A Quantum Mechanics Primer

This document recaps part of the level 2 Quantum Mechanics course, focusing on the material relevant for Quantum Mechanics 3. The various concepts and mathematical facts it covers are essential for a good understanding of Quantum Mechanics 3 and are assumed to be known from level 2. This document is not meant to be exhaustive. An in depth study of basic Quantum Mechanics would certainly go much beyond what is summarized here.

No knowledge of the material covered only in the Theoretical Physics 2 module is assumed. However, it is likely that students who took Theoretical Physics 2 will already have a greater familiarity with some of these concepts.

The "notes" typeset in small black fonts concern unessential details. What is typeset in blue with a thick bar on the left is outside the syllabus of the course and is provided for interest only.

1 Wave functions and state vectors

1.1 Wave functions

The state of a quantum system is often described by way of a function, the "wave function". In this course, we'll often use wave functions depending on the position(s) of the particle(s) forming the system. As you may remember, other representations of quantum states are also used, and we'll see examples of these from time to time.

In the so-called "position representation", the wave function for a system containing only one particle is a certain function of the coordinates of that particle: e.g., $\Psi(x, y, z, t)$ in terms of its Cartesian coordinates (t is time), or $\Psi(r, \theta, \phi, t)$ in terms of its spherical polar coordinates, or $\Psi(\mathbf{r}, t)$ in terms of its position vector $\mathbf{r}$, etc. In some applications, one makes the assumption that the system is one-dimensional, e.g., that the particle can move in the x-direction only and that the wave function does not extend in the y- and z-directions. In this case, the wave function depends only on the x-coordinate of the particle, besides time. I'll sometimes write results in terms of wave functions of a single particle in 1D, depending on x only, so as to keep the notation simple. However, it should be understood that in general the wave function depends on several spatial coordinates.

Note 1: If there are several particles in the system, as in the case of a multi-electron atom, then the wave function is a function of the coordinates of all these particles: e.g., for a two-particle system, $\Psi(x_1, y_1, z_1, x_2, y_2, z_2, t)$, where (x_1, y_1, z_1) are the coordinates of particle 1 and (x_2, y_2, z_2) are the coordinates of particle 2. (We will see examples of such multi-particle wave functions later in the course.)

Note 2: Another way of representing a quantum system, which comes handy in some applications is by way of a wave function depending on the momenta of all the particles in the system rather than on their positions. Working with such wave functions is working in the "momentum representation", by contrast to the "position representation". We will not need to work in that representation in this course.

It is essential to be aware of the following:

(i) In Classical Mechanics, in order to define the state of a system of inter-acting particles one needs to give both the position *and* the momentum or velocity of each of the particles forming the system. By contrast, the wave function of a quantum system is a function of the positions *only* (in the position representation, that is). This is not surprising, since the position and the momentum cannot be both defined with arbitrary precision. (Remember the uncertainty principle!)

(ii) Any wave function representing a possible state of a quantum system must be a solution of the time-dependent Schrödinger equation,

$$i\hbar \frac{\partial \Psi}{\partial t} = H\Psi, \tag{1}$$

where H is the Hamiltonian of the system.

(iii) The principle of superposition: see Section 1.3 below for this key concept.

(iv) The probabilistic interpretation of the wave function: For a system of one particle in 1D described by the wave function $\Psi(x, t)$, the number

$$|\Psi(x, t)|^2 \mathrm{d}x$$

is the probability that at time t the particle is in an interval of length $\mathrm{d}x$ centered at the point of position x (or, equivalently since $\mathrm{d}x$ is infinites-imally small, the probability that at time t the particle is at a position between x and $x + \mathrm{d}x$). Hence, the integral

$$\int_a^b |\Psi(x, t)|^2 \, \mathrm{d}x$$

is the probability that the particle is located within the interval $[a, b]$ (assuming that $a < b$). In 3D,

$$|\Psi(x, y, z, t)|^2 \, \mathrm{d}^3 r$$

is the probability that at time t the particle is located in an infinitesimal volume $\mathrm{d}^3 r \equiv \mathrm{d}x \, \mathrm{d}y \, \mathrm{d}z$ centered about the point of coordinates (x, y, z).

For this interpretation to make sense, it is necessary that the integral of the squared modulus of the wave function over all possible values of the coordinate(s) is the number 1 (indeed, this integral is the probability that the particle is somewhere, whatever where, and this probability is necessarily 1):

$$\int_{-\infty}^{\infty} |\Psi(x, t)|^2 \, \mathrm{d}x = 1, \tag{2}$$

or in the more general 3D case,

$$\int |\Psi(x, y, z, t)|^2 \, \mathrm{d}^3 r \equiv \int_{-\infty}^{\infty} \mathrm{d}x \int_{-\infty}^{\infty} \mathrm{d}y \int_{-\infty}^{\infty} \mathrm{d}z \, |\Psi(x, y, z, t)|^2 = 1. \tag{3}$$

Note 3:   All this carries over when the wave function describes the state of a system composed of several particles, as we will see in Part 4 of the course notes. For example, for two particles in 3D,

$$|\Psi(x_1, y_1, z_1, x_2, y_2, z_2, t)|^2 \, \mathrm{d}^3 r_1 \, \mathrm{d}^3 r_2$$

is the probability that at time t particle 1 is located in an infinitesimal volume $\mathrm{d}^3 r_1 \equiv \mathrm{d}x_1 \, \mathrm{d}y_1 \, \mathrm{d}z_1$ centered about the point of coordinates (x_1, y_1, z_1) *and* that particle 2 is located in an infinitesimal volume $\mathrm{d}^3 r_2 \equiv \mathrm{d}x_2 \, \mathrm{d}y_2 \, \mathrm{d}z_2$ centered about the point of coordinates (x_2, y_2, z_2), and the normalization condition reads

$$\int_{-\infty}^{\infty} \mathrm{d}x_1 \int_{-\infty}^{\infty} \mathrm{d}y_1 \int_{-\infty}^{\infty} \mathrm{d}z_1 \int_{-\infty}^{\infty} \mathrm{d}x_2 \int_{-\infty}^{\infty} \mathrm{d}y_2 \int_{-\infty}^{\infty} \mathrm{d}z_2$$
$$|\Psi(x_1, y_1, z_1, x_2, y_2, z_2, t)|^2 = 1. \tag{4}$$

Note 4:   Instead of insisting that wave functions are normalized according to this prescription, one can instead work with un-normalized wave functions. However, the rules to calculate probabilities are slightly more complicated in this case. To avoid these complications, we'll only use normalized wave functions throughout this course.

1.2   State vectors

Not all physical properties of quantum systems can be described by way of a wave function. For example, the spin state of an electron is described by a

column vector of two numbers (or an equivalent mathematical object), not by a certain function of its x-, y- and z-coordinates. There is much more about spin states later in these notes, but for the time being, let us just mention that the state where the electron has spin up can be represented by the column vector

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix},$$

while the state where the electron has spin down can be represented by the column vector

$$\begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Wave functions obeying Eqs. (2) or (4) and 2-component column vectors may seem to be completely different mathematical things, and of course in many respects they are. However, they have something in common: in both cases, they are "vectors" in a certain "Hilbert space". This mathematical terminology means is not essential for the course, and it suffices to know the following: At a more fundamental level than mentioned so far in these notes, the state of a quantum system can be described by a certain abstract mathematical object called, in this context, a "state vector". People often refer to the wave function, or column vector, or state vector describing the system as "the state" (e.g., as in "when the Hamiltonian acts on the state n, etc.", meaning "when the Hamiltonian acts on the quantum state represented by the wave function ψ_n etc", or similar).

1.3   The principle of superposition

The principle of superposition is at the very heart of Quantum Mechanics and has far reaching consequences.

From a mathematical point of view, it is a simple statement of the fact that any linear combination of two solutions of the time-dependent Schrödinger equation is also a solution of this equation. [By a linear combination of two wave functions $\Psi(x, t)$ and $\Phi(x, t)$, one means a sum of the form $\alpha\Psi(x, t) + \beta\Phi(x, t)$, where α and β are two real or complex constants.]

An immediate consequence of this mathematical fact is that a quantum system can be in several different states at the same time, so to speak.

For example, imagine that an atom of hydrogen, initially in the ground state (the 1s state), undergoes a collision. This collision might leave it in a state represented by a wave function formed by a linear combination of several bound states. It could be left, for instance, in a superposition of the 1s state, of the 2s state and of the 3s state. In a sense, the atom would then be simultaneously in the 1s state, in the 2s state and in the 3s state. Or to put it otherwise, if in an experiment one would check whether the atom is, say, in the 2s state after

the collision, there would then be a non-zero probability that it is found to be in that state and also a non-zero probability that it is found not to be in that state. (See Section 2.4 below about calculating such probabilities.)

Likewise, linear combinations of the two column vectors

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

represent spin states in which an electron is in a superposition of a state of spin up and a state of spin down — i.e., so to speak, states in which the electron is simultaneously in a state of spin up and a state of spin down.

2 Inner products and orthogonality

2.1 Scalar products of vectors and inner products of column vectors

Take two vectors in 3D space, e.g., a vector $\mathbf{v}_1$, and a vector $\mathbf{v}_2$, both real (their x-, y- and z-components have no imaginary parts). The scalar product of these two vectors, $\mathbf{v}_1 \cdot \mathbf{v}_2$, is the real number $v_1 v_2 \cos\theta_{12}$, where v_1 is the length of the vector $\mathbf{v}_1$, v_2 is the length of the vector $\mathbf{v}_2$ and θ_{12} is the angle between $\mathbf{v}_1$ and $\mathbf{v}_2$. In terms of the x-, y- and z-components of these two vectors, (v_{1x}, v_{1y}, v_{1z}) and (v_{2x}, v_{2y}, v_{2z}),

$$\mathbf{v}_1 \cdot \mathbf{v}_2 = v_{1x}v_{2x} + v_{1y}v_{2y} + v_{1z}v_{2z}. \tag{5}$$

This can also be written

$$\mathbf{v}_1 \cdot \mathbf{v}_2 = \begin{pmatrix} v_{1x} & v_{1y} & v_{1z} \end{pmatrix} \begin{pmatrix} v_{2x} \\ v_{2y} \\ v_{2z} \end{pmatrix} \tag{6}$$

(recall how to multiply the column of a matrix by the row of another one). Thus if we define the column vectors v_1 and v_2 as follows,

$$\mathsf{v}_1 = \begin{pmatrix} v_{1x} \\ v_{1y} \\ v_{1z} \end{pmatrix}, \qquad \mathsf{v}_2 = \begin{pmatrix} v_{2x} \\ v_{2y} \\ v_{2z} \end{pmatrix}, \tag{7}$$

then one has

$$\mathbf{v}_1 \cdot \mathbf{v}_2 = \mathsf{v}_1^\mathsf{T} \mathsf{v}_2, \tag{8}$$

where

$$\mathsf{v}_1^\mathsf{T} = \begin{pmatrix} v_{1x} & v_{1y} & v_{1z} \end{pmatrix}. \tag{9}$$

The row vector v_1^T is the transpose of the column vector v_1 (recall that transposing a matrix is swapping its columns with its rows).

In the example above, the column vectors had three components each and these components were supposed to be purely real (i.e., to have no imaginary part). However, in Quantum Mechanics one often uses column vectors with only two components or with more than three components, and the components are often complex. Generalizing the above, to any pair of N-component column vectors

$$ \mathsf{a} = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_N \end{pmatrix} \quad \text{and} \quad \mathsf{b} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_N \end{pmatrix}, \tag{10} $$

where $a_1, a_2, \ldots, a_N, b_1, b_2, \ldots, b_N$ are complex numbers, one can associate the complex number $\mathsf{a}^{T*}\mathsf{b}$, called the inner product of a and b:

$$ \mathsf{a}^{T*}\mathsf{b} \equiv \begin{pmatrix} a_1^* & a_2^* & \cdots & a_N^* \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_N \end{pmatrix} = a_1^* b_1 + a_2^* b_2 + \cdots + a_N^* b_N. \tag{11} $$

The inner product $\mathsf{a}^{T*}\mathsf{b}$ is denoted by the symbol $\langle a|b \rangle$ in the compact notation introduced by Dirac. Note the complex conjugations, which did not appear in the equations defining the scalar product of the vectors $\mathbf{v}_1$ and $\mathbf{v}_2$ we saw previously since we had assumed that the components of these two vectors were real, not complex: a^{T*} is the row vector obtained by transposing a and complex-conjugating its elements. Clearly $\langle b|a \rangle$ would be $\mathsf{b}^{T*}\mathsf{a}$, and

$$ \langle b|a \rangle = \langle a|b \rangle^*. \tag{12} $$

2.2 Inner products of wave functions

Similarly, one can define the inner product $\langle \psi|\phi \rangle$ of two wave functions $\psi(x)$ and $\phi(x)$ as the integral of the product of the complex conjugate of $\psi(x)$ with $\phi(x)$: for a single particle, in one dimension,

$$ \langle \psi|\phi \rangle = \int_{-\infty}^{\infty} \psi^*(x)\phi(x)\mathrm{d}x, \tag{13} $$

and similarly for multiple particles in three dimensions. As above, $\langle \psi|\phi \rangle$ is (in general) a complex number, and

$$ \langle \phi|\psi \rangle = \langle \psi|\phi \rangle^*. \tag{14} $$

In particular, $\langle\psi|\psi\rangle$ is the "norm" of the wave function $\psi(x)$. In this notation, Eqs. (2) and (4) both simply read

$$\langle\Psi|\Psi\rangle = 1 \tag{15}$$

(or $\langle\Psi(t)|\Psi(t)\rangle = 1$, if one wants to stress that the wave function in question is time-dependent).

2.3   Orthogonality

When a vector $\mathbf{v}_1$ and a vector $\mathbf{v}_2$ are perpendicular (orthogonal) to each other, then their scalar product is zero: $\mathbf{v}_1 \cdot \mathbf{v}_2 = 0$. Conversely, if the scalar product of two vectors is zero, then these two vectors are perpendicular to each other.

By extension, one says that two column vectors or two wave functions are orthogonal to each other if their inner product is zero, and that two quantum states are orthogonal to each other if they are represented by column vectors or wave functions orthogonal to each other.

2.4   Inner products and probabilities

As noted above, a system can be in a quantum state which is itself a superposition of several other quantum states. The following question therefore arises: Suppose that the system is represented by a wave function $\Psi(x,t)$ (or similarly for N particles in 3D) and that one tests in an experiment whether or not this system is in a certain state represented by a wave function $\Phi(x,t)$. What is the probability that the system is found to be in the latter state? According to the rules of Quantum Mechanics, this probability, P, is the modulus squared of the inner product of these two wave functions:

$$P = |\langle\Phi|\Psi\rangle|^2, \tag{16}$$

or explicitly,

$$P = \left|\int_{-\infty}^{\infty} \Phi(x,t)^*\Psi(x,t)\mathrm{d}x\right|^2. \tag{17}$$

Clearly, this probability is zero if the two states are orthogonal to each other. Thus orthogonal wave functions or state vectors represent distinct physical states (if the system is in of them, it cannot be found to be in the other one). Conversely, distinct physical states correspond to orthogonal state vectors or wave functions.

An important particular case of the above is that of a system in a superposition of several orthogonal states — e.g., assuming that the wave functions are normalized,

$$\Psi(x,t) = \sum_n c_n\Phi_n(x,t), \tag{18}$$

with
$$\langle \Phi_m | \Phi_n \rangle = \delta_{mn}. \tag{19}$$

Recall the meaning of the Kronecker delta symbol:

$$\delta_{mn} = \begin{cases} 1 & m = n, \\ 0 & m \neq n. \end{cases} \tag{20}$$

Thus Eq. (19) means that the wave functions $\Phi_n(x,t)$ are "orthonormal", that is they are both orthogonal to each other and normalized to unity (the latter meaning that $\langle \Phi_n | \Phi_n \rangle = 1$ for any n). Note that the requirement that the wave function $\Psi(x,t)$ is itself normalized implies, given Eq. (19), that the coefficients c_n must be such that

$$\sum_n |c_n|^2 = 1. \tag{21}$$

The probability that the system represented by $\Psi(x,t)$ is in a specific state forming the superposition, that represented by $\Phi_m(x,t)$ say, is $|\langle \Phi_m | \Psi \rangle|^2$. This probability is simply $|c_m|^2$, since

$$\langle \Phi_m | \Psi \rangle = \sum_n c_n \langle \Phi_m | \Phi_n \rangle = \sum_n c_n \delta_{mn} = c_m. \tag{22}$$

The result would in general be different, and neither Eq. (21) nor Eq. (22) would apply, if the wave functions $\Phi_n(x,t)$ were not orthogonal to each other.

3 Operators

3.1 What is an operator, in Mathematics?

Loosely speaking, an operator is a mathematical object which transforms a mathematical thing into a mathematical thing of the same type — for instance, transforms a function into a function, or a column vector into a column vector.

For example, the differential operator d/dx, when applied on any function $f(x)$ of the variable x, transforms $f(x)$ into the derivative of this function: $d/dx \sin(x) = d\sin(x)/dx = \cos x$; $d/dx \exp(2x) = 2\exp(2x)$, etc. Another example is a square matrix: when multiplying a column vector, a square matrix transforms this column vector into another column vector of the same size. For instance,

$$\begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} 3 \\ 5 \end{pmatrix} = \begin{pmatrix} 1 \times 3 + 2 \times 5 \\ 2 \times 3 + 1 \times 5 \end{pmatrix} = \begin{pmatrix} 13 \\ 11 \end{pmatrix}. \tag{23}$$

I denote square matrices by an upper case "sans serif" letter in these notes (for example, M), and column vectors by a lower case letter of the same style (for

example, c). The combination Mc represents the product of the column vector c by the matrix M.

Quantum Mechanics is based on a certain class of operators, called linear operators.

3.2   What is a linear operator?

As is well know, if $f(x)$ is a function of x and c is a constant (a number), then the derivative of the product $cf(x)$ is c times the derivative of $f(x)$:

$$\frac{\mathrm{d}}{\mathrm{d}x}[cf(x)] = c\frac{\mathrm{d}f}{\mathrm{d}x}. \tag{24}$$

Also, if $f(x)$ and $g(x)$ are two functions of x, then the derivative of the sum $f(x) + g(x)$ is the sum of the derivatives:

$$\frac{\mathrm{d}}{\mathrm{d}x}[f(x) + g(x)] = \frac{\mathrm{d}f}{\mathrm{d}x} + \frac{\mathrm{d}g}{\mathrm{d}x}. \tag{25}$$

Hence, for *any* (differentiable) functions $f(x)$ and $g(x)$ and for *any* constants c_1 and c_2,

$$\frac{\mathrm{d}}{\mathrm{d}x}[c_1 f(x) + c_2 g(x)] = c_1 \frac{\mathrm{d}}{\mathrm{d}x}[f(x)] + c_2 \frac{\mathrm{d}}{\mathrm{d}x}[g(x)]. \tag{26}$$

One says, in acknowledgment of this fact, that $\mathrm{d}/\mathrm{d}x$ is a "linear operator". More generally, an operator O acting on functions is said to be linear if

$$O[c_1 f(x) + c_2 g(x)] = c_1 O f(x) + c_2 O g(x) \tag{27}$$

for any pair of functions $f(x)$ and $g(x)$ and for any pair of complex constants c_1 and c_2.

The operator represented by a square matrix is also a linear operator. For instance, if f and g are two 2-component column vectors, c_1 and c_2 are two numbers and M is the matrix of Eq. (23), then one sees easily that

$$M[c_1 f + c_2 g] = c_1 M f + c_2 M g, \tag{28}$$

whatever the components of f and g are and whatever the numbers c_1 and c_2 are.

Note 1:   Not quite all operators are linear. For instance, the operator which transforms a function into the square of that function is not linear: Denoting this operator by O, we would have

$$\begin{aligned} O[c_1 f(x) + c_2 g(x)] &= [c_1 f(x) + c_2 g(x)]^2 \\ &= c_1^2 f^2(x) + 2c_1 c_2 f(x)g(x) + c_2^2 g^2(x), \quad (29) \end{aligned}$$

which is not the same as $c_1 O f(x) + c_2 O g(x)$ [which is $c_1 f^2(x) + c_2 g^2(x)$].

3.3 Please give examples of important operators in Quantum Mechanics...

- For a single particle moving along the x-axis only, the momentum operator, p_x, is $-i\hbar\,\mathrm{d}/\mathrm{d}x$ in the position representation.

- For a single particle moving in 3D, the momentum operator is $-i\hbar\nabla$. Remember that

$$\nabla = \hat{\mathbf{x}}\frac{\partial}{\partial x} + \hat{\mathbf{y}}\frac{\partial}{\partial y} + \hat{\mathbf{z}}\frac{\partial}{\partial z}, \tag{30}$$

where $\hat{\mathbf{x}}$, $\hat{\mathbf{y}}$ and $\hat{\mathbf{z}}$ are unit vectors in the x-, y- and z-directions. Thus, in 3D, the momentum operator is a vector operator; it has three components, and each of its component is itself an operator.

- For a single particle moving along the x-axis only, the position operator is x - i.e., the effect of this operator on a wave function $\psi(x)$ is to multiply it by x. In 3D, using Cartesian coordinates x, y and z, the position operator is simply $\mathbf{r}$, with $\mathbf{r} = x\hat{\mathbf{x}} + y\hat{\mathbf{y}} + z\hat{\mathbf{z}}$.

- You certainly remember to have come across the Hamiltonian operator, H: for a single particle in 1D,

$$H = -\frac{\hbar^2}{2m}\frac{\mathrm{d}^2}{\mathrm{d}x^2} + V(x), \tag{31}$$

while in 3D,

$$H = -\frac{\hbar^2}{2m}\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}\right) + V(\mathbf{r}). \tag{32}$$

In these last two equations, m is the mass of the particle and V is its potential energy. The first term in the right-hand sides can be understood as representing the kinetic energy of the particle. [Note that the kinetic energy in Classical Mechanics is $p_x^2/2m$, where p_x is the classical momentum in the x-direction, if there is no motion in the y- and z-directions, and that in Quantum Mechanics $(-\hbar^2/2m)\mathrm{d}^2/\mathrm{d}x^2$ is nothing else than $p_x^2/2m$, where p_x is the momentum operator.]

- Angular momentum operators are immensely important in Quantum Mechanics. See Section 4 below...

3.4 What does this mean, that two operators do not commute?

The content of this section is not essential for the course. It is included in these notes because of the importance of this concept in Quantum Mechanics in general, and also because commutators are referred to in Section 4.

Imagine that you have two operators, A and B. If first applying A and then applying B gives a different result than first applying B and then applying A, then one says that A and B do not commute.

The commutator of two operators A and B, which is denoted $[A, B]$, is the difference between the operators AB (where B is applied first and A second) and BA (where A is applied first and B second):

$$[A, B] = AB - BA. \tag{33}$$

Thus $[A, B] = 0$ if the operators A and B commute, whereas $[A, B] \neq 0$ if they don't.

For example, the position operator x and the momentum operator p_x do not commute: If $\psi(x, y, z)$ is a differentiable function of the independent variables x, y and z, then

$$x p_x \psi(x, y, z) = -i\hbar x \frac{\partial \psi}{\partial x} \tag{34}$$

whereas

$$
\begin{aligned}
p_x x \psi(x, y, z) &= -i\hbar \frac{\partial}{\partial x} x \psi(x, y, z) \\
&= -i\hbar \left(\frac{\partial x}{\partial x} \psi(x, y, z) + x \frac{\partial \psi}{\partial x} \right) \\
&= -i\hbar \psi(x, y, z) + x p_x \psi(x, y, z).
\end{aligned} \tag{35}
$$

Comparing Eqs. (34) and (35) one sees that

$$x p_x \psi(x, y, z) - p_x x \psi(x, y, z) = i\hbar \psi(x, y, z). \tag{36}$$

Since this result is true for any differentiable function, one can write

$$[x, p_x] = i\hbar. \tag{37}$$

By contrast, the position operator y commutes with the momentum operator p_x:

$$p_x y \psi(x, y, z) = -i\hbar \frac{\partial}{\partial x} y \psi(x, y, z) = -i\hbar \frac{\partial y}{\partial x} \psi(x, y, z) - i\hbar y \frac{\partial \psi}{\partial x} = 0 + y p_x \psi(x, y, z). \tag{38}$$

Hence $[y, p_x] = 0$ and it does not matter whether y is applied first and p_x second or the other way round.

Whether two operators representing physical quantities commute or not has a profound significance: If they don't commute, then the quantities they represent cannot be measured simultaneously with arbitrary accuracy and instead there is an uncertainty relation between them. If they do commute, then there is no obstacle (in principle) with measuring them simultaneously.

3.5   What is a Hermitian operator?

It is probably easier to start by operators represented by finite square matrices (square matrices with a finite number of rows and columns), for instance the matrix of Eq. (23). By definition, a "Hermitian matrix", M, is one that is equal to the complex conjugate of its transpose:

$$M^{T*} = M. \tag{39}$$

Remember that transposing a matrix is swapping its rows and columns. Thus, if M is Hermitian, then the elements of M, the (real or complex) numbers M_{ij}, are such that $M_{ij} = M_{ji}^*$.

Hermitian matrices have the following property: Let a and b be two column vectors and M a matrix. Multiplying the column vector a by the matrix M gives a certain column vector Ma, and taking the inner product of the latter with b gives a certain (real or complex) number, $\langle b|Ma\rangle$. In the same way, multiplying the column vector b by the matrix M gives a certain column vector Mb, and taking the inner product of the latter with a gives a certain (real or complex) number, $\langle a|Mb\rangle$. As it is relatively easy to prove, if M is a Hermitian matrix, then

$$\langle a|Mb\rangle = \langle b|Ma\rangle^* \tag{40}$$

for any pair of column vectors a and b. Conversely, if Eq. (40) is fulfilled for any pair of column vectors a and b, then M is a Hermitian matrix.

Let us look, for example, at the following matrices:

$$M_1 = \begin{pmatrix} 0 & i \\ -i & 1 \end{pmatrix} \tag{41}$$

and

$$M_2 = \begin{pmatrix} 0 & i \\ i & 1 \end{pmatrix}. \tag{42}$$

M_1 is Hermitian. Indeed, transposing this matrix gives

$$M_1^T = \begin{pmatrix} 0 & -i \\ i & 1 \end{pmatrix} \tag{43}$$

and taking the complex conjugate of M_1^T gives

$$M_1^{T*} = \begin{pmatrix} 0 & i \\ -i & 1 \end{pmatrix}, \tag{44}$$

which shows that $M_1^{T*} = M_1$. By contrast, the matrix M_2 is *not* Hermitian, since

$$M_2^{T*} = \begin{pmatrix} 0 & -i \\ -i & 1 \end{pmatrix} \neq M_2. \tag{45}$$

The matrix of Eq. (23) is another example of a Hermitian matrix.

This definition carries over for more general linear operators, such as linear operators acting on wave functions. Take for example the z-component of the orbital angular momentum operator, L_z, which, as we will see in Section 4, is simply $-i\hbar \mathrm{d}/\mathrm{d}\phi$. Here ϕ is the azimuthal angle in the usual definition of the spherical polar coordinates r, θ, ϕ. As the coordinates (r, θ, ϕ) and $(r, \theta, \phi + 2\pi)$ define the same position, a wave function $\psi(\phi)$ depending on the angle ϕ must be such that $\psi(\phi + 2\pi) = \psi(\phi)$, i.e., must be periodic in ϕ with period 2π. For any pair of such functions $\psi_a(\phi)$ and $\psi_b(\phi)$ that can be differentiated with respect to ϕ,

$$
\begin{aligned}
\int_0^{2\pi} \psi_a^*(\phi)[L_z\psi_b(\phi)]\mathrm{d}\phi &= \int_0^{2\pi} \psi_a^*(\phi)\left(-i\hbar\frac{\mathrm{d}}{\mathrm{d}\phi}\right)\psi_b(\phi)\mathrm{d}\phi \\
&= -i\hbar \int_0^{2\pi} \psi_a^*(\phi)\frac{\mathrm{d}\psi_b}{\mathrm{d}\phi}\,\mathrm{d}\phi \\
&= -i\hbar \left[\psi_a^*(\phi)\psi_b(\phi)\big|_0^{2\pi} - \int_0^{2\pi} \frac{\mathrm{d}\psi_a^*}{\mathrm{d}\phi}\psi_b(\phi)\,\mathrm{d}\phi\right]. \quad (46)
\end{aligned}
$$

The boundary term vanishes since it is assumed that $\psi_a(2\pi) = \psi_a(0)$ and that $\psi_b(2\pi) = \psi_b(0)$. Hence

$$
\begin{aligned}
\int_0^{2\pi} \psi_a^*(\phi)[L_z\psi_b(\phi)]\,\mathrm{d}\phi &= \int_0^{2\pi} \left(-i\hbar\frac{\mathrm{d}\psi_a}{\mathrm{d}\phi}\right)^* \psi_b(\phi)\,\mathrm{d}\phi \\
&= \left[\int_0^{2\pi} \psi_b^*(\phi)\left(-i\hbar\frac{\mathrm{d}}{\mathrm{d}\phi}\right)\psi_a(\phi)\,\mathrm{d}\phi\right]^*, \quad (47)
\end{aligned}
$$

which shows that

$$
\langle\psi_a|L_z\psi_b\rangle = \langle\psi_b|L_z\psi_a\rangle^*. \quad (48)
$$

Operators, such as L_z, that have the property summarized by Eq. (48) for any pair of functions for which the inner products exist (i.e., are finite), are said to be Hermitian.

Not quite all operators are Hermitian. For example, the operator $\hbar\mathrm{d}/\mathrm{d}\phi$, which is L_z without the $-i$ factor, is not Hermitian (see how the $-i$ factor makes the final sign what it should, in Eqs. (46) and (47), for Eq. (48) to be correct).

In Quantum Mechanics, measurable quantities, such as the linear momentum, the position, the energy or the angular momentum, correspond to Hermitian operators. (Time is an exception: time is merely a parameter in the theory, and there is no time operator - why this is so is a rather subtle question outside the scope of the course.) Quantum Mechanics also uses non-Hermitian operators; for example, the operator creating one quantum of excitation in an harmonic oscillator is non-Hermitian (you might or might not have encountered this operator in a previous course; if you haven't, don't worry, this is not important for us).

Before closing, I should add one brief comment concerning the notation: Instead of $\langle b|Ma\rangle$, $\langle\psi_b|L_z\psi_a\rangle$, etc, it is customary in Quantum Mechanics to add a vertical bar on the right of the operator and write instead $\langle b|M|a\rangle$, $\langle\psi_b|L_z|\psi_a\rangle$, etc. This extra vertical bar does not change the meaning of this symbol. However, it has advantages in terms of developing a more abstract theory of Quantum Mechanics (a topic not addressed here).

Note 1: The conjugate transpose is usually represented by a superscript †, or more rarely by a superscript +. For example,

$$\mathsf{M}^{\mathsf{T}*} \equiv \mathsf{M}^\dagger \equiv \mathsf{M}^+. \tag{49}$$

($\mathsf{M}^\dagger$ is read "M dagger").

Note 2: Technical note, intended only for those who have studied the mathematical theory of Hilbert spaces: for infinitely-dimensional spaces, there is a subtle distinction between "Hermitian operators" and "self-adjoint operators". This distinction is usually ignored by physicists. It does not matter for this course.

3.6 What should I know about "eigenvalues" and "eigenfunctions" (or "eigenvectors" or "eigenstates"), in general?

Let us start with the eigenvalues of an operator represented by a matrix. Imagine you have a square matrix M (real or complex). When a column vector c is multiplied by M, the result is usually a column vector differing quite a bit from c. However, in some cases the result is simply c multiplied by a certain number λ (depending on the matrix, λ may have physical dimensions, or instead be a pure number):

$$\mathsf{Mc} = \lambda\mathsf{c}. \tag{50}$$

If so, the column vector c is called an eigenvector of the matrix M and λ is the corresponding eigenvalue.

For example, in the case of the matrix of Eq. (23), one sees that

$$\begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}\begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1\times 1+2\times 1 \\ 2\times 1+1\times 1 \end{pmatrix} = \begin{pmatrix} 3 \\ 3 \end{pmatrix} = 3\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \tag{51}$$

which shows that the column vector

$$\begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

is an eigenvector of this matrix and that the number 3 is the corresponding eigenvalue. One also sees that

$$
\begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ -1 \end{pmatrix} = \begin{pmatrix} 1 \times 1 - 2 \times 1 \\ 2 \times 1 - 1 \times 1 \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \end{pmatrix} = -1 \times \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \quad (52)
$$

which shows that the column vector

$$
\begin{pmatrix} 1 \\ -1 \end{pmatrix}
$$

is also an eigenvector of this matrix and that the number -1 is the corresponding eigenvalue. It is not difficult to show that this matrix has no other eigenvalues than 3 and -1. (Please refer to your maths courses if you are not clear about how to find eigenvalues and eigenvectors.)

Similarly, for operators acting on functions: suppose that the results of acting on a certain function $u(x)$ with a linear operator O is the function $u(x)$ multiplied by a constant, λ say:

$$
Ou(x) = \lambda u(x). \quad (53)
$$

If so, $u(x)$ is called an eigenfunction of O and λ is the corresponding eigenvalue.

For example, if O is the orbital angular momentum operator L_z, Eq. (53) is simply

$$
-i\hbar \frac{du}{d\phi} = \lambda u(\phi) \quad (54)
$$

(we recall that when written in terms of the usual spherical polar coordinates, $L_z = -i\hbar d/d\phi$, with ϕ the azimuthal angle). This equation is easily solved:

$$
u(\phi) = A \exp[i(\lambda/\hbar)\phi], \quad (55)
$$

where A is an arbitrary constant. Requiring that the function $u(\phi)$ is periodic with period 2π gives a condition on λ. Indeed, $u(\phi + 2\pi) = u(\phi)$ means that

$$
A \exp[i(\lambda/\hbar)(\phi + 2\pi)] = A \exp[i(\lambda/\hbar)\phi], \quad (56)
$$

namely that

$$
\exp[i(\lambda/\hbar)2\pi] = 1, \quad (57)
$$

which implies that $\lambda = m\hbar$, $m = 0, \pm1, \pm2, \dots$ Hence, the eigenfunctions of L_z are $A \exp(im\phi)$ with $\lambda = m\hbar$, $m = 0, \pm1, \pm2, \dots$

Particularly important in Quantum Mechanics are the eigenfunctions and eigenvalues of the Hamiltonian. The eigenvalues of this operator have the physical dimensions of an energy and are often called "eigenenergies". Eq. (53), for the Hamiltonian, is nothing else than the time-independent Schrödinger equation:

$$
H\psi(x) = E\psi(x), \quad (58)
$$

and similarly for N particles in 3D.

Not all linear operators have eigenvalues. Some have infinitely many, some but a few, others none at all.

Note 1: Defining eigenfunctions and eigenvalues in a mathematically rigorous way for operators acting on functions is actually rather complicated. Not all solutions of Eqs. (53) or (58) would be called eigenfunctions. Let us take, for example, the momentum operator in 1D, p_x. For this operator, Eq. (53) would simply be

$$-i\hbar\frac{\mathrm{d}\phi}{\mathrm{d}x} = \lambda\phi(x), \tag{59}$$

whose solutions are

$$\phi(x) = A\exp(ikx), \tag{60}$$

with A and k two arbitrary constants. For each value of k, the corresponding eigenvalue, λ, is $\hbar k$. As we do not impose a periodic boundary condition here, k is not restricted to discrete values. Although the function $A\exp(ikx)$ is a solution of Eq. (59) even for complex values of k, it is only for k real that one would say that $A\exp(ikx)$ is an eigenfunction of p_x. Even for real values of k, however, there is a difficulty: since

$$|A\exp(ikx)|^2 = |A|^2 \times |\exp(ikx)|^2 = |A|^2 \times 1 = |A|^2, \tag{61}$$

which is a constant, the function $A\exp(ikx)$ cannot be normalized according to Eq. (2), which causes mathematical difficulties. Physicists usually gloss over these issues, and that's we'll do in this course.

3.7 Is an "eigenfunction" the same as a "wave function"?

No. The word eigenfunction specifically refers to a function satisfying an eigenvalue equation. The word wave function specifically refers to a function describing the state of a quantum system. Some wave functions also happen to be eigenfunctions, e.g, eigenfunctions of the Hamiltonian, but some aren't. Similarly, some eigenfunctions represent a physically realizable state while others don't.

3.8 What is a degenerate eigenvalue?

First, note that if the function $\psi(x)$ is an eigenfunction of a certain linear operator O, then so is the function obtained by multiplying $\psi(x)$ by a constant: If $\psi(x)$ satisfies Eq. (53) and c is a constant,

$$O[c\,\psi(x)] = c\,O\psi(x) = c\,\lambda\psi(x) = \lambda[c\,\psi(x)]. \tag{62}$$

If all the eigenfunctions of O corresponding to a particular eigenvalue λ differ from each other only by a constant factor, one says that this eigenvalue is "non-degenerate". In Quantum Mechanics, wave functions differing from each other only by a constant factor describe the same quantum state.

There are cases, however, where truly different eigenfunctions correspond to a same eigenvalue (we are talking here about eigenfunctions that differ from each other by more than a constant factor). One then says that this eigenvalue is "degenerate". A typical example, which you have seen when you obtained the bound state wave functions and energy levels of atomic hydrogen, is the case of the $2s_{m=0}$, $2p_{m=0}$, $2p_{m=1}$ and $2p_{m=-1}$ eigenstates of the Hamiltonian, which all belong to the same energy level. (Recall that the principal quantum number, n, is 2 in all these states, and that the orbital angular momentum quantum number, l, is 0 in the s-states and 1 in the p-states; here m is the magnetic quantum number.)

It is easy to see that any linear combination of such eigenfunctions is itself an eigenfunction corresponding to this eigenvalue. For example, if

$$O\psi_a(x) = \lambda\psi_a(x) \tag{63}$$

and

$$O\psi_b(x) = \lambda\psi_b(x), \tag{64}$$

with λ having the same value in Eqs. (63) and (64), then, if c_a and c_b are two constants, we certainly have that

$$
\begin{aligned}
O[c_a\psi_a(x) + c_b\psi_b(x)] &= c_a O\psi_a(x) + c_b O\psi_b(x) \\
&= c_a \lambda\psi_a(x) + c_b \lambda\psi_b(x) \\
&= \lambda[c_a\psi_a(x) + c_b\psi_b(x)],
\end{aligned} \tag{65}
$$

and therefore the function $c_a\psi_a(x) + c_b\psi_b(x)$ is also an eigenfunction of the operator O corresponding to the eigenvalue λ.

Conversely, any eigenfunction corresponding to a degenerate eigenvalue can be written as a linear combination of other eigenfunctions corresponding to that eigenvalue.

3.9 What's so important about eigenvalues and eigenstates of Hermitian operators?

As already noted, measurable physical quantities are represented, in Quantum Mechanics, by Hermitian operators. Amongst the latter, the Hamiltonian has a particularly important role: through the time-dependent Schrödinger equation, Eq. (1), the Hamiltonian governs the time evolution of the wave functions (and of state vectors, in general).

One essential fact about Hermitian operators is that their eigenvalues are always real (they have no imaginary part). By contrast, eigenvalues of non-Hermitian operators may be complex.

Two other essential facts are (1) that eigenvectors or eigenfunctions of a Hermitian operator which correspond to different eigenvalues are always orthogonal to each other, and (2) that the eigenvectors or eigenfunctions of a Hermitian operator form a "complete basis set" for the column vectors or functions the operator acts on. What this means is perhaps best understood by looking at a couple of examples:

- We have seen above that the column vectors

$$\begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

  and

$$\begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

  are eigenvectors of the Hermitian matrix

$$\begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix},$$

  that they correspond to two different eigenvalues (3 and -1, respectively), and that this matrix has no other eigenvalues. It is easy to see that these two eigenvectors are orthogonal:

$$\begin{pmatrix} 1 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ -1 \end{pmatrix} = 1 \times 1 + 1 \times (-1) = 0. \tag{66}$$

  That they form a complete basis set simply means, here that *any* two-component column vector can be written as their linear combination. For example,

$$\begin{pmatrix} 6 \\ 2 \end{pmatrix} = 4 \begin{pmatrix} 1 \\ 1 \end{pmatrix} + 2 \begin{pmatrix} 1 \\ -1 \end{pmatrix}. \tag{67}$$

- We have have also noted that the functions $A \exp(im\phi)$, with A an arbitrary constant, are eigenfunctions of the operator L_z for any integer value of the number m. Let us set $A = (2\pi)^{-1/2}$, which ensures that these functions are normalized to unity. One easily sees that $\exp(im\phi)/(2\pi)^{1/2}$ is orthogonal to $\exp(in\phi)/(2\pi)^{1/2}$ if $m \neq n$: Assuming that $m \neq n$,

$$\int_0^{2\pi} \exp(-im\phi) \exp(in\phi) \mathrm{d}\phi = \int_0^{2\pi} \exp[i(n-m)\phi] \mathrm{d}\phi$$

$$= \frac{1}{i(n-m)} \Big(\exp[i(n-m) \times 2\pi] - \exp[i(n-m) \times 0] \Big)$$

$$= \frac{1}{i(n-m)} (1-1) = 0. \tag{68}$$

That the set formed by these functions is complete means that any "well behaved" function $\psi(\phi)$ of the azimuthal angle ϕ such that $\psi(\phi + 2\pi) = \psi(\phi)$ can be written as a linear combination of the functions $(2\pi)^{-1/2} \exp(im\phi)$. That is, one can find coefficients c_m such that

$$\psi(\phi) = \frac{1}{\sqrt{2\pi}} \sum_{m=-\infty}^{\infty} c_m \exp(im\phi). \tag{69}$$

Clearly, this equation is nothing else that a statement of the fact that any "well behaved" periodic function can be written in the form of a Fourier series. (Please refer to the level 2 maths course for the conditions the function $\psi(\phi)$ needs to meet for Eq. (69) to be possible — for us, a "well behaved" function is one that meets these conditions.)

Note 1:   Only "linearly independent" states have to be included in the basis set. The concept of linear independence is important: wave functions or state vectors are linearly independent from each other if none of them can be written as a linear combination of the others. In particular, two wave functions or state vectors differing from each other only by a constant factor are not linearly independent.

Note 2:   To make the calculations simpler, one often prefers to work with "orthonormal" bases. A basis is said to be orthonormal if each of the eigenvectors or eigenfunctions it contains is both normalized and orthogonal to the other ones (the latter is automatic for eigenfunctions corresponding to different eigenvalues but not for eigenfunctions corresponding to a same degenerate eigenvalue).

It is also essential to know the following two rules:

1. The only possible outcome of a measurement of a physical quantity is one of the eigenvalues of the Hermitian operator representing that quantity. For example, imagine that one would measure the z-component of the orbital angular momentum, for a system composed of a single particle (e.g., an atom of hydrogen, assuming that the proton is fixed). As the eigenvalues of the corresponding operator, L_z, are $m\hbar$, where m is an integer, the only possible outcome of the measurement is an integer multiple of $\hbar$: unless the measurement would be erroneous, one would never obtain a value other than $\hbar$ times an integer (even if the system is not in a particular eigenstate of L_z before measurement). Similarly, a measurement of the energy of a quantum system can only return one of the eigenvalues of the Hamiltonian as a result, and nothing in between (even if the system is in a superposition of several energy eigenstates).

2. The probability that a certain eigenvalue is obtained as a result can be calculated by projecting the wave function of the system onto the corresponding eigenfunction (or eigenfunctions in case of degeneracy). More clearly, perhaps: Suppose that you have a system described by the wave function $\Psi(x,t)$, that you measure a physical quantity represented by a Hermitian operator O, and that you are interested in the probability that the outcome of the measurement is a certain (non-degenerate) eigenvalue λ_n of the operator, such that

$$O\psi_n(x) = \lambda_n\psi_n(x). \tag{70}$$

Then, assuming that $\langle\psi_n|\psi_n\rangle = 1$, this probability is simply $|\langle\psi_n|\Psi\rangle|^2$.

Note 3: Some sweeping statements have been made above about the completeness of the set of eigenstates of Hermitian operators. However, from a mathematical point of view, one should be prudent when dealing with operators acting on functions or operators represented by infinite matrices, and completeness is not garanteed without additional conditions on the operator. The cracks in what is mentioned above are illustrated by the following example: The position operator, x, is Hermitian by the definition given in these notes, and corresponds to a physical quantity, but it is not possible to find a function $u(x)$ such that $xu(x) = \lambda u(x)$ with λ a constant. How to turn round such difficulties is outside the scope of the course.

3.10   What is an expectation value?

The concept of "expectation value" is perhaps best grasped from the following example: Suppose that you toss a coin with someone, repeatedly, and at each toss bet one pound that that you'll get head. Thus at each toss you gain one pound if you get head and lose one pound (i.e., "gain" minus one pound) if you get tail. In probabilistic terms, the "expectation value" of your gain at each toss is the amount you gain if you get head times the probability of that outcome, plus the (negative) amount you gain if you get tail times the probability of that outcome — i.e., (1 pound) × 0.5 + (−1 pound) × 0.5 (obviously, this amounts to 0 pound, assuming that the coin is fair).

The meaning of the word is the same in Quantum Mechanics. Suppose that you measure the physical quantity described by the Hermitian operator of Eq. (70), O. Then the expectation value of that quantity, $\langle O\rangle$, is the eigenvalue λ_1 of that operator times the probability that the outcome of the experiment

is λ_1, plus the eigenvalue λ_2 times the probability that the outcome is λ_2, etc. Assuming that $\langle\psi_n|\psi_m\rangle = \delta_{nm}$ and that the system is in the state described by the wave function $\Psi(x,t)$,

$$\langle O \rangle = \sum_n \lambda_n |\langle\psi_n|\Psi\rangle|^2. \tag{71}$$

Equivalently,

$$\langle O \rangle = \langle\Psi|O|\Psi\rangle. \tag{72}$$

Suppose that one prepare the system in a state of wave function $\psi(x)$, measure the physical quantity represented by the operator O, and repeat this experiment many times. The value found for that physical quantity may vary from repeat to repeat. Over a sufficiently large number of repeats, the values found can be expected to have a mean equal to $\langle O \rangle$, within statistical error.

3.11   What about the eigenfunctions of the Hamiltonian?

Many calculations in the course will be based on the application of the above to the eigenfunctions of the Hamiltonian: Any wave function can be expanded into a linear combination of eigenfunctions of the Hamiltonian, and eigenfunctions corresponding to different eigenenergies are orthogonal to each other.

If the wave function is an eigenfunction of the Hamiltonian, e.g., in 1D for a single particle,

$$H\Psi(x,t) = E\Psi(x,t), \tag{73}$$

then one says that the system is in a "stationary state". Indeed, one then have, from Eq. (1), that

$$i\hbar\frac{\partial\Psi}{\partial t} = E\Psi(x,t), \tag{74}$$

which implies that

$$\Psi(x,t) = \exp(-iEt/\hbar)\psi(x). \tag{75}$$

In this last equation, the function $\psi(x)$, which does not depend on time, is a solution of the time-independent Schrödinger equation, Eq. (58). The corresponding quantum state is said to be "stationary" because the probabilities it describes are time-independent. Indeed, these probabilities are always given by the modulus squared of expressions containing the wave function, and these expressions factorize into a phase factor $\exp(\pm iEt/\hbar)$, whose modulus is 1, and a time-independent quantity.

A linear superposition of several stationary states corresponding to *different* eigenenergies is not stationary , though, as the time-dependence cannot be factorized out.

4 Angular momentum

4.1 Orbital angular momentum

In Classical Mechanics, the angular momentum of a particle of position vector **r** and linear momentum **p** is the vector $\mathbf{r} \times \mathbf{p}$. In Quantum Mechanics, this quantity is associated with an orbital angular momentum operator, **L**, which is defined, in a similar way as in Classical Mechanics, as the cross product of the position operator **r** and the momentum operator **p**. Thus **L** is a vector operator:

$$\mathbf{L} = L_x\hat{\mathbf{x}} + L_y\hat{\mathbf{y}} + L_z\hat{\mathbf{z}}. \tag{76}$$

L_x, L_y and L_z, the components of **L** in the x-, y- and z-directions, are themselves operators. We'll denote the dot product of **L** with itself by $\mathbf{L}^2$:

$$\mathbf{L}^2 = \mathbf{L} \cdot \mathbf{L} = L_x^2 + L_y^2 + L_z^2. \tag{77}$$

The three operators L_x, L_y and L_z can be written in terms of the Cartesian coordinates x, y, z and the three components p_x, p_y, p_z of the momentum operator in the same way as in Classical Mechanics:

$$L_x = yp_z - zp_y, \quad L_y = zp_x - xp_z, \quad L_z = xp_y - yp_x. \tag{78}$$

Recall that in the position representation, $p_x \equiv -i\hbar\partial/\partial x$, $p_y \equiv -i\hbar\partial/\partial y$ and $p_z \equiv -i\hbar\partial/\partial z$. Hence, in Cartesian coordinates,

$$
\begin{aligned}
L_x &= -i\hbar\left(y\frac{\partial}{\partial z} - z\frac{\partial}{\partial y}\right), \\
L_y &= -i\hbar\left(z\frac{\partial}{\partial x} - x\frac{\partial}{\partial z}\right), \\
L_z &= -i\hbar\left(x\frac{\partial}{\partial y} - y\frac{\partial}{\partial x}\right).
\end{aligned}
\tag{79}
$$

The operators L_x, L_y and $\mathbf{L}^2$ are rather complicated in spherical polar coordinates, but L_z takes on a rather simple form:

$$L_z = -i\hbar\frac{\partial}{\partial\phi}, \tag{80}$$

where ϕ is the azimuthal angle. Note that L_z does not depend on the radial variable r. Similarly, one finds that L_x, L_y and $\mathbf{L}^2$ do not depend on r either.

Using the above equations, one can show that

$$[L_x, L_y] = i\hbar L_z, \quad [L_y, L_z] = i\hbar L_x, \quad [L_z, L_x] = i\hbar L_y, \tag{81}$$

and that

$$[\mathbf{L}^2, L_x] = [\mathbf{L}^2, L_y] = [\mathbf{L}^2, L_z] = 0. \tag{82}$$

Thus $\mathbf{L}^2$ commutes with L_x, L_y and L_z, but L_x, L_y and L_z do not commute with each other.

One can also show that the operators L_x, L_y, L_z and $\mathbf{L}^2$ are Hermitian. $\mathbf{L}^2$ has eigenfunctions that are also eigenfunctions of L_z. ($\mathbf{L}^2$ has also eigenfunctions that are also eigenfunctions of either L_x or of L_y but not of L_z). By contrast, none of the eigenfunctions of L_x are also eigenfunctions of L_y or of L_z, and reciprocally. The eigenfunctions common to $\mathbf{L}^2$ and L_z are the functions $Y_{lm}(\theta, \phi)$, called spherical harmonics, which you are probably familiar with. The angle θ is the co-latitude angle in the usual system of spherical polar coordinates (i.e., θ is the angle between the z-axis and the position vector), and, as above, ϕ is the azimuthal angle.[1] The indexes l and m refer to the eigenvalues of L_z and of $\mathbf{L}^2$ corresponding to these eigenfunctions: The eigenvalue of L_z corresponding to the spherical harmonics $Y_{lm}(\theta, \phi)$ is $m\hbar$, and that of $\mathbf{L}^2$ is $l(l+1)\hbar^2$:

$$\begin{aligned}
L_z Y_{lm}(\theta, \phi) &= m\hbar\, Y_{lm}(\theta, \phi), \\
\mathbf{L}^2 Y_{lm}(\theta, \phi) &= l(l+1)\hbar^2\, Y_{lm}(\theta, \phi).
\end{aligned} \tag{83}$$

The quantum numbers l and m are integers. The orbital angular momentum quantum number l can take the values 0, 1, 2, $\ldots$, and for each value of l, the magnetic quantum number m can take the values $-l$, $-l+1,\ldots$, $l-1$ and l. (For example, for $l = 1$, m can only take the values -1, 0 and 1.)

Note 1: The general form of these functions is $Y_{lm}(\theta, \phi) = $ (sign) $\times$ (normalization coefficient) $\times$ (polynomial in $\sin\theta$ and $\cos\theta$) $\times \exp(im\phi)$. For example,

$$\begin{aligned}
Y_{00}(\theta, \phi) &= \sqrt{\frac{1}{4\pi}}, \\
Y_{11}(\theta, \phi) &= -\sqrt{\frac{3}{8\pi}}\, \sin\theta \exp(i\phi), \\
Y_{10}(\theta, \phi) &= \sqrt{\frac{3}{4\pi}}\, \cos\theta, \\
Y_{1\,-1}(\theta, \phi) &= \sqrt{\frac{3}{8\pi}}\, \sin\theta \exp(-i\phi).
\end{aligned}$$

(You do not need to commit these expressions to memory for this course.)

Different spherical harmonics are orthogonal to each other since they are eigenfunctions of Hermitian operators with different eigenvalues. That is,

$$\int_0^\pi d\theta \sin\theta \int_0^{2\pi} d\phi\ Y^*_{l'm'}(\theta, \phi) Y_{lm}(\theta, \phi) = 0 \qquad \text{if } l \neq l' \text{ or } m \neq m'. \tag{84}$$

[1]Recall that in this system, $x = r\sin\theta\cos\phi$, $y = r\sin\theta\sin\phi$ and $z = r\cos\theta$.

Moreover, these functions are defined in such a way that

$$\int_0^\pi d\theta \sin\theta \int_0^{2\pi} d\phi \; Y_{l'm'}^*(\theta,\phi)Y_{lm}(\theta,\phi) = 1 \qquad \text{if } l = l' \text{ and } m = m'. \tag{85}$$

(Note the $\sin\theta$ factor in these two integrals, which is part of the Jacobian of the transformation from Cartesian coordinates to spherical polar coordinates.)

Something else worth remembering is that the spherical harmonics form a complete basis set, on which any well behaved functions of the polar angles θ and ϕ can be expanded. In particular, a wave function $\Psi(r,\theta,\phi,t)$ representing the quantum state of a single particle can always be written in the following form:

$$\Psi(r,\theta,\phi,t) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} c_{lm}(r,t)Y_{lm}(\theta,\phi), \tag{86}$$

where the coefficients $c_{lm}(r,t)$ do not depend on the angles θ and ϕ.

In the frequent situation where the potential is central (i.e., where the potential V is a function of r only, not of the angles θ and ϕ), the Hamiltonian has eigenfunctions $\psi_{nlm}(r,\theta,\phi)$ that are also eigenstates of the operators $\mathbf{L}^2$ and L_z. These common eigenfunctions can be written in the following form:

$$\psi_{nlm} = R_{nl}(r)\,Y_{lm}(\theta,\phi). \tag{87}$$

For a central potential, neither the radial functions $R_{nl}(r)$ nor the eigenenergies of the Hamiltonian depend on m. Recall that you have encountered this situation when you obtained the bound state wave functions of atomic hydrogen (neglecting spin-orbit coupling, and assuming that no electric or magnetic field acts on the atom). In this context, n is the principal quantum number, l the orbital angular momentum quantum number and m the magnetic quantum number. Recall also that states with $l = 0$ are often called "s-states", $l = 1$ "p-states", $l = 2$ "d-states", $l = 3$ "f-states", etc. We'll often use this terminology in the course.

4.2 Spin

In the 1920s, physicists became aware that many quantum systems have an angular-momentum like property, distinct from the orbital angular momentum. At first, it was theorized that this property could be related to some kind of self-rotation of particles forming these systems, a bit as if electrons, protons, etc, were spinning tops. This property became referred to as "spin" for that reason. It was soon realized that the picture of a self-rotation is completely wrong — an electron is *not* a spinning top — but the word "spin" kept being used. The modern understanding of spin is that this property is a purely quantum mechanical effect with no classical analogue.

The relation with the orbital angular momentum is that spin is described by a Hermitian vector operator, $\mathbf{S}$, whose components obey the same commutation relations as the components of the orbital angular momentum operator $\mathbf{L}$. We denote by S_x, S_y and S_z the x-, y- and z-components of $\mathbf{S}$, so that

$$\mathbf{S} = S_x\hat{\mathbf{x}} + S_y\hat{\mathbf{y}} + S_z\hat{\mathbf{z}}, \tag{88}$$

and

$$\mathbf{S}^2 = \mathbf{S} \cdot \mathbf{S} = S_x^2 + S_y^2 + S_z^2. \tag{89}$$

The operators S_x, S_y, S_z and $\mathbf{S}^2$ are defined so that

$$[S_x, S_y] = i\hbar S_z, \quad [S_y, S_z] = i\hbar S_x, \quad [S_z, S_x] = i\hbar S_y, \tag{90}$$

similarly to Eq. (81), and also

$$[\mathbf{S^2}, S_x] = [\mathbf{S^2}, S_y] = [\mathbf{S^2}, S_z] = 0, \tag{91}$$

similarly to Eq. (82). Contrary to the orbital angular momentum operators, however, S_x, S_y, S_z and $\mathbf{S}^2$ are *not* constructed from the particle's position and momentum; in fact, these operators *do not depend on any spatial coordinates*. While orbital angular momentum operators can be represented by differential operators acting on functions, spin operators can be represented by square matrices acting on column vectors.

One can nonetheless prove the following general result, using algebraic methods: Given the above commutation relations, the operators $\mathbf{S}^2$ and S_z have common eigenstates, χ_{s,m_s}, such that

$$S_z\chi_{s,m_s} = m_s\hbar\,\chi_{s,m_s}, \quad \mathbf{S}^2\chi_{s,m_s} = s(s+1)\,\hbar^2\chi_{s,m_s}. \tag{92}$$

Here the number s is either an integer (0, 1, 2,...) or a half-integer (1/2, 3/2, 5/2,...) and the number m_s can take any value between $-s$ and s by step of 1. These equations are, of course, the counterpart of Eqs. (83), although here the eigenstates χ_{s,m_s} are not ordinary functions depending on spatial coordinates. Another key differences with the case of orbital angular momentum is that the spin quantum number s can be a half-integer, whereas the orbital angular momentum quantum number l is necessarily an integer. (That l cannot be a half-integer comes from the particular form of the operators L_x, L_y and L_z.)

The above is best illustrated by the important case of a spin-1/2 particle, e.g., an electron. One says that the electron is a "spin-1/2" particle because, for an electron, $s = 1/2$. The corresponding spin operators can be represented as 2×2 matrices as follows:

$$S_x = \frac{\hbar}{2}\sigma_x, \quad S_y = \frac{\hbar}{2}\sigma_y, \quad S_z = \frac{\hbar}{2}\sigma_z, \tag{93}$$

where σ_k are the so called Pauli matrices, namely

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{94}$$

Given Eqs. (93) and (94), it is easy to see that

$$\mathbf{S}^2 = \frac{3\hbar^2}{4} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \tag{95}$$

Since $s = 1/2$, the only possible values of m_s are $-1/2$ and $1/2$. The corresponding eigenstates χ_{s,m_s} are represented by 2-component column vectors denoted in this course by the letter α and β. Namely,

$$\alpha \equiv \chi_{1/2,1/2} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \beta \equiv \chi_{1/2,-1/2} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

The two states these eigenvectors represent are called the state of spin up and the state of spin down in the z-direction, respectively. For instance, in this representation, the eigenvalue equations become

$$\frac{\hbar}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \frac{1}{2}\hbar \begin{pmatrix} 1 \\ 0 \end{pmatrix} \tag{96}$$

and

$$\frac{3\hbar^2}{4} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \frac{1}{2}\left(\frac{1}{2}+1\right)\hbar^2 \begin{pmatrix} 1 \\ 0 \end{pmatrix} \tag{97}$$

for the state of spin-up, showing that indeed this column vector satisfies Eqs. (92) with $s = 1/2$ and $m_s = 1/2$. Similarly,

$$\frac{\hbar}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = -\frac{\hbar}{2} \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \tag{98}$$

which shows that the state of spin-down is an eigenstate of S_z with eigenvalue $-\hbar/2$.

Since the operator S_z has not other eigenvalues than $\hbar/2$ and $-\hbar/2$ when $s = 1/2$, a measurement of the z-component of the spin angular momentum of a spin-1/2 particle can only yield either $\hbar/2$ or $-\hbar/2$. When the particle is in the spin-state α, the probability of obtaining $\hbar/2$ is 1 and that of obtaining $-\hbar/2$ is 0, and conversely when the particle is in the spin-state β. These two spin-states are orthogonal to each other:

$$\langle \alpha | \beta \rangle \equiv \begin{pmatrix} 1 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = (1 \times 0 + 0 \times 1) = 0. \tag{99}$$

More generally, the particle can be in any linear superposition of the states α and β, e.g., in the spin-state

$$|\chi\rangle \equiv \begin{pmatrix} a \\ b \end{pmatrix} = a \begin{pmatrix} 1 \\ 0 \end{pmatrix} + b \begin{pmatrix} 0 \\ 1 \end{pmatrix},$$

where a, b are complex numbers. We will always assume that the spin states we use are normalised. This means that the square of its norm is equal to one, i.e.:

$$\langle\chi|\chi\rangle \equiv \begin{pmatrix} a^* & b^* \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = (|a|^2 + |b|^2) = 1. \tag{100}$$

For instance, the states α and β defined above are normalized: $\langle\alpha|\alpha\rangle = \langle\beta|\beta\rangle = 1$. The quantities $|a|^2$ and $|b|^2$ can then be interpreted as the probabilities that a measurement of the z-component of the spin on the state χ yields the value $\hbar/2$ or $-\hbar/2$, respectively. The sum of these two probabilities must be equal to 1, since this quantity cannot take any other values than $\hbar/2$ and $-\hbar/2$.

Note 1: These probabilies are, respectively, $|a|^2/\langle\chi|\chi\rangle$ and $|b|^2/\langle\chi|\chi\rangle$ if the state is not normalised.

How to use this formalism can also be illustrated by the calculation of the expectation value of the operator S_z for a particle of spin-1/2 in an arbitrary state $|\chi\rangle$. Assuming that $|\chi\rangle$ is normalized, the expectation value is $\langle\chi|S_z|\chi\rangle$, and

$$\begin{aligned} \langle\chi|S_z|\chi\rangle &\equiv \langle S_z\rangle \equiv \begin{pmatrix} a^* & b^* \end{pmatrix} \frac{\hbar}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} \\ &= \frac{\hbar}{2} \begin{pmatrix} a^* & b^* \end{pmatrix} \begin{pmatrix} a \\ -b \end{pmatrix} \\ &= \frac{\hbar}{2} \left(|a|^2 - |b|^2\right). \end{aligned} \tag{101}$$

5 Atomic hydrogen

5.1 The bound states of the hydrogen atom

An atom of hydrogen is composed of an electron and a proton (or, for heavy hydrogen, an electron and a deuteron or a tritium nucleus). Here we concentrate on the internal structure of this system, ignoring the motion of the atom as a whole. We also neglect spin-orbit coupling and other relativistic effects. The corresponding non-relativistic Hamiltonian is defined by the following equation:

$$H = -\frac{\hbar^2}{2\mu}\nabla^2 - \frac{1}{4\pi\epsilon_0}\frac{e^2}{r}, \tag{102}$$

where r is the distance of the atomic electron to the nucleus, ∇^2 is the square of the gradient operator with respect to the coordinates of the electron, e is the charge of the electron in absolute value ($e > 0$ throughout the course) and ϵ_0 is the permittivity of free space. The constant μ is the reduced mass of the (electron - nucleus) system; in this course, however, we will generally take μ to be equal to the mass of the electron, m, as the two differ by less than 0.1%.

Note 1: $\mu = mM/(m + M)$, where m is the mass of the electron and M is the mass of the nucleus. Since $M \gg m$, $(m+M) \approx M$ and therefore $\mu \approx m$.

Recall also that if we place the origin of the system of coordinates on the nucleus and denote by x, y and z the Cartesian coordinates of the electron, then

$$\nabla^2 \equiv \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \tag{103}$$

and $r = (x^2 + y^2 + z^2)^{1/2} = |\mathbf{r}|$, where $\mathbf{r}$ is the position vector of the electron with respect to the nucleus. The term in $1/r$ in the Hamiltonian is the potential energy of the electron due to its Coulomb interaction with the nucleus. The term in ∇^2 can be understood as corresponding to the kinetic energy of the electron.

Note 2:   It is instructive to look at the origin of this Hamiltonian. As already mentioned, the atom of hydrogen is a system of two particles, i.e., the atomic electron and the nucleus. The wave function of the whole atom is therefore a function of both the position vector of the nucleus and the position vector of the electron, $\mathbf{R}_n$ and $\mathbf{r}_{el}$, respectively. (Note that $\mathbf{r}_{el}$ is the position vector of the electron with respect to a fixed point in space, not with respect to the nucleus.)

Neglecting relativistic effects, and assuming that the atom is not interacting with other atoms or with external electric or magnetic fields, the Hamiltonian of the whole system reads

$$H_{at} = -\frac{\hbar^2}{2M}\nabla_n^2 - \frac{\hbar^2}{2m}\nabla_{el}^2 - \frac{1}{4\pi\epsilon_0}\frac{e^2}{|\mathbf{r}_{el} - \mathbf{R}_n|},$$

where M is the mass of the nucleus, m the mass of the electron, ∇_n is the gradient operator for the coordinates of the nucleus and ∇_{el} is the gradient operator for the coordinates of the electron. As in Classical Mechanics, the motion of this two-body system can be decomposed into the motion of its centre of mass (i.e., the motion of the atom as a whole) and the relative motion of the electron and the nucleus. The Hamiltonian can be separated accordingly. To this effect, one changes coordinates from the x-, y- and z-components of the position vectors $\mathbf{R}_n$ and $\mathbf{r}_{el}$ to the x-, y- and z-components of the position vectors $\mathbf{r} = \mathbf{r}_{el} - \mathbf{R}_n$ and $\mathbf{R}_{CM} = (M\mathbf{R}_n + m\mathbf{r}_{el})/(M + m)$. ($\mathbf{r}$ is the position vector of the electron relative to the nucleus, as above, and $\mathbf{R}_{CM}$ is the position vector of the centre of mass.) After calculation, one finds than in these new coordinates the full Hamiltonian is

$$H_{\text{at}} = -\frac{\hbar^2}{2(M+m)}\nabla^2_{\text{CM}} + H,$$

where H is the Hamiltonian defined by Eq. (102) and ∇_{CM} is the gradient operator for the coordinates of the centre of mass. The full Hamiltonian is thus the sum of an Hamiltonian describing the motion of the centre of mass, i.e., the motion of the atom as a whole (the kinetic energy term in ∇^2_{CM}) and an Hamiltonian describing the motion of the electron relative to the nucleus (H).

The fact that H_{at} does not include a potential energy term depending on both the coordinates of the centre of mass and on the relative coordinates shows that the motion of the centre of mass is not coupled to the motion of the electron relative to the nucleus. Therefore the former does not need to be considered in relation to studying the internal structure of the atom.

The eigenenergies and eigenfunctions of the Hamiltonian (102) can be calculated analytically. One finds an infinite set of bound states eigenfunctions $\psi_{nlm}(\mathbf{r})$, orthogonal to each other and characterized by three quantum numbers: n (the principal quantum number), l (the orbital angular momentum quantum number) and m (the magnetic quantum number, not to be confused with the mass of the electron which is also denoted by the letter m in this course).

- The principal quantum number varies from 1 to infinity: $n = 1, 2, 3,\ldots$

- For each value of n, l can take any integer value between 0 and $n-1$. Thus for $n = 1$, l can only be 0. For $n = 2$, l can be 0 or 1, etc...

- For each value of l, the magnetic quantum number m can take any integer value between $-l$ and l. Thus for $l = 0$, m can only be zero, whilst for $l = 1$, m can be -1, 0 or 1.

- The bound states are often referred to by their spectroscopic symbol: states with $l = 0$ are called s-states, $l = 1$ p-states, $l = 2$ d-states, $l = 3$ f-states, etc. For example, the $4\text{d}_{m=2}$ state is the eigenstate with $n = 4$, $l = 2$ and $m = 2$.

- The corresponding bound state eigenenergies, E_n, depend on the principal quantum number but not on l or m. The higher n is, the higher the energy of the state. The state of lowest energy (i.e., the ground state), is therefore the 1s state (the lowest possible value of n is 1, there is no other possible value of l than 0 when $n = 1$, and the only possible value of m is 0 when $l = 0$).

- The corresponding bound state wave functions, $\psi_{nlm}(\mathbf{r})$, can also be obtained analytically. In general,

$$\psi_{nlm}(\mathbf{r}) = R_{nl}(r) \times Y_{lm}(\theta, \phi), \qquad (104)$$

where $Y_{lm}(\theta, \phi)$ is a spherical harmonics depending on the electron's polar angles θ and ϕ and $R_{nl}(r)$ is a wave function depending on the radial variable r. For example, in the approximation where the reduced mass μ of the (electron - nucleus) system is taken to be equal to the electron mass,

$$\psi_{100}(\mathbf{r}) = 2\, a_0^{-3/2} \exp(-r/a_0) Y_{00}(\theta, \phi), \qquad (105)$$

where a_0 is the Bohr radius. (There is no need to learn this equation by heart for this course.)

- The energy eigenfunctions $\psi_{nlm}(\mathbf{r})$ are eigenfunctions not only of the Hamiltonian but also of the angular momentum operators $\mathbf{L}^2$ and L_z.

Note 3: Since the eigenenergies E_n do not depend on l or m, these energy levels are degenerate. It is usually convenient to take the energy eigenfunctions to be also eigenfunctions of $\mathbf{L}^2$ and L_z, each corresponding to a well defined value of l and well defined value of m, as done here. However, alternative choices are possible; i.e., one could work with energy eigenfunctions formed of linear superpositions of several different functions $\psi_{nlm}(\mathbf{r})$ belonging to the same eigenenergy E_n.

Note 4: The bound state eigenenergies of a one-particle quantum system never depend on the magnetic quantum number if the potential is central (i.e., if the force field acting on the particle is spherically symmetric). This can be understood intuitively from the fact that this quantum number distinguishes the different eigenstates of the z-component of the angular momentum operator, and that there is no reason that the choice of this particular component would have a bearing on the Physics of the problem if the atom is unperturbed (in the absence of an external electric or magnetic field, there is no privileged direction of space). That the eigenenergies of the Hamiltonian (102) do not depend on l either is a peculiar feature of that Hamiltonian.

Note 5: Besides the bound states mentioned above, in which the electron has a vanishingly small probability to be far from the nucleus, the Hamiltonian (102) also has eigenfunctions representing "unbound" or "continuum" states in which the electron can escape to infinity. We will come back to these unbound states later in the course. At this point, I just mention that the corresponding eigenenergies are positive (bound states correspond to negative eigenenergies) and are not quantized (they can take any positive value, hence the name "continuum states").

Note 6: Hydrogenic ions can be treated in the same way as atomic hydrogen. (Hydrogenic ions are atoms which have lost all their electrons but one and are thus composed of a multiply charged nucleus and a single electron, for instance He^+, Li^{2+}, etc.) The only difference is that the potential energy of the electron is a factor Z stronger, where Z is the atomic number (the number of protons in the nucleus). The corresponding Hamiltonian is

$$H = -\frac{\hbar^2}{2\mu}\nabla^2 - \frac{1}{4\pi\epsilon_0}\frac{Ze^2}{r}. \tag{106}$$

Exactly as in the case of atomic hydrogen, the eigenfunctions of this Hamiltonian can be labelled by the three quantum numbers n, l and m and the eigenenergies depend only on n.

5.2 Spin-orbit coupling

The spin of the electron was not mentioned in the previous section, and in fact, since no spin operator appears in the Hamiltonian (102), its eigenenergies do not depend on the spin state of the electron. However, this is only an approximation: physically, the orbital angular momentum of the electron is coupled to its spin by the spin-orbit interaction.

Note 1: The origin of this interaction can be understood from the following argument. An electron has a magnetic moment, which depends on its spin state. If for a second we imagine that, as in the Bohr model, the electron is a classical point charge orbiting the nucleus, then for an observer "on" the electron the nucleus is a point charge orbiting the electron. As seen from the electron, the motion of nucleus would thus form a current loop. Since current loops produce a magnetic field, and since magnetic moments couple to magnetic fields, one can expect a coupling between the spin of the electron and its orbital motion, hence between the spin of the electron and its orbital angular momentum. This intuitive reasoning is, of course, not rigorous (to start with, the eigenstates of the Hamiltonian do not represent states in which the electron "orbits" the nucleus). However, the effect does exist. It can be taken into account by adding corrective terms to the Hamiltonian. These terms are often referred to relativistic corrections, as they appear naturally and in a rigorous way in the relativistic quantum mechanical equation of motion for electrons (the Dirac equation).

Due to this coupling, the energy eigenstates of the atom are actually linear combinations of products of the form $\psi_{nlm}(\mathbf{r})\chi_{1/2,m_s}$, where the spatial wave function $\psi_{nlm}(\mathbf{r})$ is one of the eigenfunctions of the Hamiltonian (102) and $\chi_{1/2,m_s}$ is a spin state (recall that for a spin-1/2 particle such as the electron, $m_s = 1/2$ or $-1/2$, and that $\chi_{1/2,m_s}$ represents a state of spin up when $m_s = 1/2$ and a state of spin down when $m_s = -1/2$.)

The spin-orbit effects can be treated as a perturbation, with a perturbation Hamiltonian depending both on the orbital angular momentum operator $\mathbf{L}$ and on the spin operator $\mathbf{S}$. The operator $\mathbf{L}$ acts only on the spatial wave functions $\psi_{nlm}(\mathbf{r})$. The operator $\mathbf{S}$ acts only on the spin states $\chi_{1/2,m_s}$.

Once the coupling between spin and orbital angular momentum is taken into account, one finds that the energy eigenstates of the atom are eigenstates of $\mathbf{L}^2$, of $\mathbf{S}^2$, of $\mathbf{J}^2$ and of J_z. Here $\mathbf{J}$ is the total angular momentum operator, i.e.,

$$\mathbf{J} = \mathbf{L} + \mathbf{S}, \tag{107}$$

and the operator J_z is the z-component of $\mathbf{J}$.

- $\mathbf{J}$ is an angular momentum operator, as are $\mathbf{L}$ and $\mathbf{S}$.

- The eigenvalues of $\mathbf{J}^2$ are $j(j+1)\hbar^2$, where for a state (the same vector $\mathbf{r}$ as above) with orbital angular momentum l, $j = l - 1/2$ or $j = l + 1/2$. (For $l = 0$, j can only be $1/2$.)

- In the spectroscopic notation, the value of j is normally indicated as a subscript. For example, the $4d_{5/2}$ state is the state with $n = 4$, $l = 2$ and $j = 5/2$ (recall the meaning of the letters s, p, d, f,..., and note that here the subscript indicates the value of j).

- The eigenvalues of J_z are $m_j\hbar$. For an eigenstate state of $\mathbf{J}^2$ with quantum number j, the possible values of m_j range from $-j$ to j by integer step.

- The spin-orbit coupling shifts and splits the bound state energy levels into sublevels differing by the value of j (with the exception of the ground state energy, which simply shifts — it cannot split since only one value of j is possible for $n = 1$). The ensuing subdivision of the energy levels is referred to as their fine structure.

6 Time-independent perturbation theory

6.1 Introduction

The Quantum Mechanics 3 course starts with a survey of time-*dependent* perturbation theory, which is primarily concerned with the calculation of the probability of transitions induced by a weak interaction *varying* in time. The level 2 Quantum Mechanics course includes an introduction to time-*independent* perturbation theory, which is primarily concerned with the calculation of how energy levels shift and perhaps split under the effect of a weak interaction *constant* in time. We will use time-independent perturbation theory in Quantum Mechanics 3, but only briefly, when studying the energy levels of the atom of helium.

To recap, the basic problem is to calculate how the eigenenergies and energy eigenfunctions of a Hamiltonian change under the effect of a weak perturbation. More precisely, we assume that the Hamiltonian of the problem, H can be written as the sum of a "well known" Hamiltonian, H_0, and another energy operator, H', the latter describing some interaction perturbing (weakly) the system described by H_0:

$$H = H_0 + H'. \tag{108}$$

The operator H' is often called the perturbation Hamiltonian. Suppose that $E_n^{(0)}$ is one of the eigenenergies of H_0, and that this eigenenergy and the corresponding energy eigenfunction(s) can be calculated readily (that's what I meant by "well known", earlier on). The question is to find what this eigenenergy becomes under the effect of H'.

For example, take the case of an atom of hydrogen exposed to a weak static electric field oriented in the z-direction. As you may remember to have seen at level 2, the Hamiltonian describing this system is

$$H = -\frac{\hbar^2}{2\mu}\nabla^2 - \frac{1}{4\pi\epsilon_0}\frac{e^2}{r} + e\mathcal{E}z, \tag{109}$$

assuming that the electric field is homogeneous (i.e., that it has the same strength everywhere). In this equation, $\mathcal{E}$ denotes the strength of the field and, as in Section 5, e is the absolute charge of the electron ($e > 0$). Static electric fields that can be produced in a laboratory are very much weaker than the electric field induced by a proton at the distance the electron typically is in an atom of hydrogen (about 0.05 nm in the ground state). Hence, we can consider the interaction with the external electric field to be only a small perturbation, and take $H' \equiv e\mathcal{E}z$. In this case, the "unperturbed Hamiltonian", H_0, is simply the Hamiltonian of the field-free atom, as given by Eq. (102).

Coming back to the general problem of finding the eigenenergies and the corresponding energy eigenfunctions of the total Hamiltonian H, the general approach is to introduce a parameter λ and to work with the more general Hamiltonian

$$H(\lambda) = H_0 + \lambda H'. \tag{110}$$

Remember that this parameter λ is simply a number we can choose at will: $H(\lambda) \equiv H_0$ for $\lambda = 0$, and $H(\lambda) = H_0 + H'$ for $\lambda = 1$. The usefulness of writing the total Hamiltonian in this way is that it makes it possible to seek the eigenenergies and energy eigenfunctions of the perturbed Hamiltonian $H(\lambda)$ in the form of power series in λ. E.g., the eigenenergies $E_n(\lambda)$ of $H(\lambda)$ can be written as

$$E_n(\lambda) = E_n^{(0)} + \lambda E_n^{(1)} + \lambda^2 E_n^{(2)} + \cdots \tag{111}$$

and each term in that series can be calculated one by one. As mentioned above, $H(\lambda) \to H_0 + H'$ for $\lambda \to 1$; hence, the eigenenergis of the perturbed Hamiltonian

are obtained simply by setting $\lambda = 1$ in Eq. (111). Also, since $H(\lambda) \to H_0$ for $\lambda \to 0$, setting $\lambda = 0$ in Eq. (111) shows that the term $E_n^{(0)}$ is necessarily an eigenenergy of the unperturbed Hamiltonian H_0.

How to calculate the terms $E_n^{(1)}$, $E_n^{(2)}$, etc. and the corresponding energy eigenfunctions depends on whether the eigenenergy $E_n^{(0)}$ is degenerate or not (read through Section 3.8 above if you are unclear about what one means by a degenerate eigenvalue). We will start with the case where $E_n^{(0)}$ is non-degenerate, and then pass to the more complicated case of a degenerate eigenenergy.

> Note 1: There are several different formulations of time-independent perturbation theory. The one most frequently used is called Rayleigh-Schrödinger perturbation theory. It is the one summarized here and you have seen at level 2.

6.2 Perturbation of a non-degenerate energy level

Let us assume that only one eigenstate of H_0 has $E_n^{(0)}$ for eigenenergy, and that it is corresponds to a wave function $\psi_n^{(0)}(\mathbf{r})$. (All the equations quoted in this chapter generalizes easily to the case of 1D systems, systems of several particles, spin, etc.) Then the first order term in the expansion of E_n in powers of the perturbation is given by the very simple equation

$$E_n^{(1)} = \int \psi_n^{(0)*}(\mathbf{r}) \, H' \psi_n^{(0)}(\mathbf{r}) \, \mathrm{d}^3 r, \qquad (112)$$

I.e., $E_n^{(1)}$ is simply the expectation value of the perturbation Hamiltonian H' on the unperturbed state. General formulas for the coefficients $E_n^{(2)}$, $E_n^{(3)}$, etc., can also be written, but they are somewhat more complicated and we will not need them in the course.

For example, if one ignores spin, the ground state energy of atomic hydrogen is non-degenerate. How it is affected by a weak static electric field can be calculated by taking H_0 and H' as mentioned above, with the unperturbed eigenenergy $E_n^{(0)}$ the ground state energy E_0 of the field-free atom. Eq. (112) then says that the ground state energy of the atom in the presence of the field is

$$E_0 + \int \psi_{100}^*(\mathbf{r}) \, e\mathcal{E}z \, \psi_{100}(\mathbf{r}) \, \mathrm{d}^3 r + \cdots$$

where $\psi_{100}(\mathbf{r})$ is the unperturbed ground state wave function and the dots stand for terms of second or higher order in $\mathcal{E}$.

Note 1: It turns out that the integral is zero here, which means that the ground state energy of the atom depends on the field only through terms quadratic or of higher order in $\mathcal{E}$. Examples of calculations where the first order correction is non-zero are given in Griffiths and in Bransden and Joachain.

6.3 The case of a degenerate energy level

The situation is more complicated if the perturbed eigenenergy of interest, $E_n(\lambda)$, converges to a degenerate eigenenergy of H_0 for $\lambda \to 0$. That the eigenenergy $E_n^{(0)}$ is degenerate means that several different eigenstates of H_0 have that energy. Suppose, for example, that N different energy eigenfunctions of H_0 all have $E_n^{(0)}$ as eigenenergy. I.e., there are N (> 0) distinct eigenfunctions $\psi_{n;1}^{(0)}(\mathbf{r})$, $\psi_{n;2}^{(0)}(\mathbf{r}),\ldots, \psi_{n;N}^{(0)}(\mathbf{r})$ such that

$$H_0 \psi_{n;k}^{(0)}(\mathbf{r}) = E_n^{(0)} \psi_{n;k}^{(0)}(\mathbf{r}), \qquad k = 1,\ldots,N. \tag{113}$$

($E_n(0)$ is said to be "N-fold degenerate" in this case.) One should then proceed as follows:

1. First, calculate the "matrix element" of the perturbation Hamiltonian H' for each pair of unperturbed eigenfunctions; i.e., calculate the integrals

$$H'_{kk'} = \int \psi_{n;k}^{(0)*}(\mathbf{r})\, H'\, \psi_{n;k'}^{(0)}(\mathbf{r})\, \mathrm{d}^3 r, \qquad k, k' = 1,\ldots,N. \tag{114}$$

2. Arrange the N^2 matrix elements $H'_{kk'}$ into a square matrix of N rows and N columns, so that for given values of k and k' the constant $H'_{kk'}$ is the element in the k-th row and the k'-th column of the matrix.

3. Calculate the N eigenvalues of this matrix.

The N eigenvalues so obtained are the values $E_n^{(1)}$ can take for the eigenenergy or eigenenergies of $H(\lambda)$ which converge to $E_n^{(0)}$ for $\lambda \to 0$.

One may find that all the N eigenvalues of this matrix are identical. In this case, the corresponding eigenenergy of the perturbed Hamiltonian is still N-fold degenerate (to first order in H'). A more common situation, however, is that the eigenvalues differ from each other, or at least that they are not all identical. Then the effect of H' is to split the degenerate unperturbed energy level into several levels, each one corresponding to a different value of $E_n^{(1)}$.

Note 1: The possible values of $E_n^{(1)}$ are the solutions of the determinant equation

$$\begin{vmatrix} H'_{11} - E_n^{(1)} & H'_{12} & \cdots & H'_{1N} \\ H'_{21} & H'_{22} - E_n^{(1)} & \cdots & H'_{2N} \\ \vdots & \vdots & & \vdots \\ H'_{N1} & H'_{N2} & \cdots & H'_{NN} - E_n^{(1)} \end{vmatrix} = 0. \tag{115}$$

Note 2: Suppose that $E_n^{(1)}$ is one of these solutions, and that no other solution has the same value. Then, within an overall factor, there exists a unique column vector $(c_{n;k})$ such that

$$\begin{pmatrix} H_{11}' & H_{12}' & \cdots & H_{1N}' \\ H_{21}' & H_{22}' & \cdots & H_{2N}' \\ \vdots & \vdots & & \vdots \\ H_{N1}' & H_{N2}' & \cdots & H_{NN}' \end{pmatrix} \begin{pmatrix} c_{n,1} \\ c_{n,2} \\ \vdots \\ c_{n,N} \end{pmatrix} = E_n^{(1)} \begin{pmatrix} c_{n,1} \\ c_{n,2} \\ \vdots \\ c_{n,N} \end{pmatrix}. \qquad (116)$$

(Clearly, this column vector is the eigenvector of the matrix $(H_{kk'}')$ belonging to the eigenvalue $E_n^{(1)}$.) Moreover, $H(\lambda)$ also has a non-degenerate eigenenergy $E_n(\lambda)$ such that

$$H(\lambda)\psi_n(\mathbf{r}) = E_n(\lambda)\psi_n(\mathbf{r}) \qquad (117)$$

and

$$\psi_n(\mathbf{r}) \to \sum_{k=1}^{N} c_{n;k}\psi_{n;k}^{(0)}(\mathbf{r}) \quad \text{for } \lambda \to 0. \qquad (118)$$

This equation shows that the unperturbed states $\psi_{n;k}^{(0)}(\mathbf{r})$ are effectively mixed to each other by the perturbation (except of course in the case where all but one of the coefficients $c_{n;k}$ are zero).

Note 3: Contrast the above with the non-degenerate case, where *only one* eigenstate of H_0 has $E_n^{(0)}$ as eigenenergy. In this case there is no ambiguity as to which eigenstate of H_0 the perturbed eigenstate of $H(\lambda)$ reduces to in the zero-λ limit. However, when *several* eigenstates of H_0 have $E_n^{(0)}$ as eigenenergy, then it may not be clear which of them, if any in particular, is the zero-λ limit of the perturbed eigenstate. In fact, the zero-λ limit might well be a certain linear combination of eigenstates of H_0. The procedure outlined above identifies what the correct zero-λ limit is amongst this infinite number of possibilities.