \quad (6.9)$$

Hence, finally, the solutions of the time-independent Schrödinger equation in the infinite well are:

$$u_n(x) = \begin{cases} \sin\left(\frac{n\pi x}{a}\right), & n \text{ even}, \\ \cos\left(\frac{n\pi x}{a}\right), & n \text{ odd}, \end{cases} \quad (6.10)$$

where n is an integer from 1 to ∞ . We see from the explicit expression of the solutions that $n = 0$ is excluded because the wave function would vanish identically.

Physical interpretation We can now look at our solutions and discuss the physical properties that are embodied in the mathematical expressions.

The values of the energy are:

$$E_n = n^2 \frac{\pi^2 \hbar^2}{2ma^2}, \quad n = 1, 2, \dots \quad (6.11)$$

Solutions are labelled by an integer: this is an example of a discrete spectrum.

Note that the eigenfunctions u_n are also eigenstates of the momentum operator, the values of the momentum are:

$$p_n = n \frac{\pi \hbar}{a}. \quad (6.12)$$

Note that the lowest value of the energy is greater than zero. This is a consequence of the uncertainty principle. The particle is confined inside a box of size a . Hence the maximum uncertainty in the determination of its position is a , which implies that the minimum uncertainty in the value of the momentum is $\Delta p \geq 2\pi\hbar/a$. As a consequence the particle has a kinetic energy

$$E \sim (\Delta p)^2 / (2m). \quad (6.13)$$

The kinetic energy that stems from the uncertainty principle is called *zero point energy*. A striking manifestation is the existence of liquid Helium all the way down to $T = 0$. In order to have solid Helium, we need to localize the particles in a lattice. This localization entails a zero point energy, which becomes large for light particles, as shown by Eq. (6.13). In the case of Helium the zero point energy is so large that the interatomic forces can no longer constrain the particles in a lattice, and the system remains in a liquid state.

6.3 Parity

By inspecting the energy eigenstates in Eq. (6.10) we see that

$$u_n(-x) = u_n(x), \quad n \text{ odd}, \quad (6.14)$$

$$u_n(-x) = -u_n(x), \quad n \text{ even}. \quad (6.15)$$

$$(6.16)$$

Under the parity operations defined in Eq. (3.26) the eigenfunctions $u_n(x)$ are multiplied by a constant $(-1)^{n+1}$, i.e. they are eigenstate of the parity operator \mathcal{P} , with eigenvalue $(-1)^{n+1}$. The eigenvalue is called the parity of the state. States with parity $(+1)$ are called *even*, states with parity (-1) are called *odd*.

For a potential that is symmetric under parity, you can readily prove that:

$$\hat{\mathcal{P}}\hat{H}\psi(x) = \hat{H}\hat{\mathcal{P}}\psi, \quad (6.17)$$

for any wave function ψ , which is the same as saying that $[\hat{\mathcal{P}}, \hat{H}] = 0$.

As a consequence of the commutation relation, for each eigenstate of the Hamiltonian u_n , we can construct another eigenstate of \hat{H} with the *same* eigenvalue by acting with $\hat{\mathcal{P}}$ on it:

$$\begin{aligned} \hat{H}\hat{\mathcal{P}}u_n(x) &= \hat{\mathcal{P}}\hat{H}u_n(x) \\ &= \hat{\mathcal{P}}(E_n u_n(x)) \\ &= E_n \hat{\mathcal{P}}u_n(x). \end{aligned} \quad (6.18)$$

We can then construct the states

$$(1 \pm \hat{\mathcal{P}})u_n(x), \quad (6.19)$$

which are simultaneous eigenstates of \hat{H} and $\hat{\mathcal{P}}$. Notice that the result above does not necessarily imply that the energy levels are doubly-degenerate, since one of the linear combinations defined in Eq. (6.19) could vanish. The main result to remember here is the following.

If the Hamiltonian is symmetric under parity, i.e. if $[\hat{\mathcal{P}}, \hat{H}] = 0$, then there exists a set of eigenfunctions of \hat{H} that are also eigenstates of $\hat{\mathcal{P}}$. Or, equivalently, we can find a complete basis of eigenstates of \hat{H} made of either even or odd functions of the position x .

6.4 Finite potential well

The final example that we are going to discuss is the finite potential well, which corresponds to the potential:

$$V(x) = \begin{cases} -V_0, & \text{for } |x| < a/2, \\ 0, & \text{otherwise.} \end{cases} \quad (6.20)$$

Once again we have a potential that is symmetric under parity, $V(-x) = V(x)$. Therefore we can choose to look for solutions of the energy eigenvalue equation that are also parity eigenstates.

As discussed previously, energy eigenvalues must be larger than $-V_0$. This means that we can have states with $-V_0 < E < 0$; they are called *bound states*: the wave function for these states decays exponentially at large $|x|$, so that the probability of finding the particle outside the well becomes rapidly very small. On the other hand, states with $E > 0$ correspond to incident plane waves that are distorted by the potential. We will concentrate on the bound states here.

The Schrödinger equation reads:

$$\psi'' = \begin{cases} -\frac{2m}{\hbar^2} E \psi \\ -\frac{2m}{\hbar^2} (E + V_0) \psi. \end{cases} \quad (6.21)$$

The main difference between the infinite and the finite well comes from the fact that the wave function does not have to vanish outside the classically allowed region $|x| < a/2$. As we discussed before, the wave function for $|x| > a/2$ will decay exponentially.

For $-V_0 < E < 0$, the even parity solutions are of the form:

$$\psi(x) = \begin{cases} A \cos(px/\hbar), & |x| < a/2 \\ C e^{-\bar{p}x/\hbar}, & x > a/2 \\ C e^{\bar{p}x/\hbar}, & x < -a/2, \end{cases} \quad (6.22)$$

while the odd parity solutions are:

$$\psi(x) = \begin{cases} A \sin(px/\hbar), & |x| < a/2 \\ C e^{-\bar{p}x/\hbar}, & x > a/2 \\ -C e^{\bar{p}x/\hbar}, & x < -a/2. \end{cases} \quad (6.23)$$

As usual we have introduced the momenta:

$$p = \sqrt{2m(E + V_0)}, \quad \bar{p} = \sqrt{-2mE}. \quad (6.24)$$

(Remember that we are looking for bound states, and hence $E < 0$.)

Once again the values of E cannot be arbitrary, they are determined when we impose the continuity condition. Imposing continuity of the wave function and its derivative in the even parity sector, we obtain:

$$\begin{aligned} A \cos\left(\frac{pa}{2\hbar}\right) &= C e^{-\bar{p}a/(2\hbar)}, \\ -\frac{p}{\hbar} A \sin\left(\frac{pa}{2\hbar}\right) &= -\frac{\bar{p}}{\hbar} C e^{-\bar{p}a/(2\hbar)}. \end{aligned}$$

Dividing the bottom equation by the top one, we obtain the quantization condition for the energy levels:

$$p \tan\left(\frac{pa}{2\hbar}\right) = \bar{p}. \quad (6.25)$$

Similarly in the odd parity sector we obtain:

$$p \cot\left(\frac{pa}{2\hbar}\right) = -\bar{p}. \quad (6.26)$$

The graphical solution of these equations is discussed in problem sheet 3.

The main physical features of this system is that there is always at least one even bound state. For very small V_0 , i.e. for a shallow well, you can show that

$$E = -\frac{mV_0^2 a^2}{2\hbar^2}. \quad (6.27)$$

6.5 Summary

As usual, we summarize the main concepts introduced in this lecture.

- Detailed solution of the infinite potential well.
- Logical steps, mathematical tools, interpretation.
- Zero point energy.
- Potentials symmetric under parity.
- Using parity to solve the finite potential well.

Lecture 7

More dimensions

7.1 Introduction

In this lecture we generalize the concepts introduced so far to systems that evolve in more than one spatial dimension. While the generalization of the concepts is straightforward, we will encounter new features that were not present for one-dimensional systems. The three-dimensional formulation will allow us to discuss more realistic examples.

7.2 Quantum states

We discuss here the case of three spatial dimensions; positions in space are specified by e.g. three Cartesian coordinates (x, y, z) . All results can be reduced to the case of a two-dimensional space by dropping the dependence on the third coordinate z .

As in the one-dimensional case, quantum states are described by a wave function - cfr Eq. (1.2):

$$\boxed{\text{quantum state} \sim \Psi(\underline{r}, t),} \quad (7.1)$$

the wave function is now a function of the position vector: $\underline{r} = (x, y, z)$, hence it depends on three coordinates.

The modulo square of the wave function has the same probabilistic interpretation discussed in Lecture 1:

$$\boxed{|\Psi(\underline{r}, t)|^2 d\tau \text{ is the probability that a measurement of the position of the particle yields a result in the infinitesimal volume element } d\tau \text{ at } \underline{r} \text{ with } d\tau = dx dy dz \text{ in Cartesian coordinates or } r^2 \sin \theta dr d\theta d\phi \text{ in spherical polars.}}$$

Thus $|\Psi(\underline{r}, t)|^2$ is a *probability per unit volume*.

The normalisation condition becomes

$$\int_{\text{all space}} |\Psi(\underline{r}, t)|^2 d\tau = 1. \quad (7.2)$$

The physical meaning of this equation should be clear: the probability of finding the system *somewhere* in space must be equal to one.

Note that if the system under consideration is a two-dimensional system, then the position vector is a two-dimensional Euclidean vector, and all the integrals are computed over two-dimensional surfaces.

We can still use Dirac's notation to denote a state vector:

$$\Psi(\underline{r}, t) \longleftrightarrow |\Psi(t)\rangle. \quad (7.3)$$

The scalar product of two states now involves an integral over the whole volume:

$$\langle \Phi(t') | \Psi(t) \rangle = \int dt \Phi(\underline{r}, t')^* \Psi(\underline{r}, t),. \quad (7.4)$$

We can rewrite the normalization condition as:

$$\langle \Psi | \Psi \rangle = 1. \quad (7.5)$$

7.3 Observables

As far as observables are concerned, there is no conceptual difference between the one-dimensional systems that we discussed in the first part of the lectures, and systems in higher-dimensional spaces. Observables are in one-to-one correspondence with linear Hermitian operators acting on the wave functions.

Let us summarize the main features of the correspondence:

- The eigenvalues of the operator \hat{O} yield the possible outcomes of a measurement of the observable O .
- You can easily check that the orthogonality relations that we proved for the eigenfunction of a Hermitian operator in the one-dimensional case are still true.
- Likewise, the set of eigenfunctions of a Hermitian operator are a *complete* set and therefore any wave function can be expanded using these eigenfunctions as basis.

Example The eigenvalue equation for a Hermitian operator \hat{O} in a three-dimensional system is:

$$\hat{O}\psi_x(\underline{r}) = O_k\psi_k(\underline{r}). \quad (7.6)$$

Note that in this case the eigenfunctions depend on the vector \underline{r} . This is the only peculiarity of the three-dimensional case. The eigenvalues O_k are the possible values for O , the eigenfunctions $\psi_k(\underline{r})$ describe states where a measurement of O would yield O_k with probability one.

The completeness of the eigenfunctions is expressed by the fact:

$$\psi(\underline{r}) = \sum_k c_k \psi_k(\underline{r}), \quad (7.7)$$

for any generic quantum state $\psi(\underline{r})$. As before, $|c_k|^2$ yields the probability of getting the value O_k upon measuring O in the state ψ . The coefficients c_k are obtained from the scalar products:

$$c_k = \langle \psi_k | \psi \rangle, \quad (7.8)$$

where now the scalar product requires a three-dimensional integral, according to Eq. (7.4).

Using Dirac's notation, the definition of the Hermitian conjugate is formally unchanged:

$$\langle \phi | \hat{O}^\dagger | \psi \rangle = \left(\langle \psi | \hat{O} | \phi \rangle \right)^*. \quad (7.9)$$

If you want to rewrite the relation above using wave functions and integrals, you get:

$$\int d\tau \phi(\underline{r})^* \hat{O}^\dagger \psi(\underline{r}) = \left(\int d\tau \psi(\underline{r})^* \hat{O} \phi(\underline{r}) \right)^* . \quad (7.10)$$

Remember that now we need to integrate over three-dimensional space, $d\tau$ is the infinitesimal integration volume.

Example It is interesting to consider the generalization of the momentum operator defined in Eq. (2.29) for the one-dimensional system. The momentum is a three-dimensional vector \underline{P} which can be represented by its three components in a Cartesian reference frame (P_x, P_y, P_z) . Following our general prescription, to each component of the momentum we associate a Hermitean operator:

$$\hat{P}_x = -i\hbar \frac{\partial}{\partial x} , \quad (7.11)$$

$$\hat{P}_y = -i\hbar \frac{\partial}{\partial y} , \quad (7.12)$$

$$\hat{P}_z = -i\hbar \frac{\partial}{\partial z} . \quad (7.13)$$

Hence for instance:

$$\hat{P}_x \psi(\underline{r}) = -i\hbar \left. \frac{\partial \psi}{\partial x} \right|_{\underline{r}} . \quad (7.14)$$

A concise notation for the momentum operator is :

$$\underline{\hat{P}} = -i\hbar \underline{\nabla} . \quad (7.15)$$

We shall use the notation x_i , with $i = 1, 2, 3$ to denote the three components x, y, z respectively. Similarly $\frac{\partial}{\partial x_i}$ will be used to indicate $\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}$ respectively. Using a simple property of partial derivatives, we have:

$$\hat{P}_i \hat{P}_j \psi(\underline{r}) = -\hbar^2 \frac{\partial^2}{\partial x_i \partial x_j} \psi(\underline{r}) = -\hbar^2 \frac{\partial^2}{\partial x_j \partial x_i} \psi(\underline{r}) = \hat{P}_j \hat{P}_i \psi(\underline{r}) . \quad (7.16)$$

Different components of the momentum commute with each other, i.e. they can all be measured simultaneously.

Following the derivation that led to Eq. (2.39), you can readily prove the canonical commutation relations in three dimensions:

$$\left[\hat{X}_i, \hat{P}_j \right] = i\hbar \delta_{ij} . \quad (7.17)$$

Thus \hat{X} *does not* commute with \hat{P}_x , but *does* commute with \hat{P}_y and \hat{P}_z . These commutation relations lead to Heisenberg uncertainty relations involving the components of the position and the momentum of the particle:

$$\Delta X_i \cdot \Delta P_j \geq \frac{\hbar}{2} \delta_{ij}. \quad (7.18)$$

7.4 Dynamics

As for the one-dimensional systems the dynamics is determined by the Schrödinger equation:

$$i\hbar \frac{\partial}{\partial t} \Psi(\underline{r}, t) = \hat{H} \Psi(\underline{r}, t). \quad (7.19)$$

In three dimensions the hamiltonian for the system can be expressed as:

$$\begin{aligned} \hat{H} &= \hat{T} + \hat{V} \\ &= -\frac{\hbar^2}{2m} \nabla^2 + V(\hat{\underline{X}}), \end{aligned} \quad (7.20)$$

where

$$\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \quad (7.21)$$

is the Laplace operator. Thus the time evolution of a quantum state is found by solving the differential equation:

$$i\hbar \frac{\partial}{\partial t} \Psi(\underline{r}, t) = \left\{ -\frac{\hbar^2}{2m} \nabla^2 + V(\underline{r}) \right\} \Psi(\underline{r}, t). \quad (7.22)$$

As usual, we are going to solve the time-dependent Schrödinger equation by expanding the solution in eigenstates of the Hamiltonian. The latter are obtained by solving the eigenvalue problem for \hat{H} , i.e. the time-independent Schrödinger equation:

$$\hat{H} \psi_E(\underline{r}) = E \psi(\underline{r}). \quad (7.23)$$

Example The isotropic harmonic oscillator in 3 dimensions, for which the potential is

$$V(\underline{r}) = \frac{1}{2} m \omega^2 r^2 = \frac{1}{2} m \omega^2 (x^2 + y^2 + z^2) \quad (7.24)$$

The eigenvalue problem can be separated in Cartesian coordinates by writing:

$$u_n(\underline{r}) = X(x) \cdot Y(y) \cdot Z(z), \quad (7.25)$$

yielding

$$\left\{ -\frac{\hbar^2}{2m} \frac{1}{X} \frac{d^2 X}{dx^2} + \frac{1}{2} m\omega^2 x^2 \right\} + \left\{ -\frac{\hbar^2}{2m} \frac{1}{Y} \frac{d^2 Y}{dy^2} + \frac{1}{2} m\omega^2 y^2 \right\} + \left\{ -\frac{\hbar^2}{2m} \frac{1}{Z} \frac{d^2 Z}{dz^2} + \frac{1}{2} m\omega^2 z^2 \right\} = E_n. \quad (7.26)$$

Each term in braces must be equal to a constant, so we can write

$$E_{n_x} + E_{n_y} + E_{n_z} = E_n \quad (7.27)$$

with, for example,

$$\left\{ -\frac{\hbar^2}{2m} \frac{d^2 X}{dx^2} + \frac{1}{2} m\omega^2 x^2 X \right\} = E_{n_x} X, \quad (7.28)$$

which is just a 1-dimensional oscillator problem for which we know the solution:

$$E_{n_x} = \left(n_x + \frac{1}{2} \right) \hbar\omega, \quad n_x = 0, 1, 2, 3 \dots \quad (7.29)$$

$$X(x) = u_{n_x}(x) = C_{n_x} \exp(-\alpha^2 x^2/2) H_{n_x}(\alpha x). \quad (7.30)$$

Thus

$$E_n = \left(n_x + n_y + n_z + \frac{3}{2} \right) \hbar\omega, \quad n_x, n_y, n_z = 0, 1, 2, 3 \dots \quad (7.31)$$

$$\equiv \left(n + \frac{3}{2} \right) \hbar\omega, \quad n = 0, 1, 2, 3 \dots \quad (7.32)$$

7.5 Degeneracy

We see in the previous example a new feature, that we had not encountered in the one-dimensional case. The three-dimensional harmonic oscillator displays *degeneracy*, i.e. more than one eigenfunction correspond to the *same* eigenvalue.

Eq. (7.31) shows that a given value of E_n can arise in more than one way. For example:

n	n_x	n_y	n_z	g_n
0	0	0	0	1
1	1	0	0	3
	0	1	0	
	0	0	1	

g_n is the number of ways that a given value of n , and hence of the energy E_n , can arise, and is called the *degree of degeneracy*. We say, for example, that the $n = 1$ level is 3-fold degenerate, meaning that there are three distinct quantum states of the same energy, corresponding to the eigenfunctions

$$u_1(x) \cdot u_0(y) \cdot u_0(z), \quad u_0(x) \cdot u_1(y) \cdot u_0(z), \quad u_0(x) \cdot u_0(y) \cdot u_1(z). \quad (7.33)$$

7.6 Summary

As usual, we summarize the main concepts introduced in this lecture.

- Description of three-dimensional quantum states.
- Implementation of the observables as Hermitean operators.
- Momentum operator and commutation relations.
- Schrödinger equation, stationary states and separation of variables.
- Degeneracy.

Lecture 8

Angular momentum

8.1 Introduction

Now that we have introduced three-dimensional systems, we need to introduce into our quantum-mechanical framework the concept of *angular momentum*.

Recall that in classical mechanics angular momentum is defined as the vector product of position and momentum:

$$\underline{L} \equiv \underline{r} \times \underline{p} = \begin{vmatrix} \underline{i} & \underline{j} & \underline{k} \\ x & y & z \\ p_x & p_y & p_z \end{vmatrix}. \quad (8.1)$$

Note that the angular momentum is itself a *vector*. The three Cartesian components of the angular momentum are:

$$L_x = y p_z - z p_y, \quad L_y = z p_x - x p_z, \quad L_z = x p_y - y p_x. \quad (8.2)$$

8.2 Angular momentum operator

For a quantum system the angular momentum is an observable, we can *measure* the angular momentum of a particle in a given quantum state. According to the postulates that we have spelled out in previous lectures, we need to associate to each observable a Hermitean operator. We have already defined the operators \hat{X} and \hat{P} associated respectively to the position and the momentum of a particle. Therefore we can define the *operator*

$$\hat{L} \equiv \hat{X} \times \hat{P}, \quad (8.3)$$

where $\hat{P} = -i\hbar\nabla$. Note that in order to define the angular momentum, we have used the definitions for the position and momentum operators *and* the expression for the angular momentum in classical mechanics. Eq. (8.3) yields explicit expressions for the components of the angular momentum as differential operators:

$$\hat{L}_x = -i\hbar \left(y \frac{\partial}{\partial z} - z \frac{\partial}{\partial y} \right), \quad \hat{L}_y = -i\hbar \left(z \frac{\partial}{\partial x} - x \frac{\partial}{\partial z} \right), \quad \hat{L}_z = -i\hbar \left(x \frac{\partial}{\partial y} - y \frac{\partial}{\partial x} \right). \quad (8.4)$$

Eq. (8.4) can be economically rewritten as:

$$\hat{L}_i = -i\hbar \varepsilon_{ijk} x_j \frac{\partial}{\partial x_k}, \quad (8.5)$$

where we have to sum over the repeated indices.

Mathematical aside

In Eq. (8.5) we have used the same convention introduced in Lecture 7; we use:

$$x_1 = x, \quad x_2 = y, \quad x_3 = z, \quad (8.6)$$

to denote the three components of the position vector. The same convention is also used for the partial derivatives:

$$\frac{\partial}{\partial x_1} = \frac{\partial}{\partial x}, \quad \frac{\partial}{\partial x_2} = \frac{\partial}{\partial y}, \quad \frac{\partial}{\partial x_3} = \frac{\partial}{\partial z}. \quad (8.7)$$

In general the components of a vector \underline{V} can be labeled as:

$$V_1 = V_x, \quad V_2 = V_y, \quad V_3 = V_z. \quad (8.8)$$

The symbol ε_{ijk} denotes the totally antisymmetric unit tensor:

$$\varepsilon_{123} = \varepsilon_{231} = \varepsilon_{312} = 1, \quad \text{cyclic indices} \quad (8.9)$$

$$\varepsilon_{213} = \varepsilon_{132} = \varepsilon_{321} = -1, \quad \text{anticyclic indices} \quad (8.10)$$

Out of twenty-seven components, only the six above are actually different from zero. Check that you understand Eq. (8.5).

The following relations are useful:

$$\varepsilon_{ikl}\varepsilon_{imn} = \delta_{km}\delta_{ln} - \delta_{kn}\delta_{lm}, \quad (8.11)$$

$$\varepsilon_{ikl}\varepsilon_{ikm} = 2\delta_{lm}, \quad (8.12)$$

$$\varepsilon_{ikl}\varepsilon_{ikl} = 6. \quad (8.13)$$

Using the canonical commutation relations, Eq. (7.17), we can easily prove that:

$$[\hat{L}_x, \hat{L}_y] = i\hbar\hat{L}_z, \quad [\hat{L}_y, \hat{L}_z] = i\hbar\hat{L}_x, \quad [\hat{L}_z, \hat{L}_x] = i\hbar\hat{L}_y. \quad (8.14)$$

The proof of this statement is left as an exercise in problem sheet 4. Once again, it is useful to get familiar with the more compact notation:

$$\boxed{[\hat{L}_i, \hat{L}_j] = i\hbar\varepsilon_{ijk}\hat{L}_k}. \quad (8.15)$$

Example Instead of using the canonical commutation relations, we can derive the commutation relations between the components L_i using their representation as differential operators.

$$\begin{aligned}\hat{L}_x\hat{L}_y &= -\hbar^2 \left(y \frac{\partial}{\partial z} - z \frac{\partial}{\partial y} \right) \left(z \frac{\partial}{\partial x} - x \frac{\partial}{\partial z} \right) \\ &= -\hbar^2 \left\{ y \frac{\partial}{\partial x} + yz \frac{\partial^2}{\partial z \partial x} - yx \frac{\partial^2}{\partial z^2} - z^2 \frac{\partial^2}{\partial y \partial x} + zx \frac{\partial^2}{\partial y \partial z} \right\},\end{aligned}$$

whilst

$$\begin{aligned}\hat{L}_y\hat{L}_x &= -\hbar^2 \left(z \frac{\partial}{\partial x} - x \frac{\partial}{\partial z} \right) \left(y \frac{\partial}{\partial z} - z \frac{\partial}{\partial y} \right) \\ &= -\hbar^2 \left\{ zy \frac{\partial^2}{\partial x \partial z} - z^2 \frac{\partial^2}{\partial x \partial y} - xy \frac{\partial^2}{\partial z^2} + xz \frac{\partial^2}{\partial z \partial y} + x \frac{\partial}{\partial y} \right\}\end{aligned}$$

Noting the usual properties of partial derivatives

$$\frac{\partial^2}{\partial x \partial z} = \frac{\partial^2}{\partial z \partial x}, \quad \text{etc} \quad (8.16)$$

we obtain on subtraction the desired result:

$$[\hat{L}_x, \hat{L}_y] = \hbar^2 \left(x \frac{\partial}{\partial y} - y \frac{\partial}{\partial x} \right) = i\hbar \hat{L}_z. \quad (8.17)$$

Note that the Cartesian components of the angular momentum **do not** commute with each other. Following our previous discussion on compatible observables, this means that the components are **not** compatible observables. We cannot measure, for instance, L_x and L_y simultaneously, and we do not have a basis of common eigenfunctions of the two operators. Physically, this also implies that measuring one component of the angular momentum modifies the probability of finding a given result for the other two.

Angular momentum plays a central role in discussing *central potentials*, i.e. potentials that only depend on the radial coordinate r . It will also prove useful to have expression for the operators \hat{L}_x , \hat{L}_y and \hat{L}_z in spherical polar coordinates. Using the expression for the Cartesian coordinates as functions of the spherical ones, and the chain rule for the derivative, yields

$$\begin{aligned}\hat{L}_x &= i\hbar \left(\sin \phi \frac{\partial}{\partial \theta} + \cot \theta \cos \phi \frac{\partial}{\partial \phi} \right) \\ \hat{L}_y &= i\hbar \left(-\cos \phi \frac{\partial}{\partial \theta} + \cot \theta \sin \phi \frac{\partial}{\partial \phi} \right) \\ \hat{L}_z &= -i\hbar \frac{\partial}{\partial \phi}.\end{aligned}$$

8.3 Square of the angular momentum

Let us now introduce an operator that represents the square of the magnitude of the angular momentum:

$$\hat{L}^2 \equiv \hat{L}_x^2 + \hat{L}_y^2 + \hat{L}_z^2 = \sum_{i=1}^3 \hat{L}_i^2, \quad (8.18)$$

or, in spherical polar coordinates

$$\hat{L}^2 \equiv -\hbar^2 \left[\frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial}{\partial \theta} \right) + \frac{1}{\sin^2 \theta} \frac{\partial^2}{\partial \phi^2} \right]. \quad (8.19)$$

The importance of this observable is that *it is compatible with any of the Cartesian components of the angular momentum*;

$$[\hat{L}^2, \hat{L}_x] = [\hat{L}^2, \hat{L}_y] = [\hat{L}^2, \hat{L}_z] = 0. \quad (8.20)$$

Sample proof. Consider for instance the commutator $[\hat{L}^2, \hat{L}_z]$:

$$\begin{aligned} [\hat{L}^2, \hat{L}_z] &= [\hat{L}_x^2 + \hat{L}_y^2 + \hat{L}_z^2, \hat{L}_z] \quad \text{from the definition of } \hat{L}^2 \\ &= [\hat{L}_x^2, \hat{L}_z] + [\hat{L}_y^2, \hat{L}_z] + [\hat{L}_z^2, \hat{L}_z] \\ &= [\hat{L}_x^2, \hat{L}_z] + [\hat{L}_y^2, \hat{L}_z] \quad \text{since } \hat{L}_z \text{ commutes with itself} \\ &= \hat{L}_x \hat{L}_x \hat{L}_z - \hat{L}_z \hat{L}_x \hat{L}_x + \hat{L}_y \hat{L}_y \hat{L}_z - \hat{L}_z \hat{L}_y \hat{L}_y. \end{aligned}$$

We can use the commutation relation $[\hat{L}_z, \hat{L}_x] = i\hbar \hat{L}_y$ to rewrite the first term on the RHS as

$$\hat{L}_x \hat{L}_x \hat{L}_z = \hat{L}_x \hat{L}_z \hat{L}_x - i\hbar \hat{L}_x \hat{L}_y,$$

and the second term as

$$\hat{L}_z \hat{L}_x \hat{L}_x = \hat{L}_x \hat{L}_z \hat{L}_x + i\hbar \hat{L}_y \hat{L}_x.$$

In a similar way, we can use $[\hat{L}_y, \hat{L}_z] = i\hbar \hat{L}_x$ to rewrite the third term as

$$\hat{L}_y \hat{L}_y \hat{L}_z = \hat{L}_y \hat{L}_z \hat{L}_y + i\hbar \hat{L}_y \hat{L}_x,$$

and the fourth term

$$\hat{L}_z \hat{L}_y \hat{L}_y = \hat{L}_y \hat{L}_z \hat{L}_y - i\hbar \hat{L}_x \hat{L}_y,$$

thus, on substituting in we find that

$$[\hat{L}^2, \hat{L}_z] = -i\hbar \hat{L}_x \hat{L}_y - i\hbar \hat{L}_y \hat{L}_x + i\hbar \hat{L}_y \hat{L}_x + i\hbar \hat{L}_x \hat{L}_y = 0.$$

QED

8.4 Eigenfunctions

The compatibility theorem tells us that \hat{L}^2 and \hat{L}_z thus have *simultaneous eigenfunctions*. These turn out to be the *spherical harmonics*, $Y_\ell^m(\theta, \phi)$. In particular, the eigenvalue equation for \hat{L}^2 is

$$\hat{L}^2 Y_\ell^m(\theta, \phi) = \ell(\ell + 1)\hbar^2 Y_\ell^m(\theta, \phi), \quad (8.21)$$

where $\ell = 0, 1, 2, 3, \dots$ and

$$Y_\ell^m(\theta, \phi) = (-1)^m \left[\frac{2\ell + 1}{4\pi} \frac{(\ell - m)!}{(\ell + m)!} \right]^{1/2} P_\ell^m(\cos \theta) \exp(im\phi), \quad (8.22)$$

with $P_\ell^m(\cos \theta)$ known as the *associated Legendre polynomials*. Some examples of spherical harmonics will be given below.

The eigenvalue $\ell(\ell + 1)\hbar^2$ is *degenerate*; there exist $(2\ell + 1)$ eigenfunctions corresponding to a given ℓ and they are distinguished by the label m which can take any of the $(2\ell + 1)$ values

$$m = \ell, \ell - 1, \dots, -\ell, \quad (8.23)$$

In fact it is easy to show that m labels the eigenvalues of \hat{L}_z . Since

$$Y_\ell^m(\theta, \phi) \sim \exp(im\phi), \quad (8.24)$$

we obtain directly that

$$\hat{L}_z Y_\ell^m(\theta, \phi) \equiv -i\hbar \frac{\partial}{\partial \phi} Y_\ell^m(\theta, \phi) = m\hbar Y_\ell^m(\theta, \phi), \quad (8.25)$$

confirming that the spherical harmonics are also eigenfunctions of \hat{L}_z with eigenvalues $m\hbar$.

Mathematical aside

A few examples of spherical harmonics are

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

8.5 Physical interpretation

We have arrived at the important conclusion that *angular momentum is quantised*. The square of the magnitude of the angular momentum can only assume one of the discrete set of values

$$\ell(\ell + 1)\hbar^2, \quad \ell = 0, 1, 2, \dots$$

and the z -component of the angular momentum can only assume one of the discrete set of values

$$m\hbar, \quad m = \ell, \ell - 1, \dots, -\ell$$

for a given value of ℓ .

ℓ and m are called the *angular momentum quantum number* and the *magnetic quantum number* respectively.

Finally a piece of jargon: we refer to a particle in a state with angular momentum quantum number ℓ as *having angular momentum ℓ* , rather than saying, more clumsily but accurately, that it has angular momentum of magnitude $\sqrt{\ell(\ell + 1)}\hbar$.

Lecture 9

Angular momentum - 2

9.1 Spectroscopic Notation

We saw in the last section that the only allowed values of the angular momentum quantum number, ℓ , are

$$\ell = 0, 1, 2, 3, \dots$$

For historical reasons, each possible ℓ value is denoted by a letter of the alphabet according to the following sequence:

s, p, d, f, g, h, \dots and so on alphabetically

Thus, for example, an $\ell = 0$ state is referred to as an s -wave state, and so on. The reason for the non-alphabetic order of the first four letters is that they are taken from the names of four series of spectral lines, called *sharp, principal, diffuse and fine*, observed in the spectra of alkali atoms such as Sodium.

ℓ	0	1	2	3	4	5
notation	s	p	d	f	g	h

Table 9.1: Spectroscopic notation for the first few ℓ values.

9.2 Dirac's notation

We can introduce a Dirac notation for the spherical harmonics. To each function $Y_\ell^m(\theta, \phi)$ we associate a ket $|\ell, m\rangle$, which represents a vector in an infinite-dimensional space.

The scalar product of two vectors can be written using the bra-ket notation:

$$\langle \ell', m' | \ell, m \rangle = \int_0^\pi \sin \theta d\theta \int_0^{2\pi} d\phi Y_{\ell'}^{m'}(\theta, \phi)^* Y_\ell^m(\theta, \phi). \quad (9.1)$$

The eigenvalue equations for \hat{L}^2 and \hat{L}_z can be written using Dirac's notation:

$$\hat{L}_z |\ell, m\rangle = \hbar m |\ell, m\rangle, \quad (9.2)$$

$$\hat{L}^2 |\ell, m\rangle = \hbar^2 \ell(\ell + 1) |\ell, m\rangle. \quad (9.3)$$

The orthonormality of the spherical harmonics is concisely written as:

$$\langle \ell', m' | \ell, m \rangle = \delta_{\ell', \ell} \delta_{m', m}. \quad (9.4)$$

The completeness of the spherical harmonics amounts to the fact that *any* function $f(\theta, \phi)$ has a unique expansion:

$$f(\theta, \phi) = \sum_{\ell, m} c_{\ell, m} Y_\ell^m(\theta, \phi), \quad (9.5)$$

where:

$$c_{\ell,m} = \langle \ell, m | f \rangle. \quad (9.6)$$

9.3 Experimental Evidence for Quantisation

Since the results that we have just obtained seem so bizarre, let us consider the experimental evidence supporting the contention that the component of angular momentum in a particular direction, say the z -direction, cannot take any old value but is restricted to be one of $(2\ell + 1)$ possible values.

We will consider in barest outline an experiment known as the Stern-Gerlach experiment, in which a beam of neutral paramagnetic particles (atoms) is deflected by a inhomogeneous magnetic field. The apparatus is schematically represented in Fig. 9.1. The inhomogeneous magnetic field between the poles sketched in the figure causes each particle in the beam to be deflected by an amount proportional to its magnetic moment. The original experiment was first performed by Stern and Gerlach in 1922 in Frankfurt. Note that Stern at that time was an assistant to Max Born, before the latter came to Edinburgh.

Let the vertical direction be the z -direction, then the deflection at the screen is proportional to the z -component of the magnetic moment, which in turn is proportional to the z -component of the angular momentum of each particle. Thus the deflection is a measure of the observable L_z .

In classical mechanics we would expect a continuous range of deflections, corresponding to the random orientations of the magnetic moments of the particles emerging from the oven. Instead the beam appears to be split into a number of discrete components. This phenomenon is sometimes referred to as *spatial quantisation*. The number of beams is $(2\ell + 1)$, so that if the theory discussed so far is correct there should always be an odd number of beams.

Actually, the original experiment of Stern and Gerlach observed just two beams, which would correspond to $\ell = \frac{1}{2}$, not one of the allowed values! We shall return to discuss the significance of this observation later. In the meantime we simply note that the experiment supports the idea that the component of angular momentum in a given direction can only take certain discrete values.

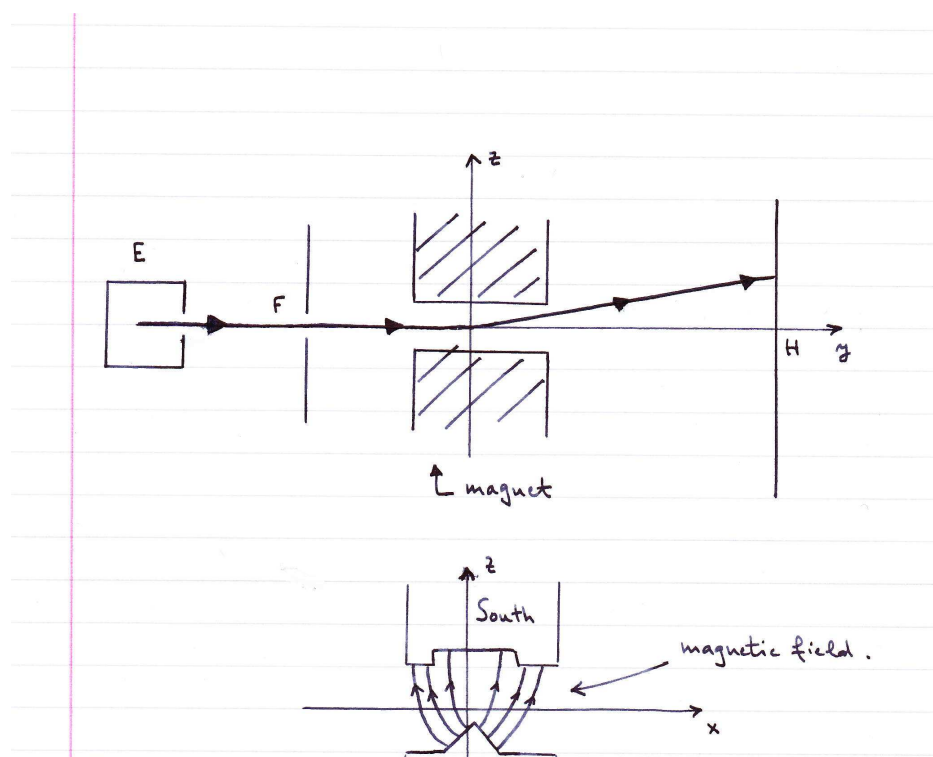


Figure 9.1: A beam of atoms leaves the oven E , is collimated by the slit F and is split by the inhomogeneous magnetic field between the pole pieces, before being detected at the screen. The deflection of the beam depends on the angular momentum of the particles.

9.4 Summary

As usual, we summarize the main concepts introduced in the last two lectures.

- Definition of the angular momentum. Angular momentum as a differential operator.
- Commutation relations between the components of the angular momentum.
- Square of the angular momentum, commutation relations.
- Eigenvalue problem for \hat{L}^2 and \hat{L}_z .
- Spectroscopic notation.
- Experimental evidence.

Lecture 10

Central potential

10.1 Introduction

We are now ready to study a generic class of three-dimensional physical systems. They are the systems that have a central potential, i.e. a potential energy that depends only on the distance r from the origin: $V(\underline{r}) = V(r)$. If we use spherical coordinates to parametrize our three-dimensional space, a central potential does *not* depend on the angular variables θ and ϕ .

An example of central potential is the Coulomb potential between electrically charged particles:

$$V(r) = -\frac{Ze^2}{4\pi\epsilon_0 r}. \quad (10.1)$$

10.2 Stationary states.

As usual, the dynamics of the system is encoded in the solutions of the time-independent Schrödinger equation:

$$\left[-\frac{\hbar^2}{2\mu} \nabla^2 + V(r) \right] u(\underline{r}) = Eu(\underline{r}). \quad (10.2)$$

We have denoted the mass of the particle by μ in order to avoid confusion with the magnetic quantum number m , which will appear in the solutions below.

It is clearly convenient to use spherical coordinates, and write the Laplacian as:

$$\nabla^2 = \frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial}{\partial r} \right) + \frac{1}{r^2 \sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial}{\partial \theta} \right) + \frac{1}{r^2 \sin^2 \theta} \frac{\partial^2}{\partial \phi^2}. \quad (10.3)$$

The key observation to solve the eigensystem in Eq. (10.2) is that we can use Eq. (8.19) and rewrite the Laplacian as:

$$\nabla^2 = \frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial}{\partial r} \right) - \frac{1}{\hbar^2 r^2} \hat{L}^2, \quad (10.4)$$

and thus the eigenvalue equation becomes:

$$\frac{\hbar^2}{2\mu} \left[-\frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial}{\partial r} \right) + \frac{1}{\hbar^2 r^2} \hat{L}^2 \right] u(r, \theta, \phi) = (E - V(r)) u(r, \theta, \phi). \quad (10.5)$$

Separation of variables. We have already seen that the eigenfunctions of the operator \hat{L}^2 are the spherical harmonics $Y_l^m(\theta, \phi)$. Therefore it makes sense to look for a solution of the time-independent Schrödinger equation by separating the solution into:

$$u(r, \theta, \phi) = R(r) Y_l^m(\theta, \phi). \quad (10.6)$$

We have written the general solution $u(\underline{r})$ as the product of a radial function $R(r)$, which depends only on the radius r , times the spherical harmonics. The latter encode all the angular dependence of the solutions $u(\underline{r})$.

Using Eq. (8.21), we can rewrite Eq. (10.5) as an ordinary differential equation for the function $R(r)$:

$$\frac{1}{r^2} \frac{d}{dr} \left(r^2 \frac{dR}{dr} \right) - \frac{\ell(\ell+1)}{r^2} R + \frac{2\mu}{\hbar^2} [E - V(r)] R = 0. \quad (10.7)$$

Note that the magnetic quantum number m does not enter in the equation for the radial wave function R . Hence each energy level will have a $(2\ell + 1)$ -fold degeneracy.

Equation for the radial wave function. Let us now discuss the solution for the radial part of the equation. The radial equation is simplified by the substitution:

$$R(r) = \frac{\chi(r)}{r}; \quad (10.8)$$

inserting Eq. (10.8) into Eq. (10.7) yields:

$$\frac{d^2\chi}{dr^2} + \left[\frac{2m}{\hbar^2} (E - V(r)) - \frac{\ell(\ell+1)}{r^2} \right] \chi(r) = 0. \quad (10.9)$$

We have obtained a one-dimensional eigenvalue problem: Eq. (10.9) is the time-independent Schrödinger equation for a one-dimensional system in the potential:

$$V_l(r) = V(r) + \frac{\hbar^2}{2\mu} \frac{\ell(\ell+1)}{r^2}. \quad (10.10)$$

This is simply the sum of the three-dimensional potential $V(r)$ and a “centrifugal” term:

$$\frac{1}{2\mu} \frac{\hat{L}^2}{r^2}. \quad (10.11)$$

Boundary condition for the radial wave function. If the potential energy is finite, we know that the wave function ψ must also be finite. Therefore the radial part $R(r)$ has to be finite everywhere including the origin. For this to happen we need

$$\chi(0) = 0. \quad (10.12)$$

Actually this condition turns out to be true also for a potential energy that diverges as $r \rightarrow 0$.

Hence the motion of a quantum system in a central potential can be reduced to a motion of a one-dimensional system in the region $r > 0$ - remember that r is the radial coordinate and hence is always positive.

The normalization condition for the radial wave function is:

$$\int_0^\infty |R(r)|^2 r^2 dr = \int_0^\infty |\chi(r)|^2 dr = 1. \quad (10.13)$$

10.3 Physical interpretation.

The solution of the one-dimensional problem in Eq. (10.9) is entirely determined by the value of the energy E . Given that the angular part is given by the spherical harmonics Y_ℓ^m , we obtain that the three-dimensional wave function is entirely determined by the values of E, ℓ, m .

The eigenstates can be written as:

$$u_{n\ell m}(\underline{r}) = R_{n\ell}(r)Y_\ell^m(\theta, \phi), \quad (10.14)$$

where n indicates the allowed values of the energy E as usual. As already noted above, the magnetic quantum number does not enter in the radial equation, and therefore the solution $R_{n\ell}(r)$ only depends on the two indices n and ℓ .

Let us discuss the behaviour of the solution R near the origin for a potential such that:

$$\lim_{r \rightarrow 0} (V(r)r^2) = 0, \quad (10.15)$$

i.e. for a potential that diverges in the origin *at most* as $1/r^2$. We shall look for a solution of the form:

$$R(r) = \text{const} \cdot r^s. \quad (10.16)$$

Substituting Eq. (10.16) into Eq. (10.7), and neglecting terms that vanish in the limit $r \rightarrow 0$, we find:

$$s(s+1) = \ell(\ell+1). \quad (10.17)$$

The latter has two solutions for s :

$$s = \ell, \quad \text{or } s = -(\ell+1). \quad (10.18)$$

Clearly the solution with $s = -(\ell+1)$ is divergent at the origin $r = 0$ and therefore does not satisfy the boundary condition for the radial wave function. Hence, close to the origin, the solutions with angular momentum ℓ are proportional to r^ℓ .

The probability density for a particle to be at a distance r from the origin is given by:

$$|R|^2 r^2 \simeq r^{2(\ell+1)}. \quad (10.19)$$

The larger the angular momentum, the more rapidly the probability goes to zero at $r = 0$.

10.4 Quantum rotator

Let us conclude this lecture with a simple example of physical relevance: the quantum rotator. The quantum rotator is a quantum system made of two particles of mass m_1 and m_2 separated

by a fixed distance r_e . It is a simple but effective description of the rotational degrees of freedom of a diatomic molecule. By solving the time-independent Schrödinger equation we are going to find the rotational energy levels, i.e. the energy levels of the molecule in the limit where we neglect the vibrational energy.

Let us choose the centre-of-mass of the molecule as the origin of the reference frame. The state of the system is completely specified by two angles θ and ϕ that specify the orientation of the axis of the molecule with respect to the axis of the reference frame, as illustrated in Fig. 10.1.

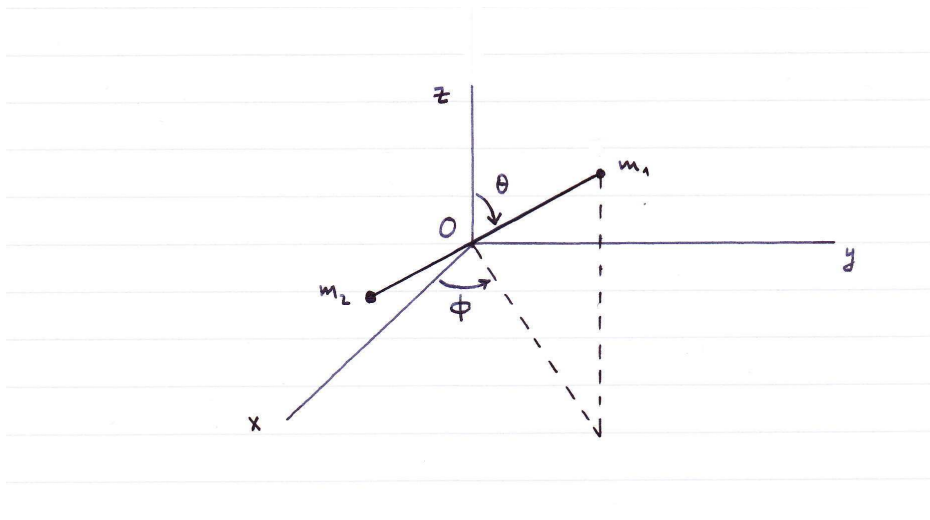


Figure 10.1: Rotator in three-dimensional space. The length of the axis between the two atoms is fixed to be r_e . The origin of the reference frame is chosen to coincide with the centre-of-mass of the system, denoted by O in the figure. The direction of the axis of the rotator is completely specified by the two angles θ and ϕ .

The distance from the origin to the first mass is denoted by r_1 , and similarly the distance from the origin to the second mass is r_2 . Since we are neglecting the vibrational degrees of freedom both r_1 and r_2 are constant, and clearly $r_1 + r_2 = r_e$. Since we have chosen the origin of the reference frame to be the centre of mass of the system we also have:

$$m_1 r_1 = m_2 r_2, \quad (10.20)$$

and therefore:

$$r_1/m_2 = r_2/m_1 = r_e/(m_1 + m_2). \quad (10.21)$$

The moment of inertia of this system is simply:

$$I = m_1 r_1^2 + m_2 r_2^2 \equiv \mu r_e^2, \quad (10.22)$$

where we have introduced the reduced mass:

$$\frac{1}{\mu} = \frac{1}{m_1} + \frac{1}{m_2}. \quad (10.23)$$

In classical mechanics the angular momentum of the system is given by:

$$|\underline{L}| = I\omega_R, \quad (10.24)$$

where ω_R is the angular velocity of the system. The energy can be expressed as:

$$H = \frac{1}{2}I\omega_R^2 = \frac{L^2}{2I} = \frac{L^2}{2\mu r_e^2}. \quad (10.25)$$

Eq. (10.25) is the starting point to define the Hamiltonian for the quantum rotator. According to the general principles defined at the beginning of the course, we define the Hamiltonian operator as:

$$\hat{H} = \frac{\hat{L}^2}{2\mu r_e^2}, \quad (10.26)$$

where \hat{L}^2 is the operator associated to the square of the angular momentum - see Eq. (8.19). The reduced mass μ and the radius of the molecule r_e are constants that define the physical system under study: different diatomic molecules have different reduced masses, or sizes.

Note that the wave function of the system does not depend on r since we are neglecting the vibrations of the molecule. So the state of the quantum system is described by a wave function $\psi(\theta, \phi)$ that depends only on the angular variables.

The Hamiltonian for the diatomic molecule is proportional to \hat{L}^2 ; we have already computed the eigenvalues and the eigenfunctions of \hat{L}^2 when we discussed the angular momentum. The stationary states are:

$$\hat{H}Y_\ell^m(\theta, \phi) = \frac{\ell(\ell+1)\hbar^2}{2\mu r_e^2}Y_\ell^m(\theta, \phi). \quad (10.27)$$

The constant $B = \hbar/(4\pi\mu r_e^2)$ is usually called the “rotational constant”. It has the dimensions of a frequency (**check this!**). The energy levels are therefore:

$$E_\ell = Bh\ell(\ell+1). \quad (10.28)$$

More details about the quantum rotator can be found in problem sheet 5. We shall see that this rather simple model allow us to make physical predictions about the behaviour of diatomic molecules.

10.5 Summary

As usual, we summarize the main concepts introduced in this lecture.

- Time-independent equation for a system in a central potential.
- Separation of variables.
- Reduction to a one-dimensional problem, effective potential, boundary condition.
- Solutions for the stationary states and their generic properties.
- Quantum rotator.

Lecture 11

The Hydrogen atom

11.1 Introduction

The H atom is a bound state of a proton and an electron. The masses of the two particles are respectively:

$$m_p = 1.7 \times 10^{-27} \text{ kg}, \quad (11.1)$$

$$m_e = 0.91 \times 10^{-30} \text{ kg}. \quad (11.2)$$

They have opposite charges, q and $-q$, with:

$$q = 1.6 \times 10^{-19} \text{ C}. \quad (11.3)$$

The ratio of the two masses is:

$$m_p/m_e = 1836.15267247(80); \quad (11.4)$$

it is known to 5 parts per billion!!

The interaction between the two particles is due to electromagnetism; in a nonrelativistic formulation we can therefore model the H atom as a particle of reduced mass m :

$$\frac{1}{\mu} = \frac{1}{m_p} + \frac{1}{m_e} \approx \frac{1}{0.995m_e}, \quad (11.5)$$

in a Coulomb potential:

$$V(r) = -\frac{q^2}{4\pi\epsilon_0 r} \equiv -\frac{e^2}{r}. \quad (11.6)$$

Given that the proton mass is much larger than the electron one ($m_e/m_p \approx 1/1836$), the reduced mass of the ep system is very close to the electron mass. The distance r that appears in the expression for the Coulomb potential is the distance between the electron and the proton. We can identify the origin of our reference frame with the position of the proton; the potential is clearly symmetric under rotations around the origin. The parameters that specify the physical system are the reduced mass μ and the electron charge e .

The H atom is an example of motion in a central force, and we are going to use the formalism developed in Lecture 10 to find the stationary states of this system.

11.2 Stationary states

The Hamiltonian for the system is:

$$\hat{H} = -\frac{\hbar^2}{2\mu} \nabla^2 - \frac{e^2}{r}. \quad (11.7)$$

The time-independent Schrödinger equation

$$\hat{H}\psi(\underline{r}) = E\psi(\underline{r}), \quad (11.8)$$

is solved as usual by separation of variables. Using for the solution $\psi(\underline{r})$ the Ansatz:

$$\psi(r, \theta, \phi) = \frac{\chi_{nl}(r)}{r} Y_{\ell}^m(\theta, \phi), \quad (11.9)$$

we obtain the radial one-dimensional equation:

$$\left[-\frac{\hbar^2}{2\mu} \frac{d^2}{dr^2} + \frac{\hbar^2 \ell(\ell+1)}{2\mu r^2} - \frac{e^2}{r} \right] \chi_{nl}(r) = E_{nl} \chi_{nl}(r). \quad (11.10)$$

Eq. (11.10) is simply Eq. (10.7) for the specific case of the Coulomb potential. The effective potential in this case is:

$$V_l(r) = -\frac{e^2}{r} + \frac{\hbar^2 \ell(\ell+1)}{2\mu r^2}. \quad (11.11)$$

The boundary condition for the radial function $\chi_{nl}(r)$ is:

$$\chi_{nl}(0) = 0. \quad (11.12)$$

The effective potential is sketched in Fig. 11.1.

It is useful to describe the H atom in terms of a characteristic length:

$$a_0 = \frac{\hbar^2}{\mu e^2} \approx 0.52 \text{ \AA}, \quad (11.13)$$

and energy:

$$E_I = \frac{\mu e^4}{2\hbar^2} \approx 13.6 \text{ eV}. \quad (11.14)$$

Let us denote by E_{nl} the energy levels, and introduce the dimensionless variables:

$$\rho = r/a_0, \quad \lambda_{nl} = \sqrt{-E_{nl}/E_I}. \quad (11.15)$$

Eq. (11.10) can be written as:

$$\left[-\frac{\hbar^2}{2\mu} \frac{1}{a_0^2} \frac{d^2}{d\rho^2} + \frac{\hbar^2 \ell(\ell+1)}{2\mu a_0^2 \rho^2} - \frac{e^2}{a_0 \rho} - E_{nl} \right] \chi_{nl}(a_0 \rho) = 0. \quad (11.16)$$

The latter can be rewritten as:

$$\left[\frac{d^2}{d\rho^2} - \frac{\ell(\ell+1)}{\rho^2} + \frac{2\mu e^2 a_0}{\hbar^2} \frac{1}{\rho} + \frac{2\mu a_0^2}{\hbar^2} E_{n,\ell} \right] u_{nl}(\rho) = 0, \quad (11.17)$$

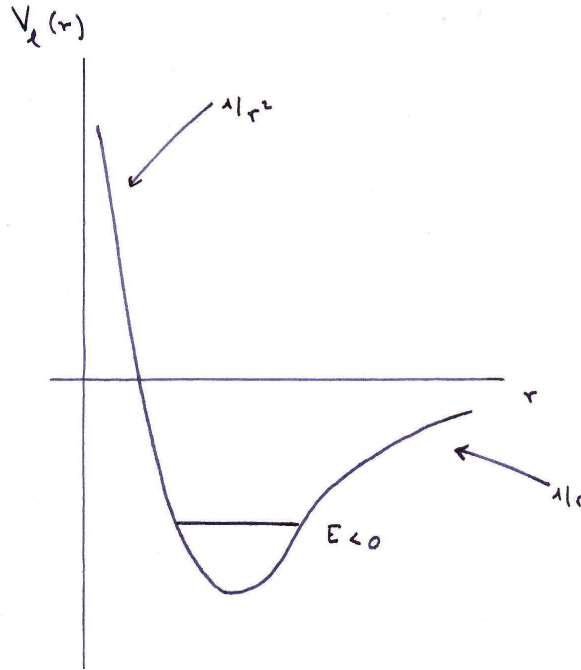


Figure 11.1: Effective potential for the one-dimensional radial Schrödinger equation for a system with total angular momentum l . We look for solutions of the time-independent Schrödinger equation with negative energy E corresponding to bound states of the proton/electron system.

and finally

$$\left[\frac{d^2}{d\rho^2} - \frac{\ell(\ell+1)}{\rho^2} + \frac{2}{\rho} - \lambda_{nl}^2 \right] u_{n,\ell}(\rho) = 0. \quad (11.18)$$

We need to solve Eq. (11.18) in order to find the eigenvalues of the Hamiltonian E_{nl} and the corresponding radial wave functions $\chi_{nl}(r)$. Remember that the angular part of the wave functions is given by the spherical harmonics.

11.3 Solution of the radial equation

In this section we shall discuss some technical details related to the solution of the radial equation (11.18). We wish to keep the mathematical details of the solution separated from the physical interpretation, which will be discussed in the next section.

In order to shine a light on the form of the solution, we can start by considering its limiting behaviours as $\rho \rightarrow 0$ and $\rho \rightarrow \infty$.

Let us first discuss the large-distance regime. As ρ is increased, the centrifugal and Coulomb potentials tend to zero and become unimportant in Eq. (11.18), which becomes:

$$\left[\frac{d^2}{d\rho^2} - \lambda_{nl}^2 \right] u_{nl}(\rho) = 0. \quad (11.19)$$

The solutions to this latter equation are simply:

$$u_{nl}(\rho) = \exp(\pm \lambda_{nl} \rho), \quad (11.20)$$

and the solution that grows exponentially must be discarded because it yields a non-normalizable wave function.

Of course we cannot completely neglect the potential terms, so we will look for a complete solution of the form:

$$u_{nl}(\rho) = e^{-\lambda_{nl} \rho} \eta_{nl}(\rho). \quad (11.21)$$

Eq. (11.18) becomes:

$$\eta_{nl}'' - 2\lambda_{nl} \eta_{nl}' + \left(-\frac{\ell(\ell+1)}{\rho^2} + \frac{2}{\rho} \right) \eta_{nl} = 0, \quad (11.22)$$

where the prime symbol “ \prime ” denotes the differentiation with respect to ρ . The boundary condition for u_n translates into a boundary condition for η_{nl} , namely $\eta_{nl}(0) = 0$.

As $\rho \rightarrow 0$ we know from the discussion in Lecture 10 that $u_{nl} \sim \rho^{\ell+1}$. Therefore we can look for a solution for η_{nl} expanded as a power series in ρ :

$$\eta_{nl}(\rho) = \rho^{\ell+1} \sum_{q=0}^{\infty} c_q \rho^q. \quad (11.23)$$

Inserting this Ansatz into Eq. (11.22) we obtain:

$$\begin{aligned} \sum_q (q+l+1)(q+l) c_q \rho^{q+\ell-1} - 2\lambda_{nl} (q+l+1) c_q \rho^{q+\ell} + \\ + 2c_q \rho^{q+\ell} - \ell(\ell+1) c_q \rho^{q+\ell-1} = 0, \end{aligned} \quad (11.24)$$

and hence:

$$\sum_q q(q+2\ell+1) c_q \rho^{q+\ell-1} - 2[\lambda_{nl}(q+l+1) - 1] c_q \rho^{q+l} = 0. \quad (11.25)$$

Shifting the summation index, $q \rightarrow q - 1$, in the second term of the sum above, we obtain:

$$\sum_q \{q(q + 2\ell + 1)c_q - 2[\lambda_{nl}(q + \ell) - 1]c_{q-1}\} \rho^{q+\ell-1} = 0. \quad (11.26)$$

Since the last equality must hold for all values of q , we deduce:

$$q(q + 2\ell + 1)c_q - 2[\lambda_{nl}(q + \ell) - 1]c_{q-1} = 0. \quad (11.27)$$

Eq. (11.27) is a recursion relation between the coefficients of the Taylor expansion of $\eta_{nl}/r\hbar\omega^{l+1}$. It is crucial to note that for large q :

$$\frac{c_q}{c_{q-1}} \underset{q \rightarrow \infty}{\sim} \frac{2\lambda_{nl}}{q}, \quad \text{i.e. } c_q \sim \frac{(2\lambda_{nl})^q}{q!}. \quad (11.28)$$

The asymptotic behaviour for c_q would yield a solution:

$$\eta_{nl}(\rho) \sim \rho^{\ell+1} e^{2\lambda_{nl}\rho}, \quad (11.29)$$

which in turn yields a wave function u_{nl} that is not normalizable.

Therefore we *must* have $c_q = 0$ for some finite value of q , that we will denote $q = n_r > 0$. According to Eq. (11.27) this can only happen if:

$$\lambda_{nl} = \frac{1}{n_r + \ell}. \quad (11.30)$$

Then the expansion in Eq. (11.23) only contains a finite number of terms, i.e. it is simply a polynomial in ρ of finite order n_r .

We can see from Eq. (11.30) that this condition implies that the energy eigenvalues are quantized. Remember that λ_{nl} is related to the eigenvalues of the Hamiltonian via Eq. (11.15).

11.4 Physical interpretation

The computation in the previous section shows that the eigenvalues of the Hamiltonian for the H atom are:

$$E_{nl} = \frac{-E_I}{(n_r + \ell)^2}, \quad (11.31)$$

where ℓ is the angular momentum of the state, and $n_r > 0$ is an integer. We see that for the H atom the value of the energy does not depend on n_r and ℓ , but only on their sum $n = n_r + \ell$. The integer n is called the *principal quantum number*, its value characterizes the so-called an *electron shells*.

We can rewrite Eq. (11.14) as:

$$E_I = \frac{1}{2}\alpha^2\mu c^2, \quad (11.32)$$

where α is the *fine-structure constant*:

$$\alpha = \frac{e^2}{\hbar c} \simeq \frac{1}{137}. \quad (11.33)$$

Note that $\mu \simeq m_e$, and therefore μc^2 is the rest energy of the electron. Eq. (11.32) shows that the typical scale of the energy levels in the H atom is 10^{-4} the rest energy of the electron. This justifies the nonrelativistic treatment of the H atom that we have used here. Clearly there *are* relativistic corrections; however they are small effect, typically of order α , and can be studied in perturbation theory.

Using Eq. (11.13) we can express the energy levels as:

$$E_n = -\frac{e^2}{2n^2 a_0}, \quad n = n_r + l. \quad (11.34)$$

This the formula put forward by Bohr in 1913, before Quantum Mechanics was fully developed.

Eq. (11.34) shows clearly that for a given principal quantum number n , ℓ can take the values $\ell = 0, 1, \dots, n-1$ (corresponding respectively to $n_r = n, n-1, \dots, 1$). Since for each value of ℓ we have a $(2\ell + 1)$ degeneracy, the total degeneracy of the level E_n is:

$$\sum_{\ell=0}^{n-1} (2\ell + 1) = n(n-1) + n = n^2. \quad (11.35)$$

The polynomials w_{nl} are called the *associated Laguerre polynomials*, their explicit expression can be found in standard textbooks on special functions ¹. The full solution for the eigenfunctions of the energy is:

$$\psi_{nlm}(\underline{r}) = R_{nl}(r)Y_{\ell}^m(\theta, \phi). \quad (11.36)$$

The first few radial functions are:

$$R_{1,0} = 2a_0^{-3/2}e^{-r/a_0}, \quad (11.37)$$

$$R_{2,0} = \frac{1}{2\sqrt{2}}a_0^{-3/2}\left(2 - \frac{r}{a_0}\right)e^{-r/(2a_0)}, \quad (11.38)$$

$$R_{2,1} = \frac{1}{2\sqrt{6}}a_0^{-3/2}\frac{r}{a_0}e^{-r/(2a_0)}. \quad (11.39)$$

It is worthwhile to conclude this Lecture with a few remarks.

¹see e.g. M. Abramowitz and I.A. Stegun, *Handbook of mathematical functions*.

- The formula for the energy levels reproduces the Bohr spectrum, in agreement with the experimental data. The degeneracy of each level can only be obtained by the proper quantum-mechanical description that we have presented.
- The Bohr radius a_0 is the typical spatial extension of the ground state.
- The treatment described here can be applied to any hydrogen-like atom, i.e. an atom with an electron and a nucleus of charge Zq . Simply replace everywhere $e^2 \rightarrow Ze^2$.
- It is interesting to compute the expectation value of the momentum. Since the typical size of the atom is a_0 , we deduce from the Heisenberg uncertainty relation that $\sqrt{\langle p^2 \rangle} \sim \mu e^2 / \hbar$. Thus we obtain:

$$v \sim p/\mu \simeq e^2/\hbar = \alpha c \simeq \frac{1}{137}c \ll c. \quad (11.40)$$

We see *a posteriori* that motion of the electron is nonrelativistic, and hence our nonrelativistic description is accurate. Relativistic corrections are expected to be $\mathcal{O}(v/c) \simeq \mathcal{O}(\alpha)$.

- The ionization energy is the energy needed to extract the electron from the bound state and is E_I . This is known as the *Rydberg energy*.
- The electron in the H atom can go from one shell to a lower one by emitting a photon. The series of transitions from principal number $n \geq 2$ to $n = 1$ is called the *Lyman series*². The transitions are named by Greek letters: the transition from $n = 2$ to $n = 1$ is called Lyman- α , from 3 to 1 Lyman- β , etc. Likewise the transitions from $n \geq 3$ to $n = 2$ form the *Balmer series*. These transitions, which were discovered before Quantum Mechanics was established, have been observed with great precision and provide a strong experimental evidence in favour of Quantum Mechanics.

11.5 Summary

As usual, we summarize the main concepts introduced in this lecture.

²A physics department at Harvard University is named after Lyman.

- Setup of the problem. Hamiltonian for the H atom.
- Separation of variables. One-dimensional equation for the radial wave function.
- Effective potential, boundary condition.
- Solutions for the stationary states, quantization of the energy levels.
- Discussion of the mathematical results. Physical picture.

Lecture 12

The harmonic oscillator

12.1 Introduction

In this chapter, we are going to find explicitly the eigenfunctions and eigenvalues for the time-independent Schrödinger equation for the one-dimensional harmonic oscillator. We have already described the solutions in Chap. 3.

Recall that the TISE for the 1-dimensional quantum harmonic oscillator is

$$\left[\frac{\hat{P}^2}{2m} + \frac{1}{2}m\omega^2\hat{X}^2 \right] u_n(x) = E_n u_n(x),$$

which we write in Dirac notation as

$$\hat{H} |n\rangle = E_n |n\rangle.$$

We have denoted by $|n\rangle$ the ket associated to the eigenfunctions $u_n(x)$.

12.2 Factorizing the Hamiltonian

The Hamiltonian for the harmonic oscillator is:

$$\hat{H} = \frac{\hat{P}^2}{2m} + \frac{1}{2}m\omega^2\hat{X}^2. \quad (12.1)$$

Let us factor out $\hbar\omega$, and rewrite the Hamiltonian as:

$$\hat{H} = \hbar\omega \left[\frac{\hat{P}^2}{2m\hbar\omega} + \frac{m\omega}{2\hbar}\hat{X}^2 \right]. \quad (12.2)$$

Checking the dimensions of the constants, you can readily verify that:

$$[\hbar\omega] = \text{energy}, \quad [2m\omega\hbar] = \text{momentum}^2, \quad \left[\frac{2\hbar}{m\omega} \right] = \text{length}^2. \quad (12.3)$$

Introducing the dimensionless quantities:

$$\hat{\xi} = \sqrt{\frac{m\omega}{2\hbar}} \hat{X}, \quad (12.4)$$

$$\hat{\eta} = \frac{\hat{P}}{\sqrt{2m\hbar\omega}}, \quad (12.5)$$

the Hamiltonian becomes:

$$\hat{H} = \hbar\omega \left[\hat{\eta}^2 + \hat{\xi}^2 \right]. \quad (12.6)$$

The operators $\hat{\xi}$ and $\hat{\eta}$ are simply the position and the momentum operators rescaled by some real constants; therefore both of them are Hermitean. Their commutation relation can be easily computed using the canonical commutation relations:

$$[\hat{\xi}, \hat{\eta}] = \frac{1}{2\hbar} [\hat{X}, \hat{P}] = \frac{i}{2}. \quad (12.7)$$

If $\hat{\xi}$ and $\hat{\eta}$ were commuting variables, we would be tempted to factorize the Hamiltonian as:

$$\hat{H} = \hbar\omega (\hat{\xi} + i\hat{\eta}) (\hat{\xi} - i\hat{\eta}). \quad (12.8)$$

We must be careful here, because the operators do not commute. So let us introduce:

$$\begin{cases} \hat{a} = \hat{\xi} + i\hat{\eta}, \\ \hat{a}^\dagger = \hat{\xi} - i\hat{\eta}; \end{cases} \quad (12.9)$$

the expressions for \hat{a} and \hat{a}^\dagger in terms of \hat{X} and \hat{P} are:

$$\begin{cases} \hat{a} = \sqrt{\frac{m\omega}{2\hbar}} \hat{X} + \frac{i}{\sqrt{2m\omega\hbar}} \hat{P} \\ \hat{a}^\dagger = \sqrt{\frac{m\omega}{2\hbar}} \hat{X} - \frac{i}{\sqrt{2m\omega\hbar}} \hat{P} \end{cases}.$$

We can then compute

$$\hat{a}\hat{a}^\dagger = \hat{\xi}^2 + i[\hat{\eta}, \hat{\xi}] + \hat{\eta}^2, \quad (12.10)$$

$$\hat{a}^\dagger\hat{a} = \hat{\xi}^2 - i[\hat{\eta}, \hat{\xi}] + \hat{\eta}^2. \quad (12.11)$$

Summing the two equations above:

$$\hat{H} = \frac{\hbar\omega}{2} (\hat{a}\hat{a}^\dagger + \hat{a}^\dagger\hat{a}) \quad (12.12)$$

Subtracting the same two equations yields the commutation relation between \hat{a} and \hat{a}^\dagger :

$$[\hat{a}, \hat{a}^\dagger] = 1. \quad (12.13)$$

This commutation relations plays an important role in the rest of this chapter.

An alternative, and more useful, expression for \hat{H} is

$$\hat{H} = \left(\hat{a}^\dagger\hat{a} + \frac{1}{2} \right) \hbar\omega. \quad (12.14)$$

12.3 Creation and annihilation

We are now going to find the eigenvalues of \hat{H} using the operators \hat{a} and \hat{a}^\dagger . First let us compute the commutators $[\hat{H}, \hat{a}]$ and $[\hat{H}, \hat{a}^\dagger]$:

$$[\hat{H}, \hat{a}] = \left[\left(\hat{a}^\dagger \hat{a} + \frac{1}{2} \right) \hbar\omega, \hat{a} \right] = \hbar\omega [\hat{a}^\dagger \hat{a}, \hat{a}] \quad \text{since} \quad \left[\frac{1}{2}, \hat{a} \right] = 0.$$

Now

$$[\hat{a}^\dagger \hat{a}, \hat{a}] = \hat{a}^\dagger \hat{a} \hat{a} - \hat{a} \hat{a}^\dagger \hat{a} = [\hat{a}^\dagger, \hat{a}] \hat{a} = -\hat{a},$$

so that we obtain

$$\boxed{[\hat{H}, \hat{a}] = -\hbar\omega \hat{a}}. \quad (12.15)$$

Similarly

$$[\hat{H}, \hat{a}^\dagger] = \left[\left(\hat{a}^\dagger \hat{a} + \frac{1}{2} \right) \hbar\omega, \hat{a}^\dagger \right] = \hbar\omega [\hat{a}^\dagger \hat{a}, \hat{a}^\dagger] \quad \text{since} \quad \left[\frac{1}{2}, \hat{a}^\dagger \right] = 0,$$

and

$$[\hat{a}^\dagger \hat{a}, \hat{a}^\dagger] = \hat{a}^\dagger \hat{a} \hat{a}^\dagger - \hat{a}^\dagger \hat{a}^\dagger \hat{a} = \hat{a}^\dagger [\hat{a}, \hat{a}^\dagger] = \hat{a}^\dagger,$$

so that we obtain

$$\boxed{[\hat{H}, \hat{a}^\dagger] = \hbar\omega \hat{a}^\dagger}. \quad (12.16)$$

Let us now compute:

$$\hat{H}(\hat{a}|n\rangle) = \hat{a}\hat{H}|n\rangle + [\hat{H}, \hat{a}]|n\rangle, \quad (12.17)$$

$$= E_n \hat{a}|n\rangle - \hbar\omega \hat{a}|n\rangle, \quad (12.18)$$

$$= (E_n - \hbar\omega)(\hat{a}|n\rangle). \quad (12.19)$$

We have found an eigenvalue equation: it states that $\hat{a}|n\rangle$ is an eigenfunction of \hat{H} belonging to the eigenvalue $(E_n - \hbar\omega)$, unless $\hat{a}|n\rangle \equiv 0$. We say that the operator \hat{a} is a *lowering operator*; its action on an energy eigenstate is to turn it into another energy eigenstate of lower energy. It is also called an *annihilation operator*, because it removes one quantum of energy $\hbar\omega$ from the system.

Similarly it is straightforward to show that

$$\boxed{\hat{H}\hat{a}^\dagger|n\rangle = (E_n + \hbar\omega)\hat{a}^\dagger|n\rangle},$$

which says that $\hat{a}^\dagger|n\rangle$ is an eigenfunction of \hat{H} belonging to the eigenvalue $(E_n + \hbar\omega)$, unless $\hat{a}^\dagger|n\rangle \equiv 0$. We say that the operator \hat{a}^\dagger is a *raising operator*; its action on an energy eigenstate is to turn it into another energy eigenstate of higher energy. It is also called an *creation operator*, because it adds one quantum of energy $\hbar\omega$ to the system.

We can summarise these results by denoting the states of energy $E_n \pm \hbar\omega$ by $|n \pm 1\rangle$ and writing

$$\boxed{\hat{a} |n\rangle = c_n |n-1\rangle \quad \text{and} \quad \hat{a}^\dagger |n\rangle = d_n |n+1\rangle},$$

where c_n and d_n are constants of proportionality (NOT eigenvalues!) and

$$\begin{aligned} \hat{H} |n-1\rangle &= E_{n-1} |n-1\rangle = (E_n - \hbar\omega) |n-1\rangle \\ \hat{H} |n+1\rangle &= E_{n+1} |n+1\rangle = (E_n + \hbar\omega) |n+1\rangle. \end{aligned}$$

12.4 Eigensystem

12.4.1 Eigenvalues

It should be clear that repeated application of the lowering operator, \hat{a} , generates states of successively lower energy ad infinitum *unless there is a state of lowest energy*; application of the operator to such a state must yield zero identically (because otherwise we would be able to generate another state of lower energy still, a contradiction).

Is there such a state? The answer is yes because the Hamiltonian can only have positive eigenvalues. Consider the expectation value of \hat{H} in an arbitrary state $|\Psi\rangle$:

$$\langle \hat{H} \rangle = \left\langle \frac{\hat{p}^2}{2m} \right\rangle + \left\langle \frac{1}{2} m \omega^2 x^2 \right\rangle,$$

and both terms on the right hand side are non-negative. Thus there cannot be any states of negative energy.

We denote the state of lowest energy, or ground state, by $|0\rangle$. Then since there cannot be a state of lower energy,

$$\hat{a} |0\rangle = 0.$$

Applying the Hamiltonian to this state we see that

$$\hat{H} |0\rangle = \hbar\omega \left(\hat{a}^\dagger \hat{a} + \frac{1}{2} \right) |0\rangle = \frac{1}{2} \hbar\omega |0\rangle \equiv E_0 |0\rangle.$$

Thus we have found the ground state energy: $E_0 = \frac{1}{2} \hbar\omega$. Application of the raising operator to the ground state generates the state $|1\rangle$ with energy $E_1 = \frac{3}{2} \hbar\omega$, whilst n applications of the raising operator generates the state $|n\rangle$ with energy $(n + \frac{1}{2}) \hbar\omega$, so that

$$\boxed{E_n = (n + \frac{1}{2}) \hbar\omega \quad n = 0, 1, 2, 3, \dots},$$

which is the previously quoted result for the energy eigenvalues of the 1-dimensional oscillator!

12.4.2 Normalisation of Eigenstates

Requiring that both $|n\rangle$ and $|n-1\rangle$ be normalised enables us to determine the constant of proportionality c_n . Consider

$$\begin{aligned}\langle n|\hat{a}^\dagger\hat{a}|n\rangle &= c_n \langle n|\hat{a}^\dagger|n-1\rangle && \text{from property of } \hat{a} \\ &= c_n \langle n-1|\hat{a}|n\rangle^* && \text{from definition of } \dagger \\ &= c_n c_n^* \langle n-1|n-1\rangle^* && \text{from property of } \hat{a} \\ &= |c_n|^2 && \text{since } \langle n-1|n-1\rangle^* = 1.\end{aligned}$$

We can evaluate the left-hand side if we note that $\hat{a}^\dagger\hat{a} = \left(\hat{H}/\hbar\omega\right) - \frac{1}{2}$, giving

$$\langle n|\hat{a}^\dagger\hat{a}|n\rangle = n \langle n|n\rangle = n.$$

Thus $|c_n|^2 = n$ and if we choose the phase so that c_n is real we can write

$$c_n = \sqrt{n}.$$

A similar calculation shows that

$$\langle n|\hat{a}\hat{a}^\dagger|n\rangle = |d_n|^2 = (n+1),$$

so that if we again choose the phase so that d_n is real we obtain

$$d_n = \sqrt{n+1}.$$

In summary then we have

$$\boxed{\hat{a}|n\rangle = \sqrt{n}|n-1\rangle \quad \text{and} \quad \hat{a}^\dagger|n\rangle = \sqrt{n+1}|n+1\rangle}$$

12.4.3 Wave functions

Finally let us show that we can reproduce the analytic expression for the eigenfunctions of the energy.

The ground state is defined by the relation:

$$\hat{a}|0\rangle = 0. \tag{12.20}$$

We can rewrite the equation above as a differential operator acting on the wave function of the ground state $u_0(x)$:

$$\hat{a}u_0(x) = \left[\sqrt{\frac{m\omega}{2\hbar}} \hat{X} + \frac{i}{\sqrt{2m\omega\hbar}} \hat{P} \right] u_0(x) \tag{12.21}$$

$$= \left[\sqrt{\frac{m\omega}{2\hbar}} x + \frac{\hbar}{\sqrt{2m\omega}} \frac{d}{dx} \right] u_0(x) \tag{12.22}$$

$$= 0. \tag{12.23}$$

Hence:

$$\frac{d}{dx}u_0(x) = -\sqrt{\frac{2m\omega}{\hbar}}\sqrt{\frac{m\omega}{2\hbar}}xu_0(x) \quad (12.24)$$

$$= -\frac{m\omega}{\hbar}xu_0(x) \quad (12.25)$$

$$= -\alpha^2xu_0(x), \quad (12.26)$$

where $\alpha^2 = m\omega/\hbar$. The solution of the equation above is the Gaussian that we have already seen in Chap. 3:

$$u_0(x) = C_0 \exp[-\alpha^2x^2/2]. \quad (12.27)$$

Every other eigenfunctions is obtained by repeatedly applying the creation operator \hat{a}^\dagger to ground state:

$$u_n(x) = \frac{1}{\sqrt{n!}}\left(\hat{a}^\dagger\right)^n u_0(x). \quad (12.28)$$

Remember that \hat{a}^\dagger is just a differential operator acting on wave functions. Check that you can reproduce the wave functions for the first and second excited states of the harmonic oscillator.

12.5 Summary

As usual, we summarize the main concepts introduced in this lecture.

- Raising and lowering operators; factorization of the Hamiltonian.
- Commutation relations and interpretation of the raising and lowering operators.
- Existence of the ground states, construction and normalization of the excited states. Eigenvalues of the Hamiltonian.
- Construction of the wave functions.

Lecture 13

Angular momentum - Reloaded

13.1 Introduction

In previous lectures we have introduced the angular momentum starting from the classical expression:

$$\underline{L} = \underline{r} \times \underline{p}, \quad (13.1)$$

and have defined a quantum mechanical operator by replacing \underline{r} and \underline{p} with the corresponding operators: Eq. (13.1) then defines a triplet of differential operators acting on the wave functions. The eigenvalue equations for \hat{L}^2 and \hat{L}_z can be written as differential equations, whose solutions yield the eigenvalues and the eigenfunctions of the angular momentum.

In this lecture we are going to follow a different approach, and derive the quantization of angular momentum directly from the commutation relations of the components of \underline{L} . This approach is more generic and does not rely on the specific realization of the angular momentum as a differential operator. We will use in this context the symbol $\hat{\underline{J}}$ to denote the angular momentum. Remember that $\hat{\underline{J}}$ is a vector, i.e. it is a triplet of operators. In Cartesian coordinates $\hat{\underline{J}} = (\hat{J}_x, \hat{J}_y, \hat{J}_z)$, and the commutation relations are: Recall that the commutation relations are

$$\boxed{[\hat{J}_x, \hat{J}_y] = i\hbar\hat{J}_z \quad [\hat{J}_y, \hat{J}_z] = i\hbar\hat{J}_x \quad [\hat{J}_z, \hat{J}_x] = i\hbar\hat{J}_y}. \quad (13.2)$$

The square of the angular momentum is represented by the operator

$$\hat{J}^2 \equiv \hat{J}_x^2 + \hat{J}_y^2 + \hat{J}_z^2, \quad (13.3)$$

with the property that

$$[\hat{J}^2, \hat{J}_x] = [\hat{J}^2, \hat{J}_y] = [\hat{J}^2, \hat{J}_z] = 0. \quad (13.4)$$

The Compatibility Theorem tells us that, for example, the operators \hat{J}^2 and \hat{J}_z have simultaneous eigenstates. We denote these common eigenstates by $|\lambda, m\rangle$. Looking back at the results obtained in the previous lectures, these are the kets associated to the spherical harmonics $Y_\ell^m(\theta, \phi)$. We can write:

$$\hat{J}^2 |\lambda, m\rangle = \lambda\hbar^2 |\lambda, m\rangle \quad (13.5)$$

$$\hat{J}_z |\lambda, m\rangle = m\hbar |\lambda, m\rangle \quad (13.6)$$

so that the eigenvalues of \hat{J}^2 are denoted by $\lambda\hbar^2$ whilst those of \hat{J}_z are denoted by $m\hbar$.

13.2 The Eigenvalue Spectra of \hat{J}^2 and \hat{J}_z

We now address the problem of finding the allowed values of λ and m . We introduce *raising and lowering operators*:

$$\boxed{\hat{J}_\pm \equiv \hat{J}_x \pm i\hat{J}_y}, \quad (13.7)$$

and calculate the commutators with \hat{J}^2 and \hat{J}_z . Since \hat{J}^2 commutes with both \hat{J}_x and \hat{J}_y we get immediately that

$$[\hat{J}^2, \hat{J}_\pm] = 0, \quad (13.8)$$

whilst

$$[\hat{J}_z, \hat{J}_+] = [\hat{J}_z, \hat{J}_x] + i[\hat{J}_z, \hat{J}_y] = i\hbar \hat{J}_y + i(-i\hbar \hat{J}_x) = \hbar \hat{J}_+, \quad (13.9)$$

and

$$[\hat{J}_z, \hat{J}_-] = [\hat{J}_z, \hat{J}_x] - i[\hat{J}_z, \hat{J}_y] = i\hbar \hat{J}_y - i(-i\hbar \hat{J}_x) = -\hbar \hat{J}_-. \quad (13.10)$$

The structure of these commutation relations should remind you of the harmonic oscillator problem: we can show that \hat{J}_+ and \hat{J}_- act as *raising and lowering operators*, not for the total energy but for the z -component of the angular momentum.

Consider the action of the commutator in Eq. (13.9) on an eigenstate $|\lambda, m\rangle$:

$$[\hat{J}_z, \hat{J}_+]|\lambda, m\rangle \equiv (\hat{J}_z \hat{J}_+ - \hat{J}_+ \hat{J}_z)|\lambda, m\rangle = \hbar \hat{J}_+|\lambda, m\rangle, \quad (13.11)$$

but we can use the eigenvalue equation $\hat{J}_z|\lambda, m\rangle = m\hbar|\lambda, m\rangle$ to rewrite this as

$$\boxed{\hat{J}_z \hat{J}_+|\lambda, m\rangle = (m+1)\hbar \hat{J}_+|\lambda, m\rangle}, \quad (13.12)$$

which says that $\hat{J}_+|\lambda, m\rangle$ is also an eigenstate of \hat{J}_z but with eigenvalue $(m+1)\hbar$, unless $\hat{J}_+|\lambda, m\rangle \equiv 0$. Thus the operator \hat{J}_+ acts as a raising operator for the z -component of angular momentum.

Similarly the second commutator can be used to show that

$$\boxed{\hat{J}_z \hat{J}_-|\lambda, m\rangle = (m-1)\hbar \hat{J}_-|\lambda, m\rangle}, \quad (13.13)$$

which says that $\hat{J}_-|\lambda, m\rangle$ is also an eigenstate of \hat{J}_z but with eigenvalue $(m-1)\hbar$, unless $\hat{J}_-|\lambda, m\rangle \equiv 0$. Thus the operator \hat{J}_- acts as a lowering operator for the z -component of angular momentum.

Notice that since $[\hat{J}^2, \hat{J}_\pm] = 0$ we have

$$\hat{J}^2(\hat{J}_\pm|\lambda, m\rangle) = \hat{J}_\pm(\hat{J}^2|\lambda, m\rangle) = \lambda\hbar^2(\hat{J}_\pm|\lambda, m\rangle), \quad (13.14)$$

so that the states generated by the action of \hat{J}_\pm are still eigenstates of \hat{J}^2 belonging to the same eigenvalue $\lambda\hbar^2$. Thus we can write

$$\begin{aligned} \hat{J}_+|\lambda, m\rangle &= c_+\hbar|\lambda, m+1\rangle \\ \hat{J}_-|\lambda, m\rangle &= c_-\hbar|\lambda, m-1\rangle \end{aligned}$$

where c_\pm are constants of proportionality.

We now observe that, for a given λ , $m^2 \leq \lambda$ so that m must have both a *maximum value*, m_{\max} , and a *minimum value*, m_{\min} .

Proof:

$$\begin{aligned} (\hat{J}^2 - \hat{J}_z^2) |\lambda, m\rangle &= (\hat{J}_x^2 + \hat{J}_y^2) |\lambda, m\rangle \quad \text{implying that} \\ (\lambda - m^2) \hbar^2 |\lambda, m\rangle &= (\hat{J}_x^2 + \hat{J}_y^2) |\lambda, m\rangle \end{aligned}$$

Taking the scalar product with $\langle \lambda, m |$ yields

$$(\lambda - m^2) \hbar^2 = \langle \hat{J}_x^2 + \hat{J}_y^2 \rangle \geq 0,$$

so that

$$\lambda - m^2 \geq 0 \quad \text{or} \quad -\sqrt{\lambda} \leq m \leq \sqrt{\lambda}. \quad (13.15)$$

Hence the spectrum of \hat{J}_z is bounded above and below, for a given λ . We can deduce that

$$\begin{aligned} \hat{J}_+ |\lambda, m_{\max}\rangle &= 0, \quad \text{and} \\ \hat{J}_- |\lambda, m_{\min}\rangle &= 0 \end{aligned}$$

To proceed further, we need a couple of identities for \hat{J}^2 which follow from the definitions of \hat{J}_{\pm} . Noting that

$$\begin{aligned} \hat{J}_+ \hat{J}_- &= \hat{J}_x^2 + \hat{J}_y^2 + i\hat{J}_y \hat{J}_x - i\hat{J}_x \hat{J}_y = \hat{J}_x^2 + \hat{J}_y^2 + \hbar \hat{J}_z, \\ \hat{J}_- \hat{J}_+ &= \hat{J}_x^2 + \hat{J}_y^2 - i\hat{J}_y \hat{J}_x + i\hat{J}_x \hat{J}_y = \hat{J}_x^2 + \hat{J}_y^2 - \hbar \hat{J}_z, \end{aligned}$$

we can write

$$\boxed{\begin{aligned} \hat{J}^2 &\equiv \hat{J}_+ \hat{J}_- - \hbar \hat{J}_z + \hat{J}_z^2 \quad \text{or, equally,} \\ \hat{J}^2 &\equiv \hat{J}_- \hat{J}_+ + \hbar \hat{J}_z + \hat{J}_z^2 \end{aligned}}. \quad (13.16)$$

Applying the first of these to the state of minimum m , we find

$$\begin{aligned} \hat{J}^2 |\lambda, m_{\min}\rangle &= (\hat{J}_+ \hat{J}_- - \hbar \hat{J}_z + \hat{J}_z^2) |\lambda, m_{\min}\rangle \\ &= (-m_{\min} \hbar^2 + m_{\min}^2 \hbar^2) |\lambda, m_{\min}\rangle, \quad \text{since } \hat{J}_- |\lambda, m_{\min}\rangle = 0 \\ &= m_{\min}(m_{\min} - 1) \hbar^2 |\lambda, m_{\min}\rangle \\ &\equiv \lambda \hbar^2 |\lambda, m_{\min}\rangle. \end{aligned}$$

Thus we deduce that

$$\lambda = m_{\min}(m_{\min} - 1).$$

Similarly, using the second of the two identities for \hat{J}^2

$$\begin{aligned}\hat{J}^2 |\lambda, m_{\max}\rangle &= \left(\hat{J}_- \hat{J}_+ + \hbar \hat{J}_z + \hat{J}_z^2 \right) |\lambda, m_{\max}\rangle \\ &= (m_{\max} \hbar^2 + m_{\max}^2 \hbar^2) |\lambda, m_{\max}\rangle, \quad \text{since } \hat{J}_+ |\lambda, m_{\max}\rangle = 0 \\ &= m_{\max}(m_{\max} + 1) \hbar^2 |\lambda, m_{\max}\rangle \\ &\equiv \lambda \hbar^2 |\lambda, m_{\max}\rangle\end{aligned}$$

and hence we obtain a second expression for λ :

$$\lambda = m_{\max}(m_{\max} + 1).$$

Usually, m_{\max} is denoted by j and so we write this as

$$\lambda = j(j + 1) = m_{\min}(m_{\min} - 1).$$

This is a quadratic equation for m_{\min} :

$$m_{\min}^2 - m_{\min} - j^2 - j = 0,$$

which can be factorised:

$$(m_{\min} + j)(m_{\min} - j - 1) = 0,$$

and we see that, since $m_{\min} \leq j$ by definition, the only acceptable root is

$$m_{\min} = -j.$$

Now since m_{\max} and m_{\min} differ by some integer, k , say, we can write

$$m_{\max} - m_{\min} = k, \quad k = 0, 1, 2, 3, \dots,$$

or

$$j - (-j) \equiv 2j = k, \quad k = 0, 1, 2, 3, \dots,$$

so that the allowed values of j are

$$\boxed{j = 0, \frac{1}{2}, 1, \frac{3}{2}, 2, \dots}. \quad (13.17)$$

For a given value of j , we see that m ranges over the values

$$\boxed{j, j - 1, \dots, -j + 1, -j} \quad (13.18)$$

a total of $(2j + 1)$ values.

13.3 Nomenclature

From the results presented above we can draw the following conclusions.

- The eigenvalues of \hat{J}^2 are $j(j+1)\hbar^2$, where j is one of the allowed values

$$j = 0, \frac{1}{2}, 1, \frac{3}{2}, 2, \dots$$

- Since $\lambda = j(j+1)$, we can equally well label the simultaneous eigenstates of \hat{J}^2 and \hat{J}_z by j rather than by λ so that

$$\begin{aligned} \hat{J}^2 |j, m\rangle &= j(j+1)\hbar^2 |j, m\rangle & \text{and} \\ \hat{J}_z |j, m\rangle &= m\hbar |j, m\rangle \end{aligned}$$

- For a given value of j , there are $(2j+1)$ possible eigenvalues of \hat{J}_z , denoted by $m\hbar$, where m runs from j to $-j$ in integer steps.
- The set of $(2j+1)$ states $\{|j, m\rangle\}$ is called a *multiplet*.

13.4 Normalization

We need to determine the constants of proportionality c_{\pm} that appear in the equations

$$\begin{aligned} \hat{J}_+ |j, m\rangle &= c_+ \hbar |j, m+1\rangle \\ \hat{J}_- |j, m\rangle &= c_- \hbar |j, m-1\rangle \end{aligned}$$

Let us consider

$$\begin{aligned} \langle j, m | \hat{J}_- \hat{J}_+ |j, m\rangle &= c_+ \hbar \langle j, m | \hat{J}_- |j, m+1\rangle \\ &= c_+ \hbar \langle j, m+1 | (\hat{J}_-)^{\dagger} |j, m\rangle^*, \quad \text{from the definition of } \dagger \\ &= c_+ \hbar \langle j, m+1 | \hat{J}_+ |j, m\rangle \quad \text{since } (\hat{J}_-)^{\dagger} = \hat{J}_+ \\ &= c_+ c_+^* \hbar^2 \langle j, m+1 | j, m+1\rangle \\ &= |c_+|^2 \hbar^2, \quad \text{from orthonormality.} \end{aligned}$$

But we can evaluate the left hand side by making use of the identity Eq. (13.16):

$$\hat{J}_- \hat{J}_+ = \hat{J}^2 - \hat{J}_z^2 - \hbar \hat{J}_z,$$

yielding

$$\begin{aligned}
 \langle j, m | \hat{J}_- \hat{J}_+ | j, m \rangle &= \langle j, m | \hat{J}^2 - \hat{J}_z^2 - \hbar \hat{J}_z | j, m \rangle \\
 &= \langle j, m | j(j+1)\hbar^2 - m^2\hbar^2 - m\hbar^2 | j, m \rangle \\
 &= [j(j+1) - m(m+1)]\hbar^2, \quad \text{from orthonormality.}
 \end{aligned}$$

Thus we obtain

$$|c_+|^2 = j(j+1) - m(m+1).$$

In similar fashion,

$$\begin{aligned}
 \langle j, m | \hat{J}_+ \hat{J}_- | j, m \rangle &= c_- \hbar \langle j, m | \hat{J}_+ | j, m-1 \rangle \\
 &= c_- \hbar \langle j, m-1 | (\hat{J}_+)^{\dagger} | j, m \rangle^*, \quad \text{from the definition of } \dagger \\
 &= c_- \hbar \langle j, m-1 | \hat{J}_- | j, m \rangle, \quad \text{since } (\hat{J}_+)^{\dagger} = \hat{J}_- \\
 &= c_- c_-^* \hbar^2 \langle j, m-1 | j, m-1 \rangle \\
 &= |c_-|^2 \hbar^2, \quad \text{from orthonormality.}
 \end{aligned}$$

together with the other identity

$$\hat{J}_+ \hat{J}_- = \hat{J}^2 - \hat{J}_z^2 + \hbar \hat{J}_z,$$

yields

$$|c_-|^2 = j(j+1) - m(m-1).$$

Condon-Shortley Phase Convention

Taking c_{\pm} to be real and positive gives

$$c_{\pm} = \sqrt{j(j+1) - m(m \pm 1)}. \quad (13.19)$$

13.5 Matrix Representations

For a given j , the quantities $\langle j, m' | \hat{J}_z | j, m \rangle$ are known as the *matrix elements* of \hat{J}_z . We can calculate what they are as follows:

$$\langle j, m' | \hat{J}_z | j, m \rangle = m\hbar \langle j, m' | j, m \rangle = m\hbar \delta_{m', m}, \quad (13.20)$$

where we have used the orthonormality properties of the basis. Why matrix elements? As we have already seen, we can regard the labels m' and m as *labelling the rows and columns*,

respectively, of a matrix. Given that m' and m can each take $(2j + 1)$ values, the matrix in question is $(2j + 1) \times (2j + 1)$.

Similarly, the matrix elements of the raising and lowering operators are given by

$$\langle j, m' | \hat{J}_{\pm} | j, m \rangle = c_{\pm} \hbar \langle j, m' | j, m \pm 1 \rangle = \sqrt{j(j+1) - m(m \pm 1)} \hbar \delta_{m', m \pm 1}$$

Check that for $j = 1$ you recover the results obtained in Q4 of Problem Sheet 5.

13.6 Summary

As usual, we summarize the main concepts introduced in this lecture.

- Properties of the angular momentum from commutation relations.
- Derivation of the eigenvalue spectrum.
- Eigenstates and normalization.
- Matrix representations.

Lecture 14

Spin

14.1 Introduction

Using the commutation relations of the components of the angular momentum we have found that the allowed eigenvalues for \hat{J}^2 are $\hbar^2 j(j+1)$, where $j = 0, \frac{1}{2}, 1, \frac{3}{2}, \dots$. For each value of j , the eigenvalues of J_z are $\hbar m$, with $m = -j, -j+1, \dots, j-1, j$.

Comparing with the solutions of the eigensystem discussed in lecture 8, we see that we have found more solutions than there are in Eq. (8.21). Eq. (8.21) is a partial differential equation in θ and ϕ . The solutions of this equation are the spherical harmonics $Y_\ell^m(\theta, \phi)$, Eq. (8.22). We can see from the explicit expression for the spherical harmonics that the ϕ dependence is simply:

$$Y_\ell^m(\theta, \phi) \propto \exp(im\phi), \quad (14.1)$$

as expected, since the spherical harmonics are *also* eigenfunctions of $L_z = -i\hbar \frac{\partial}{\partial \phi}$. Since we required that wave functions must be single-valued, the spherical harmonics must be periodic in ϕ with period 2π :

$$Y_\ell^m(\theta, \phi) = Y_\ell^m(\theta, \phi + 2\pi). \quad (14.2)$$

Eq. (14.2) requires m to be integer, and hence j must also be an integer.

In order to understand the physical meaning of the solutions with half-integer j , let us investigate their properties in more detail.

14.2 Matrix representation

For $j = \frac{1}{2}$, we can compute the matrix elements: $\langle \frac{1}{2}, m' | \hat{J}_i | \frac{1}{2}, m \rangle$; the possible values for m' and m are: $m' = \frac{1}{2}$ or $-\frac{1}{2}$ and $m = \frac{1}{2}$ or $-\frac{1}{2}$. If we choose the convention that the row and column labels start with the largest value of the magnetic quantum number and decrease, so that the first row (column) is labelled by $m'(m) = \frac{1}{2}$ and the second row (column) is labelled by $m'(m) = -\frac{1}{2}$, we find the following 2×2 matrix:

$$\frac{\hbar}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

We say that this matrix *represents* the operator \hat{J}_z in the $j = \frac{1}{2}$ multiplet.

The only non-zero element for the matrix representing \hat{J}_+ is when $m' = \frac{1}{2}$ and $m = -\frac{1}{2}$, for which

$$c_+ = \sqrt{\frac{1}{2}(\frac{1}{2} + 1) + \frac{1}{2}(\frac{1}{2})} = 1,$$

and hence the matrix is

$$\hat{J}_+ \longrightarrow \hbar \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix},$$

whilst the only non-zero element for the matrix representing \hat{J}_- is when $m' = -\frac{1}{2}$ and $m = \frac{1}{2}$, for which

$$c_- = \sqrt{\frac{1}{2}(\frac{1}{2} + 1) + \frac{1}{2}(\frac{1}{2})} = 1,$$

also, giving

$$\hat{J}_- \longrightarrow \hbar \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}.$$

From these two matrices it is easy to construct the matrices representing \hat{J}_x and \hat{J}_y , since

$$\begin{aligned} \hat{J}_x &= \frac{1}{2}(\hat{J}_+ + \hat{J}_-), \\ \hat{J}_y &= \frac{1}{2i}(\hat{J}_+ - \hat{J}_-). \end{aligned}$$

Thus

$$\hat{J}_x \longrightarrow \frac{\hbar}{2} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \hat{J}_y \longrightarrow \frac{\hbar}{2} \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}.$$

You can readily verify that these 2×2 matrices satisfy the angular momentum commutation relations from which we started. We say, therefore, that they *provide a matrix representation* of the angular momentum operators.

The set of three numerical 2×2 matrices which appear above in the matrix representations of \hat{J}_x , \hat{J}_y and \hat{J}_z are known as the *Pauli spin matrices* and are usually denoted as follows:

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}; \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}; \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

Collectively, we can write

$$\hat{J} \longrightarrow \frac{1}{2}\hbar\sigma,$$

meaning $\hat{J}_x \longrightarrow \frac{1}{2}\hbar\sigma_x$, etc. Often we will just write $=$ instead of \longrightarrow , but you should remember that this is just one possible choice for representing the operators.

The Pauli matrices have the following property, which you can easily verify

$$\sigma_x^2 = \sigma_y^2 = \sigma_z^2 = 1,$$

where 1 denotes the unit 2×2 matrix.

14.3 Eigenvectors

It is trivial to show that the matrix σ_z has eigenvectors which are just two-component column matrices:

$$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = - \begin{pmatrix} 0 \\ 1 \end{pmatrix},$$

so that the eigenvalue equations for \hat{J}_z are

$$\frac{1}{2}\hbar \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \frac{1}{2}\hbar \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \frac{1}{2}\hbar \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = -\frac{1}{2}\hbar \begin{pmatrix} 0 \\ 1 \end{pmatrix},$$

and we see that \hat{J}_z has eigenvalues $\pm\frac{1}{2}\hbar$ as it should for a system with $j = \frac{1}{2}$.

Furthermore, if we construct the matrix representing \hat{J}^2 , we see that it has these same two column matrices as eigenvectors with a common eigenvalue $j(j+1)\hbar^2 \equiv \frac{3}{4}\hbar^2$:

$$\hat{J}^2 \equiv \hat{J}_x^2 + \hat{J}_y^2 + \hat{J}_z^2 = \frac{1}{4}\hbar^2 \left[\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}^2 + \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}^2 + \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}^2 \right] = \frac{3}{4}\hbar^2 \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

and

$$\frac{3}{4}\hbar^2 \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \frac{3}{4}\hbar^2 \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \frac{3}{4}\hbar^2 \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \frac{3}{4}\hbar^2 \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

We can thus identify the two column matrices with the two simultaneous eigenstates of the operators \hat{J}^2 and \hat{J}_z :

$$|j = \frac{1}{2}, m = \frac{1}{2}\rangle \longrightarrow \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad |j = \frac{1}{2}, m = -\frac{1}{2}\rangle \longrightarrow \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

We can also see that an arbitrary state $|\psi\rangle$ with $j = \frac{1}{2}$ may be represented as a linear combination of these two states since they span the two-dimensional space of 2-component column matrices:

$$|\psi\rangle = \sum_{m=-\frac{1}{2}}^{\frac{1}{2}} c_m |j = \frac{1}{2}, m\rangle$$

is represented by

$$\begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix} = c_{\frac{1}{2}} \begin{pmatrix} 1 \\ 0 \end{pmatrix} + c_{-\frac{1}{2}} \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

14.4 Scalar Products

The rule for scalar products is different when using a matrix representation; it doesn't involve any integration. The Dirac kets we have seen are represented by column matrices of rank $(2j+1)$; the corresponding conjugates, or Dirac bras, are represented by row matrices of the same rank. The rule is that if

$$|\psi\rangle \longrightarrow \begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix} \quad \text{then} \quad \langle\psi| \longrightarrow (\psi_1^* \quad \psi_2^*).$$

The scalar product of two states $|\psi\rangle$ and $|\phi\rangle$ is then defined to be

$$\langle\phi|\psi\rangle \equiv (\phi_1^* \quad \phi_2^*) \begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix} = \phi_1^* \psi_1 + \phi_2^* \psi_2.$$

Thus, for example, for a normalised state

$$\langle\psi|\psi\rangle = \psi_1^* \psi_1 + \psi_2^* \psi_2 = |\psi_1|^2 + |\psi_2|^2 = 1$$

The orthonormality property of the eigenvectors is also obvious:

$$\begin{pmatrix} 1 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = 1, \quad \begin{pmatrix} 0 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = 1, \quad \begin{pmatrix} 1 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = 0, \quad \begin{pmatrix} 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = 0,$$

and these may be used to project out the coefficients c_m in the expansion of the arbitrary state $|\psi\rangle$.

14.5 Eigenvectors of \hat{J}_x

So far we have concentrated on the eigenvalues and eigenstates of \hat{J}_z , but what of the other Cartesian components of angular momentum? It is clear that, since there is nothing special about the z direction, we should also expect that measuring say the x component of the angular momentum for a system with $j = \frac{1}{2}$ can only yield either $\pm\frac{1}{2}\hbar$. Let us verify this. The matrix representing \hat{J}_x is $\frac{1}{2}\hbar\sigma_x$ so we need to find the eigenvalues and eigenvectors of the 2×2 matrix σ_x . Let us write

$$\sigma_x \chi = \rho \chi \quad \text{with} \quad \chi = \begin{pmatrix} \chi_1 \\ \chi_2 \end{pmatrix},$$

where ρ denotes an eigenvalue and χ the corresponding eigenvector. We find the eigenvalues by rewriting this as

$$(\sigma_x - \rho 1) \chi = 0,$$

where 1 denotes the unit 2×2 matrix. This is a pair of simultaneous equations for χ_1 and χ_2 , which only have a non-trivial solution if the determinant of the 2×2 coefficient matrix on the left-hand side is singular. The condition for this is

$$\det(\sigma_x - \rho 1) = \begin{vmatrix} -\rho & 1 \\ 1 & -\rho \end{vmatrix} = 0,$$

which yields

$$\rho^2 - 1 = 0 \quad \text{implying that} \quad \rho = \pm 1.$$

The eigenvalues of \hat{J}_x are thus $\pm \frac{1}{2}\hbar$ as anticipated.

More generally, the eigenvalues of \hat{J}_x are written $m_x \hbar$. In this case, where $j = \frac{1}{2}$, we have $m_x = \pm \frac{1}{2}$.

Let us now find the eigenvectors corresponding to the two eigenvalues.

The case $\rho = 1$ The equation for the eigenvectors becomes:

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} \chi_1 \\ \chi_2 \end{pmatrix} = \begin{pmatrix} \chi_1 \\ \chi_2 \end{pmatrix} \Rightarrow \begin{pmatrix} \chi_2 \\ \chi_1 \end{pmatrix} = \begin{pmatrix} \chi_1 \\ \chi_2 \end{pmatrix} \Rightarrow \chi_2 = \chi_1.$$

We can pick any 2-component column matrix which satisfies this condition. In particular, a suitably normalised eigenvector which represents the state with $j = \frac{1}{2}$ and $m_x = \frac{1}{2}$ is

$$|j = \frac{1}{2}, m_x = \frac{1}{2}\rangle \longrightarrow \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

The case $\rho = -1$ In close analogy with the computation above, let us write:

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} \chi_1 \\ \chi_2 \end{pmatrix} = - \begin{pmatrix} \chi_1 \\ \chi_2 \end{pmatrix} \Rightarrow \begin{pmatrix} \chi_2 \\ \chi_1 \end{pmatrix} = - \begin{pmatrix} \chi_1 \\ \chi_2 \end{pmatrix} \Rightarrow \chi_2 = -\chi_1.$$

Thus a suitably normalised eigenvector which represents the state with $j = \frac{1}{2}$ and $m_x = -\frac{1}{2}$ is

$$|j = \frac{1}{2}, m_x = -\frac{1}{2}\rangle \longrightarrow \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

Comments Let us briefly comment on the solutions found above.

- The eigenvectors corresponding to $m_x = \frac{1}{2}$ and $m_x = -\frac{1}{2}$ are orthogonal, as they must be:

$$\frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix} = 0.$$

- The eigenvectors of \hat{J}_x are expressible as linear combinations of the eigenvectors of \hat{J}_z :

$$\frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \end{pmatrix},$$

and

$$\frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \end{pmatrix} - \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Thus if a system with $j = \frac{1}{2}$ is in an eigenstate of \hat{J}_x , for example with $m_x = \frac{1}{2}$, then the probability that a measurement of \hat{J}_z yields the result $m = \frac{1}{2}$ is $|\frac{1}{\sqrt{2}}|^2 = \frac{1}{2}$.

What has emerged from this analysis is that we can consider systems with $j = 1/2$ as having *intrinsic* angular momentum, which has nothing to do with the orbital motion of the particle about some point. It is a property of the system in their own rest frame, and can be seen as an internal degree of freedom of the particle. Hence the wave function describing the state of the system must also depend on an index m labelling the values of the internal degrees of freedom. For the case $j = 1/2$, m can take two values, and therefore the wave functions have two components as discussed above. For the general case of spin j , the wave functions have $2j + 1$ components. This intrinsic angular momentum is known as *spin* and doesn't really have any classical analogue. Electrons, protons, neutrons and many of the more unstable particles have spin $\frac{1}{2}$.

The theory that we have developed for $j = \frac{1}{2}$ provides the framework for analysing the quantum mechanics of spin $\frac{1}{2}$ particles. Conventionally we write $s = \frac{1}{2}$ rather than $j = \frac{1}{2}$ when discussing such particles. The spin angular momentum operator is written \hat{S} . \hat{S}_z has eigenvalues $m_s \hbar$ with $m_s = \pm \frac{1}{2}$. Often these two states, with $m_s = \pm \frac{1}{2}$, are referred to as 'spin up' and 'spin down' respectively.

Of course, the wavefunction of a spin- $\frac{1}{2}$ particle also has a spatial dependence so the complete specification of the state is of the form

$$\psi = \psi_1(\underline{r}) \alpha + \psi_2(\underline{r}) \beta = \begin{pmatrix} \psi_1(\underline{r}) \\ \psi_2(\underline{r}) \end{pmatrix},$$

where

$$\alpha \equiv \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \beta \equiv \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

The probability interpretation is then a straightforward generalisation:

$|\psi_i(\underline{r})|^2 d\tau$ is the probability of finding the particle in the volume $d\tau$ at \underline{r} with z -component of spin $\frac{1}{2}\hbar$ if $i = 1$ or spin $-\frac{1}{2}\hbar$ if $i = 2$.

14.6 The Stern-Gerlach Experiment

We can now understand the result of the original Stern-Gerlach experiment, which was conducted with a beam of silver atoms and found two emergent beams, corresponding to $j = \frac{1}{2}$.

Let us now consider a more elaborate experiment involving not one but several Stern-Gerlach magnets, which we use to make *successive* measurements of various components of angular momentum.

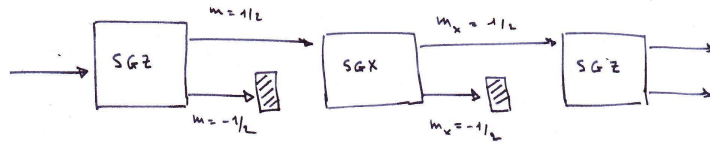


Figure 14.1: A beam of atoms, each with $j = \frac{1}{2}$, travelling along the y -axis passes through a sequence of Stern-Gerlach magnets whose mean fields are oriented along either the z -direction (SGZ) or the x -direction (SGX). The shaded boxes represent absorbers.

We assume that we can neglect any interaction between the particles in the beam. The two beams emerging from the first magnet have $m = \frac{1}{2}$ and $m = -\frac{1}{2}$, respectively, but only the former is allowed to proceed to the second magnet. Thus *we know that each particle entering the second magnet is in the state* $|j = \frac{1}{2}, m = \frac{1}{2}\rangle$, represented by the column matrix

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

The second magnet, which has its mean field aligned with the x -direction, serves to measure the x -component of angular momentum. We can predict the outcome by expanding the state $|j = \frac{1}{2}, m = \frac{1}{2}\rangle$ in eigenstates of \hat{J}_x and finding the probability amplitudes for the two possible outcomes, $m_x = \frac{1}{2}$ and $m_x = -\frac{1}{2}$. Thus

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix} = a \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} + b \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

We find the amplitudes a and b by orthogonal projection in the usual way:

$$a = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \frac{1}{\sqrt{2}},$$

$$b = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \frac{1}{\sqrt{2}}.$$

We have then for the desired probabilities

$$\text{probability of getting } m_x = \frac{1}{2} \quad \text{is } |a|^2 = \frac{1}{2}$$

$$\text{probability of getting } m_x = -\frac{1}{2} \quad \text{is } |b|^2 = \frac{1}{2}.$$

Since each particle is therefore equally likely to be found with $m_x = \frac{1}{2}$ or $m_x = -\frac{1}{2}$, equal numbers, on average, go into each of the two emergent beams and so the two beams will have *equal intensity*.

Regeneration What happens if we select only those particles with $m_x = \frac{1}{2}$ emerging from the second magnet and allow them to impinge on a third magnet whose mean field is aligned with the z -direction? This is the situation illustrated in Fig. 14.1. We are, in effect, remeasuring \hat{J}_z by means of the third apparatus. We know that the state of particles entering the third magnet is $|j = \frac{1}{2}, m_x = \frac{1}{2}\rangle$ and we can expand this state in terms of the complete set of eigenstates of \hat{J}_z . The expansion coefficients will be the probability amplitudes required to compute the probabilities of getting the two possible outcomes $m = \frac{1}{2}$ and $m = -\frac{1}{2}$ when we measure \hat{J}_z for each particle:

$$\frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

and we see that the desired amplitudes are both $\frac{1}{\sqrt{2}}$, so giving equal probabilities for the two outcomes. The remarkable feature of this result is that the probability of getting $m = -\frac{1}{2}$ is non-zero *despite our having eliminated the beam with $m = -\frac{1}{2}$ which emerged from the first magnet!* This phenomenon is referred to as *regeneration*. It has arisen here because the second measurement, of the x -component of angular momentum, was *incompatible with the first measurement*, of the z -component.

General Remarks More generally, if the second apparatus is aligned so that its mean field lies not in the x -direction, but in the $x - z$ plane at an angle θ to the z -axis, then it measures the component of angular momentum not along the x -direction but along the direction of a unit vector

$$\underline{n} = \sin \theta \underline{i} + \cos \theta \underline{k},$$

where \underline{i} and \underline{k} are the usual Cartesian unit vectors in the x - and z -directions respectively.

The relevant eigenstates are then those of the matrix

$$\underline{\sigma} \cdot \underline{n} = \sigma_x \sin \theta + \sigma_z \cos \theta = \begin{pmatrix} \cos \theta & \sin \theta \\ \sin \theta & -\cos \theta \end{pmatrix}.$$

14.7 Summary

As usual, we summarize the main concepts introduced in this lecture.

- Half-integer values of the angular momentum. Orbital angular momentum can only have integer values of j .
- Matrix representation of the angular momentum for $j = 1/2$.
- Two-dimensional complex space of states with $j = 1/2$.
- Spin as an internal degree of freedom.
- Examples.

Lecture 15

Addition of angular momenta

15.1 Introduction

We now turn to the problem of dealing with two angular momenta. For example, we might wish to consider an electron which has both an intrinsic spin and some orbital angular momentum, as in a real hydrogen atom. Or we might have a system of two electrons and wish to know what possible values the total spin of the system can take. Classically, the angular momentum is a vector quantity, and the total angular momentum is simply $\underline{J} = \underline{J}_1 + \underline{J}_2$. The maximum and minimum values that \underline{J} can take correspond to the case where either \underline{J}_1 and \underline{J}_2 are parallel, so that the magnitude of \underline{J} is $|\underline{J}_1| + |\underline{J}_2|$ or antiparallel when it has magnitude $||\underline{J}_1| - |\underline{J}_2||$.

This lecture discusses the addition of angular momenta for a quantum system.

15.2 Total angular momentum operator

In the quantum case, the total angular momentum is represented by the operator

$$\hat{J} \equiv \hat{J}_1 + \hat{J}_2.$$

We assume that \hat{J}_1 and \hat{J}_2 are *independent angular momenta*, meaning each satisfies the usual angular momentum commutation relations

$$[\hat{J}_{nx}, \hat{J}_{ny}] = i\hbar \hat{J}_{nz}, \quad \text{etc.}, \quad [\hat{J}_n^2, \hat{J}_{ni}] \quad \text{etc.}$$

where $n = 1, 2$ labels the individual angular momenta, $i = x, y, z$ and etc. stands for cyclic permutations. *Furthermore, any component of \hat{J}_1 commutes with any component of \hat{J}_2 :*

$$[\hat{J}_{1i}, \hat{J}_{2k}] = 0, \quad i, k = x, y, z,$$

so that the two angular momenta are *compatible*. It follows that the four operators $\hat{J}_1^2, \hat{J}_{1z}, \hat{J}_2^2, \hat{J}_{2z}$ are *mutually commuting* and so must possess a common eigenbasis. This common eigenbasis is known as the *uncoupled basis* and is denoted $\{|j_1, m_1, j_2, m_2\rangle\}$ in Dirac notation. It has the following properties:

$$\begin{array}{l} \hat{J}_1^2 |j_1, m_1, j_2, m_2\rangle = j_1(j_1 + 1)\hbar^2 |j_1, m_1, j_2, m_2\rangle \\ \hat{J}_{1z} |j_1, m_1, j_2, m_2\rangle = m_1\hbar |j_1, m_1, j_2, m_2\rangle \\ \hat{J}_2^2 |j_1, m_1, j_2, m_2\rangle = j_2(j_2 + 1)\hbar^2 |j_1, m_1, j_2, m_2\rangle \\ \hat{J}_{2z} |j_1, m_1, j_2, m_2\rangle = m_2\hbar |j_1, m_1, j_2, m_2\rangle \end{array}.$$

It is easy to establish that the total angular momentum operators satisfy the usual commutation relations

$$[\hat{J}_x, \hat{J}_y] = i\hbar \hat{J}_z, \quad \text{etc.}, \quad [\hat{J}^2, \hat{J}_i] = 0.$$

As for any angular momentum operator then, \hat{J}^2 has eigenvalues $j(j+1)\hbar^2$ whilst the operator corresponding to the z -component of the total angular momentum has eigenvalues $m\hbar$ with m running between j and $-j$ in integer steps for a given j .

15.3 Addition Theorem

The question which then arises is, given two angular momenta, corresponding to angular momentum quantum numbers j_1 and j_2 respectively, what are the allowed values of the total angular momentum quantum number, j ? The answer is provided by the *Angular Momentum Addition Theorem*:

The allowed values of the total angular momentum quantum number j , given two angular momenta corresponding to quantum numbers j_1 and j_2 are:

$$j = j_1 + j_2, j_1 + j_2 - 1, \dots, |j_1 - j_2|$$

and for each of these values of j , m takes on the $(2j + 1)$ values

$$m = j, j - 1, \dots, -j.$$

The proof of this theorem is beyond the scope of these lectures, and is deferred to more advanced courses.

It is easy to show that \hat{J}^2 commutes with \hat{J}_1^2 and \hat{J}_2^2 but *not* with \hat{J}_{1z} or with \hat{J}_{2z} by writing

$$\hat{J}^2 \equiv (\hat{\underline{J}}_1 + \hat{\underline{J}}_2)^2 = \{\hat{J}_1^2 + \hat{J}_2^2 + 2\hat{\underline{J}}_1 \cdot \hat{\underline{J}}_2\}.$$

The dot product contains the x and y components of the two angular momenta which do not commute with the respective z components.

The operator \hat{J}_z commutes with \hat{J}_1^2 and \hat{J}_2^2 and so the set of four operators $\hat{J}^2, \hat{J}_z, \hat{J}_1^2, \hat{J}_2^2$ are *also* a mutually commuting set of operators with a common eigenbasis known as the *coupled basis*, denoted $\{|j, m, j_1, j_2\rangle\}$ and satisfying

$$\begin{aligned} \hat{J}^2 |j, m, j_1, j_2\rangle &= j(j+1)\hbar^2 |j, m, j_1, j_2\rangle \\ \hat{J}_z |j, m, j_1, j_2\rangle &= m\hbar |j, m, j_1, j_2\rangle \\ \hat{J}_1^2 |j, m, j_1, j_2\rangle &= j_1(j_1+1)\hbar^2 |j, m, j_1, j_2\rangle \\ \hat{J}_2^2 |j, m, j_1, j_2\rangle &= j_2(j_2+1)\hbar^2 |j, m, j_1, j_2\rangle \end{aligned}.$$

These are states of *definite total angular momentum and definite z component of total angular momentum* but *not* in general states with definite J_{1z} or J_{2z} . In fact, they are expressible as linear combinations of the states of the uncoupled basis, with coefficients known as *Clebsch-Gordan coefficients*, which you can find tabulated in many textbooks.

Example We consider the important case of two spin- $\frac{1}{2}$ particles for which the spin quantum numbers are $s_1 = \frac{1}{2}$ and $s_2 = \frac{1}{2}$ respectively. According to the theorem, the total spin quantum number s takes on the values $s_1 + s_2 \equiv 1$ and $|s_1 - s_2| \equiv 0$ only.

Thus two electrons can have a total spin of 1 or 0 only: these states of definite total spin are referred to as *triplet* and *singlet* states respectively, because in the former, there are three possible values of the spin magnetic quantum number, $m_s = 1, 0, -1$, whereas in the latter there is only one such value, $m_s = 0$.

The states of the uncoupled basis are

$$\alpha_1 \alpha_2, \quad \alpha_1 \beta_2, \quad \beta_1 \alpha_2, \quad \beta_1 \beta_2,$$

where the subscripts 1 and 2 refer to electrons 1 and 2 respectively. The operators \hat{S}_1^2 and \hat{S}_{1z} act only on the parts labelled 1, whilst \hat{S}_2^2 and \hat{S}_{2z} act only on the parts labelled 2.

It should be clear that since $\alpha_1 \alpha_2$ has $m_{s_1} = \frac{1}{2}$ and $m_{s_2} = \frac{1}{2}$ it must have $m_s = 1$, that is, total z -component of spin \hbar , and *can therefore only be $s = 1$ and not $s = 0$* . This is an example of what is known as a *stretched state*: it has the maximum possible value of the z component of total angular momentum (spin) and must therefore be a member of both the coupled and uncoupled basis: $\alpha_1 \alpha_2 \equiv |s = 1, m_s = 1, s_1 = \frac{1}{2}, s_2 = \frac{1}{2}\rangle$.

A similar argument shows that $\beta_1 \beta_2$ has $m_s = -1$ and thus also can only be $s = 1$, so that $\beta_1 \beta_2 \equiv |s = 1, m_s = -1, s_1 = \frac{1}{2}, s_2 = \frac{1}{2}\rangle$. The remaining two states of the coupled basis are, however, non-trivial linear combinations of the two remaining states of the uncoupled basis:

$$\begin{aligned} |s = 1, m_s = 0, s_1 = \frac{1}{2}, s_2 = \frac{1}{2}\rangle &= \frac{1}{\sqrt{2}}[\alpha_1 \beta_2 + \beta_1 \alpha_2] \\ |s = 0, m_s = 0, s_1 = \frac{1}{2}, s_2 = \frac{1}{2}\rangle &= \frac{1}{\sqrt{2}}[\alpha_1 \beta_2 - \beta_1 \alpha_2]. \end{aligned}$$

Proof:

We apply the lowering operator, \hat{S}_- for the z component of total spin to the stretched state:

$$\hat{S}_- \alpha_1 \alpha_2 \equiv (\hat{S}_{1-} + \hat{S}_{2-}) \alpha_1 \alpha_2.$$

The left-hand side we write as

$$\begin{aligned} \hat{S}_- |s = 1, m_s = 1, s_1 = \frac{1}{2}, s_2 = \frac{1}{2}\rangle &= \sqrt{1(1+1) - 1(1-1)} \hbar |s = 1, m_s = 0, s_1 = \frac{1}{2}, s_2 = \frac{1}{2}\rangle \\ &= \sqrt{2} \hbar |s = 1, m_s = 0, s_1 = \frac{1}{2}, s_2 = \frac{1}{2}\rangle, \end{aligned}$$

using the usual properties of the lowering operator, whilst we write the right-hand side as

$$(\hat{S}_{1-} + \hat{S}_{2-}) |s_1 = \frac{1}{2}, m_{s_1} = \frac{1}{2}, s_2 = \frac{1}{2}, m_{s_2} = \frac{1}{2}\rangle,$$

and note that

$$\begin{aligned} \hat{S}_{1-} |s_1 = \frac{1}{2}, m_{s_1} = \frac{1}{2}, s_2 = \frac{1}{2}, m_{s_2} = \frac{1}{2}\rangle &= \\ \sqrt{\frac{1}{2}(\frac{1}{2} + 1) - \frac{1}{2}(\frac{1}{2} - 1)} \hbar |s_1 = \frac{1}{2}, m_{s_1} = -\frac{1}{2}, s_2 = \frac{1}{2}, m_{s_2} = \frac{1}{2}\rangle &= \\ = \hbar |s_1 = \frac{1}{2}, m_{s_1} = -\frac{1}{2}, s_2 = \frac{1}{2}, m_{s_2} = \frac{1}{2}\rangle &\equiv \hbar \beta_1 \alpha_2, \end{aligned}$$

whilst

$$\begin{aligned} \hat{S}_{2-} |s_1 = \frac{1}{2}, m_{s_1} = \frac{1}{2}, s_2 = \frac{1}{2}, m_{s_2} = \frac{1}{2}\rangle &= \\ \sqrt{\frac{1}{2}(\frac{1}{2} + 1) - \frac{1}{2}(\frac{1}{2} - 1)} \hbar |s_1 = \frac{1}{2}, m_{s_1} = \frac{1}{2}, s_2 = \frac{1}{2}, m_{s_2} = -\frac{1}{2}\rangle & \\ = \hbar |s_1 = \frac{1}{2}, m_{s_1} = \frac{1}{2}, s_2 = \frac{1}{2}, m_{s_2} = -\frac{1}{2}\rangle &\equiv \hbar \alpha_1 \beta_2. \end{aligned}$$

Equating the two sides then yields the stated result:

$$|s = 1, m_s = 0, s_1 = \frac{1}{2}, s_2 = \frac{1}{2}\rangle = \frac{1}{\sqrt{2}}[\alpha_1 \beta_2 + \beta_1 \alpha_2].$$

The remaining member of the coupled basis must be a linear combination of $\alpha_1 \beta_2$ and $\beta_1 \alpha_2$ orthogonal to this, which we can take to be

$$|s = 0, m_s = 0, s_1 = \frac{1}{2}, s_2 = \frac{1}{2}\rangle = \frac{1}{\sqrt{2}}[\alpha_1 \beta_2 - \beta_1 \alpha_2].$$

15.4 Summary

As usual, we summarize the main concepts introduced in this lecture.

- Angular momentum operator for a system of two particles.
- Choice of sets of commuting observables.
- Addition theorem.
- Examples

Lecture 16

Identical Particles

16.1 Introduction

There are many systems in nature that are made of several particles of the same species. These particles all have the same mass, charge, and spin. For instance the electrons in an atom are identical particles. Identical particles cannot be distinguished by measuring their properties. This is also true for classical particles. In classical mechanics we can always follow the trajectory of each individual particle, i.e. their time evolution in space. The trajectories identify each particle in classical mechanics, making identical particles distinguishable.

In quantum mechanics the concept of trajectory does not exist and identical particles are *indistinguishable*. Let us consider for simplicity a system of two identical particles. The state of the system is described by a wave function:

$$\psi(\xi_1, \xi_2), \quad \xi = \{\underline{x}, \sigma\}, \quad (16.1)$$

where \underline{x} yields the position of the particle, and σ yields the z -component of the spin of the particle, if the latter is different from zero.

The state with the two particles exchanged is described by the wave function:

$$\psi(\xi_2, \xi_1). \quad (16.2)$$

If the two particles are identical, the two functions represent the *same* quantum state, and therefore:

$$\psi(\xi_1, \xi_2) = e^{i\alpha} \psi(\xi_2, \xi_1). \quad (16.3)$$

Repeating the exchange of the two particles we find:

$$e^{2i\alpha} = 1 \implies e^{i\alpha} = \pm 1. \quad (16.4)$$

Hence the wave function of a system of two identical particles must be either symmetric or antisymmetric under the exchange of the two particles.

The Spin-Statistics Theorem

Systems of identical particles with integer spin ($s = 0, 1, 2, \dots$), known as *bosons*, have wave functions which are symmetric under interchange of any pair of particle labels. The wave function is said to obey *Bose-Einstein* statistics.

Systems of identical particles with half-odd-integer spin ($s = \frac{1}{2}, \frac{3}{2}, \dots$), known as *fermions*, have wave functions which are antisymmetric under interchange of any pair of particle labels. The wave function is said to obey *Fermi-Dirac* statistics.

This law was discovered by Wolfgang Pauli and is supported by experimental evidence.

16.2 A first look at Helium

In the simplest model of the helium atom, the Hamiltonian is

$$\hat{H} = \hat{H}_1 + \hat{H}_2 + \frac{e^2}{4\pi\epsilon_0|r_1 - r_2|}$$

where

$$\hat{H}_i = \frac{\hat{p}_i^2}{2\mu} - \frac{2e^2}{4\pi\epsilon_0 r_i}$$

Note that it is *symmetric* under permutation of the indices 1 & 2 which label the two electrons. This must be the case if the two electrons are *identical* or *indistinguishable*: it cannot matter which particle we label 1 and which we label 2.

This observation is quite general: the same argument holds for identical particles other than electrons and can be extended to systems of more than two identical particles.

Let us write the symmetry condition concisely as:

$$\hat{H}(1, 2) = \hat{H}(2, 1)$$

Suppose that

$$\hat{H}(1, 2)\psi(1, 2) = E\psi(1, 2)$$

then interchanging the labels 1 & 2 gives

$$\hat{H}(2, 1)\psi(2, 1) = E\psi(2, 1)$$

but using the symmetry property of \hat{H} means that

$$\hat{H}(1, 2)\psi(2, 1) = E\psi(2, 1)$$

so we conclude that $\psi(1, 2)$ and $\psi(2, 1)$ are both eigenfunctions belonging to the *same* eigenvalue, E , as is any linear combination of $\psi(1, 2)$ and $\psi(2, 1)$. In particular, the normalised symmetric and antisymmetric combinations

$$\psi_{\pm} \equiv \frac{1}{\sqrt{2}}\{\psi(1, 2) \pm \psi(2, 1)\}$$

are eigenfunctions belonging to the eigenvalue, E .

If we introduce a particle interchange operator, P_{12} , with the property that

$$P_{12}\psi(1, 2) = \psi(2, 1)$$

then the symmetric and antisymmetric combinations are eigenfunctions of P_{12} with eigenvalues ± 1 respectively:

$$P_{12}\psi_{\pm} = \pm\psi_{\pm}$$

Since ψ_{\pm} are simultaneous eigenfunctions of \hat{H} and P_{12} it follows that $[\hat{H}, P_{12}] = 0$, as you can easily verify from the above equations, and that the symmetry of the wavefunction is a constant of the motion.

16.3 Two-electron wave function

In the previous lecture, we constructed the states of the coupled representation for two spin- $\frac{1}{2}$ electrons, the three triplet states:

$$\begin{aligned}\chi_{1,1} &= \alpha_1\alpha_2 \\ \chi_{1,0} &= \frac{1}{\sqrt{2}}\{\alpha_1\beta_2 + \beta_1\alpha_2\} \\ \chi_{1,-1} &= \beta_1\beta_2\end{aligned}$$

and the singlet state:

$$\chi_{0,0} = \frac{1}{\sqrt{2}}\{\alpha_1\beta_2 - \beta_1\alpha_2\}$$

where we have used a simplified notation for the states of the coupled basis:

$$\chi_{s,m_s} \equiv |s, m_s, s_1, s_2\rangle.$$

Notice that the triplet states are *symmetric* under interchange of the labels 1 and 2, whereas the singlet state is *antisymmetric*. If we are to satisfy the Spin-Statistics Theorem, this has implications for the symmetry of the spatial wavefunctions that we combine with the spin functions to give the full wavefunction of the 2-electron system. The 2-electron wavefunction will have the general form

$$\Psi(1, 2) = \psi(\underline{r}_1, \underline{r}_2) \cdot \chi$$

The symmetry properties of the various factors are as follows:

	symmetry of χ_s	symmetry of ψ	symmetry of Ψ
$s = 0$ (singlet)	a	s	a
$s = 1$ (triplet)	s	a	a

Thus the spatial wavefunction must be *antisymmetric* if the two electrons are in a *spin triplet state* but *symmetric* if they are in a *spin singlet state*.

16.4 More on the He atom

Suppose for the moment that we neglect spin and also neglect the mutual Coulomb repulsion between the two electrons. That is, we treat the two electrons as moving independently in the Coulomb field of the nucleus. The Hamiltonian then reduces to

$$\hat{H} = \hat{H}_1 + \hat{H}_2 \quad \text{where} \quad \hat{H}_i = \frac{\hat{p}_i^2}{2\mu} - \frac{2e^2}{4\pi\epsilon_0 r_i}$$

We already know what the eigenfunctions and eigenvalues for \hat{H}_1 and \hat{H}_2 are, namely

$$\begin{aligned}\hat{H}_1 u_{n_1 \ell_1 m_1}(\mathbf{r}_1) &= E_{n_1} u_{n_1 \ell_1 m_1}(\mathbf{r}_1) \\ \hat{H}_2 u_{n_2 \ell_2 m_2}(\mathbf{r}_2) &= E_{n_2} u_{n_2 \ell_2 m_2}(\mathbf{r}_2)\end{aligned}$$

so it is easy to see that \hat{H} has eigenfunctions which are just products of the 1-electron eigenfunctions:

$$\begin{aligned}\hat{H} u_{n_1 \ell_1 m_1}(\mathbf{r}_1) u_{n_2 \ell_2 m_2}(\mathbf{r}_2) &\equiv (\hat{H}_1 + \hat{H}_2) u_{n_1 \ell_1 m_1}(\mathbf{r}_1) u_{n_2 \ell_2 m_2}(\mathbf{r}_2) \\ &= (E_{n_1} + E_{n_2}) u_{n_1 \ell_1 m_1}(\mathbf{r}_1) u_{n_2 \ell_2 m_2}(\mathbf{r}_2) \equiv E_n u_{n_1 \ell_1 m_1}(\mathbf{r}_1) u_{n_2 \ell_2 m_2}(\mathbf{r}_2)\end{aligned}$$

Thus the energy eigenvalues are given by

$$E_n = E_{n_1} + E_{n_2} \quad \text{where} \quad E_{n_i} = -\frac{\mu}{2\hbar^2} \left(\frac{Ze^2}{4\pi\epsilon_0} \right)^2 \frac{1}{n_i^2}$$

The Ground State:

In this crude model the ground state energy is just

$$E_{n=1} = E_{n_1=1} + E_{n_2=1} = 2 E_{n_1=1}.$$

Setting $Z = 2$ in the Bohr formula thus yields for the ground state energy:

$$E_1 = 8 \times (-13.6 \text{ eV}) = -108.8 \text{ eV}$$

to be compared with the experimentally measured value of -78.957 eV .

The ground state spatial wavefunction has $n_1 = n_2 = 1$ and $\ell_1 = \ell_2 = m_1 = m_2 = 0$ and is thus

$$u_{100}(\mathbf{r}_1) u_{100}(\mathbf{r}_2).$$

Each electron is in a $1s$ state; we say that the *electronic configuration* is $(1s)^2$.

If we now worry about spin, we remember that the total wavefunction is a product of a spatial wavefunction and a spin wavefunction of the correct symmetry. But the spatial wavefunction is *symmetric* and can thus *only be combined with a spin singlet spin function* to give an overall antisymmetric 2-electron wavefunction;

$$\Psi(\text{ground state}) = u_{100}(\mathbf{r}_1) u_{100}(\mathbf{r}_2) \chi_{0,0}.$$

Notice that, since $\ell_1 = \ell_2 = 0$, the total orbital angular momentum quantum number of the ground state configuration is $\ell = \ell_1 + \ell_2 = 0$. Thus the ground state has zero orbital and spin angular momentum, and hence zero total angular momentum.

The First Excited States:

The first excited states correspond to one electron being excited to a $2s$ or $2p$ state, with the other remaining in a $1s$ state. The electronic configurations are denoted by $(1s)(2s)$ and $(1s)(2p)$ respectively. The (degenerate) energy eigenvalue can again be obtained from the Bohr formula with $Z = 2$:

$$E_{n=2} = E_{n_1=1} + E_{n_2=2} = E_{n_1=2} + E_{n_2=1} = 5 \times -13.6 \text{ eV} = -68.0 \text{ eV}$$

In this case it is possible to construct spatial wavefunctions which are either *symmetric* or *antisymmetric*. The overall antisymmetric combinations are then:

$$\begin{aligned} \Psi^{\text{singlet}}(1, 2) &= \frac{1}{\sqrt{2}} \{u_{100}(\underline{r}_1) u_{2\ell m_\ell}(\underline{r}_2) + u_{2\ell m_\ell}(\underline{r}_1) u_{100}(\underline{r}_2)\} \chi_{0,0} \\ \Psi^{\text{triplet}}(1, 2) &= \frac{1}{\sqrt{2}} \{u_{100}(\underline{r}_1) u_{2\ell m_\ell}(\underline{r}_2) - u_{2\ell m_\ell}(\underline{r}_1) u_{100}(\underline{r}_2)\} \chi_{1, m_s} \end{aligned}$$

16.5 Pauli exclusion principle

The results that we have just obtained for the independent particle approximation to the helium atom illustrate a more general result, related to the Spin-Statistics Theorem and known as *the Pauli Exclusion Principle* which states

No two identical fermions can be in the same quantum state

For example,

- in the ground state, we see that although both electrons have $n = 1$ and $\ell = m_\ell = 0$ i.e. both are in a $1s$ state, they are in a spin singlet state, which means that if one electron is in the spin state α , the other must be in the state β : the two electrons cannot have an identical set of quantum numbers; if both were in the spin state α , the 2-electron spin state would be a triplet state, which is ruled out by the Spin-Statistics Theorem;
- in any excited state, both electrons can be in the spin state α , corresponding to the triplet state, but then the spatial wavefunction is forced to be antisymmetric, so that the quantum numbers n, ℓ, m_ℓ , of the two electrons have to differ - otherwise the spatial wavefunction would vanish identically!

No such restriction applies to identical bosons; any number of identical bosons can occupy the same quantum state.

Note that the correlation between spin and statistics has been postulated in the nonrelativistic context used in this course. The spin-statistic theorem can actually be derived in a relativistic formulation of quantum mechanics. It is a consequence of the principles of special relativity, quantum mechanics, and the positivity of the energy.

16.6 Summary

As usual, we summarize the main concepts introduced in this lecture.

- Indistinguishable particles in quantum mechanics.
- Symmetry of the wave function, example of the He atom.
- Two-electron wave function, combining spin and spatial wave functions.
- More on the He atom.
- Pauli exclusion principle.

Lecture 17

Perturbation Theory

17.1 Introduction

So far we have concentrated on systems for which we could find exactly the eigenvalues and eigenfunctions of the Hamiltonian, like e.g. the harmonic oscillator, the quantum rotator, or the hydrogen atom.

However the vast majority of systems in Nature cannot be solved exactly, and we need to develop appropriate tools to deal with them.

Perturbation theory is extremely successful in dealing with those cases that can be modelled as a “small deformation” of a system that we can solve exactly.

Let us translate the above statement into a precise mathematical framework. We are going to consider systems that have an Hamiltonian:

$$\hat{H} = \hat{H}_0 + \epsilon \hat{V}, \quad (17.1)$$

where \hat{H}_0 is the Hamiltonian of the unperturbed system, ϵ is a small parameter, and \hat{V} is the potential describing the perturbation. We shall assume that the perturbation V is independent of time.

Let us also assume that we can solve the time-independent Schrödinger equation for \hat{H}_0 , i.e. that we know its eigenvalues and eigenfunctions:

$$\hat{H}_0 \psi^{(n)}(x) = E^{(n)} \psi^{(n)}(x). \quad (17.2)$$

For simplicity we start by considering the case where all the unperturbed levels $E^{(n)}$ are not degenerate.

17.2 Perturbative solution

Let us discuss the solution of the time-independent Schrödinger equation for the full Hamiltonian H . The eigenvalue equation reads:

$$\hat{H} \psi(x) = E \psi(x). \quad (17.3)$$

Since ϵ is a small parameter, we shall expand the solution of Eq. (17.3) as a Taylor series in ϵ :

$$\psi(x) = \psi_0(x) + \epsilon \psi_1(x) + \epsilon^2 \psi_2(x) + \dots, \quad (17.4)$$

$$E = E_0 + \epsilon E_1 + \epsilon^2 E_2 + \dots. \quad (17.5)$$

Plugging Eqs. (17.4) and (17.5) into Eq. (17.3), we obtain:

$$\begin{aligned} \left(\hat{H}_0 + \epsilon \hat{V} \right) \left(\psi_0(x) + \epsilon \psi_1(x) + \epsilon^2 \psi_2(x) + \dots \right) = \\ \left(E_0 + \epsilon E_1 + \epsilon^2 E_2 + \dots \right) \left(\psi_0(x) + \epsilon \psi_1(x) + \epsilon^2 \psi_2(x) + \dots \right). \end{aligned} \quad (17.6)$$

We can now solve Eq. (17.6) order by order in ϵ .

Order ϵ^0 At order ϵ^0 we find:

$$(H_0 - E_0)\psi_0 = 0, \quad (17.7)$$

and therefore ψ_0 has to be one of the unperturbed eigenfunctions $\psi^{(n)}$, and E_0 must be the corresponding unperturbed eigenvalue $E^{(n)}$. Since we assumed that the unperturbed levels are nondegenerate the choice of $\psi^{(n)}$ is unique.

Order ϵ^1 At order ϵ we have:

$$(\hat{H}_0 - E_0)\psi_1 + (\hat{V} - E_1)\psi_0 = 0. \quad (17.8)$$

Taking the scalar product of Eq. (17.8) with ψ_0 yields:

$$\langle \psi_0 | \hat{H}_0 | \psi_1 \rangle + \langle \psi_0 | \hat{V} | \psi_0 \rangle = E_0 \langle \psi_0 | \psi_1 \rangle + E_1 \langle \psi_0 | \psi_0 \rangle. \quad (17.9)$$

Since \hat{H}_0 is Hermitean, the first term on the LHS of Eq. (17.9) cancels with the first term on the RHS, and we are left with:

$$E_1 = \frac{\langle \psi_0 | \hat{V} | \psi_0 \rangle}{\langle \psi_0 | \psi_0 \rangle}. \quad (17.10)$$

To first order in the perturbation parameter ϵ , the shift of the n -th energy eigenvalue is given by Eq. (17.10). Hence the eigenvalue of the Hamiltonian is:

$$E = E^{(n)} + \epsilon \frac{\langle \psi^{(n)} | \hat{V} | \psi^{(n)} \rangle}{\langle \psi^{(n)} | \psi^{(n)} \rangle} + \mathcal{O}(\epsilon^2). \quad (17.11)$$

This is a very useful result, since it allows us to compute the perturbed energy levels starting from the unperturbed ones.

Let us now consider the scalar product of Eq. (17.8) with $\psi^{(m)}$, for $\psi^{(m)} \neq \psi_0$:

$$\langle \psi^{(m)} | \hat{H}_0 | \psi_1 \rangle + \langle \psi^{(m)} | \hat{V} | \psi_0 \rangle = E_0 \langle \psi^{(m)} | \psi_1 \rangle + E_1 \langle \psi^{(m)} | \psi_0 \rangle. \quad (17.12)$$

Using the fact that:

$$\langle \psi^{(m)} | \psi_0 \rangle = 0, \quad (17.13)$$

we obtain:

$$\langle \psi^{(m)} | \psi_1 \rangle = \frac{\langle \psi^{(m)} | \hat{V} | \psi_0 \rangle}{E_0 - E^{(m)}}. \quad (17.14)$$

Hence at first order in ϵ we have:

$$\psi(x) = \psi^{(n)}(x) + \epsilon \sum_{m \neq n} \frac{\langle \psi^{(m)} | \hat{V} | \psi^{(n)} \rangle}{E^{(n)} - E^{(m)}} \psi^{(m)}(x) + \mathcal{O}(\epsilon^2). \quad (17.15)$$

Using Dirac's notation, we can rewrite the solution above as:

$$|\psi\rangle = |\psi^{(n)}\rangle + \epsilon \sum_{m \neq n} \frac{\langle \psi^{(m)} | \hat{V} | \psi^{(n)} \rangle}{E^{(n)} - E^{(m)}} |\psi^{(m)}\rangle + \mathcal{O}(\epsilon^2). \quad (17.16)$$

Note that to first order in ϵ the solution in Eq. (17.16) is already normalized:

$$\langle \psi | \psi \rangle = \langle \psi^{(n)} | \psi^{(n)} \rangle + \epsilon \sum_{m \neq n} \left[\frac{\langle \psi^{(m)} | \hat{V} | \psi^{(n)} \rangle}{E^{(n)} - E^{(m)}} \langle \psi^{(n)} | \psi^{(m)} \rangle + \text{c.c.} \right] + \mathcal{O}(\epsilon^2) \quad (17.17)$$

$$= 1 + \mathcal{O}(\epsilon^2). \quad (17.18)$$

Example A particle moves in the 1-dimensional potential

$$V(x) = \infty, \quad |x| > a, \quad V(x) = V_0 \cos(\pi x/2a), \quad |x| \leq a$$

Calculate the ground-state energy to first order in perturbation theory.

Here we take the unperturbed Hamiltonian, \hat{H}_0 , to be that of the infinite square well, for which we already know the eigenvalues and eigenfunctions:

$$E^{(n)} = \frac{\pi^2 \hbar^2 n^2}{8ma^2}, \quad u^{(n)} = \frac{1}{\sqrt{a}} \begin{cases} \cos \\ \sin \end{cases} \frac{n\pi x}{2a}; \quad n \begin{cases} \text{odd} \\ \text{even} \end{cases}$$

The perturbation \hat{H}' is $V_0 \cos(\pi x/2a)$, which is small provided $V_0 \ll E^{(2)} - E^{(1)}$.

To first order, then,

$$\Delta E \equiv E_1 = H'_{11} = \int_{-\infty}^{\infty} u^{(1)} \hat{H}' u^{(1)} dx = \frac{V_0}{a} \int_{-a}^a \cos^3 \frac{\pi x}{2a} dx$$

Evaluating the integral is straightforward and yields the result

$$\Delta E = \frac{8V_0}{3\pi} = 0.85 V_0$$

Iterative solution At order ϵ^L the eigenvalue equation yields:

$$(\hat{H}_0 - E_0)\psi_L + (\hat{V} - E_1)\psi_{L-1} - \sum_{K=2}^L E_K \psi_{L-K} = 0. \quad (17.19)$$

Taking the same scalar products described above, we find:

$$E_L = \langle \psi_0 | \hat{V} | \psi_{L-1} \rangle, \quad (17.20)$$

which yields the correction of order ϵ^L to the unperturbed energy level.

Following the computation above we also obtain:

$$\langle \psi^{(m)} | \psi_L \rangle = \frac{\langle \psi^{(m)} | V | \psi_{L-1} \rangle}{E^{(n)} - E^{(m)}} - \frac{1}{E^{(n)} - E^{(m)}} \sum_{K=1}^{L-1} E_K \langle \psi^{(m)} | \psi_{L-K} \rangle. \quad (17.21)$$

Using Eq. (17.21) for $L = 2$ we find the second-order correction to the n -th energy level:

$$E_2 = \langle \psi_0 | \hat{V} | \psi_1 \rangle = \sum_{m \neq n} \frac{\langle \psi^{(n)} | \hat{V} | \psi^{(m)} \rangle \langle \psi^{(m)} | \hat{V} | \psi^{(n)} \rangle}{E^{(n)} - E^{(m)}}. \quad (17.22)$$

17.3 Degenerate levels

Equation (17.15) shows that the correction to the energy eigenfunctions at first order in perturbation theory is small only if

$$\frac{\langle \psi^{(m)} | \hat{V} | \psi^{(n)} \rangle}{E^{(n)} - E^{(m)}} \simeq 1. \quad (17.23)$$

If the energy splitting between the unperturbed levels is small compared to the matrix element in the numerator, then the perturbation becomes large, and the approximation breaks down. In particular, if there are degenerate levels, the denominator is singular, and the solution is not applicable.

Let us see how we can deal with a g_0 -fold degenerate level of the unperturbed Hamiltonian. We shall denote P the projector onto such level, and Q the projector orthogonal to this level.

The first-order equation:

$$(\hat{H}_0 - E_0)\psi_1 + (\hat{V} - E_1)\psi_0 = 0 \quad (17.24)$$

can be projected using P onto the space spanned by the degenerate states:

$$P(\hat{V} - E_1)\psi_0 = 0. \quad (17.25)$$

Choosing a basis for the space of degenerate levels, we can write ψ_0 as:

$$\psi_0 = \sum_{i=1}^{g_0} c_i \phi_i, \quad (17.26)$$

and then rewrite Eq. (17.25):

$$\langle \phi_i | \hat{V} | \phi_j \rangle c_j = E_1 c_i, \quad (17.27)$$

i.e. E_1 is an eigenvalue of the matrix $V_{ij} = \langle \phi_i | \hat{V} | \phi_j \rangle$. This equation has g_0 roots (not necessarily distinct), and generalizes Eq. (17.10) to the case of degenerate levels. If the eigenvalues are indeed all distinct, then the degeneracy is completely lifted. If some of the eigenvalues are equal, the degeneracy is only partially lifted.

Example A well-known example of degenerate perturbation theory is the Stark effect, i.e. the separation of levels in the H atom due to the presence of an electric field. Let us consider the $n = 2$ level, which has a 4-fold degeneracy:

$$|2s\rangle, |2p, 0\rangle, |2p, +1\rangle, |2p, -1\rangle. \quad (17.28)$$

The electric field is chosen in the z -direction, hence the perturbation can be written as:

$$V = -ez\mathcal{E}, \quad (17.29)$$

where \mathcal{E} is the magnitude of the electric field.

We need to compute the matrix V_{ij} in the subspace of the unperturbed states of the H atom with $n = 2$. This is a 4×4 Hermitean matrix.

Note that the perturbation V is odd under parity, and therefore it has non-vanishing matrix elements only between states of opposite parity. Since the eigenstates of the H atom are eigenstates of L^2 and L_z , we find that only the matrix elements between s and p states can be different from zero.

Moreover, V commutes with L_z and therefore only matrix elements between states with the same value of L_z are different from zero.

So we have proved that the only non-vanishing matrix elements are $\langle 2s|\hat{V}|2p, 0\rangle$ and its Hermitean conjugate. Hence the matrix V is given by:

$$\begin{pmatrix} 0 & 3e\mathcal{E}a_0 & 0 & 0 \\ 3e\mathcal{E}a_0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad (17.30)$$

where a_0 is the Bohr radius. We see that the external field only removes the degeneracy between the $|2s\rangle$, and the $|2p, 0\rangle$ states; the states $|2p, \pm 1\rangle$ are left unchanged.

The two other levels are split:

$$E = E_2 \pm 3ea_0\mathcal{E}. \quad (17.31)$$

17.4 Applications

There are numerous applications of perturbation theory, which has proven to be a very effective tool to gain quantitative information on the dynamics of a system whenever a small expansion parameter can be identified.

Here we discuss briefly two examples.

17.4.1 Ground state of Helium

We can now attempt to incorporate the effect of the inter-electron Coulomb repulsion by treating it as a perturbation. We write the Hamiltonian as

$$\hat{H} = \hat{H}_0 + \hat{H}'$$

where

$$\hat{H}_0 = \hat{H}_1 + \hat{H}_2 \quad \text{and} \quad \hat{H}' = \frac{e^2}{4\pi\epsilon_0|r_1 - r_2|}$$

The ground state wavefunction that we wrote down earlier is an eigenfunction of the unperturbed Hamiltonian, \hat{H}_0 ;

$$\Psi(\text{ground state}) = u_{100}(r_1) u_{100}(r_2) \chi_{0,0}.$$

To compute the first order correction to the ground state energy, we have to evaluate the expectation value of the perturbation, \hat{H}' , with respect to this wavefunction;

$$\Delta E_1 = \frac{e^2}{4\pi\epsilon_0} \int u_{100}^*(r_1) u_{100}^*(r_2) \chi_{0,0}^* \frac{1}{r_{12}} u_{100}(r_1) u_{100}(r_2) \chi_{0,0} d\tau_1 d\tau_2$$

The scalar product of $\chi_{0,0}$ with its conjugate = 1, since it is normalised. Putting in the explicit form of the hydrogenic wavefunction from Lecture 10

$$u_{100}(r) = \frac{1}{\sqrt{\pi}} (Z/a_0)^{3/2} \exp(-Zr/a_0)$$

thus yields the expression

$$\Delta E_1 = \frac{e^2}{4\pi\epsilon_0} \left(\frac{Z^3}{\pi a_0^3} \right)^2 \int \frac{1}{r_{12}} \exp\{-2Z(r_1 + r_2)/a_0\} d\tau_1 d\tau_2$$

Amazingly, this integral can be evaluated analytically. See, for example, Bransden and Joachain, *Introduction to Quantum Mechanics*, pp 465-466. The result is

$$\Delta E_1 = \frac{5}{4} Z Ry = \frac{5}{2} Ry = 34 eV$$

giving for the first-order estimate of the ground state energy

$$E_1 = -108.8 + 34 eV = -74.8 eV = -5.5 Ry$$

to be compared with the experimentally-measured value of $-78.957 eV$.

17.4.2 Spin-orbit effects in hydrogenic atoms

Classically, an electron of mass M and charge $-e$ moving in an orbit with angular momentum \underline{L} would have a magnetic moment

$$\underline{\mu} = -\frac{e}{2M} \underline{L}$$

suggesting that in the quantum case,

$$\hat{\underline{\mu}} = -\frac{e}{2M} \hat{\underline{L}} \quad \text{and} \quad \hat{\mu}_z = -\frac{e}{2M} \hat{L}_z$$

The eigenvalues of $\hat{\mu}_z$ are thus given by

$$-\frac{e\hbar}{2M} m_\ell \equiv -\mu_B m_\ell,$$

where the quantity μ_B is known as the *Bohr magneton*.

Similarly, there is a magnetic moment associated with the intrinsic spin of the electron;

$$\hat{\mu}_z = -\frac{g_s e}{2M} \hat{S}_z$$

where the constant, g_s , cannot be determined from classical arguments, but is predicted to be 2 by relativistic quantum theory and is found experimentally to be very close to 2.

The interaction between the orbital and spin magnetic moments of the electron introduces an extra term into the Hamiltonian of the form

$$\hat{H}_{S-O} = f(r) \hat{\underline{L}} \cdot \hat{\underline{S}}$$

where

$$f(r) = \frac{1}{2M^2 c^2 r} \frac{dV(r)}{dr}$$

We can attempt to treat this extra term by the methods of perturbation theory, by taking the unperturbed Hamiltonian to be

$$\hat{H}_0 = \frac{\hat{p}^2}{2M} + V(r) = \frac{\hat{p}^2}{2M} - \frac{Ze^2}{4\pi\epsilon_0 r}$$

Cautionary Note In our derivation of the first-order formula for the shift in energy induced by a perturbation, we assumed that there were no degeneracies in the energy eigenvalue spectrum and noted that the method could break down in the presence of degeneracies.

- In general, when considering the effects of a perturbation on a degenerate level, *it is necessary to use degenerate state perturbation theory*, which we briefly discussed above.

- There are, however, important exceptions to this rule. In particular, *if the perturbation \hat{H}' , is diagonal with respect to the degenerate states, the non-degenerate theory can be used to compute the energy shifts.*

In the case of the spin-orbit interaction in the hydrogenic atom, we know that the degeneracy of a level with given n and ℓ is $(2\ell + 1) \times 2$, since, for a given ℓ , there are $(2\ell + 1)$ possible values of m_ℓ and 2 possible values of m_s .

However, if we choose to work with states of the coupled basis $|n, j, m_j, \ell, s\rangle$, rather than with the states of the uncoupled basis $|n, \ell, m_\ell, s, m_s\rangle$, we can use non-degenerate theory. Firstly, we note that we can rewrite the spin-orbit term as follows:

$$\hat{H}_{S-O} = f(r) \hat{L} \cdot \hat{S} = \frac{1}{2} f(r) \{ \hat{J}^2 - \hat{L}^2 - \hat{S}^2 \}$$

using the fact that $\hat{J}^2 \equiv (\hat{L} + \hat{S})^2 = \hat{L}^2 + \hat{S}^2 + 2\hat{L} \cdot \hat{S}$.

Noting that

$$\{ \hat{J}^2 - \hat{L}^2 - \hat{S}^2 \} |n, j, m_j, \ell, s\rangle = \{ j(j+1) - \ell(\ell+1) - s(s+1) \} \hbar^2 |n, j, m_j, \ell, s\rangle$$

we see that the expectation value of \hat{H}' in the unperturbed basis is

$$\langle n, j, m_j, \ell, s | \hat{H}_{S-O} | n, j, m_j, \ell, s \rangle = \frac{1}{2} \{ j(j+1) - \ell(\ell+1) - s(s+1) \} \hbar^2 \langle f(r) \rangle$$

Since $f(r)$ is independent of the angular variables θ, ϕ and of the spin, the expectation value of $f(r)$ may be written

$$\langle f(r) \rangle = \frac{Ze^2}{8\pi\epsilon_0 M^2 c^2} \int_0^\infty \frac{1}{r^3} |R_{n\ell}(r)|^2 r^2 dr$$

The integral can be evaluated exactly using the hydrogenic radial functions and gives:

$$\left\langle \frac{1}{r^3} \right\rangle_{n\ell} = \frac{Z^3}{a_0^3} \frac{1}{n^3 \ell (\ell + \frac{1}{2}) (\ell + 1)}$$

Now $s = \frac{1}{2}$ for an electron, so that j can have two values for a given ℓ , namely, $j = (\ell + \frac{1}{2})$ and $j = (\ell - \frac{1}{2})$, except in the case $\ell = 0$, which means that a state of given n and ℓ separates into a doublet when the spin-orbit interaction is present.

Term Notation There is yet another piece of notation used widely in the literature, the so-called *term notation*. The states that arise in coupling orbital angular momentum ℓ and spin s to give total angular momentum j are denoted:

$$\boxed{{}^{(2S+1)}L_J}$$

where L denotes the letter corresponding to the ℓ value in the usual way, and the factor $(2S + 1)$ is the spin multiplicity i.e. the number of allowed values of m_s .

17.5 Summary

As usual, we summarize the main concepts introduced in this lecture.

- Perturbations of a system.
- Solution by perturbative expansion.
- Shifted energy levels and wave functions.
- Examples