

PH3062 Quantum mechanics 2

Lecture notes *

Candlemas semester 2020

1 Review of exact solutions; degeneracy

This lecture mainly summarises some results from Quantum Mechanics 1 that we shall be using extensively in the first part of the course.

1.1 Eigenvalues and eigenkets

An observable A is represented in quantum mechanics by a Hermitian operator \hat{A} . Hermitian means that

$$\hat{A}^\dagger = \hat{A}. \quad (1)$$

The eigenvalues of \hat{A} are defined by

$$\hat{A}\phi_n = a_n\phi_n, \quad (2)$$

where ϕ_n is the eigenfunction corresponding to the eigenvalue a_n . There may be more than one such eigenfunction — see below. The eigenvalues a_n represent the complete set of possible results if A is measured. The fact that \hat{A} is Hermitian guarantees that the eigenvalues are all real, which if they are to be the results of measurement appears to be a good thing.

The eigenfunctions of a Hermitian operator are (or can be chosen to be) orthonormal, i.e.

$$\int_{\text{all space}} \phi_n^* \phi_m d\tau = \delta_{nm} \equiv \begin{cases} 1 & n = m, \\ 0 & n \neq m. \end{cases} \quad (3)$$

The integral on the left-hand side is called the **overlap** of the eigenfunctions ϕ_n and ϕ_m . Orthogonality means that the overlap of different eigenfunctions is zero; normalisation means that the self-overlap of an eigenfunction is unity. These words combine to make ‘orthonormal’.

*The majority of the texts have been written by Christopher Hooley; Sections 3.1, 3.2 and 12.5 by Brendon Lovett, Checkpoints with solutions, sections 3.3-3.5, 12.4, 15.4, 15.5 and minor amendments throughout by Antje Kohnle.

The eigenfunctions of a Hermitian operator form a **complete basis**. This means that any wave function can be expanded as a weighted sum of such eigenfunctions,

$$\psi = \sum_n c_n \phi_n. \quad (4)$$

If ψ is normalised, its self-overlap is unity, i.e.

$$\int_{\text{all space}} \psi^* \psi d\tau = \int_{\text{all space}} |\psi|^2 d\tau = 1. \quad (5)$$

Substituting in the decomposition (4), we find

$$\int_{\text{all space}} \left(\sum_n c_n^* \phi_n^* \right) \left(\sum_m c_m \phi_m \right) d\tau = 1. \quad (6)$$

Expanding the brackets on the left-hand side, and bringing the sums and coefficients to the front,

$$\sum_n \sum_m c_n^* c_m \int_{\text{all space}} \phi_n^* \phi_m d\tau = 1. \quad (7)$$

But since the eigenfunctions are orthonormal, the overlap is non-zero only when $n = m$. Hence the sum over m contains a single term, and is not really a sum at all; we may therefore delete it, and set $m = n$, whereupon the overlap becomes unity:

$$\sum_n c_n^* c_n = 1, \quad \text{i.e.} \quad \sum_n |c_n|^2 = 1. \quad (8)$$

Interpretation: $|c_n|^2$ is the probability of finding outcome a_n when a measurement of A is made on a system with wave function ψ . Since *some* result *must* be obtained, the sum of these probabilities over all possible results is unity.

1.2 Commutators

The **commutator** of two operators \hat{A} and \hat{B} is defined by

$$[\hat{A}, \hat{B}] \equiv \hat{A}\hat{B} - \hat{B}\hat{A}. \quad (9)$$

It represents the extent to which measuring A interferes with the value of B . For example, if I have a particle in an eigenstate of momentum,

$$\phi_{\mathbf{p}} = N e^{i\mathbf{p} \cdot \mathbf{x}}. \quad (10)$$

A position measurement collapses this to a delta-function, so that immediately after the measurement

$$\psi = \delta^{(3)}(\mathbf{x} - \mathbf{x}_0), \quad (11)$$

where \mathbf{x}_0 is the position at which I found the particle. This is clearly no longer a momentum eigenfunction, since it is not of the form (10); hence my measurement of position has interfered with the value of momentum. We would thus expect a non-zero commutator. Explicit calculation shows that

$$[\hat{x}, \hat{p}] = i\hbar \quad (12)$$

in one dimension. In higher dimensions, we have

$$[\hat{x}_j, \hat{p}_k] = i\hbar\delta_{jk}, \quad (13)$$

i.e. measuring the j th position co-ordinate only interferes with the j th momentum. ($j = 1, 2, \dots, d$, where d is the dimension of space. This is usually three, but things do get crazy sometimes.)

✓ Rewrite in terms of commutators (c is a constant; \hat{A} , \hat{B} , \hat{C} are operators): i) $[c\hat{A}, \hat{B}]$;

ii) $[\hat{A} + \hat{B}, \hat{C}]$ ¹

✓ What is $[\hat{A}, \hat{A}^n]$ where n is a positive integer? ²

1.3 Eigenfunctions and eigenvalues for particular problems

This subsection is a digest of the eigenfunctions and eigenvalues for the three main exactly solvable problems you have seen so far.

1.3.1 The infinitely deep square well

In $d = 1$, the Hamiltonian is

$$\hat{H} = \frac{\hat{p}^2}{2m} + V(x), \quad (14)$$

with

$$V(x) = \begin{cases} 0 & 0 < x < L, \\ +\infty & \text{otherwise.} \end{cases} \quad (15)$$

The exact eigenfunctions are

$$\phi_n(x) = \begin{cases} \sqrt{\frac{2}{L}} \sin\left(\frac{n\pi x}{L}\right) & 0 < x < L, \\ 0 & \text{otherwise,} \end{cases} \quad (16)$$

where $n = 1, 2, 3, \dots$. The corresponding eigenenergies are

$$E_n = \frac{\pi^2 \hbar^2}{2mL^2} n^2. \quad (17)$$

¹i) $[c\hat{A}, \hat{B}] = c[\hat{A}, \hat{B}]$; ii) $[\hat{A} + \hat{B}, \hat{C}] = [\hat{A}, \hat{C}] + [\hat{B}, \hat{C}]$

² $[\hat{A}, \hat{A}^n] = \hat{A}^{n+1} - \hat{A}^{n+1} = 0$

1.3.2 The simple harmonic oscillator

In $d = 1$, the Hamiltonian is

$$\hat{H} = \frac{\hat{p}^2}{2m} + \frac{1}{2}m\omega^2 x^2. \quad (18)$$

The exact eigenfunctions are

$$\phi_n(x) = N H_n(y) e^{-\alpha x^2}, \quad (19)$$

where $n = 0, 1, 2, \dots$, $H_n(y)$ is the n th Hermite polynomial, $y = \sqrt{\frac{m\omega}{\hbar}}x$ and $\alpha = m\omega/2\hbar$. The first two Hermite polynomials are respectively 1 and y . The corresponding eigenenergies are

$$E_n = \hbar\omega \left(n + \frac{1}{2} \right). \quad (20)$$

✓ Write down an expression for the energy eigenvalues of a two-dimensional harmonic oscillator. ³

1.3.3 The hydrogen atom

In $d = 3$, the Hamiltonian is

$$\hat{H} = \frac{\hat{\mathbf{p}}^2}{2m} - \frac{e^2}{4\pi\epsilon_0 r}. \quad (21)$$

The exact eigenfunctions are labelled by three quantum numbers: n , l , and m . The principal quantum number n may take any integer value from 1 upwards; the angular momentum quantum number l may take any integer value from 0 to $n - 1$; and the azimuthal quantum number m may take any integer value from $-l$ to l . The exact eigenfunctions may be written in spherical polar co-ordinates as

$$\phi_{nlm}(r, \theta, \phi) = R_{nl}(r) Y_{lm}(\theta, \phi), \quad (22)$$

where $Y_{lm}(\theta, \phi)$ is a **spherical harmonic**. Explicit forms of the spherical harmonics, and for the radial parts $R_{nl}(r)$, may be found in any good quantum mechanics textbook.

1.4 Degeneracy

The notion of **degeneracy** will be very important in what follows. It refers to the situation where two or more distinct eigenfunctions of an operator — often the Hamiltonian, \hat{H} — have the same eigenvalue. In one-dimensional problems, degeneracy does not usually occur. However, in higher dimensions it becomes rife.

For example, let's take the solution of the infinitely deep square well problem in $d = 2$. Because the Hamiltonian is **separable**, the eigenfunctions are just products of the $d = 1$ eigenfunctions, and the eigenvalues are just sums of the $d = 1$ eigenvalues. Thus the exact eigenfunctions are

$$\phi_{n_x n_y}(x, y) = \begin{cases} \frac{2}{L} \sin\left(\frac{n_x \pi x}{L}\right) \sin\left(\frac{n_y \pi y}{L}\right) & 0 < x < L \text{ and } 0 < y < L, \\ 0 & \text{otherwise.} \end{cases} \quad (23)$$

³ $E_{n_x n_y} = E_{n_x} + E_{n_y} = \hbar\omega(n_x + \frac{1}{2}) + \hbar\omega(n_y + \frac{1}{2}) = \hbar\omega(n_x + n_y + 1)$

Note that there are now two quantum numbers, n_x and n_y , which may be chosen independently to take any integer value from 1 upwards. **Every distinct choice of the pair (n_x, n_y) corresponds to a distinct eigenfunction.** The eigenenergies are

$$E_{n_x n_y} = \frac{\pi^2 \hbar^2}{2mL^2} (n_x^2 + n_y^2). \quad (24)$$

A table of the first few possible choices for n_x and n_y , along with the associated values of $n_x^2 + n_y^2$,

n_x	n_y	$n_x^2 + n_y^2$
1	1	2
1	2	5
2	1	5
2	2	8
1	3	10
3	1	10

clearly shows that there are now some energies that correspond to two different eigenfunctions. These energy levels are said to be **degenerate**. In this case, the energy level $E = \pi^2 \hbar^2 / mL^2$ is non-degenerate, while the energy level $E = 5\pi^2 \hbar^2 / 2mL^2$ is doubly degenerate (or equivalently has degeneracy $g = 2$).

✓ Consider the hydrogen atom electron. What is the degeneracy with respect to the energy $E_2 = -\frac{R_H}{4}$? ⁴

⁴If $n = 2$, then $l = 0, m = 0$ or $l = 1, m = -1, 0, 1$. Thus E_2 is 4-fold degenerate. $\phi_{200}, \phi_{210}, \phi_{211}, \phi_{21,-1}$ all have energy $E_2 = -\frac{R_H}{4}$.

2 Dirac notation and examples of its use

In this lecture, we shall develop a new notation for quantum mechanics, known as Dirac notation. For systems whose degrees of freedom are spatial (such as a particle moving in an oscillator potential), it is optional but advantageous. For systems whose degrees of freedom are internal (such as a quantum spin), however, it is essential.

2.1 The vector analogy

What we do in quantum mechanics with eigenfunctions of Hermitian operators is very similar to what we do in classical mechanics with vectors. Here's a table that will hopefully make that clear. The functions ϕ_n are the eigenfunctions of some Hermitian operator \hat{B} .

Quantum mechanics	Vectors
Decomposition of a wave function as a sum of eigenfunctions $\psi = \sum_n c_n \phi_n = c_1 \phi_1 + c_2 \phi_2 + \dots$	Expression of a vector in a basis $\mathbf{A} = a_x \mathbf{e}_x + a_y \mathbf{e}_y + a_z \mathbf{e}_z$
Orthonormality of eigenfunctions $\int \phi_n^* \phi_m d\tau = \delta_{nm}$	Orthonormality of basis vectors $\mathbf{e}_i \cdot \mathbf{e}_j = \delta_{ij}$
Normalisation of a wave function $\int \psi^* \psi d\tau = 1$ $\sum_n c_n ^2 = 1$	Unit length of a vector $\mathbf{A} \cdot \mathbf{A} = 1$ $a_x^2 + a_y^2 + a_z^2 = 1$

This strongly suggests that quantum mechanics is really some sort of vector theory. To make the identification exact, we must interpret the state of the system as a vector in this abstract space, the eigenstates as a complete set of a basis vectors, and the overlap

$$\int \psi^* \phi d\tau \quad (25)$$

as the equivalent of a dot product.

2.2 Dirac notation

Dirac notation takes this analogy seriously, and tries to make our notation look more like it would if it were describing vectors in a space. Here are the basics:

- The vectors in this space are called **kets**, and they are written like this:

$$|\phi\rangle. \quad (26)$$

Their complex conjugates are called **bras**, and they are written like this:

$$\langle\psi|. \quad (27)$$

The label inside is arbitrary: it just tells you which state the ket or bra represents. An eigenket of the Hamiltonian (or another Hermitian operator) is usually just labelled directly with its quantum number(s), e.g.

$$|n\rangle. \quad (28)$$

- The ‘eigen’-ideas carry across unchanged. Hence if a ket obeys the following equation

$$\hat{A}|n\rangle = a_n|n\rangle \quad (29)$$

then it is said to be an **eigenket** of the operator \hat{A} . Similarly, if we conjugate this equation, we obtain

$$\langle n|\hat{A}^\dagger = a_n^*\langle n|; \quad (30)$$

$\langle n|$ is then called an **eigenbra** of \hat{A}^\dagger . (Note the inversion of order on the left-hand side. It’s like taking the transpose of a matrix: $(AB)^T = B^T A^T$.) If \hat{A} is Hermitian, then $\hat{A}^\dagger = \hat{A}$ and a_n is real, and we have

$$\langle n|\hat{A} = a_n\langle n|. \quad (31)$$

- The dot product (25) is written

$$\langle\psi|\phi\rangle. \quad (32)$$

(A bra followed by a ket, you see: it’s a bra-ket = bracket.) Hence the normalisation requirement is written in Dirac notation as

$$\langle\psi|\psi\rangle = 1. \quad (33)$$

The expansion in eigenstates for a general wave function ψ is written as

$$|\psi\rangle = \sum_n c_n |n\rangle = \sum_n \langle n|\psi\rangle |n\rangle \quad (34)$$

as the expansion coefficient for a wave function $\psi(x)$ is $c_n = \int_{-\infty}^{\infty} \phi_n^* \psi dx$.

- A matrix element such as

$$\int \psi^* \hat{A} \phi d\tau \quad (35)$$

would be written

$$\langle \psi | \hat{A} | \phi \rangle, \quad (36)$$

and hence in particular an expectation value like

$$\langle \hat{A} \rangle_\psi = \int \psi^* \hat{A} \psi d\tau \quad (37)$$

would be written

$$\langle \hat{A} \rangle_\psi = \langle \psi | \hat{A} | \psi \rangle. \quad (38)$$

✓ Write in Dirac notation: i) normalization of a 3D wave function: $\iiint_{all\ space} \psi^* \psi d\tau = 1$ with volume element $d\tau$ ii) orthonormality of 1D eigenfunctions $\int_{-\infty}^{\infty} \phi_n^* \phi_m dx = \delta_{nm}$ iii) property of the adjoint operator: $\int_{-\infty}^{\infty} (\hat{A}\psi)^* \phi dx = \int_{-\infty}^{\infty} \psi^* \hat{A}^+ \phi dx$ ⁵
 ✓ If $\langle \psi | \phi \rangle = C$, what is $\langle \phi | \psi \rangle$? ⁶

- It is also possible to define an operator using bras and kets. As seen above, if we put them together in bra-ket order we get a dot (or ‘inner’) product, whose result is just a number. But if we put them together in ket-bra order (an ‘outer’ product), we get an operator:

$$\hat{B} = |\psi\rangle\langle\phi|. \quad (39)$$

This must be an operator, because when it acts on a ket it gives another ket:

$$\hat{B}|\eta\rangle = |\psi\rangle\langle\phi|\eta\rangle = c|\psi\rangle, \quad (40)$$

since the overlap $\langle\phi|\eta\rangle \equiv c$ is just a number. Similarly, when it acts on a bra it gives another bra:

$$\langle\kappa|\hat{B} = \langle\kappa|\psi\rangle\langle\phi| = g\langle\phi|, \quad (41)$$

where again $\langle\kappa|\psi\rangle \equiv g$ is just a number.

✓ Consider the operators $\hat{A} = |0\rangle\langle 1|$ and $\hat{B} = |1\rangle\langle 0|$, where $|0\rangle$ and $|1\rangle$ are orthonormal states. Is $\hat{A}\hat{B}$ equal to $\hat{B}\hat{A}$? ⁷

- The spatial wave function that we have been used to dealing with is written as follows in Dirac notation:

$$\psi(x) = \langle x | \psi \rangle, \quad (42)$$

where $\langle x |$ is an eigenbra of the position operator, i.e.

$$\langle x | \hat{x} = x \langle x |. \quad (43)$$

⁵i) $\langle \psi | \psi \rangle = 1$ ii) $\langle n | m \rangle = \delta_{nm}$ iii) $\langle \hat{A}\psi | \phi \rangle = \langle \psi | \hat{A}^+ | \phi \rangle$

⁶If $\langle \psi | \phi \rangle = C$, then $\langle \phi | \psi \rangle = \langle \psi | \phi \rangle^* = C^*$ (the complex conjugate of C)

⁷ $\hat{A}\hat{B} = |0\rangle\langle 1|1\rangle\langle 0| = |0\rangle\langle 0|$ is not equal to $\hat{B}\hat{A} = |1\rangle\langle 0|0\rangle\langle 1| = |1\rangle\langle 1|$

This makes it clear that the position basis is merely one of many questions we can ask the wave function. We could alternatively ask about momentum, obtaining the momentum-space wave function the same way:

$$\tilde{\psi}(p) = \langle p | \psi \rangle. \quad (44)$$

Here $\langle p |$ is an eigenbra of the momentum operator, i.e.

$$\langle p | \hat{p} = p \langle p |. \quad (45)$$

✓ Is $\langle \psi | x \rangle$ equal to $\int_{-\infty}^{\infty} \psi^*(x) x dx$? ⁸

✓ What is $\langle x | n \rangle$? ⁹

✓ What is $\langle x | p \rangle$? ¹⁰

2.3 The completeness identity; relationships between different wave functions

There is an important identity in Dirac notation, which is called the **completeness relation**. It says that

$$\hat{P} \equiv \sum_n |n\rangle \langle n| = \hat{1}, \quad (46)$$

where $|n\rangle$ are the eigenkets of any Hermitian operator, and $\hat{1}$ is the identity operator.

2.3.1 Proof

First let the operator \hat{P} act on an eigenket. In this case, the result is easily established:

$$\hat{P}|m\rangle = \sum_n |n\rangle \underbrace{\langle n|m\rangle}_{=\delta_{nm}} = |m\rangle. \quad (47)$$

But if \hat{P} is the identity operator for any of the eigenkets, it must be the identity operator for every ket:

$$\hat{P}|\psi\rangle = \hat{P} \sum_m c_m |m\rangle \quad (48)$$

$$= \sum_m c_m \hat{P}|m\rangle \quad (49)$$

$$= \sum_m c_m |m\rangle = |\psi\rangle \quad (50)$$

⁸No, as the ket $|x\rangle$ stands for a position eigenfunction, not “multiply by x ” (the ket just contains a label for the state). Thus $\langle \psi | x \rangle = \int_{-\infty}^{\infty} \psi^*(x') \delta(x' - x) dx' = \psi^*(x) = \langle x | \psi \rangle^*$

⁹The “ n ” in the ket stands for an energy eigenfunction $\phi_n(x)$ (the ket just contains a label for the state). We know $\langle x | \psi \rangle = \psi(x)$ so $\langle x | n \rangle = \int_{-\infty}^{\infty} \delta(x' - x) \phi_n(x') dx' = \phi_n(x)$

¹⁰The “ p ” in the ket stands for a momentum eigenfunction $\phi_p(x)$. We know $\langle x | \psi \rangle = \psi(x)$ so $\langle x | p \rangle = \int_{-\infty}^{\infty} \delta(x' - x) \phi_p(x') dx' = \phi_p(x)$. We know the explicit form of the momentum eigenfunctions: they are plane waves $\phi_p(x) = A e^{ipx/\hbar}$ (you can verify that $\hat{p}\phi_p(x) = -i\hbar \frac{\partial}{\partial x} \phi_p(x) = p\phi_p(x)$).

where we have made use of equation (47) in the third step. Hence \hat{P} in (46) is the identity operator. This so-called **decomposition of the identity** is very useful in manipulating Dirac-notation expressions.

✓ The operator $\hat{P}_n = |n\rangle\langle n|$ is called the **projection operator**. Assuming an orthonormal basis with a general wave function $|\psi\rangle = \sum_n c_n |n\rangle$, what is $\hat{P}_j |\psi\rangle$?¹¹

✓ Show that $\hat{P}_j^2 |\psi\rangle = \hat{P}_j |\psi\rangle$. Why does this result make physical sense? ¹²

2.3.2 Application

As an example of this notation, let us establish the relationship between the position wave function $\psi(x)$ in (42) and the momentum wave function $\tilde{\psi}(p)$ in (44). The momentum wave function is defined by

$$\tilde{\psi}(p) = \langle p | \psi \rangle. \quad (51)$$

We may insert the decomposition of the identity on the right-hand side:

$$\tilde{\psi}(p) = \langle p | \hat{1} | \psi \rangle \quad (52)$$

$$= \langle p | \sum_n |n\rangle\langle n| \psi \rangle \quad (53)$$

$$= \sum_n \langle p | n \rangle \langle n | \psi \rangle. \quad (54)$$

Now let us take the eigenkets $|n\rangle$ in this decomposition to be position eigenkets (which we're at liberty to do). The sum becomes an integral because x can take a continuous range of values:

$$\tilde{\psi}(p) = \int dx \langle p | x \rangle \langle x | \psi \rangle. \quad (55)$$

But we recognise the last factor on the right-hand side as the position wave function $\psi(x)$:

$$\tilde{\psi}(p) = \int dx \langle p | x \rangle \psi(x). \quad (56)$$

Furthermore,

$$\langle p | x \rangle = (\langle x | p \rangle)^*, \quad (57)$$

and $\langle x | p \rangle$ is just the position wave function of a momentum eigenstate, which we know to be $\frac{1}{\sqrt{2\pi\hbar}} e^{ipx/\hbar}$. Hence

$$\tilde{\psi}(p) = \frac{1}{\sqrt{2\pi\hbar}} \int dx e^{-ipx/\hbar} \psi(x). \quad (58)$$

This is, of course, just the usual Fourier-transform relationship between the position and momentum wave functions.

¹¹ $\hat{P}_j |\psi\rangle = |j\rangle\langle j| \sum_n c_n |n\rangle = \sum_n c_n \langle j | n \rangle |j\rangle = \sum_n c_n \delta_{jn} |j\rangle = c_j |j\rangle$. Thus P_j projects $|\psi\rangle$ into state $|j\rangle$.

¹² $\hat{P}_j^2 |\psi\rangle = \hat{P}_j \hat{P}_j |\psi\rangle = \hat{P}_j c_j |j\rangle = c_j \hat{P}_j |j\rangle = c_j |j\rangle\langle j | j \rangle = c_j |j\rangle = \hat{P}_j |\psi\rangle$. The first projector projects $|\psi\rangle$ into state $|j\rangle$, so that the second projector does not change the state and thus has no effect.

3 Matrix representation of quantum mechanics

3.1 Vector representation of states

The vector analogy in section 2.1 gives us a very useful way of representing both quantum states and operators.

Bearing in mind their orthogonality, we will represent the kets of basis states as vector objects in the following way:

$$|\lambda_1\rangle = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \quad |\lambda_2\rangle = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \quad |\lambda_3\rangle = \begin{pmatrix} 0 \\ 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix} \quad \dots \quad |\lambda_n\rangle = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix} \quad (59)$$

with the bras represented through the (conjugate) transpose, for example:

$$\langle\lambda_1| = (1 \ 0 \ 0 \ \dots \ 0) \quad \langle\lambda_2| = (0 \ 1 \ 0 \ \dots \ 0) \quad \dots \quad (60)$$

(you can easily check that we immediately obtain $\langle\lambda_i|\lambda_j\rangle = \delta_{ij}$.) A generalized state is then simply:

$$|\psi\rangle = \sum_{i=1}^n c_i |\lambda_i\rangle = \begin{pmatrix} c_1 \\ c_2 \\ c_3 \\ \vdots \\ c_n \end{pmatrix} \quad \text{or} \quad \langle\psi| = (c_1^* \ c_2^* \ c_3^* \ \dots \ c_n^*). \quad (61)$$

We note that then $\langle\psi|\psi\rangle = \sum_{i=1}^n |c_i|^2 = 1$ as required.

We may also define an *outer product* of two states $|\psi\rangle = \sum_{i=1}^n c_i |\lambda_i\rangle$ and $|\phi\rangle = \sum_{i=1}^n d_i |\lambda_i\rangle$ by

$$|\psi\rangle\langle\phi| = \sum_{r=1}^n \sum_{s=1}^n c_r d_s^* |\lambda_r\rangle\langle\lambda_s| = \begin{pmatrix} c_1 d_1^* & c_1 d_2^* & c_1 d_3^* & \dots & c_1 d_n^* \\ c_2 d_1^* & c_2 d_2^* & c_2 d_3^* & \dots & c_2 d_n^* \\ c_3 d_1^* & c_3 d_2^* & c_3 d_3^* & \dots & c_3 d_n^* \\ \vdots & & & & \ddots \\ c_n d_1^* & c_n d_2^* & c_n d_3^* & \dots & c_n d_n^* \end{pmatrix}. \quad (62)$$

✓ Assume a quantum system with two orthonormal states $|0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $|1\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$.

Write in vector/matrix notation and evaluate:

i) $-i|0\rangle + i|1\rangle$ ii) $-i\langle 0|$ iii) $\langle 0|1\rangle$ iv) $-i|0\rangle\langle 1|$ ¹³

$$\begin{aligned} \text{13i) } -i|0\rangle + i|1\rangle &= -i \begin{pmatrix} 1 \\ 0 \end{pmatrix} + i \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -i \\ i \end{pmatrix} & \text{ii) } -i\langle 0| &= -i (1 \ 0) = (-i \ 0) \\ \text{iii) } \langle 0|1\rangle &= (1 \ 0) \begin{pmatrix} 0 \\ 1 \end{pmatrix} = 0 & \text{iv) } -i|0\rangle\langle 1| &= -i \begin{pmatrix} 1 \\ 0 \end{pmatrix} (0 \ 1) = -i \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & -i \\ 0 & 0 \end{pmatrix} \end{aligned}$$

3.2 Matrix representation of operators: Special case

Now let's consider the Hermitian operator \hat{A} , which acts on the n dimensional Hilbert space we have just described, and has the n orthogonal eigenstates $|\lambda_i\rangle$. Each has a corresponding real eigenvalue λ_i . In other words:

$$\hat{A}|\lambda_i\rangle = \lambda_i|\lambda_i\rangle \quad (63)$$

for all $i \in \{1 \dots n\}$. Thus, in this special case the $|\lambda_i\rangle$ are the eigenkets of \hat{A} .

In fact, we might just as well write our operator in the following form

$$\hat{A} = \sum_{r=1}^n \lambda_r |\lambda_r\rangle \langle \lambda_r| \quad (64)$$

since by the orthogonality relation $\langle \lambda_r | \lambda_i \rangle = \delta_{ri}$ it is clear that substituting (64) into the LHS of (63) that we obtain the RHS of (63) for all possible i .¹⁴

It is now straightforward to represent our operator in a matrix form, in the basis $|\lambda_i\rangle$. We use the vector representations of $|\lambda_i\rangle$ and $\langle \lambda_i|$ and form the outer products. For example, the first term in the sum is

$$\lambda_1 |\lambda_1\rangle \langle \lambda_1| = \lambda_1 \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \end{pmatrix} = \begin{pmatrix} \lambda_1 & 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & \dots & 0 \\ \vdots & & & \ddots & 0 \\ 0 & 0 & 0 & \dots & 0 \end{pmatrix}. \quad (65)$$

Each of the n terms in the sum then gives a different term on the diagonal, and the sum of all the terms gives is

$$\mathbf{A}_\lambda = \begin{pmatrix} \lambda_1 & 0 & 0 & \dots & 0 \\ 0 & \lambda_2 & 0 & \dots & 0 \\ 0 & 0 & \lambda_3 & \dots & 0 \\ \vdots & & & \ddots & 0 \\ 0 & 0 & 0 & \dots & \lambda_n \end{pmatrix}. \quad (66)$$

We have named the matrix \mathbf{A}_λ since it is the representation of the operator A with respect to the basis $|\lambda\rangle$. The matrix is obviously diagonal in this basis - its eigenvalues and eigenvectors are precisely those corresponding to our operator \hat{A} , as we expected by construction. Thus we see: **An operator is always diagonal in its own basis. The diagonal elements are the eigenvalues of the operator (i.e., the possible measurement outcomes of the observable that the operator represents) ordered in the same way as the corresponding eigenvectors. Eigenvectors are unit vectors in their own basis.**

✓ What are the possible energy measurement outcomes assuming a Hamiltonian for a 2D orthonormal system with matrix $\mathbf{H} = \begin{pmatrix} -\varepsilon & 0 \\ 0 & \varepsilon \end{pmatrix}$ ¹⁵

¹⁴Of course, we can use any index for the summation; we have chosen r in this case since to prove the relation (64) one needs to substitute into (63), and it is then important to use a different index to i to avoid confusion.

¹⁵The possible energy measurement outcomes are the eigenvalues of \mathbf{H} , so $-\varepsilon$ and ε .

3.3 Matrix representation of operators: General case

Often in quantum mechanical problems it will be most convenient to write an operator \hat{A} down in an orthonormal basis that is not necessarily its eigenbasis. Consider a vector space for simplicity that only has two dimensions with orthonormal basis $|1\rangle$ and $|2\rangle$. Examples would be a spin 1/2 particle, or an atom with a ground state and only one excited state. Note that now $|1\rangle$ and $|2\rangle$ are *not* necessarily eigenkets of \hat{A} .

An operator is then represented by a 2×2 matrix

$$\mathbf{A} = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}. \quad (67)$$

If we apply \hat{A} to the first basis vector, we obtain

$$\hat{A}|1\rangle = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} A_{11} \\ A_{21} \end{pmatrix}. \quad (68)$$

Thus, we can extract one of the elements of the matrix using

$$\langle 1|\hat{A}|1\rangle = \begin{pmatrix} 1 & 0 \end{pmatrix} \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = A_{11} \quad (69)$$

Similarly,

$$\langle 1|\hat{A}|2\rangle = \begin{pmatrix} 1 & 0 \end{pmatrix} \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = A_{12}, \quad (70)$$

$$\langle 2|\hat{A}|1\rangle = \begin{pmatrix} 0 & 1 \end{pmatrix} \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = A_{21}, \quad (71)$$

and

$$\langle 2|\hat{A}|2\rangle = \begin{pmatrix} 0 & 1 \end{pmatrix} \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = A_{22}, \quad (72)$$

Thus, the matrix

$$A = \begin{pmatrix} \langle 1|\hat{A}|1\rangle & \langle 1|\hat{A}|2\rangle \\ \langle 2|\hat{A}|1\rangle & \langle 2|\hat{A}|2\rangle \end{pmatrix}. \quad (73)$$

We call this inner product of a bra, operator and ket such as $\langle 1|\hat{A}|2\rangle$ a **matrix element**. More generally, in a vector space with more than two dimensions, the matrix representation of an operator is

$$A = \begin{pmatrix} A_{11} & A_{12} & A_{13} & \dots \\ A_{21} & A_{22} & A_{23} & \dots \\ A_{31} & A_{32} & A_{33} & \dots \\ \dots & \dots & \dots & \ddots \end{pmatrix} \quad (74)$$

where the matrix elements $A_{ij} = \langle i|\hat{A}|j\rangle$ and the basis states are as in (59). **Hermitian operators are represented by Hermitian matrices.** If an operator is Hermitian, then $\langle i|\hat{A}|j\rangle = \langle \hat{A}i|j\rangle = \langle j|\hat{A}|i\rangle^*$. Thus, the matrix element $A_{ij} = A_{ji}^*$. This implies that the matrix \mathbf{A} equals the complex conjugate of its transpose, i.e. it is Hermitian.

3.4 Operators as a sum of outer products

Using our basis vectors (59), we can see that the outer product of $|i\rangle$ and $\langle j|$ gives a matrix that is zero everywhere excepting element ij which equals 1. Using again the example of a two-dimensional space with orthonormal basis $|1\rangle$ and $|2\rangle$ we have

$$|1\rangle\langle 1| = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad (75)$$

$$|1\rangle\langle 2| = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad (76)$$

$$|2\rangle\langle 1| = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \quad (77)$$

$$|2\rangle\langle 2| = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \begin{pmatrix} 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}. \quad (78)$$

Thus, the matrix

$$\mathbf{A} = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} = A_{11}|1\rangle\langle 1| + A_{12}|1\rangle\langle 2| + A_{21}|2\rangle\langle 1| + A_{22}|2\rangle\langle 2|, \quad (79)$$

and inserting the expression for the matrix element,

$$\hat{A} = \sum_{i,j=1}^2 A_{ij}|i\rangle\langle j| = \sum_{i,j=1}^2 |i\rangle\langle i|\hat{A}|j\rangle\langle j|. \quad (80)$$

We can generalize these results to n dimensions. **Every operator \hat{A} can be written as a sum of outer products:**

$$\hat{A} = \sum_{i,j=1}^n A_{ij}|i\rangle\langle j| = \sum_{i,j=1}^n |i\rangle\langle i|\hat{A}|j\rangle\langle j|. \quad (81)$$

3.5 Diagonalization of operators

Assume we have given the matrix representation of an operator \mathbf{A} and want to find the possible results of a measurement of the corresponding observable. These possible results are the eigenvalues of the operator, and the eigenvectors are the corresponding quantum states. In order to find the eigenvectors and eigenvalues of the operator we must diagonalize the matrix representing the operator. To do this, we use the usual procedure of solving the characteristic equation by setting the determinant

$$|\mathbf{A} - \lambda\mathbf{I}| = 0. \quad (82)$$

In this expression, \mathbf{I} is the identity matrix. This gives an equation in the parameter λ whose roots are the eigenvalues. Once these eigenvalues are found, they are individually inserted into the eigenvalue equation to find the eigenvectors. The eigenvectors need to be normalized given that they represent quantum states.

4 First-order time-independent perturbation theory

In this lecture, we begin our investigations of one of the major topics of this course: **perturbation theory**. The basic idea is to construct approximate solutions to problems that are not exactly solvable, using our stock of exactly solvable problems (IDSW, SHO, and hydrogen atom) as a starting point.

4.1 Perturbations

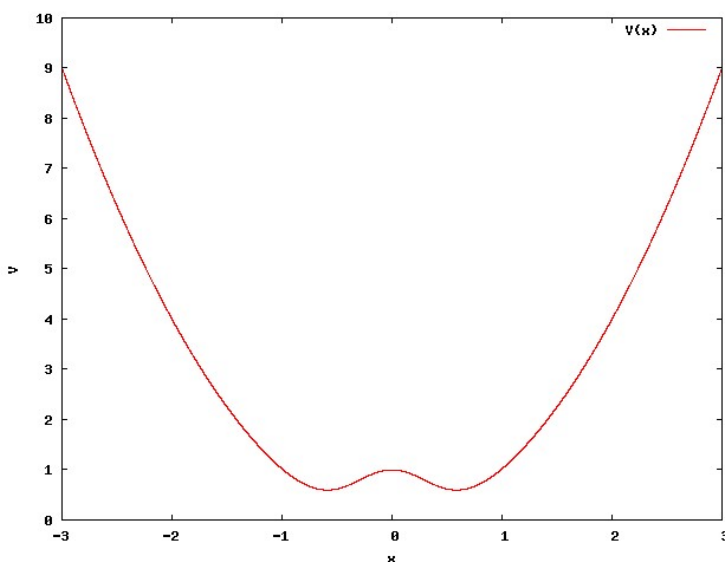


Figure 1: The simple harmonic oscillator potential with an added central dimple.

Consider the potential shown in Fig. 1. Its analytic form is

$$V(x) = \frac{1}{2}m\omega^2 x^2 + Ce^{-\alpha x^2}, \quad (83)$$

where m and ω have their usual meanings, and C and α are constants. The first term of (83) is just the usual harmonic oscillator potential, while the second is a **perturbation**: an added piece that stops the problem from being exactly solvable. However, in this case the perturbation is rather small¹⁶, so we might think that the eigenvalues and eigenkets for a particle in the potential (83) will not be too different from those of a particle in a simple harmonic oscillator potential. Perturbation theory is based on this idea: that the full eigenkets and eigenvalues will not be vastly different from the unperturbed ones.

This specific example is an instance of a general kind of problem in quantum mechanics. We have some Hamiltonian, \hat{H}_0 , for which we know all the eigenvalues and eigenkets. This is called the **unperturbed Hamiltonian**. Then a bit is added to it, changing the shape of the problem, but only slightly. The resulting Hamiltonian is written

$$\hat{H} = \hat{H}_0 + \lambda \hat{V}; \quad (84)$$

¹⁶We'll come back later to exactly what 'small' means in this context.

here \hat{H}_0 is the unperturbed Hamiltonian, while \hat{V} is the perturbation. λ is a parameter that makes the smallness of the perturbation explicit; it is mainly useful for keeping track of things during the calculation, and is set to 1 at the end. If the perturbation \hat{V} does not depend on time, we are doing **time-independent perturbation theory**; if it does depend on time, we are doing **time-dependent perturbation theory**. We shall consider both types in this course, but we shall do the time-independent case first.

In the example above, the unperturbed Hamiltonian is

$$\hat{H}_0 = \frac{\hat{p}^2}{2m} + \frac{1}{2}m\omega^2 x^2, \quad (85)$$

while the perturbation is

$$\hat{V} = Ce^{-\alpha x^2}. \quad (86)$$

4.2 Time-independent perturbation theory: basic approach

The full Hamiltonian, \hat{H} , is a Hermitian operator. Hence it must have a set of eigenkets $|n\rangle$ and eigenvalues E_n , which satisfy the **full Schrödinger equation**

$$\hat{H}|n\rangle = E_n|n\rangle. \quad (87)$$

It is these eigenkets and eigenvalues that we are trying to find. The method is as follows:

- Our starting point is that we know the eigenkets $|n^{(0)}\rangle$ and eigenvalues $E_n^{(0)}$ of the unperturbed Hamiltonian \hat{H}_0 ; these satisfy the **unperturbed Schrödinger equation**

$$\hat{H}_0|n^{(0)}\rangle = E_n^{(0)}|n^{(0)}\rangle. \quad (88)$$

We shall use these unperturbed eigenkets as a basis. Since any ket can be decomposed in such a basis, it must be possible to write the full eigenkets $|n\rangle$ in terms of the unperturbed ones $|n^{(0)}\rangle$:

$$|n\rangle = \sum_m c_{nm}(\lambda)|m^{(0)}\rangle. \quad (89)$$

The coefficients $c_{nm}(\lambda)$ are called **mixing coefficients**. Naturally they depend on λ , the strength of the perturbation.

- Since λ is a small parameter, we expand the eigenenergies E_n and the mixing coefficients c_{nm} in terms of powers of λ :

$$E_n = E_n^{(0)} + \lambda E_n^{(1)} + \lambda^2 E_n^{(2)} + \dots \quad (90)$$

$$c_{nm} = \delta_{nm} + \lambda c_{nm}^{(1)} + \lambda^2 c_{nm}^{(2)} + \dots \quad (91)$$

The first term on the right-hand side of (90) is fixed by the requirement that when $\lambda = 0$ the full eigenenergies revert to the unperturbed ones. The same requirement for the eigenkets gives the first term on the right-hand side of (91), which guarantees that

$$\lim_{\lambda \rightarrow 0} |n\rangle = |n^{(0)}\rangle. \quad (92)$$

- The procedure of perturbation theory is to substitute (90) and (91) into (87), and thus to determine $E_n^{(a)}$ (the corrections to the unperturbed eigenenergies) and $c_{nm}^{(a)}$ (the corrections to the unperturbed mixing coefficients). Clearly there are an infinite number of such corrections, so the exact result will never be obtained; but in the usual spirit of Maclaurin expansions, we shall keep only as many terms as we need to reach our desired level of accuracy in any given problem.

✓ Assume n starts with 1, and set $\lambda = 1$ (as we will do at the end of the derivation). Write using the notation given: i) the unperturbed energy of the first-excited state; ii) the first-order energy correction of the first-excited state; iii) the perturbed energy of the first-excited state to second order; iv) the second-order energy correction of the ground state ¹⁷

✓ Using the same assumptions, write using the notation given: i) the unperturbed first-excited state energy eigenfunction; ii) the first-order correction to the first-excited state eigenfunction; iii) the perturbed first excited state eigenfunction to second order ¹⁸

4.3 Time-independent perturbation theory: derivation of the corrections

Let us, then, apply the procedure outlined above. Substituting (90) and (91) into (87), we obtain

$$\hat{H}|n\rangle = E_n|n\rangle, \quad (93)$$

$$\therefore (\hat{H}_0 + \lambda\hat{V})|n\rangle = E_n|n\rangle, \quad (94)$$

$$\therefore (\hat{H}_0 + \lambda\hat{V}) \sum_m c_{nm}|m^{(0)}\rangle = E_n \sum_m c_{nm}|m^{(0)}\rangle, \quad (95)$$

$$\therefore \sum_m c_{nm} (\hat{H}_0 + \lambda\hat{V})|m^{(0)}\rangle = \sum_m c_{nm} E_n|m^{(0)}\rangle. \quad (96)$$

Using (88), we can replace \hat{H}_0 by its eigenvalue $E_m^{(0)}$ on the left-hand side:

$$\sum_m c_{nm} E_m^{(0)}|m^{(0)}\rangle + \lambda \sum_m c_{nm} \hat{V}|m^{(0)}\rangle = \sum_m c_{nm} E_n|m^{(0)}\rangle. \quad (97)$$

This is the key equation, from which everything else is derived.

¹⁷ i) $E_2^{(0)}$ ii) $E_2^{(1)}$ iii) $E_2 \approx E_2^{(0)} + E_2^{(1)} + E_2^{(2)}$ iv) $E_1^{(2)}$

¹⁸i) $|2^{(0)}\rangle$ or $\phi_2^{(0)}$ ii) $|2^{(1)}\rangle = \sum_{k \neq 2} c_{2k}^{(1)}|k^{(0)}\rangle$ or $\phi_2^{(1)} = \sum_{k \neq 2} c_{2k}^{(1)}\phi_k^{(0)}$ iii) $|2\rangle \approx |2^{(0)}\rangle + |2^{(1)}\rangle + |2^{(2)}\rangle = |2^{(0)}\rangle + \sum_{k \neq 2} c_{2k}^{(1)}|k^{(0)}\rangle + \sum_{k \neq 2} c_{2k}^{(2)}|k^{(0)}\rangle$ or $\phi_2 \approx \phi_2^{(0)} + \phi_2^{(1)} + \phi_2^{(2)} = \phi_2^{(0)} + \sum_{k \neq 2} c_{2k}^{(1)}\phi_k^{(0)} + \sum_{k \neq 2} c_{2k}^{(2)}\phi_k^{(0)}$.

The summations already assume (as we will find) that the coefficients c_{nk} from first order onwards are only non-zero between different states with $n \neq k$.

4.3.1 Perturbative expansion in λ

Now we insert the perturbative expansions (90) and (91) into (97):

$$\begin{aligned} \sum_m (\delta_{nm} + \lambda c_{nm}^{(1)} + \lambda^2 c_{nm}^{(2)} + \dots) E_m^{(0)} |m^{(0)}\rangle + \lambda \sum_m (\delta_{nm} + \lambda c_{nm}^{(1)} + \lambda^2 c_{nm}^{(2)} + \dots) \hat{V} |m^{(0)}\rangle \\ = \sum_m (\delta_{nm} + \lambda c_{nm}^{(1)} + \lambda^2 c_{nm}^{(2)} + \dots) (E_n^{(0)} + \lambda E_n^{(1)} + \lambda^2 E_n^{(2)} + \dots) |m^{(0)}\rangle. \end{aligned} \quad (98)$$

4.3.2 Terms of order λ^0

Since λ is a small parameter, the coefficients of each power of λ must agree separately between the left- and right-hand sides of (98). Let us first read off the coefficients of λ^0 , i.e. those terms with no powers of λ in at all:

$$\sum_m \delta_{nm} E_m^{(0)} |m^{(0)}\rangle = \sum_m \delta_{nm} E_n^{(0)} |m^{(0)}\rangle, \quad (99)$$

$$\therefore E_n^{(0)} |n^{(0)}\rangle = E_n^{(0)} |n^{(0)}\rangle, \quad (100)$$

which is trivially satisfied.

4.3.3 Terms of order λ^1 : first-order perturbation theory

Reading off the coefficients of λ^1 from (98), we find

$$\sum_m c_{nm}^{(1)} E_m^{(0)} |m^{(0)}\rangle + \sum_m \delta_{nm} \hat{V} |m^{(0)}\rangle = \sum_m (\delta_{nm} E_n^{(1)} + c_{nm}^{(1)} E_n^{(0)}) |m^{(0)}\rangle. \quad (101)$$

Eliminating the sums from the terms containing delta-symbols, this becomes

$$\sum_m c_{nm}^{(1)} E_m^{(0)} |m^{(0)}\rangle + \hat{V} |n^{(0)}\rangle = E_n^{(1)} |n^{(0)}\rangle + \sum_m c_{nm}^{(1)} E_n^{(0)} |m^{(0)}\rangle. \quad (102)$$

To extract the unknowns $c_{nm}^{(1)}$ and $E_n^{(1)}$ one by one, we multiply from the left by $\langle k^{(0)}|$:

$$\sum_m c_{nm}^{(1)} E_m^{(0)} \langle k^{(0)} | m^{(0)} \rangle + \langle k^{(0)} | \hat{V} | n^{(0)} \rangle = E_n^{(1)} \langle k^{(0)} | n^{(0)} \rangle + \sum_m c_{nm}^{(1)} E_n^{(0)} \langle k^{(0)} | m^{(0)} \rangle. \quad (103)$$

Using the orthonormality of the unperturbed eigenkets,

$$\langle k^{(0)} | m^{(0)} \rangle = \delta_{km}, \quad (104)$$

we find

$$\sum_m c_{nm}^{(1)} E_m^{(0)} \delta_{km} + \langle k^{(0)} | \hat{V} | n^{(0)} \rangle = E_n^{(1)} \delta_{kn} + \sum_m c_{nm}^{(1)} E_n^{(0)} \delta_{km}, \quad (105)$$

and eliminating the sums from the terms containing delta-symbols yields

$$c_{nk}^{(1)} E_k^{(0)} + \langle k^{(0)} | \hat{V} | n^{(0)} \rangle = E_n^{(1)} \delta_{kn} + c_{nk}^{(1)} E_n^{(0)}. \quad (106)$$

We now distinguish two cases:

- (i) $k = n$. In this case, the first term on the left-hand side of (106) cancels the final term on the right-hand side, leaving

$$E_n^{(1)} = \langle n^{(0)} | \hat{V} | n^{(0)} \rangle. \quad (107)$$

This is the expression for the **first-order correction to the eigenenergies**. Note that it can be calculated provided that we know the unperturbed eigenkets (which we do) and the operator corresponding to the perturbation (which we do).

- (ii) $k \neq n$. In this case, the first term on the right-hand side of (106) vanishes because the delta-symbol is zero for $k \neq n$. This leaves

$$c_{nk}^{(1)} E_k^{(0)} + \langle k^{(0)} | \hat{V} | n^{(0)} \rangle = c_{nk}^{(1)} E_n^{(0)}, \quad (108)$$

which may be rearranged to read

$$c_{nk}^{(1)} = \frac{\langle k^{(0)} | \hat{V} | n^{(0)} \rangle}{E_n^{(0)} - E_k^{(0)}}. \quad (109)$$

This is the expression for the **first-order correction to the mixing coefficients**.

You should learn the expressions (107) and (109): we shall be using them extensively.

✓ i) What is $\langle 3^{(0)} | \hat{V} | 3^{(0)} \rangle$? ii) What is $\frac{\langle 2^{(0)} | \hat{V} | 3^{(0)} \rangle}{E_3^{(0)} - E_2^{(0)}}$? ¹⁹

✓ What is $c_{nn}^{(0)}$, $c_{nn}^{(1)}$, $c_{nn}^{(2)}$, etc? ²⁰

✓ Consider the constant perturbation $\hat{V} = V_0$. Determine $E_n^{(1)}$ and $c_{nk}^{(1)}$. Determine the perturbed energy eigenfunction $|n\rangle$ to first order. ²¹

¹⁹i) $E_3^{(1)}$, the first-order energy correction of the perturbed state ϕ_3 . ii) the mixing coefficient $c_{32}^{(1)}$ giving the first-order contribution of the unperturbed $\phi_2^{(0)}$ to the perturbed ϕ_3 .

²⁰ $c_{nn}^{(0)} = 1$ and $c_{nn}^{(k)} = 0$ for all $k > 0$. The mixing coefficients only act between different states.

²¹ $E_n^{(1)} = \langle n^{(0)} | \hat{V} | n^{(0)} \rangle = V_0 \langle n^{(0)} | n^{(0)} \rangle = V_0$ and $c_{nk}^{(1)} = 0$ as $\langle k^{(0)} | \hat{V} | n^{(0)} \rangle = V_0 \langle k^{(0)} | n^{(0)} \rangle = 0$ for $k \neq n$. Thus, in first-order approximation $|n\rangle = |n^{(0)}\rangle + \sum_{k \neq n} c_{nk}^{(1)} |k^{(0)}\rangle = |n^{(0)}\rangle$.

5 First-order perturbation theory: example

In this lecture, we shall use (107) and (109) to calculate the first-order corrections to the eigenenergies and mixing coefficients for a specific example.

5.1 Unperturbed Hamiltonian

Let us take as the unperturbed Hamiltonian the infinitely deep square well in one dimension, i.e.

$$\hat{H} = \frac{\hat{p}^2}{2m} + V_{\text{IDSW1}}(x), \quad (110)$$

where

$$V_{\text{IDSW1}}(x) = \begin{cases} 0 & 0 < x < L, \\ +\infty & \text{otherwise.} \end{cases} \quad (111)$$

We know that the exact eigenfunctions of this problem are

$$\phi_n^{(0)} = \begin{cases} \sqrt{\frac{2}{L}} \sin\left(\frac{n\pi x}{L}\right) & 0 < x < L, \\ 0 & \text{otherwise,} \end{cases} \quad (112)$$

with eigenenergies

$$E_n^{(0)} = \frac{\pi^2 \hbar^2}{2mL^2} n^2, \quad (113)$$

where $n = 1, 2, 3, \dots$

5.2 Perturbation

Let us add to this potential a small perturbation of the form

$$V(x) = \begin{cases} Cx & 0 < x < L, \\ 0 & \text{otherwise,} \end{cases} \quad (114)$$

where C is a small positive constant. A graph of the perturbed potential ($V_{\text{IDSW1}} + V$) is shown in Fig. 2. Note that this potential corresponds to a classical force given by

$$F = -\frac{dV}{dx} = -C, \quad (115)$$

i.e. a small constant force pushing the particle to the left. In physical reality this could be provided by (for example) an electric field if the particle were charged, in which case we would have $C = qE$.

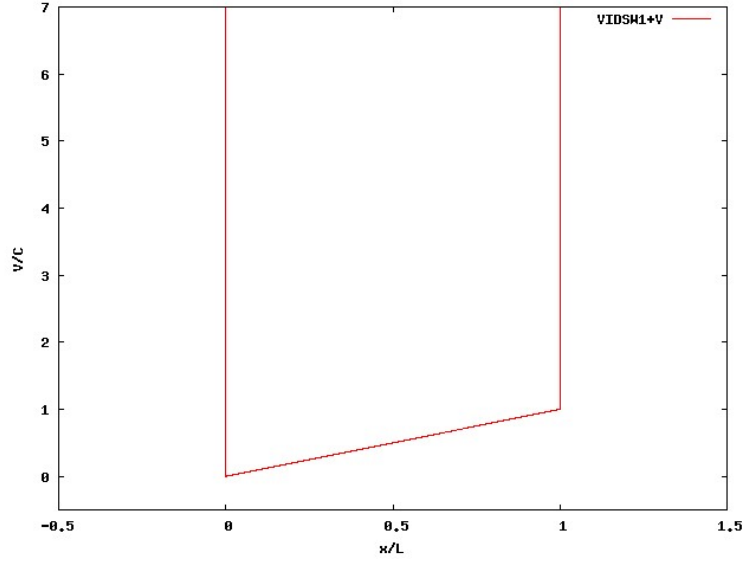


Figure 2: A perturbed version of the infinitely deep square well potential.

5.3 First-order corrections to the eigenenergies

Now let us calculate the first-order corrections to the eigenenergies (113), using (107):²²

$$E_n^{(1)} = \langle n^{(0)} | \hat{V} | n^{(0)} \rangle \quad (116)$$

$$= \int_{-\infty}^{\infty} (\phi_n^{(0)})^* V(x) \phi_n^{(0)} dx \quad (117)$$

$$= \frac{2C}{L} \int_0^L x \sin^2 \left(\frac{n\pi x}{L} \right) dx. \quad (118)$$

To deal with the squared sine, we use the trigonometric identity

$$\sin^2 \theta = \frac{1}{2} (1 - \cos 2\theta), \quad (119)$$

which yields

$$E_n^{(1)} = \frac{C}{L} \int_0^L x \left(1 - \cos \left(\frac{2\pi n x}{L} \right) \right) dx \quad (120)$$

$$= \frac{C}{L} \int_0^L x dx - \frac{C}{L} \int_0^L x \cos \left(\frac{2\pi n x}{L} \right) dx \quad (121)$$

²²Note that, although we have been developing perturbation theory in Dirac notation, for specific problems it is often easier to work in the old position-basis notation, since that's the basis in which we know the wave functions.

$$= \frac{C}{L} \left[\frac{x^2}{2} \right]_0^L - \frac{C}{L} \int_0^L x \cos \left(\frac{2\pi n x}{L} \right) dx \quad (122)$$

$$= \frac{CL}{2} - \frac{C}{L} \int_0^L x \cos \left(\frac{2\pi n x}{L} \right) dx. \quad (123)$$

The remaining integral can be carried out by parts, with $u = x$ and $\frac{dv}{dx} = \cos \left(\frac{2\pi n x}{L} \right)$. Hence

$$\int_0^L x \cos \left(\frac{2\pi n x}{L} \right) dx = \left[x \frac{L}{2\pi n} \sin \left(\frac{2\pi n x}{L} \right) \right]_0^L - \frac{L}{2\pi n} \int_0^L \sin \left(\frac{2\pi n x}{L} \right) dx. \quad (124)$$

However, the sine in the first term vanishes at both $x = 0$ and $x = L$.²³ Hence

$$\int_0^L x \cos \left(\frac{2\pi n x}{L} \right) dx = -\frac{L}{2\pi n} \int_0^L \sin \left(\frac{2\pi n x}{L} \right) dx \quad (125)$$

$$= -\frac{L}{2\pi n} \left[-\frac{L}{2\pi n} \cos \left(\frac{2\pi n x}{L} \right) \right]_0^L \quad (126)$$

$$= \frac{L^2}{4\pi^2 n^2} (\cos 2\pi n - \cos 0). \quad (127)$$

But $\cos 2\pi n = 1$ and $\cos 0 = 1$, so this vanishes. Hence the first-order corrections to the eigenenergies are just

$$E_n^{(1)} = \frac{CL}{2}. \quad (128)$$

✓To first order, what is the perturbed second-excited state energy for this sytem? ²⁴

5.4 First-order corrections to the mixing coefficients

Its effect on the eigenenergies may be trivial, but the perturbation will nonetheless alter the eigenfunctions, even at first order. The main ingredients of the corrections to the mixing coefficients are the matrix elements

$$\langle k^{(0)} | \hat{V} | n^{(0)} \rangle = \frac{2C}{L} \int_0^L x \sin \left(\frac{k\pi x}{L} \right) \sin \left(\frac{n\pi x}{L} \right) dx. \quad (129)$$

This product of sines can likewise be dealt with using a trig. identity:

$$2 \sin A \sin B = \cos(A - B) - \cos(A + B). \quad (130)$$

²³This is obvious, in fact, since that sine is just the unperturbed eigenfunction of state $\phi_{2n}^{(0)}$, and we know that all of the unperturbed eigenfunctions vanish at the edges of the well.

²⁴ $E_3 = E_3^{(0)} + E_3^{(1)} = \frac{9\pi^2 \hbar^2}{2mL^2} + \frac{CL}{2}$

Hence

$$\langle k^{(0)} | \hat{V} | n^{(0)} \rangle = \frac{C}{L} \int_0^L x \cos\left(\frac{(k-n)\pi x}{L}\right) dx - \frac{C}{L} \int_0^L x \cos\left(\frac{(k+n)\pi x}{L}\right) dx. \quad (131)$$

Now, both of the integrals on the right-hand side are of the form

$$\int_0^L x \cos\left(\frac{q\pi x}{L}\right) dx = \underbrace{\left[\frac{L}{q\pi} x \sin\left(\frac{q\pi x}{L}\right) \right]_0^L}_{=0} - \frac{L}{q\pi} \int_0^L \sin\left(\frac{q\pi x}{L}\right) dx \quad (132)$$

$$= -\frac{L}{q\pi} \left[-\frac{L}{q\pi} \cos\left(\frac{q\pi x}{L}\right) \right]_0^L \quad (133)$$

$$= \frac{L^2}{q^2\pi^2} (\cos q\pi - \cos 0) \quad (134)$$

$$= \begin{cases} -\frac{2L^2}{q^2\pi^2} & q \text{ odd,} \\ 0 & q \text{ even.} \end{cases} \quad (135)$$

If $k+n$ is even, then $k-n$ is also even, so either both of the terms in (131) are zero or both are non-zero. Hence

$$\langle k^{(0)} | \hat{V} | n^{(0)} \rangle = \begin{cases} -\frac{2CL}{(k-n)^2\pi^2} + \frac{2CL}{(k+n)^2\pi^2} & k+n \text{ odd,} \\ 0 & k+n \text{ even.} \end{cases} \quad (136)$$

Inserting this into (109), we obtain

$$c_{nk}^{(1)} = \frac{\langle k^{(0)} | \hat{V} | n^{(0)} \rangle}{E_n^{(0)} - E_k^{(0)}} \quad (137)$$

$$= \begin{cases} \frac{1}{\frac{\pi^2\hbar^2}{2mL^2}n^2 - \frac{\pi^2\hbar^2}{2mL^2}k^2} \left(-\frac{2CL}{(k-n)^2\pi^2} + \frac{2CL}{(k+n)^2\pi^2} \right) & k+n \text{ odd} \\ 0 & k+n \text{ even} \end{cases} \quad (138)$$

$$= \begin{cases} \frac{4mCL^3}{\pi^4\hbar^2} \frac{1}{n^2 - k^2} \left(-\frac{1}{(k-n)^2} + \frac{1}{(k+n)^2} \right) & k+n \text{ odd} \\ 0 & k+n \text{ even.} \end{cases} \quad (139)$$

✓For the perturbed first-excited state eigenfunction for this system, which first-order mixing coefficients will be non-zero? ²⁵

✓To first order, write down an expression for the perturbed first-excited state eigenfunction ϕ_2 for this system. ²⁶

²⁵For $c_{2k}^{(1)}$, we have $n = 2$, so that k must be odd so that $k+n$ is odd. Thus, $c_{21}^{(1)}$, $c_{23}^{(1)}$, $c_{25}^{(1)}$, ... are non-zero.

²⁶ $\phi_2 = \phi_2^{(0)} + \phi_2^{(1)} = \phi_2^{(0)} + \sum_{k \text{ odd}} c_{2k}^{(1)} \phi_k^{(0)} = \sqrt{\frac{2}{L}} \sin\left(\frac{2\pi x}{L}\right) + \sum_{k \text{ odd}} c_{2k}^{(1)} \sqrt{\frac{2}{L}} \sin\left(\frac{k\pi x}{L}\right)$

6 Perturbation theory: interpretation

Last time we derived the first-order energy corrections

$$E_n^{(1)} = \frac{CL}{2} \quad (140)$$

and first-order corrections to the mixing coefficients

$$c_{nk}^{(1)} = \begin{cases} \frac{4mCL^3}{\pi^4 \hbar^2} \frac{1}{n^2 - k^2} \left(-\frac{1}{(k-n)^2} + \frac{1}{(k+n)^2} \right) & k+n \text{ odd} \\ 0 & k+n \text{ even} \end{cases} \quad (141)$$

when a small perturbation $V(x) = Cx$ is added to the IDSW1 Hamiltonian. Let us try to interpret these.

6.1 Interpretation

6.1.1 Energy corrections

Notes:

- The energy correction (140) is independent of n , i.e. the first-order energy correction is the same for each eigenstate.
- The energy correction is just the average value of $V(x)$ across the well.

Why is this? It's easy to see if we rewrite the perturbation as

$$V(x) = \begin{cases} \underbrace{\frac{CL}{2}}_{V_1} + \underbrace{C\left(x - \frac{L}{2}\right)}_{V_2} & 0 < x < L \\ 0 & \text{otherwise.} \end{cases} \quad (142)$$

The first term is just a constant. This amounts to a redefinition of the zero of potential energy, and will therefore just increase all the eigenenergies by $\frac{CL}{2}$ without altering the eigenfunctions. The second term is odd under a reflection about the centre of the well. However, all the probability densities $(\phi_n^{(0)})^2$ are even under such a reflection. So when we form the integral

$$\langle n^{(0)} | \hat{V}_2 | n^{(0)} \rangle = \int_{-\infty}^{\infty} \underbrace{V_2(x)}_{\text{odd}} \underbrace{(\phi_n^{(0)})^2}_{\text{even}} dx, \quad (143)$$

the integrand is odd under reflection about the centre of the well, and thus the integral vanishes. As you gain more experience in perturbation theory, you will learn to recognise such things at a glance, thereby avoiding a lot of tedious calculation!

✓ A function f is odd with respect to a point a if $f(a+x) = -f(a-x)$ and even if $f(a+x) = +f(a-x)$. Show that i) “odd \times even = odd” and ii) “odd \times odd = even”.²⁷
 ✓ Is $V(x) = Cx$ itself even or odd with respect to the middle of the infinite well, assuming the walls are at $x = 0$ and $x = L$?²⁸

6.1.2 Mixing coefficient corrections

Now let us turn to the mixing coefficients (141). There are two things to note about them:

- Half of them are zero. Specifically, an unperturbed state with an odd quantum number (e.g. $n = 1$) only gets mixed with unperturbed states with even quantum numbers ($k = 2, 4, 6, \dots$). This is normally expressed by saying that the mixing between $n = 1$ and $k = 3$ (for example) is forbidden — or, since this rule applies only at first order, **first forbidden**. The rules telling us which states can mix with which others are called **selection rules**.
- The magnitude of $c_{nk}^{(1)}$ decreases as n and k get further apart. Not only does it decrease: it decreases like $1/k^5$ (prove this if you can!), which is quite rapid. So it should be possible to get a reasonable estimate of the perturbed eigenfunctions by considering only those mixing coefficients for which n and k are close to each other.

The obvious question is: how do the selection rules discussed in the first bullet point come about? Again, the above decomposition of the potential (142) will be useful here.

Let us look again at the expression for $\langle k^{(0)} | \hat{V} | n^{(0)} \rangle$ — the main ingredient of $c_{nk}^{(1)}$ — in the (old) position representation:

$$\langle k^{(0)} | \hat{V} | n^{(0)} \rangle = \int_0^L [\phi_k^{(0)}]^* V \phi_n^{(0)} dx \quad (144)$$

$$= \int_0^L [\phi_k^{(0)}]^* (V_1 + V_2) \phi_n^{(0)} dx \quad (145)$$

$$= \underbrace{\frac{CL}{2} \int_0^L [\phi_k^{(0)}]^* \phi_n^{(0)} dx}_{=0 \text{ (orthogonality)}} + \int_0^L [\phi_k^{(0)}]^* V_2 \phi_n^{(0)} dx \quad (146)$$

$$= \int_0^L [\phi_k^{(0)}]^* V_2 \phi_n^{(0)} dx. \quad (147)$$

But the three terms in the integrand here all have definite symmetry under reflection about the centre of the well. V_2 is odd (see above). The wave functions are (confusingly) even if n is odd, and odd if n is even. So there are two cases:

²⁷i) If f is odd and g is even then $f(a+x)g(a+x) = -f(a-x)g(a-x)$ so fg is odd.
 ii) If f, g both odd then $f(a+x)g(a+x) = -f(a-x)(-g(a-x)) = +f(a-x)g(a-x)$ so fg is even.
²⁸No, $V(x) = Cx$ is neither even nor odd about $\frac{L}{2}$ as $C(\frac{L}{2} + x) \neq \pm C(\frac{L}{2} - x)$.

- n and k have the same parity (either both odd or both even). In this case, the functions $\phi_k^{(0)}$ and $\phi_n^{(0)}$ have the same parity, so their product is even. But then the integrand is an even function $([\phi_k^{(0)}]^* \phi_n^{(0)})$ times an odd function (V_2), and hence it vanishes upon integration.
- n and k have opposite parity. In this case, the functions $\phi_k^{(0)}$ and $\phi_n^{(0)}$ have opposite parity, so their product is odd. But then the integrand is an odd function $([\phi_k^{(0)}]^* \phi_n^{(0)})$ times an odd function (V_2), which gives a non-zero answer upon integration.

This illustrates an important principle: *the fact that certain mixings are forbidden arises from a mismatch between the symmetry of the perturbation and the symmetry of the wave functions.* We shall see more of this when we consider the hydrogen atom in lecture 7.

Finally, let us look graphically at how the perturbed ground-state wave function differs from the unperturbed one. For simplicity, let us consider only the dominant mixing coefficient $c_{12}^{(1)}$, which from (141) is given by

$$c_{12}^{(1)} = \frac{4mCL^3}{\pi^4 \hbar^2} \frac{1}{1-4} \left(-\frac{1}{1} + \frac{1}{9} \right) = \frac{4mCL^3}{\pi^4 \hbar^2} \frac{8}{27} = \frac{2^5}{3^3} \frac{mCL^3}{\pi^4 \hbar^2}. \quad (148)$$

We don't care too much about the magnitude of this (except that it's small), but the sign is important. Given this mixing coefficient, and ignoring the others, the perturbed ground-state wave function is given approximately by

$$\phi_1(x) \approx \phi_1^{(0)}(x) + \frac{2^5}{3^3} \frac{mCL^3}{\pi^4 \hbar^2} \phi_2^{(0)}(x) \quad (149)$$

$$= \sqrt{\frac{2}{L}} \left\{ \sin\left(\frac{\pi x}{L}\right) + \frac{2^5}{3^3} \frac{mCL^3}{\pi^4 \hbar^2} \sin\left(\frac{2\pi x}{L}\right) \right\}. \quad (150)$$

The two pieces of the right-hand side, along with their sum, are plotted in Fig. 3. What has happened? The perturbation has caused the 'centre of mass' of the wave function to move slightly to the left. Looking at Fig. 2 from lecture 4, we see that this shift is towards the area where the potential is lower. So by tilting the bottom of the well, we have 'tipped' the particle towards the left-hand side.²⁹

²⁹Note that this is the *opposite* of what would happen for a classical particle. A classical particle in such a potential would move faster where the potential was lower, and therefore spend *less* time there.

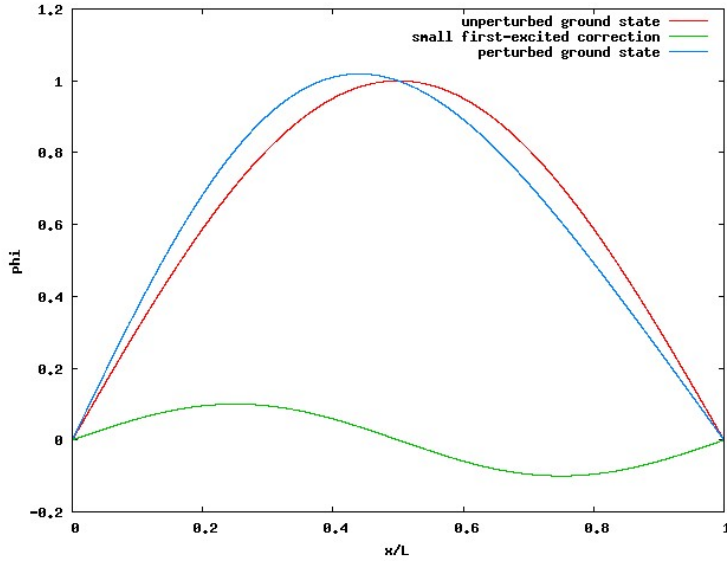


Figure 3: The unperturbed ground-state wave function, the small piece of first-excited state wave function with which it is mixed, and the resulting approximation to the perturbed ground-state wave function.

✓ Consider a *two-dimensional* infinite square well with spatial parts of the energy eigenfunctions $\phi_{n_x n_y}(x, y)$, energy eigenvalues $E_{n_x n_y}$ and a perturbation \hat{V} . Write down expressions for i) the unperturbed ground state energy; ii) the first-order energy correction to the ground state energy; iii) the unperturbed ground state energy eigenfunction; iv) the first-order correction to the ground state energy eigenfunction ³⁰

$$\begin{array}{llll}
 \text{30 i) } E_{11}^{(0)} & \text{ii) } E_{11}^{(1)} = \langle 11^{(0)} | \hat{V} | 11^{(0)} \rangle & \text{iii) } | 11^{(0)} \rangle \text{ or } \phi_{11}^{(0)} & \text{iv) } | 11^{(1)} \rangle = \sum_{k_x \neq 1} \sum_{k_y \neq 1} c_{11, k_x k_y}^{(1)} | k_x k_y^{(0)} \rangle \\
 \text{or } \phi_{11}^{(1)} = \sum_{k_x \neq 1} \sum_{k_y \neq 1} c_{11, k_x k_y}^{(1)} \phi_{k_x k_y}^{(0)} & & &
 \end{array}$$

7 Perturbation theory: the degenerate case

7.1 Principle

You may already have noticed that there is a problem with the formula for $c_{nk}^{(1)}$ when the unperturbed energy spectrum contains degenerate levels. Let us look again at that formula:

$$c_{nk}^{(1)} = \frac{\langle k^{(0)} | \hat{V} | n^{(0)} \rangle}{E_n^{(0)} - E_k^{(0)}}. \quad (151)$$

It contains the difference $E_n^{(0)} - E_k^{(0)}$ in the denominator. If the spectrum is non-degenerate, this is fine, since then the fact that $n \neq k$ implies that $E_n^{(0)} \neq E_k^{(0)}$. But if the spectrum contains degenerate levels, then it is possible that $E_n^{(0)} = E_k^{(0)}$ even when $n \neq k$. If this happens, our corrections become divergent — which is a bad thing, considering that they were supposed to be small.

This problem can be dealt with by a simple trick. First, note that the matrix element $\langle k^{(0)} | \hat{V} | n^{(0)} \rangle$ appears in the numerator of (151). Hence, if this matrix element vanished, things would be all right.³¹ But how do we make the matrix element vanish? Surely it's not under our control?

The answer is that in the degenerate case, we *can* control it. This is because when the states are degenerate the choice of eigenkets is not unique. For example, consider the following linear combination of the unperturbed kets:

$$|\tilde{n}^{(0)}\rangle = \alpha|n^{(0)}\rangle + \beta|k^{(0)}\rangle, \quad (152)$$

$$|\tilde{k}^{(0)}\rangle = \gamma|n^{(0)}\rangle + \delta|k^{(0)}\rangle. \quad (153)$$

It is easy to show that the new kets are also eigenkets of \hat{H}_0 , i.e. that

$$\hat{H}_0|\tilde{n}^{(0)}\rangle = E_n^{(0)}|\tilde{n}^{(0)}\rangle, \quad \hat{H}_0|\tilde{k}^{(0)}\rangle = E_k^{(0)}|\tilde{k}^{(0)}\rangle. \quad (154)$$

(Remember that the two energies here are equal: that's why we're having this problem in the first place!)

✓ Show that $\hat{H}_0|\tilde{n}^{(0)}\rangle = E_n^{(0)}|\tilde{n}^{(0)}\rangle$, the first of the relations (154).³²

We can freely choose α , β , γ , and δ , subject to the following constraints:

1. $|\tilde{n}^{(0)}\rangle$ should be normalised;
2. $|\tilde{k}^{(0)}\rangle$ should be normalised;
3. $\langle \tilde{k}^{(0)} | \tilde{n}^{(0)} \rangle$ should be zero.

These are three constraints on four unknowns, which means that we still have one parameter to play with. Thus we can choose our linear combination so that the fourth 'constraint'

³¹This isn't obvious. Looking at the expression (151), all we can say is that the fraction would then be $\frac{0}{0}$, which is undefined. However, if the analysis is repeated starting from equation (97), it can be shown that the corrections do indeed vanish in this case.

³² $\hat{H}_0|\tilde{n}^{(0)}\rangle = \alpha\hat{H}_0|n^{(0)}\rangle + \beta\hat{H}_0|k^{(0)}\rangle = \alpha E_n^{(0)}|n^{(0)}\rangle + \beta E_k^{(0)}|k^{(0)}\rangle = E_n^{(0)}(\alpha|n^{(0)}\rangle + \beta|k^{(0)}\rangle) = E_n^{(0)}|\tilde{n}^{(0)}\rangle.$

4. $\langle \tilde{k}^{(0)} | \hat{V} | \tilde{n}^{(0)} \rangle$ should be zero

is also satisfied. Then we can just do our perturbation theory starting from the unperturbed basis $\{|\tilde{k}^{(0)}\rangle, |\tilde{n}^{(0)}\rangle\}$ rather than $\{|k^{(0)}\rangle, |n^{(0)}\rangle\}$: the troublesome matrix elements will vanish, and the normal formulas (151) can be used without danger.

7.2 Degenerate perturbation theory: Example

Let us see an example of how this works. To have a degenerate unperturbed spectrum, we'll have to be in more than one dimension; let's take the simplest $d > 1$ problem, the infinitely deep square well in two dimensions (IDSW2). The ground state here is non-degenerate, but the first-excited states are degenerate:

$$\left. \begin{aligned} \phi_{12}(x, y) &= \begin{cases} \frac{2}{L} \sin\left(\frac{\pi x}{L}\right) \sin\left(\frac{2\pi y}{L}\right) & 0 < x < L \text{ and } 0 < y < L, \\ 0 & \text{otherwise} \end{cases} \\ \phi_{21}(x, y) &= \begin{cases} \frac{2}{L} \sin\left(\frac{2\pi x}{L}\right) \sin\left(\frac{\pi y}{L}\right) & 0 < x < L \text{ and } 0 < y < L, \\ 0 & \text{otherwise} \end{cases} \end{aligned} \right\} E_{12}^{(0)} = E_{21}^{(0)} = \frac{5\pi^2 \hbar^2}{2mL^2}. \quad (155)$$

Let us now apply the perturbation³³

$$V(x, y) = Cxy. \quad (156)$$

The diagonal matrix elements are quite easy to calculate:

$$V_{12,12} = \int_0^L dx \int_0^L dy Cxy \frac{4}{L^2} \sin^2\left(\frac{\pi x}{L}\right) \sin^2\left(\frac{2\pi y}{L}\right) \quad (157)$$

$$= C \left(\int_0^L dx \frac{2}{L} x \sin^2\left(\frac{\pi x}{L}\right) \right) \left(\int_0^L dy \frac{2}{L} y \sin^2\left(\frac{2\pi y}{L}\right) \right). \quad (158)$$

The first bracket on the right-hand side is just the expectation value of x in one of the IDSW1 eigenstates, and so is the second bracket. By symmetry, each of these is $L/2$. The same reasoning would clearly apply to the other diagonal matrix element, and so

$$V_{12,12} = V_{21,21} = \frac{CL^2}{4}. \quad (159)$$

The off-diagonal elements (the ones that will give us trouble) are slightly trickier:

$$V_{12,21} = \frac{4C}{L^2} \int_0^L dx \int_0^L dy \sin\left(\frac{\pi x}{L}\right) \sin\left(\frac{2\pi y}{L}\right) xy \sin\left(\frac{2\pi x}{L}\right) \sin\left(\frac{\pi y}{L}\right) \quad (160)$$

³³It's worth working out what classical force field this corresponds to; you are encouraged to do so.

$$= \frac{4C}{L^2} \left[\int_0^L dx x \sin\left(\frac{\pi x}{L}\right) \sin\left(\frac{2\pi x}{L}\right) \right]^2. \quad (161)$$

The integral inside the square brackets may be simplified using the usual trigonometric identity

$$2 \sin A \sin B = \cos(A - B) - \cos(A + B), \quad (162)$$

yielding

$$V_{12,21} = \frac{C}{L^2} \left[\int_0^L x \cos\left(\frac{\pi x}{L}\right) dx - \int_0^L x \cos\left(\frac{3\pi x}{L}\right) dx \right]^2. \quad (163)$$

Both of these integrals are of the same form, namely

$$\int_0^L x \cos\left(\frac{p\pi x}{L}\right) dx \quad (164)$$

with p odd. Using integration by parts,

$$\int_0^L x \cos\left(\frac{p\pi x}{L}\right) dx = \underbrace{\left[\frac{L}{p\pi} x \sin\left(\frac{p\pi x}{L}\right) \right]_0^L}_{=0} - \frac{L}{p\pi} \int_0^L \sin\left(\frac{p\pi x}{L}\right) dx \quad (165)$$

$$= -\frac{L}{p\pi} \left[-\frac{L}{p\pi} \cos\left(\frac{p\pi x}{L}\right) \right]_0^L \quad (166)$$

$$= \frac{L^2}{p^2\pi^2} (\cos p\pi - \cos 0) \quad (167)$$

$$= -\frac{2L^2}{p^2\pi^2} \quad (168)$$

for the case where p is odd. Substituting this back into (163) with $p = 1$ and $p = 3$ respectively, we find

$$V_{12,21} = \frac{C}{L^2} \left[-\frac{2L^2}{\pi^2} + \frac{2L^2}{9\pi^2} \right]^2 = \frac{256CL^2}{81\pi^4}. \quad (169)$$

The integrals for $V_{21,12}$ are exactly the same.

So to summarise our results in matrix form, the matrix \mathbf{V} representing the perturbation in the degenerate (two-dimensional) subspace is

$$\mathbf{V} = \begin{pmatrix} V_{12,12} & V_{12,21} \\ V_{21,12} & V_{21,21} \end{pmatrix} = \frac{CL^2}{4} \begin{pmatrix} 1 & \frac{1024}{81\pi^4} \\ \frac{1024}{81\pi^4} & 1 \end{pmatrix}. \quad (170)$$

So we are looking for combinations of the unperturbed states that are not mixed by the perturbation. But the vectors that are not mixed by the action of a matrix are its

eigenvectors, and the diagonal matrix elements in the new basis (which give the first-order energy corrections) are its **eigenvalues**. So all we need to do is find the eigenvalues and eigenvectors of (170): the eigenvalues will give the first-order energy corrections in the new basis, and the eigenvectors will give the combinations of the unperturbed kets that we should have been using all along.

Now we find the eigenvectors and eigenvalues of \mathbf{V} . The eigenvalues are obtained, as usual, by setting the determinant

$$|\mathbf{V} - \lambda \mathbf{I}| = 0. \quad (171)$$

In this case, this gives

$$\begin{vmatrix} \frac{CL^2}{4} - \lambda & \frac{256CL^2}{81\pi^4} \\ \frac{256CL^2}{81\pi^4} & \frac{CL^2}{4} - \lambda \end{vmatrix} = 0 \quad (172)$$

which implies

$$\left(\frac{CL^2}{4} - \lambda\right)^2 = \left(\frac{256CL^2}{81\pi^4}\right)^2, \quad (173)$$

so that

$$\frac{CL^2}{4} - \lambda = \pm \frac{256CL^2}{81\pi^4}, \quad (174)$$

and hence the eigenvalues are

$$\lambda_+ = \frac{CL^2}{4} + \frac{256CL^2}{81\pi^4}, \quad \lambda_- = \frac{CL^2}{4} - \frac{256CL^2}{81\pi^4}. \quad (175)$$

The eigenvectors of \mathbf{V} are obtained by inserting each eigenvalue successively into

$$\mathbf{V} \begin{pmatrix} x \\ y \end{pmatrix} = \lambda \begin{pmatrix} x \\ y \end{pmatrix} \quad (176)$$

or equivalently

$$(\mathbf{V} - \lambda \mathbf{I}) \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad (177)$$

and solving for the two unknowns x and y .

Inserting λ_+ into equation (177) gives

$$\begin{pmatrix} \frac{CL^2}{4} - \left(\frac{CL^2}{4} + \frac{256CL^2}{81\pi^4}\right) & \frac{256CL^2}{81\pi^4} \\ \frac{256CL^2}{81\pi^4} & \frac{CL^2}{4} - \left(\frac{CL^2}{4} + \frac{256CL^2}{81\pi^4}\right) \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad (178)$$

and thus

$$\frac{256CL^2}{81\pi^4} \begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad (179)$$

This implies that $-x + y = 0$ so that $y = x$ and

$$\mathbf{e}_+ = N \begin{pmatrix} 1 \\ 1 \end{pmatrix}. \quad (180)$$

Note that the two equations in (179) are linearly dependent. The second equation for the two unknowns x and y is the normalization condition

$$\sqrt{|x|^2 + |y|^2} = 1. \quad (181)$$

Applying the normalization condition to \mathbf{e}_+ gives

$$\sqrt{N^2 + N^2} = \sqrt{2}N = 1 \quad (182)$$

and thus $N = \frac{1}{\sqrt{2}}$. Thus, the normalized eigenvector

$$\mathbf{e}_+ = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}. \quad (183)$$

Inserting λ_- into equation (177) gives

$$\begin{pmatrix} \frac{CL^2}{4} - \left(\frac{CL^2}{4} - \frac{256CL^2}{81\pi^4} \right) & \frac{256CL^2}{81\pi^4} \\ \frac{256CL^2}{81\pi^4} & \frac{CL^2}{4} - \left(\frac{CL^2}{4} - \frac{256CL^2}{81\pi^4} \right) \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad (184)$$

and thus

$$\frac{256CL^2}{81\pi^4} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \quad (185)$$

This implies that $x + y = 0$ so that $y = -x$ and

$$\mathbf{e}_- = N \begin{pmatrix} 1 \\ -1 \end{pmatrix}. \quad (186)$$

The first component by convention is chosen to be real and positive. Thus, the normalized eigenvector

$$\mathbf{e}_- = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}. \quad (187)$$

We have found that the eigenvectors of \mathbf{V} are given by

$$\mathbf{e}_+ = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \mathbf{e}_- = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}. \quad (188)$$

Hence the combinations of the basis functions (155) that we should use for our perturbation theory are

$$\tilde{\phi}_{12}^{(0)} = \frac{1}{\sqrt{2}} \left(\phi_{12}^{(0)} + \phi_{21}^{(0)} \right), \quad (189)$$

$$\tilde{\phi}_{21}^{(0)} = \frac{1}{\sqrt{2}} \left(\phi_{12}^{(0)} - \phi_{21}^{(0)} \right). \quad (190)$$

If we were to choose the basis $\tilde{\phi}_{12}^{(0)}$ and $\tilde{\phi}_{21}^{(0)}$ in this order, then the matrix \mathbf{V} representing \hat{V} in the two-dimensional degenerate subspace would be diagonal, with the eigenvalues as the diagonal elements (see section 3.2). The matrix

$$\mathbf{V} = \begin{pmatrix} \langle \tilde{\phi}_{12}^{(0)} | \hat{V} | \tilde{\phi}_{12}^{(0)} \rangle & \langle \tilde{\phi}_{12}^{(0)} | \hat{V} | \tilde{\phi}_{21}^{(0)} \rangle \\ \langle \tilde{\phi}_{21}^{(0)} | \hat{V} | \tilde{\phi}_{12}^{(0)} \rangle & \langle \tilde{\phi}_{21}^{(0)} | \hat{V} | \tilde{\phi}_{21}^{(0)} \rangle \end{pmatrix} = \begin{pmatrix} \lambda_+ & 0 \\ 0 & \lambda_- \end{pmatrix}. \quad (191)$$

We see from this representation that the **eigenvalues of \mathbf{V} in the degenerate subspace are the first-order energy corrections in the new basis**, as $\lambda_+ = \langle \tilde{\phi}_{12}^{(0)} | \hat{V} | \tilde{\phi}_{12}^{(0)} \rangle$ and $\lambda_- = \langle \tilde{\phi}_{21}^{(0)} | \hat{V} | \tilde{\phi}_{21}^{(0)} \rangle$ (compare with the general formula for the first-order energy correction $E_n^{(1)} = \langle n^{(0)} | \hat{V} | n^{(0)} \rangle$). We can also see that **the new basis states are not mixed by the perturbation, as the problematic (off-diagonal) matrix elements are zero**, namely

$$\langle \tilde{\phi}_{21}^{(0)} | \hat{V} | \tilde{\phi}_{12}^{(0)} \rangle = \langle \tilde{\phi}_{12}^{(0)} | \hat{V} | \tilde{\phi}_{21}^{(0)} \rangle = 0 \quad (192)$$

Equation (192) implies that there is zero contribution of $\tilde{\phi}_{12}^{(0)}$ to the perturbed $\tilde{\phi}_{21}$, and zero contribution of $\tilde{\phi}_{21}^{(0)}$ to the perturbed $\tilde{\phi}_{12}$, as the relevant first-order mixing coefficients are zero. Note however that there will in general be non-zero contributions from other unperturbed eigenfunctions to the perturbed $\tilde{\phi}_{12}$ and $\tilde{\phi}_{21}$, just none in the degenerate subspace.

We see that the perturbation **lifts the degeneracy** between the levels that were originally at $\frac{5\pi^2\hbar^2}{2mL^2}$, giving (to first order in C) perturbed energy levels at

$$E_{\text{pert1}} \approx \frac{5\pi^2\hbar^2}{2mL^2} + \frac{CL^2}{4} - \frac{256CL^2}{81\pi^4}; \quad (193)$$

$$E_{\text{pert2}} \approx \frac{5\pi^2\hbar^2}{2mL^2} + \frac{CL^2}{4} + \frac{256CL^2}{81\pi^4}. \quad (194)$$

✓ Are $\frac{CL^2}{4} \pm \frac{256CL^2}{81\pi^4}$ also the first-order energy corrections of the original basis states ϕ_{12} and ϕ_{21} , given that the eigenvalues of a matrix are the same irrespective of which basis one chooses? ³⁴

✓ How do you know which energy correction $\frac{CL^2}{4} \pm \frac{256CL^2}{81\pi^4}$ belongs to which basis state $\tilde{\phi}_{12}^{(0)}$ and $\tilde{\phi}_{21}^{(0)}$, or is their assignment arbitrary? ³⁵

³⁴No, they are not! We can not proceed with perturbation theory using the original basis states ϕ_{12} and ϕ_{21} , so that first-order energy corrections are not defined for these states.

³⁵The assignment is not arbitrary - we need to be careful in linking up an eigenvalue with its corresponding eigenvector. Thus, $\lambda_+ = \frac{CL^2}{4} + \frac{256CL^2}{81\pi^4}$ is the energy correction to $\tilde{\phi}_{12}^{(0)}$, as this basis state is given by \mathbf{e}_+ .

8 Perturbation theory: second-order energy corrections, and time-dependent perturbations

In this lecture, we shall look at a couple of final topics in perturbation theory: properties of the energy corrections at second order, and what happens when the perturbation depends on time.

8.1 Second-order energy corrections

We see from the example in lecture 4 that the first-order corrections to the mixing coefficients, $c_{nk}^{(1)}$, contain much physical information about how the quantum system responds to the perturbation. The same cannot be said, however, for the first-order energy corrections: in the lecture 4 example these gave merely a constant upward displacement of the entire energy spectrum. It is therefore frequently desirable to go to second order. As you showed in your tutorial work, the result is

$$E_n^{(2)} = \sum_{k \neq n} \frac{|\langle k^{(0)} | \hat{V} | n^{(0)} \rangle|^2}{E_n^{(0)} - E_k^{(0)}}. \quad (195)$$

A few notes on this formula:

- The matrix element that forms its core is the same one that appears in $c_{nk}^{(1)}$. Therefore, in a sense, the physics that appears at first order in the mixing coefficients (selection rules, etc.) doesn't appear until second order in the energies.
- A sum is involved, and this often has an infinite number of terms. However, (a) the denominator makes the terms get smaller as n and k get further apart, and (b) the matrix element in the numerator usually gets smaller as well. So the sum is typically strongly convergent, and can often be well approximated by taking only the few terms where k is close to n .
- If n is the ground-state, then $E_n^{(2)} \leq 0$. This follows since the numerator of the summand is always positive, and for the ground state the denominator is always negative. This shows a very important result: *the second-order correction always lowers the ground-state energy.*

You should **learn the expressions** for $E_n^{(1)}$, $c_{nk}^{(1)}$, and $E_n^{(2)}$. They are the most often used formulas in perturbation theory, and will be recurring many times in the course.

✓ Consider the constant perturbation $\hat{V} = V_0$ in the infinite square well. What is $E_n^{(2)}$? ³⁶

8.2 Time-dependent perturbations

8.2.1 Introduction

So far, we have studied perturbations that were independent of time. Thus the notion of eigenstates still applied, and we could meaningfully look for the eigenstates and

³⁶For $\hat{V} = V_0$, we find $E_n^{(2)} = 0$, as $\langle k^{(0)} | \hat{V} | n^{(0)} \rangle = V_0 \langle k^{(0)} | n^{(0)} \rangle = 0$ for $k \neq n$.

eigenenergies of the new Hamiltonian. However, it is frequently the case that the applied perturbation does vary with time. An example from the early days of quantum mechanics is the exposure of an atom to a beam of light: the electric field of the light beam takes the form³⁷

$$\mathbf{E}(t) = \Gamma \hat{\mathbf{z}} \cos \omega t, \quad (196)$$

and hence the potential energy of the electron in the field is likewise time-dependent:

$$V(z, t) = - \int \mathbf{F} \cdot d\mathbf{r} = e \int \mathbf{E} \cdot d\mathbf{r} = e\Gamma z \cos \omega t. \quad (197)$$

What can we say about the response of the atom to this light field? In particular (remembering that this is perturbation theory), what can we say about the response of the atom to this light field *when Γ is small*?

8.2.2 Time-dependent perturbation theory: rewriting the Schrödinger equation

The general kind of problem we're looking at is of the form

$$\hat{H} = \hat{H}_0 + \lambda f(t) \hat{V}. \quad (198)$$

Here \hat{H}_0 is an unperturbed Hamiltonian, whose eigenkets $|n^{(0)}\rangle$ and eigenenergies $E_n^{(0)}$ we know. λ is the usual small parameter, used for book-keeping during our calculations and set to one at the end. $f(t)$ gives the time-dependence of the perturbation, but is not an operator; \hat{V} is an operator, but is time-independent. In the above example of the hydrogen atom, we would have³⁸

$$\hat{H}_0 = -\frac{\hbar^2}{2m} \nabla^2 - \frac{e^2}{4\pi\epsilon_0 r}, \quad (199)$$

$$f(t) = \cos \omega t, \quad (200)$$

$$\hat{V} = e\Gamma \hat{z}. \quad (201)$$

✓What would be $f(t)$ for a perturbation that turns on at time t_0 and then remains constant? ³⁹

Since the perturbation is time-dependent, we must solve the time-dependent Schrödinger equation,

$$i\hbar \frac{\partial}{\partial t} |\psi\rangle = \hat{H} |\psi\rangle. \quad (202)$$

We can still use the unperturbed eigenkets as a basis in which to express our solution. However, the appropriate superposition to take now varies with time:

$$|\psi(t)\rangle = \sum_n c_n(t) \underbrace{e^{-iE_n^{(0)}t/\hbar} |n^{(0)}\rangle}_{=|n^{(0)}(t)\rangle}. \quad (203)$$

³⁷Here we have taken the polarisation of the light to be such that its electric field points along the z -direction.

³⁸Obviously we could have divided the constants e and Γ differently between $f(t)$ and \hat{V} , but this has no effect on the final result.

³⁹ $f(t) = \theta(t - t_0)$ where $\theta(t - t_0)$ is the Heaviside step function (equals 1 for $t - t_0 \geq 0$ and 0 else)

Notice that we have used the full time-dependent form of the basis kets.

Now we insert (203) into (202), obtaining the following:

$$i\hbar \frac{\partial}{\partial t} \left(\sum_n c_n(t) e^{-iE_n^{(0)}t/\hbar} |n^{(0)}\rangle \right) = \sum_n c_n(t) e^{-iE_n^{(0)}t/\hbar} \hat{H} |n^{(0)}\rangle \quad (204)$$

$$= \sum_n c_n(t) e^{-iE_n^{(0)}t/\hbar} \left(\hat{H}_0 + \lambda f(t) \hat{V} \right) |n^{(0)}\rangle \quad (205)$$

$$= \sum_n c_n(t) e^{-iE_n^{(0)}t/\hbar} E_n^{(0)} |n^{(0)}\rangle + \lambda f(t) \sum_n c_n(t) e^{-iE_n^{(0)}t/\hbar} \hat{V} |n^{(0)}\rangle. \quad (206)$$

The time-derivative on the left-hand side may be taken explicitly using the product rule:

$$\frac{\partial}{\partial t} \left(\sum_n c_n(t) e^{-iE_n^{(0)}t/\hbar} |n^{(0)}\rangle \right) = \sum_n \dot{c}_n(t) e^{-iE_n^{(0)}t/\hbar} |n^{(0)}\rangle - \frac{i}{\hbar} \sum_n c_n(t) E_n^{(0)} e^{-iE_n^{(0)}t/\hbar} |n^{(0)}\rangle. \quad (207)$$

Thus (206) becomes

$$i\hbar \sum_n \dot{c}_n(t) e^{-iE_n^{(0)}t/\hbar} |n^{(0)}\rangle + \sum_n c_n(t) E_n^{(0)} e^{-iE_n^{(0)}t/\hbar} |n^{(0)}\rangle = \sum_n c_n(t) e^{-iE_n^{(0)}t/\hbar} E_n^{(0)} |n^{(0)}\rangle + \lambda f(t) \sum_n c_n(t) e^{-iE_n^{(0)}t/\hbar} \hat{V} |n^{(0)}\rangle. \quad (208)$$

The second term on the left-hand side cancels the first term on the right-hand side, leaving

$$i\hbar \sum_n \dot{c}_n(t) e^{-iE_n^{(0)}t/\hbar} |n^{(0)}\rangle = \lambda f(t) \sum_n c_n(t) e^{-iE_n^{(0)}t/\hbar} \hat{V} |n^{(0)}\rangle. \quad (209)$$

Now we can get rid of the kets altogether. Multiplying from the left by a bra $\langle k^{(0)}|$ yields

$$i\hbar \sum_n \dot{c}_n(t) e^{-iE_n^{(0)}t/\hbar} \langle k^{(0)} | n^{(0)} \rangle = \lambda f(t) \sum_n c_n(t) e^{-iE_n^{(0)}t/\hbar} \underbrace{\langle k^{(0)} | \hat{V} | n^{(0)} \rangle}_{\equiv V_{kn}}; \quad (210)$$

using the orthonormality of the unperturbed states, $\langle k^{(0)} | n^{(0)} \rangle = \delta_{kn}$, this becomes

$$i\hbar \dot{c}_k(t) e^{-iE_k^{(0)}t/\hbar} = \lambda f(t) \sum_n c_n(t) e^{-iE_n^{(0)}t/\hbar} V_{kn}. \quad (211)$$

Dividing both sides by $i\hbar e^{-iE_k^{(0)}t/\hbar}$ gives

$$\frac{dc_k}{dt} = -\frac{i\lambda}{\hbar} f(t) \sum_n c_n(t) e^{-iE_n^{(0)}t/\hbar} V_{kn}, \quad (212)$$

where we have defined the unperturbed energy difference

$$E_{nk}^{(0)} \equiv E_n^{(0)} - E_k^{(0)}. \quad (213)$$

8.2.3 Interpretation

Equation (212) contains the solution to our problem: it gives the time-evolution of the coefficients in the superposition (203). A few observations:

- When λ is set to zero (i.e. the perturbation is turned off), the right-hand side of (212) vanishes. Thus the coefficients c_n in (203) become time-independent. This just reflects the fact that, in the absence of a time-dependent perturbation, the state remains the same (apart from its ‘trivial’ $e^{-iEt/\hbar}$ time-dependence) forever.
- The variation of c_k will lead to a variation of $|c_k|^2$, the probability of finding the system in state $|k^{(0)}\rangle$ at time t . This can be interpreted as the perturbation’s provoking **transitions** of the system between its unperturbed eigenstates.
- Notice that a transition between state n and state k occurs only if the matrix element V_{nk} is non-zero. Thus the notion of **selection rules** applies in this case just as it did in time-independent perturbation theory.
- Since (212) is a set of first-order differential equations — one for each coefficient c_k — it requires initial conditions to be specified. These correspond to specifying all the coefficients c_k at time $t = 0$, i.e. to specifying the state that the system was in before the perturbation was applied. This is frequently taken to be one of the unperturbed eigenstates, which leads to the simple initial condition

$$c_k(0) = \delta_{km} = \begin{cases} 1 & k = m, \\ 0 & k \neq m, \end{cases} \quad (214)$$

where $|m^{(0)}\rangle$ is the unperturbed eigenstate that the system is in before the perturbation is applied.

✓ Assume a system starts out in the ground state $\phi_1^{(0)}$. What expression gives the probability for finding the system in the first-excited state at time $t_0 > 0$? ⁴⁰

⁴⁰The probability for finding the system in $\phi_2^{(0)}$ at time t_0 is given by $|c_2(t_0)|^2$.

9 Time-dependent perturbation theory: the ‘golden rule’

In this lecture, we shall continue to analyse the response of quantum systems to time-dependent perturbations. Starting from (212), we shall specialise to a particular form of time-dependence, namely the **monochromatic** (single-frequency) form

$$f(t) = \cos \omega t. \quad (215)$$

Furthermore, we shall assume that before the application of the perturbation our quantum system is in one of its eigenstates. This means that the coefficients at time $t = 0$ satisfy

$$c_n(t = 0) = c_n^{(0)} = \delta_{ni} = \begin{cases} 1 & n = i, \\ 0 & \text{otherwise,} \end{cases} \quad (216)$$

where i (which stands for ‘initial’) is the label of the state in which the system begins.

9.1 The ‘golden rule’: derivation

Now we apply the perturbation. What happens? We see from (212) that the coefficients of states $f \neq i$ will start to rise from zero; physically, this corresponds to the perturbation’s provoking transitions from the initial state to those other states. Let us calculate the rate at which such transitions occur.

The rate is given by

$$\Gamma_{if}(t) = \frac{|c_f|^2}{t}, \quad (217)$$

where f stands for ‘final’. There will initially be some transient behaviour, where the populations fluctuate wildly, but once the perturbation has been ‘on’ for a while we expect the system to settle down. Thus we study the long-time limit of (217):

$$\Gamma_{if} \equiv \lim_{t \rightarrow \infty} \frac{|c_f|^2}{t}. \quad (218)$$

Our procedure will be to find $c_f(t)$ by solving (212) at first order in perturbation theory, then to substitute our result into (218) and obtain the transition rate.

✓What is the unit of Γ_{if} ? ⁴¹

From (212), the differential equation for $c_f(t)$ is

$$\frac{dc_f}{dt} = -\frac{i\lambda}{\hbar} f(t) \sum_n c_n(t) e^{-iE_{nf}^{(0)}t/\hbar} V_{fn}. \quad (219)$$

Now let us substitute our perturbative expansion of the population coefficients,

$$c_n(t) = c_n^{(0)} + \lambda c_n^{(1)} + O(\lambda^2); \quad (220)$$

⁴¹s⁻¹ or “probability per second to transition from state $|i^{(0)}\rangle$ to state $|f^{(0)}\rangle$ ”

or, inserting the value of $c_n^{(0)}$ from (216),

$$c_n(t) = \delta_{ni} + \lambda c_n^{(1)} + O(\lambda^2). \quad (221)$$

Inserting this into (219), and keeping only the terms with a single power of λ , we obtain

$$\frac{dc_f^{(1)}}{dt} = -\frac{i}{\hbar} f(t) \sum_n c_n^{(0)} e^{-iE_{nf}^{(0)}t/\hbar} V_{fn}. \quad (222)$$

But $c_n^{(0)} = \delta_{ni}$, so only one term in the sum survives:

$$\frac{dc_f^{(1)}}{dt} = -\frac{i}{\hbar} f(t) e^{-iE_{if}^{(0)}t/\hbar} V_{fi}. \quad (223)$$

This reflects the physical fact that, for a small perturbation, most of the transitions that populate the state f come directly from the initial state i .

Now we solve this equation for the monochromatic case. Substituting in $f(t) = \cos \omega t$, we find

$$\frac{dc_f^{(1)}}{dt} = -\frac{i}{\hbar} \cos \omega t e^{-iE_{if}^{(0)}t/\hbar} V_{fi} \quad (224)$$

$$= -\frac{i}{2\hbar} (e^{i\omega t} + e^{-i\omega t}) e^{-iE_{if}^{(0)}t/\hbar} V_{fi} \quad (225)$$

$$= -\frac{i}{2\hbar} e^{-i((E_{if}^{(0)}/\hbar)-\omega)t} V_{fi} - \frac{i}{2\hbar} e^{-i((E_{if}^{(0)}/\hbar)+\omega)t} V_{fi}. \quad (226)$$

This may be straightforwardly integrated with respect to time, to give

$$c_f^{(1)}(t) = -\frac{i}{2\hbar} V_{fi} \left\{ \int e^{-i((E_{if}^{(0)}/\hbar)-\omega)t} dt + \int e^{-i((E_{if}^{(0)}/\hbar)+\omega)t} dt \right\} \quad (227)$$

$$= -\frac{i}{2\hbar} V_{fi} \left\{ \frac{i}{(E_{if}^{(0)}/\hbar) - \omega} e^{-i((E_{if}^{(0)}/\hbar)-\omega)t} + \frac{i}{(E_{if}^{(0)}/\hbar) + \omega} e^{-i((E_{if}^{(0)}/\hbar)+\omega)t} \right\} + C \quad (228)$$

where C is a constant of integration. Its value is determined by the fact that $c_f^{(1)} = 0$ at $t = 0$; the resulting expression is

$$c_f^{(1)}(t) = -\frac{i}{2\hbar} V_{fi} \left\{ \frac{i}{(E_{if}^{(0)}/\hbar) - \omega} \left(e^{-i((E_{if}^{(0)}/\hbar)-\omega)t} - 1 \right) + \frac{i}{(E_{if}^{(0)}/\hbar) + \omega} \left(e^{-i((E_{if}^{(0)}/\hbar)+\omega)t} - 1 \right) \right\}. \quad (229)$$

The next step (brace yourself!) is to take the square modulus of this, and insert the result into the expression (218) to obtain the transition rate from i to f . The complex conjugate of (229) is

$$[c_f^{(1)}(t)]^* = \frac{i}{2\hbar} V_{fi}^* \left\{ -\frac{i}{(E_{if}^{(0)}/\hbar) - \omega} \left(e^{i((E_{if}^{(0)}/\hbar)-\omega)t} - 1 \right) - \frac{i}{(E_{if}^{(0)}/\hbar) + \omega} \left(e^{i((E_{if}^{(0)}/\hbar)+\omega)t} - 1 \right) \right\}; \quad (230)$$

multiplying this by (229), we obtain

$$\begin{aligned}
|c_f^{(1)}(t)|^2 &= \frac{|V_{fi}|^2}{4\hbar^2} \left\{ \frac{1}{((E_{if}^{(0)}/\hbar) - \omega)^2} \left(e^{i((E_{if}^{(0)}/\hbar) - \omega)t} - 1 \right) \left(e^{-i((E_{if}^{(0)}/\hbar) - \omega)t} - 1 \right) \right. \\
&\quad + \frac{1}{((E_{if}^{(0)}/\hbar) + \omega)^2} \left(e^{i((E_{if}^{(0)}/\hbar) + \omega)t} - 1 \right) \left(e^{-i((E_{if}^{(0)}/\hbar) + \omega)t} - 1 \right) \\
&\quad + \frac{2}{((E_{if}^{(0)}/\hbar) + \omega)((E_{if}^{(0)}/\hbar) - \omega)} \left[\left(e^{-i((E_{if}^{(0)}/\hbar) - \omega)t} - 1 \right) \left(e^{i((E_{if}^{(0)}/\hbar) + \omega)t} - 1 \right) \right. \\
&\quad \left. \left. + \left(e^{-i((E_{if}^{(0)}/\hbar) + \omega)t} - 1 \right) \left(e^{i((E_{if}^{(0)}/\hbar) - \omega)t} - 1 \right) \right] \right\}. \tag{231}
\end{aligned}$$

This is evidently completely horrifying. However, all the terms have basically the same structure, so let's just see what happens with the first one. If we define $\alpha = (E_{if}^{(0)}/\hbar) - \omega$, the first term inside the brackets above becomes

$$T_1 = \frac{1}{\alpha^2} (e^{i\alpha t} - 1) (e^{-i\alpha t} - 1). \tag{232}$$

To get the transition rate — see (218) — we divide this by time and then take the long-time limit. So we really want

$$R_1 \equiv \lim_{t \rightarrow \infty} \frac{T_1}{t} = \frac{1}{\alpha^2} \lim_{t \rightarrow \infty} \left[\frac{(e^{i\alpha t} - 1)(e^{-i\alpha t} - 1)}{t} \right]. \tag{233}$$

Expanding out the brackets, we obtain

$$R_1 = \frac{1}{\alpha^2} \lim_{t \rightarrow \infty} \frac{1}{t} [2 - e^{i\alpha t} - e^{-i\alpha t}] \tag{234}$$

$$= \frac{1}{\alpha^2} \lim_{t \rightarrow \infty} \frac{1}{t} [2 - 2 \cos \alpha t] \tag{235}$$

$$= \frac{1}{\alpha^2} \lim_{t \rightarrow \infty} \frac{1}{t} \left[4 \sin^2 \left(\frac{\alpha t}{2} \right) \right]. \tag{236}$$

Remarkably, this is one of the representations of the Dirac delta function!:

$$\lim_{t \rightarrow \infty} \left(\frac{\sin^2(\gamma t)}{\gamma^2 t} \right) = \pi \delta(\gamma). \tag{237}$$

Inserting this result with $\gamma = \alpha/2$, we obtain

$$R_1 = \pi \delta \left(\frac{\alpha}{2} \right) = 2\pi \delta(\alpha). \tag{238}$$

Hence the first term in the rate equation (231) is given by

$$\Gamma_1 = \frac{|V_{fi}|^2}{4\hbar^2} R_1 = \frac{\pi |V_{fi}|^2}{2\hbar^2} \delta((E_{if}^{(0)}/\hbar) - \omega). \tag{239}$$

Taking a factor of $1/\hbar$ out of the delta-function using the usual rules, this may be written

$$\Gamma_1 = \frac{\pi |V_{fi}|^2}{2\hbar} \delta(E_{if}^{(0)} - \hbar\omega). \tag{240}$$

The second term in (231) gives exactly the same contribution, but with ω changed to $-\omega$; the remaining terms go to zero in the long-time limit. One finally obtains the transition rate to be

$$\Gamma_{if} = \frac{\pi}{2\hbar} |V_{fi}|^2 \left\{ \delta(E_{if}^{(0)} - \hbar\omega) + \delta(E_{if}^{(0)} + \hbar\omega) \right\}. \quad (241)$$

This is the so-called **golden rule** for the transition rate from unperturbed eigenstate i to unperturbed eigenstate f under the action of a monochromatic perturbation. It was first derived by Dirac. It is sometimes called the Fermi golden rule, but this is a misattribution — Fermi was simply the first author to include it in a textbook on quantum mechanics.

✓ Assume a time-dependent perturbation $V_0 x \cos(\omega t)$. What is $V_{fi} = V_{21}$? ⁴²

✓ What two conditions are necessary for Γ_{if} to be non-zero in formula (241) for a given $|i^{(0)}\rangle$, $|f^{(0)}\rangle$ and perturbation \hat{V} ? ⁴³

✓ Assume a perturbation had some other time dependence than $f(t) = \cos(\omega t)$, such as $f(t) = e^{-t/\tau}$ for $t > 0$ and 0 otherwise. Which formula would you start with to redo the derivation of the transition rate Γ_{if} ? ⁴⁴

9.2 Interpretation

A few notes on this remarkable formula:

- The transition rate is proportional to the square modulus of the matrix element $V_{fi} = \langle f^{(0)} | \hat{V} | i^{(0)} \rangle$. Thus, if the matrix element vanishes, there are no transitions between states i and f — at least in first order in perturbation theory. So the selection rules we derived in time-independent perturbation theory also apply here.
- The delta-functions enforce conservation of energy. The first term in (241) is non-zero only when

$$E_i^{(0)} - E_f^{(0)} = \hbar\omega; \quad (242)$$

for this to be true, the final-state energy must be lower than the initial-state energy. This term therefore corresponds to **stimulated emission**. The second term, by contrast, is non-zero only when

$$E_f^{(0)} - E_i^{(0)} = \hbar\omega; \quad (243)$$

here the final-state energy is higher than the initial-state energy, so this term corresponds to **absorption**.

- Despite the fact that we have modelled the light field classically — as just an oscillating electric field — the quantum of energy $\hbar\omega$ has appeared in the calculation! This might make you wonder whether the photoelectric effect really demonstrates that light is ‘made of’ photons.

⁴² $\langle 2^{(0)} | V_0 x | 1^{(0)} \rangle$. Beware: the transition matrix element only includes the spatial part of the perturbation, not the temporal part $f(t)$! The temporal part $f(t) = \cos(\omega t)$ disappeared via integration over t and the long-time limit, see formulas (224) onwards.

⁴³The transition matrix element $V_{fi} \neq 0$, and the driving frequency of the perturbation must be such that $E_{if}^{(0)}$ equals $\hbar\omega$ or $-\hbar\omega$.

⁴⁴Formula (224), inserting $e^{-t/\tau}$ rather than $\cos(\omega t)$.

- The delta-functions appear in (241) only as a result of our having taken the long-time limit. This is natural: energy can be conserved only if it is well defined. But at short times, because of the energy-time uncertainty relation $\Delta E \Delta t \geq \hbar/2$, the energy is not well defined. Therefore, a process that violates (classical) energy conservation may take place, provided that it does not last longer than

$$\Delta t = \frac{\hbar}{2\Delta E}, \quad (244)$$

where ΔE is the process's departure from the classically allowed energy. Such processes are called **virtual**: they are very important for the interpretation of quantum mechanics.

10 Intrinsic angular momentum (‘spin’): a theoretical motivation

In Quantum Mechanics 1, as part of the project of solving the Schrödinger equation for the hydrogen atom, you met the angular momentum operators. In this lecture, we shall rederive their eigenvalues using only their commutation algebra. This will hint at an intriguing possibility: *half-integer* values of the quantum number l . We shall see in lecture 12 that these are indeed realised in nature, in the form of *intrinsic angular momentum*, or ‘*spin*’.

10.1 Commutation relations for angular momentum operators

In classical mechanics, the angular momentum of a particle is defined as

$$\mathbf{L} = \mathbf{r} \times \mathbf{p}, \quad (245)$$

where $\mathbf{r} = (x, y, z)$ is the particle’s position and $\mathbf{p} = (p_x, p_y, p_z)$ is its momentum. By the correspondence principle, the quantum operator for angular momentum therefore obeys the same equation:

$$\hat{\mathbf{L}} = \hat{\mathbf{r}} \times \hat{\mathbf{p}}; \quad (246)$$

or, writing that out in components,⁴⁵

$$\begin{pmatrix} \hat{L}^x \\ \hat{L}^y \\ \hat{L}^z \end{pmatrix} = \begin{pmatrix} \hat{y}\hat{p}_z - \hat{z}\hat{p}_y \\ \hat{z}\hat{p}_x - \hat{x}\hat{p}_z \\ \hat{x}\hat{p}_y - \hat{y}\hat{p}_x \end{pmatrix}. \quad (247)$$

Notice that the second and third rows are simply cyclic permutations of the first, obtained by making the shifts $xyz \rightarrow yzx \rightarrow zxy$ in turn.

Hence we may derive the commutation relations of these operators⁴⁶:

$$[L^x, L^y] = [yp_z - zp_y, zp_x - xp_z] \quad (248)$$

$$= [yp_z, zp_x] - [zp_y, zp_x] - [yp_z, xp_z] + [zp_y, xp_z], \quad (249)$$

using the fact that commutators ‘expand’ in the same way that brackets do. Now, since the only things that don’t commute are a co-ordinate and its own momentum, we see immediately that the second and third terms on the right-hand side are zero. This leaves

$$[L^x, L^y] = [yp_z, zp_x] + [zp_y, xp_z] \quad (250)$$

$$= y \underbrace{[p_z, z]}_{=-i\hbar} p_x + x \underbrace{[z, p_z]}_{=i\hbar} p_y \quad (251)$$

⁴⁵Note that we are now writing the component as a superscript (\hat{L}^x) rather than a subscript (\hat{L}_x). This is because, when we start talking about multiple particles, we shall need the subscript slot to refer to the particle number.

⁴⁶We will stop writing the hats on them now, just so as not to clutter up the formulae — but they are all still operators!

$$= i\hbar(xp_y - yp_x) \quad (252)$$

$$= i\hbar L^z. \quad (253)$$

The remaining two commutation relations,

$$[L^y, L^z] = i\hbar L^x \quad \text{and} \quad (254)$$

$$[L^z, L^x] = i\hbar L^y \quad (255)$$

are obtained from (253) by cyclic permutation of xyz .

✓ Determine the commutator of the following operators: i) $[xp_y, zp_y]$ ii) $[xp_y, yp_z]$ ⁴⁷

In fact, relations (253) to (255) don't have anything particular to do with quantum mechanics. They reflect the fundamental fact that rotations in three dimensions around different axes don't commute with each other. (Try it with a matchbox, or a book!) Indeed they are, as a group theorist would recognise, the Lie algebra of $SO(3)$, the group of length-preserving rotations in three dimensions. But that's for another course...

10.2 Eigenstates

Having derived the algebra of the operators, we now look for states in which some of them take definite values. Since no component commutes with the other two, only one of the components can take a definite value at a given time. Let's choose (arbitrarily) L^z to be the one that takes the definite value. However, there is another operator,

$$\mathbf{L}^2 \equiv (L^x)^2 + (L^y)^2 + (L^z)^2, \quad (256)$$

which commutes with all three components,

$$[\mathbf{L}^2, L^j] = 0 \quad j = x, y, z. \quad (257)$$

It can therefore take a definite value at the same time as L^z . We interpret \mathbf{L}^2 as the squared length of the angular momentum vector, and L^z as its projection along the z -axis.

An eigenstate can thus be labelled by the values that \mathbf{L}^2 and L^z take in that state. Let's call these labels l and m respectively, in terms of which⁴⁸

$$\mathbf{L}^2 |lm\rangle = \alpha \hbar^2 |lm\rangle; \quad (258)$$

$$L^z |lm\rangle = m \hbar |lm\rangle. \quad (259)$$

Now we shall use the algebra (253) to deduce the possible values of the quantum numbers l and m . We shall see in section 10.4 that $\alpha = l(l+1)$.

⁴⁷i) $[xp_y, zp_y] = x[p_y, p_y]z = 0$ ii) $[xp_y, yp_z] = x[p_y, y]p_z = -i\hbar xp_z$

⁴⁸We have introduced the factors of \hbar so that l and m will just be dimensionless numbers.

10.3 Ladder operators

We know that the operators \mathbf{L}^2 and L^z transform a given eigenket into itself, by (258) and (259). It follows that the operators L^x and L^y must transform a given eigenket into something else. Our ‘way in’ to the problem involves investigating what L^x and L^y do to the eigenkets $|lm\rangle$.

In fact, it turns out that it is simpler to consider not the action of L^x and L^y , but of the **ladder operators** L^+ and L^- , which we define as

$$L^+ \equiv L^x + iL^y; \quad (260)$$

$$L^- \equiv L^x - iL^y. \quad (261)$$

It is not hard to show, using (253) and its cyclic partners, that these operators obey the following commutation relations:

$$[L^z, L^+] = \hbar L^+; \quad (262)$$

$$[L^z, L^-] = -\hbar L^-; \quad (263)$$

$$[L^+, L^-] = 2\hbar L^z. \quad (264)$$

Now let’s put these ladder operators to work. Consider the ‘new’ state

$$|?\rangle \equiv L^+|lm\rangle. \quad (265)$$

(The notation ‘?’ is simply intended to indicate that at the moment we don’t know any of the new state’s properties, so we don’t know what it should be labelled with.) To explore the properties of $|?\rangle$, let us try measuring its values of \mathbf{L}^2 and L^z :

$$\mathbf{L}^2|?\rangle = \mathbf{L}^2 L^+|lm\rangle \quad (266)$$

$$= \left(\underbrace{[\mathbf{L}^2, L^+]}_{=0} + L^+ \mathbf{L}^2 \right) |lm\rangle \quad (267)$$

$$= L^+ \mathbf{L}^2 |lm\rangle \quad (268)$$

$$= L^+ (l(l+1)\hbar^2) |lm\rangle \quad (269)$$

$$= (l(l+1)\hbar^2) (L^+|lm\rangle) \quad (270)$$

$$= (l(l+1)\hbar^2) |?\rangle. \quad (271)$$

Hence the new state $|?\rangle$ is an eigenstate of \mathbf{L}^2 , with eigenvalue $l(l+1)\hbar^2$. What about its value of L^z ? A similar calculation to (271) yields the following:

$$L^z|?\rangle = L^z L^+|lm\rangle \quad (272)$$

$$= \left(\underbrace{[L^z, L^+]}_{=\hbar L^+} + L^+ L^z \right) |lm\rangle \quad (273)$$

$$= \hbar L^+|lm\rangle + L^+ L^z|lm\rangle \quad (274)$$

$$= \hbar|?\rangle + L^+ (m\hbar|lm\rangle) \quad (275)$$

$$= (m+1)\hbar|?\rangle, \quad (276)$$

proving that $|?\rangle$ is an eigenstate of L^z with eigenvalue $(m+1)\hbar$. Now that we know its eigenvalues, we can label the new state with them:

$$|?\rangle \equiv L^+|lm\rangle = C_1|l, m+1\rangle. \quad (277)$$

(Notice that we *don't* know that the new state $|?\rangle$ is normalised — in fact it isn't! — so we have included an unknown constant C_1 to reflect this fact.) It is easy to show, using similar reasoning, that

$$L^-|lm\rangle = C_2|l, m-1\rangle. \quad (278)$$

In words: the **raising operator** L^+ increases the value of L^z by \hbar , while the **lowering operator** L^- reduces it by \hbar . Neither of these operators alters the value of \mathbf{L}^2 . If you'd like a rough analogy, we might say that the raising and lowering operators change the direction of the angular momentum vector \mathbf{L} (and hence its z -projection), but do not change its length.

✓ Determine $L^z(L^+)^2|lm\rangle$.⁴⁹

✓ Hence, what is $(L^+)^2|lm\rangle$?⁵⁰

10.4 Terminating the ladder; allowed values of l

It seems as if, by repeatedly acting with L^+ , we can make the value of m as large as we like. But this can't be right, because the z -projection of a vector should never be greater than its total length!; we must have that

$$(L^z)^2 \leq \mathbf{L}^2. \quad (279)$$

The only way to resolve this apparent contradiction is to suppose that there is some state which L^+ annihilates completely, terminating the ladder of states. Let's call the value of m in this state m_{\max} :

$$L^+|l, m_{\max}\rangle = 0. \quad (280)$$

We may use this information to calculate the value of α , since

$$\mathbf{L}^2|l, m_{\max}\rangle = \{(L^z)^2 + (L^x)^2 + (L^y)^2\}|l, m_{\max}\rangle \quad (281)$$

$$= \left\{ (L^z)^2 + \frac{1}{2}L^+L^- + \frac{1}{2}L^-L^+ \right\} |l, m_{\max}\rangle. \quad (282)$$

Now, we know what the first and third terms in the brackets do to the state, but we don't know how the second term acts on it. To resolve this problem, let us use the commutation relation

$$[L^+, L^-] = 2\hbar L^z \quad \rightarrow \quad L^+L^- = 2\hbar L^z + L^-L^+ \quad (283)$$

⁴⁹ $L^z L^+ (L^+ |lm\rangle) = ([L^z, L^+] + L^+ L^z) (L^+ |lm\rangle) = \hbar L^+ (L^+ |lm\rangle) + L^+ (L^z L^+ |lm\rangle) = \hbar L^+ (L^+ |lm\rangle) + L^+ (m+1)\hbar L^+ |lm\rangle = (m+2)\hbar (L^+)^2 |lm\rangle$

⁵⁰ $(L^+)^2 |lm\rangle = C_3 |l, m+2\rangle$ with some constant C_3 .

to write

$$\mathbf{L}^2|l, m_{\max}\rangle = \{(L^z)^2 + \hbar L^z + L^- L^+\} |l, m_{\max}\rangle \quad (284)$$

$$= (m_{\max}^2 \hbar^2 + m_{\max} \hbar^2) |l, m_{\max}\rangle \quad (285)$$

$$= m_{\max} (m_{\max} + 1) \hbar^2 |l, m_{\max}\rangle, \quad (286)$$

proving that the value of α is in fact $m_{\max}(m_{\max} + 1)$. It is conventional to use l to denote this maximum value of m , and to label the states accordingly with l and m .

✓What will be the second ladder termination condition? ⁵¹

10.5 Summary, and a puzzle

We have shown that the eigenstates of the angular momentum operators form sets (or ‘ladders’) $|l, m\rangle$, where

$$\mathbf{L}^2|l, m\rangle = l(l+1)\hbar^2|l, m\rangle; \quad (287)$$

$$L^z|l, m\rangle = m\hbar|l, m\rangle. \quad (288)$$

The value of l is the same for all states in a given ladder, while the value of m increases by one for every rung we go up the ladder. The highest allowed value of m is l ; the lowest allowed value of m is $-l$.

Therefore, since it must be possible to go from $m = l$ to $m = -l$ in integer steps, we deduce that $2l$ (the number of steps) must be an integer. Clearly if l is an integer, this is satisfied. These $l = 1$ ladders represent the orbital angular momentum of (for example) the electron in the hydrogen atom: $l = 0$ for the s -states, $l = 1$ for the p -states, $l = 2$ for the d -states, and so on.

But it seems from the algebra above that l could also be a half-integer. This wouldn’t be allowable if the particle had a spatial wave function, because the azimuthal part $e^{im\phi}$ would not be single-valued for half-integer m . But could these half-integer angular momenta show up in nature to describe some kind of ‘non-orbital’ angular momentum? Read on...

✓By considering the values of m , show that 1.9 would not be a possible value for l . ⁵²

✓Show that the azimuthal part $e^{im\phi}$ of the spherical harmonics (the $|l, m\rangle$) requires integer m and thus does not allow half-integer m . ⁵³

⁵¹ $L^-|l, m_{\min}\rangle = 0$. This condition leads to $m_{\min} = -l$.

⁵² m goes from $-l$ to l in integer steps, but -1.9, -0.9, 0.1, 1.1 does not fulfill this condition for m .

⁵³The azimuthal part must be singly valued, so that $e^{im\phi} = e^{im(\phi+2\pi)} = e^{im\phi}e^{im2\pi}$. This implies that $e^{im2\pi} = \cos(2m\pi) + i\sin(2m\pi) = 1$ and thus m must be an integer, not a half-integer.

11 Intrinsic angular momentum (‘spin’): the Stern-Gerlach experiment

In lecture 11, we derived the eigenvalues of the angular momentum operators from first principles. We noted that the algebra is consistent with the existence of half-integer values of l , as well as the usual integer ones. Are these realised in nature?

11.1 The Stern-Gerlach experiment

To answer this question, we need to ask what intrinsic angular momentum would look like. Stern and Gerlach reasoned semi-classically, as follows: Intrinsic angular momentum should correspond, roughly, to the particle spinning on its own axis⁵⁴. A charged sphere spinning on its axis constitutes a current distribution, and thus should act like a magnetic dipole $\boldsymbol{\mu}$. To be quantitative,

$$\boldsymbol{\mu} = g\mu_B\mathbf{S}, \quad (289)$$

where g and μ_B are constants (called respectively the g -factor and the Bohr magneton), and \mathbf{S} is the angular momentum vector corresponding to the rotation. Notice that we’re now calling this \mathbf{S} rather than \mathbf{L} . Conventionally, we use the notation \mathbf{S} for spin and \mathbf{L} for orbital angular momentum; but they have the same units and obey the same quantum mechanical algebra. As an extension of this convention, we label the eigenstates of orbital angular momentum $|l, m\rangle$, where l is an integer, but we label the eigenstates of spin angular momentum $|S, m_S\rangle$, where S is a half-integer. Try not to get confused by this!

This spin, like any magnetic moment, will couple to a magnetic field. Indeed, if placed in a field \mathbf{B} , its energy will be given by the usual electromagnetic formula

$$E = -\boldsymbol{\mu} \cdot \mathbf{B} = -\mu^z B = -g\mu_B B S^z, \quad (290)$$

where to get the third expression we have chosen the field to lie along the z -axis, and to get the fourth we have used (289). Classically, S^z may take a continuum of values; quantum mechanically, it is restricted to the values allowed by the algebra. In the simplest non-trivial case, $S = 1/2$, these values would just be $\pm\hbar/2$.

Since the energy is proportional to the magnetic field, a gradient in the magnetic field will create a physical force on the particles. If S^z is positive, it will push them in the positive z -direction; if S^z is negative, it will push them in the negative z -direction. Thus a magnetic field gradient should separate the particles according to their value of S^z . This is the operating principle of the Stern-Gerlach experiment, which is shown in Fig. 4(a).

The results are shown in Fig. 4(b). Contrary to classical expectation, we do not observe a continuous distribution of the beam; rather, we see two distinguishable spots, corresponding to the two different eigenvalues of S^z ($\hbar/2$ and $-\hbar/2$) allowed for quantum-mechanical particles.

✓For spin quantum number $S = 3/2$, what values of m_S are possible? What values of S^z are possible? ⁵⁵

⁵⁴You are warned against taking this analogy too seriously! — however, it’s certainly helpful for designing experiments.

⁵⁵ S and L obey the same relations. Thus, the spin magnetic quantum number m_S goes from $-S$ to S in integer steps. Thus, $m_S = -\frac{3}{2}, -\frac{1}{2}, \frac{1}{2}, \frac{3}{2}$. From $S^z = m_S\hbar$ it follows that $S^z = -\frac{3}{2}\hbar, -\frac{1}{2}\hbar, \frac{1}{2}\hbar, \frac{3}{2}\hbar$.

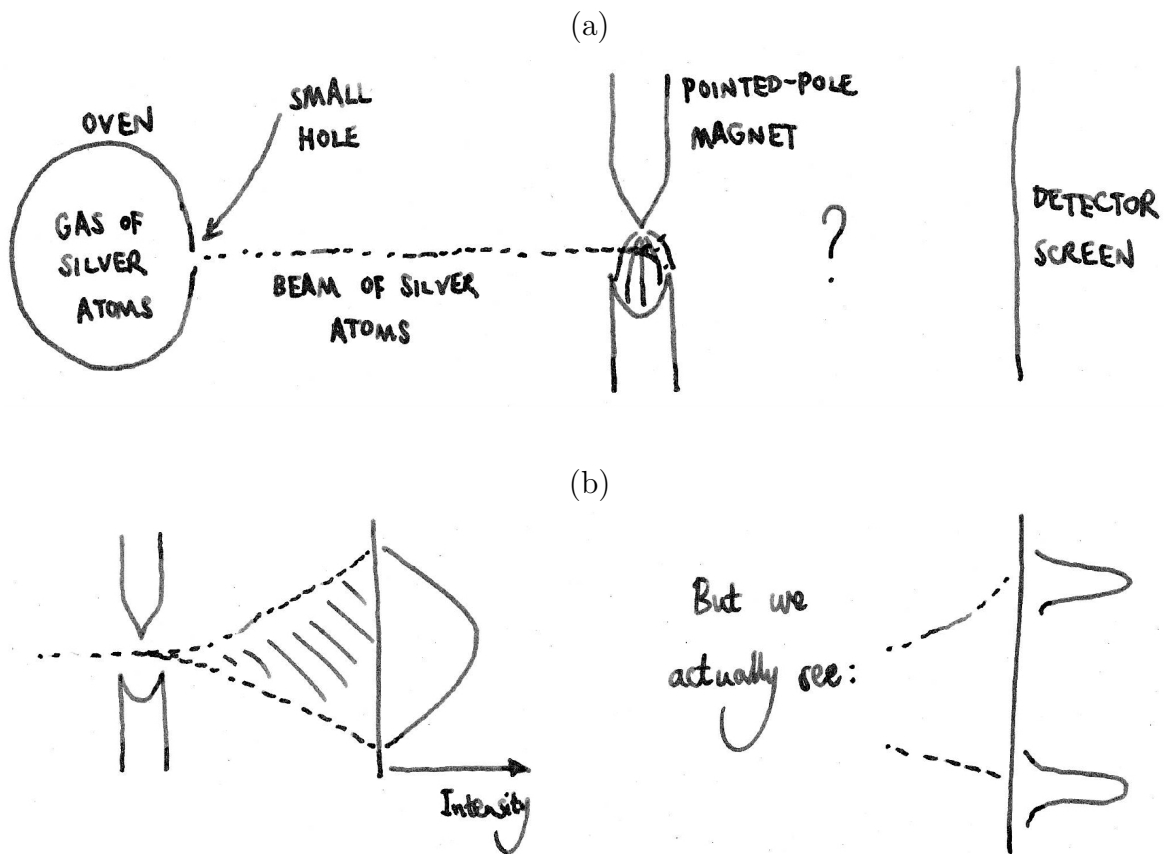


Figure 4: (a) A schematic of the Stern-Gerlach experiment. A classical mixture of spin- $1/2$ particles is prepared in a hot oven. It is collimated into a beam, and passed through a pointed-pole magnet. (b) The results of the experiment. Classically, one would expect a continuous spread of the beam (left), corresponding to the continuum of possible values of S^z ; in reality (right) two distinguishable spots appear on the detector screen, corresponding to the two allowed values of S^z .

✓What would you see in Fig. 4(b) if the beam particles had spin quantum number $S = 3/2$? ⁵⁶

11.2 Stern-Gerlach apparatuses in series

As an illustration of the properties of measurement in quantum mechanics, let us now ask what happens if we pass a beam of spin-1/2 particles through a series of Stern-Gerlach apparatuses with different orientations.

Consider first the systems shown in Fig. 5(i). The label ‘CM’ on the input beam means ‘classical mixture’, of the kind one would get from the oven in the original version of the experiment, and the 1 is a measure of its intensity. As we have already seen, passing this beam through a Stern-Gerlach apparatus with the field gradient in the z -direction (represented schematically here as a box labelled ‘ z ’) separates the beam into two equally intense components. We label these \uparrow_z and \downarrow_z , and write their intensities below them.

We might think that, since the z -projection of \mathbf{S} is near-maximal for the particles in these two beams, the x -projection would be near zero. But that would be to ignore quantum mechanics! We could re-do the algebra of lecture 11 for S^x rather than S^z , and we would reach exactly the same conclusion: that S^x is only allowed to take two values for a spin-1/2 particle, those being $\hbar/2$ and $-\hbar/2$. So the second Stern-Gerlach apparatus in Fig. 5(i) must separate the incoming beam into two beams as the first one did. These beams now correspond to the spin’s being respectively up or down along the x -axis. Since the x and z axes are orthogonal, these beams must — by symmetry — have equal intensities.

We might go further, as shown in Fig. 5(ii), and pass the \downarrow_x beam through a third apparatus. Since all the particles reaching this third apparatus came from the \downarrow_z beam of the first apparatus, we might think that they would all come out in the \downarrow_z beam of this one. But no! S^x and S^z are incompatible observables (their operators don’t commute), which means that the intervening S^x -measurement has destroyed any information about the z -projection. Hence the third S.-G. apparatus separates the beam into two beams of equal intensity, just as if it were operating on a classical mixture.

But what about the situation shown in Fig. 5(iii)? This is somewhere intermediate between the two simple cases:

1. $\theta = 0$. In this case, the second apparatus is also oriented in the positive z -direction, and thus all particles (having been pre-filtered into the $|\downarrow_z\rangle$ state) will emerge in the lower beam.
2. $\theta = \pi/2$. In this case, the second apparatus would be oriented in the positive x -direction, and thus (from the reasoning above) the intensities of the two emerging beams would be equal.

But how do we calculate the relative intensities when the angle is in between these two values? To do this, we need to develop a mathematical representation of the situation.

⁵⁶For $S = 3/2$, you would see four distinguishable spots on the detector screen, corresponding to the four allowed values of S^z .

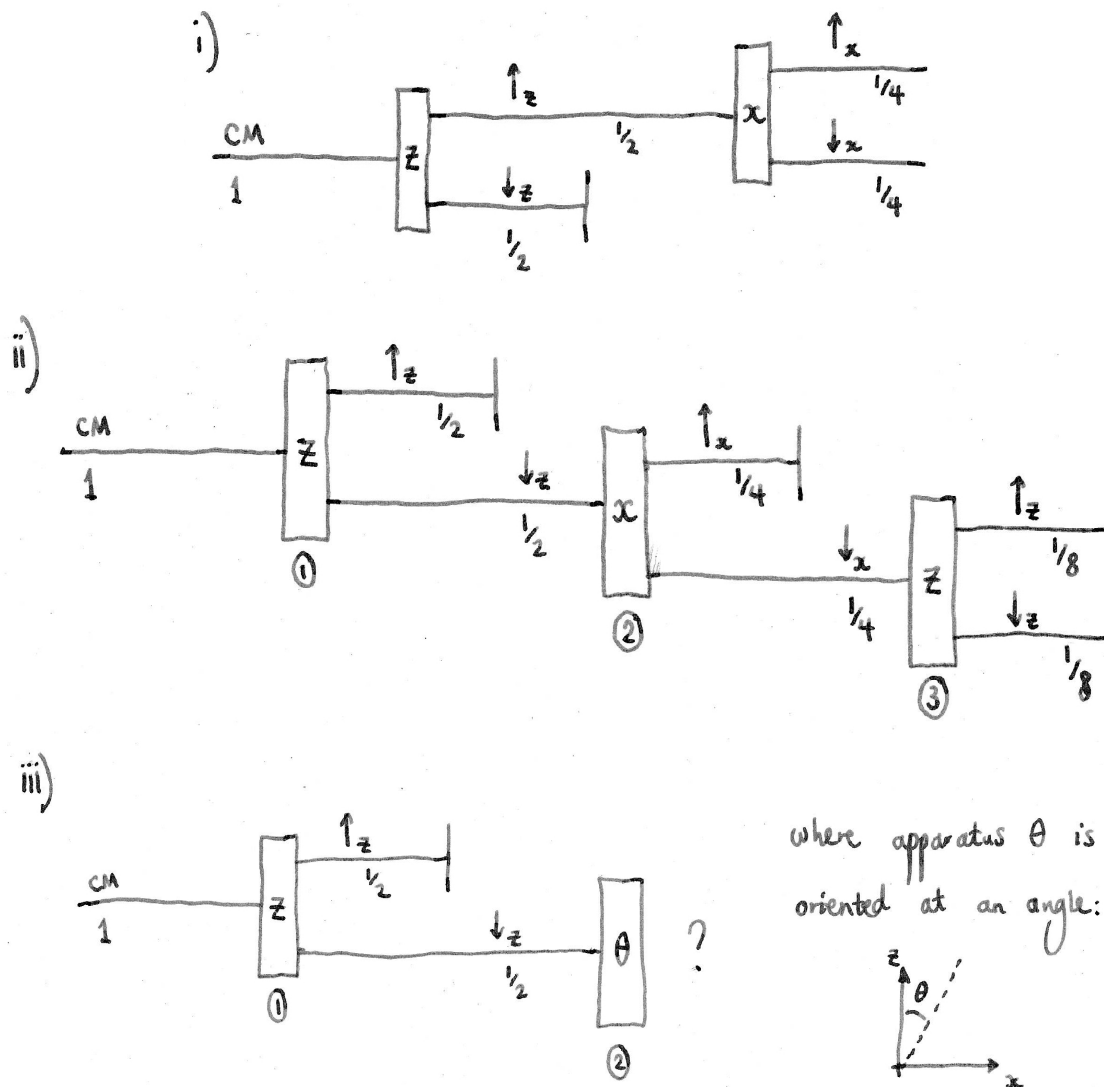


Figure 5: Multiple Stern-Gerlach apparatuses. (i) Measurement along one axis gives no information about an orthogonal one. (ii) Measurement along one axis destroys pre-existing information about an orthogonal one. (iii) A Stern-Gerlach apparatus along a non-principal axis.

11.3 The Pauli matrices

As we discussed at the beginning of the course, quantum mechanics is essentially a vector theory, with the eigenkets playing the rôle of basis vectors in an abstract space. It is useful, when dealing with systems with a very small number of basis states (two, in this case) to take this analogy very seriously. Let us represent the two possible projections of spin along the z -axis as two basis vectors in a two-dimensional space:

$$|\uparrow_z\rangle : \begin{pmatrix} 1 \\ 0 \end{pmatrix}; \quad |\downarrow_z\rangle : \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \quad (291)$$

An operator acts on a ket to give another ket. Hence if we represent the kets by vectors, we must represent operators by matrices. The following matrix representation is commonly used:

$$\mathbf{S}_x = \frac{\hbar}{2}\boldsymbol{\sigma}_x, \quad \mathbf{S}_y = \frac{\hbar}{2}\boldsymbol{\sigma}_y, \quad \mathbf{S}_z = \frac{\hbar}{2}\boldsymbol{\sigma}_z, \quad (292)$$

where the matrices

$$\boldsymbol{\sigma}_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \boldsymbol{\sigma}_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \boldsymbol{\sigma}_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad (293)$$

are called the **Pauli matrices**.

✓Using i) Dirac notation and ii) vector/matrix notation, write an expression for the expectation value of \hat{S}^y with respect to the spin $1/2$ state $|\uparrow_z\rangle$.⁵⁷

✓Using the ket notation $|S, m_S\rangle$, write down a basis for spin $S = 3/2$.⁵⁸

✓Do the Pauli matrices multiplied by $\frac{\hbar}{2}$ represent the spin component operators \mathbf{S}_x , \mathbf{S}_y , \mathbf{S}_z for a spin $S = 3/2$ particle?⁵⁹

⁵⁷i) $\langle \hat{S}^y \rangle = \langle \uparrow_z | \hat{S}^y | \uparrow_z \rangle$ ii) $\langle \hat{S}^y \rangle = \begin{pmatrix} 1 & 0 \end{pmatrix} \frac{\hbar}{2} \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix}$

⁵⁸ $|\frac{3}{2}, -\frac{3}{2}\rangle, |\frac{3}{2}, -\frac{1}{2}\rangle, |\frac{3}{2}, \frac{1}{2}\rangle, |\frac{3}{2}, \frac{3}{2}\rangle$

⁵⁹No, they do not! The spin $3/2$ vector space is 4-dimensional (there are 4 basis states). Thus, the spin-component operators will be represented by 4×4 matrices. The Pauli matrices are only for $S = 1/2$.

12 Stern-Gerlach apparatus at an arbitrary angle

In this lecture, we shall explore further the representation of spin-1/2 operators using Pauli matrices. We shall thereby derive an expression for the intensities of the two beams for a Stern-Gerlach apparatus at an arbitrary angle between the z - and x -axes.

12.1 Properties of the Pauli matrices

It is not difficult to check that the matrices $\mathbf{S}_a = \frac{\hbar}{2}\boldsymbol{\sigma}_a$ ($a = x, y, z$) obey the commutation relations for angular momentum operators:

$$[\mathbf{S}_x, \mathbf{S}_y] = \frac{\hbar^2}{4} [\boldsymbol{\sigma}_x, \boldsymbol{\sigma}_y] \quad (294)$$

$$= \frac{\hbar^2}{4} (\boldsymbol{\sigma}_x \boldsymbol{\sigma}_y - \boldsymbol{\sigma}_y \boldsymbol{\sigma}_x) \quad (295)$$

$$= \frac{\hbar^2}{4} \left\{ \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} - \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \right\} \quad (296)$$

$$= \frac{\hbar^2}{4} \left\{ \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix} - \begin{pmatrix} -i & 0 \\ 0 & i \end{pmatrix} \right\} \quad (297)$$

$$= \frac{\hbar^2}{2} \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix} \quad (298)$$

$$= i\hbar \underbrace{\frac{1}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}}_{=\boldsymbol{\sigma}_z} \quad (299)$$

$$= i\hbar \mathbf{S}_z. \quad (300)$$

Furthermore, the total squared-spin operator

$$\hat{\mathbf{S}}^2 \equiv (\hat{S}_x^2 + \hat{S}_y^2 + \hat{S}_z^2) \quad (301)$$

becomes, in the matrix representation,

$$\hat{\mathbf{S}}^2 : \mathbf{S}_x^2 + \mathbf{S}_y^2 + \mathbf{S}_z^2 \quad (302)$$

$$= \frac{\hbar^2}{4} \{ \boldsymbol{\sigma}_x^2 + \boldsymbol{\sigma}_y^2 + \boldsymbol{\sigma}_z^2 \} \quad (303)$$

$$= \frac{\hbar^2}{4} \left\{ \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} + \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} + \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \right\} \quad (304)$$

$$= \frac{\hbar^2}{4} \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right\} \quad (305)$$

$$= \frac{3\hbar^2}{4} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad (306)$$

$$= \frac{1}{2} \left(\frac{1}{2} + 1 \right) \hbar^2 \mathbf{I}, \quad (307)$$

in agreement with the general expression $S(S+1)\hbar^2$ for the eigenvalue of $\hat{\mathbf{S}}^2$.

✓What are the eigenvalues of \mathbf{S}^2 for a spin $S = 3/2$ particle? ⁶⁰

✓What would be the matrix representation of \mathbf{S}^2 for a spin $S = 3/2$ particle? ⁶¹

12.2 Eigenvectors of the spin projection operator in two principal directions

The states with definite spin-projection (up or down) along a certain axis are represented by the eigenvectors of the appropriate matrix. For example, to find the eigenvectors representing $|\uparrow_z\rangle$ and $|\downarrow_z\rangle$, we just find the eigenvectors of \mathbf{S}_z . The eigenvalues of \mathbf{S}_z are given by

$$\begin{vmatrix} \frac{\hbar}{2} - \lambda & 0 \\ 0 & -\frac{\hbar}{2} - \lambda \end{vmatrix} = 0; \quad (308)$$

expanding the determinant, this gives

$$\left(\frac{\hbar}{2} - \lambda\right) \left(-\frac{\hbar}{2} - \lambda\right) = 0, \quad (309)$$

and hence

$$\lambda = \pm \frac{\hbar}{2}. \quad (310)$$

(Of course, we already knew this; these eigenvalues of \mathbf{S}_z follow directly from the commutation algebra (300), as we showed in lecture 11.) To find the eigenvector associated with the positive eigenvalue, we solve the equation

$$\begin{pmatrix} \frac{\hbar}{2} - \lambda & 0 \\ 0 & -\frac{\hbar}{2} - \lambda \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \quad (311)$$

The top line is just $0 = 0$, while the bottom one tells us that $y = 0$. Hence the eigenvector associated with $\lambda = \hbar/2$ is

$$\mathbf{e}_+^z = \begin{pmatrix} 1 \\ 0 \end{pmatrix}. \quad (312)$$

Similarly, one may show that the eigenvector associated with $\lambda = -\hbar/2$ is

$$\mathbf{e}_-^z = \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \quad (313)$$

This is no surprise, of course, since we built our matrix representation on the basis that the states $|\uparrow_z\rangle$ and $|\downarrow_z\rangle$ were represented this way.

As another example, consider the eigenstates of the spin-projection along the x axis. These may be found similarly, but this time by finding the eigenvalues and eigenvectors of \mathbf{S}_x . This time

$$\begin{vmatrix} -\lambda & \frac{\hbar}{2} \\ \frac{\hbar}{2} & -\lambda \end{vmatrix} = 0; \quad (314)$$

⁶⁰The eigenvalues are $S(S+1)\hbar^2 = \frac{15}{4}\hbar^2$.

⁶¹ $\mathbf{S}^2 = \frac{15}{4}\hbar^2\mathbf{I}$, where \mathbf{I} is the 4×4 identity matrix.

expanding the determinant gives

$$\lambda^2 - \frac{\hbar^2}{4} = 0, \quad (315)$$

yielding the eigenvalues

$$\lambda = \pm \frac{\hbar}{2}. \quad (316)$$

(Same as \mathbf{S}_z ! — but then since the z direction was totally arbitrary, we must get exactly the same possible projections along any axis, including x .) The eigenvectors are found just as above; for the eigenvalue $\lambda = \hbar/2$, we have

$$\begin{pmatrix} -\frac{\hbar}{2} & \frac{\hbar}{2} \\ \frac{\hbar}{2} & -\frac{\hbar}{2} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \quad (317)$$

Both lines tell us that $x = y$, and hence the normalised eigenvector corresponding to the state $|\uparrow_x\rangle$ is

$$\mathbf{e}_+^x = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}. \quad (318)$$

The equivalent calculation for $\lambda = -\hbar/2$ gives the eigenvector representing $|\downarrow_x\rangle$ as

$$\mathbf{e}_-^x = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}. \quad (319)$$

Notice that \mathbf{e}_+^x may be written

$$\mathbf{e}_+^x = \frac{1}{\sqrt{2}} \left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\} = \frac{1}{\sqrt{2}} (\mathbf{e}_+^z + \mathbf{e}_-^z). \quad (320)$$

Translating this statement back into the ket language, it reads

$$|\uparrow_x\rangle = \frac{1}{\sqrt{2}} (|\uparrow_z\rangle + |\downarrow_z\rangle). \quad (321)$$

Hence the overlap between the states $|\uparrow_z\rangle$ and $|\uparrow_x\rangle$ is

$$\langle\uparrow_z|\uparrow_x\rangle = \frac{1}{\sqrt{2}} \left(\underbrace{\langle\uparrow_z|\uparrow_z\rangle}_{=1} + \underbrace{\langle\uparrow_z|\downarrow_z\rangle}_{=0} \right) = \frac{1}{\sqrt{2}}. \quad (322)$$

The square modulus of this,

$$|\langle\uparrow_z|\uparrow_x\rangle|^2 = \frac{1}{2}, \quad (323)$$

gives the probability that a particle entering a z -direction Stern-Gerlach apparatus in the state $|\uparrow_x\rangle$ will emerge in the upper beam.

✓ Find the eigenvalues of \mathbf{S}_y for a spin $S = 1/2$ particle. ⁶²

⁶² $|\mathbf{S}_y - \lambda \mathbf{I}| = 0$ gives $\begin{vmatrix} -\lambda & -i\frac{\hbar}{2} \\ i\frac{\hbar}{2} & -\lambda \end{vmatrix} = \lambda^2 - (-i\frac{\hbar}{2})(i\frac{\hbar}{2}) = \lambda^2 - \frac{\hbar^2}{4} = 0$ and thus $\lambda = \pm \frac{\hbar}{2}$.

12.3 Stern-Gerlach apparatus at angles θ and ϕ

We are now in a position to turn to our question of the transmission probability of a $|\uparrow_z\rangle$ beam through a Stern-Gerlach apparatus oriented at a polar angle θ to the positive z -direction in the x - z plane and an azimuthal angle ϕ in the x - y plane (an angle $\phi = 90^\circ$ corresponds to the $+y$ direction). The operator for the spin projection along this direction is

$$\mathbf{n} \cdot \hat{\mathbf{S}}, \quad (324)$$

where \mathbf{n} is a unit vector in the direction in question. In this case,

$$\mathbf{n} = \begin{pmatrix} \sin \theta \cos \phi \\ \sin \theta \sin \phi \\ \cos \theta \end{pmatrix}, \quad (325)$$

and hence

$$\hat{S}_n = \sin \theta \cos \phi \hat{S}_x + \sin \theta \sin \phi \hat{S}_y + \cos \theta \hat{S}_z. \quad (326)$$

In the matrix representation, this becomes

$$\mathbf{S}_n = \sin \theta \cos \phi \frac{\hbar}{2} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} + \sin \theta \sin \phi \frac{\hbar}{2} \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} + \cos \theta \frac{\hbar}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad (327)$$

$$= \begin{pmatrix} \frac{\hbar}{2} \cos \theta & \frac{\hbar}{2} \sin \theta e^{-i\phi} \\ \frac{\hbar}{2} \sin \theta e^{+i\phi} & -\frac{\hbar}{2} \cos \theta \end{pmatrix}. \quad (328)$$

Now we must find the eigenvalues and eigenvectors of this matrix. Its eigenvalues are given by

$$\left| \begin{array}{cc} \frac{\hbar}{2} \cos \theta - \lambda & \frac{\hbar}{2} \sin \theta e^{-i\phi} \\ \frac{\hbar}{2} \sin \theta e^{+i\phi} & -\frac{\hbar}{2} \cos \theta - \lambda \end{array} \right| = 0; \quad (329)$$

expanding the determinant, this becomes

$$\lambda^2 - \frac{\hbar^2}{4} \cos^2 \theta - \frac{\hbar^2}{4} \sin^2 \theta = 0. \quad (330)$$

Noting that $\cos^2 \theta + \sin^2 \theta = 1$, we may rewrite this as

$$\lambda^2 = \frac{\hbar^2}{4}, \quad (331)$$

which yields

$$\lambda = \pm \frac{\hbar}{2}, \quad (332)$$

just as above.

Now let us find the eigenvector corresponding to $\lambda = \hbar/2$. That is given by the equation

$$\begin{pmatrix} \frac{\hbar}{2} (\cos \theta - 1) & \frac{\hbar}{2} \sin \theta e^{-i\phi} \\ \frac{\hbar}{2} \sin \theta e^{+i\phi} & -\frac{\hbar}{2} (\cos \theta + 1) \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \quad (333)$$

The top line tells us that

$$x (\cos \theta - 1) + y \sin \theta e^{-i\phi} = 0, \quad (334)$$

i.e. that

$$y \sin \theta e^{-i\phi} = x (1 - \cos \theta). \quad (335)$$

Using $\sin 2\alpha = 2 \sin \alpha \cos \alpha$ and $\cos 2\alpha = 1 - 2 \sin^2 \alpha$, this may be rewritten

$$2y \sin \frac{\theta}{2} \cos \frac{\theta}{2} e^{-i\phi} = 2x \sin^2 \frac{\theta}{2}, \quad (336)$$

or, cancelling the common factors,

$$y \cos \frac{\theta}{2} e^{-i\phi} = x \sin \frac{\theta}{2}. \quad (337)$$

Clearly the choice

$$x = \cos \frac{\theta}{2}, \quad (338)$$

$$y = e^{+i\phi} \sin \frac{\theta}{2} \quad (339)$$

solves this equation, and by luck we don't even need an extra normalisation constant:

$$\mathbf{e}_+^n = \begin{pmatrix} \cos \frac{\theta}{2} \\ e^{+i\phi} \sin \frac{\theta}{2} \end{pmatrix}. \quad (340)$$

Back in the ket language, this reads

$$|\uparrow_n\rangle = \cos \frac{\theta}{2} |\uparrow_z\rangle + e^{+i\phi} \sin \frac{\theta}{2} |\downarrow_z\rangle. \quad (341)$$

If $\phi = 0$, then we obtain

$$|\uparrow_\theta\rangle = \cos \frac{\theta}{2} |\uparrow_z\rangle + \sin \frac{\theta}{2} |\downarrow_z\rangle. \quad (342)$$

Hence the overlap between the up-beam from a θ -apparatus with $\phi = 0$ and the up-beam from a z -apparatus is

$$\langle \uparrow_\theta | \uparrow_z \rangle = \cos \frac{\theta}{2}, \quad (343)$$

making the transmission probability in the upper beam

$$|\langle \uparrow_\theta | \uparrow_z \rangle|^2 = \cos^2 \left(\frac{\theta}{2} \right). \quad (344)$$

✓ Repeat the calculation $|\langle \uparrow_\theta | \uparrow_z \rangle|^2 = \cos^2 \left(\frac{\theta}{2} \right)$ from formulas (342) to (344) using vector notation. ⁶³

✓ What is $\langle \uparrow_n | \downarrow_z \rangle$? ⁶⁴

$$^{63} |\langle \uparrow_\theta | \uparrow_z \rangle|^2 = \left| \begin{pmatrix} \cos \frac{\theta}{2} & \sin \frac{\theta}{2} \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right|^2 = \cos^2 \left(\frac{\theta}{2} \right)$$

$$^{64} \langle \uparrow_n | \downarrow_z \rangle = \begin{pmatrix} \cos \frac{\theta}{2} & e^{-i\phi} \sin \frac{\theta}{2} \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = e^{-i\phi} \sin \left(\frac{\theta}{2} \right)$$

12.4 How can we picture the spin state of an electron? The Bloch sphere

How can we picture the spin state of an electron with its spin in an arbitrary direction, not specifically along a coordinate axis? Classically, the state of a spinning particle is represented on the surface of a sphere, with the vector from the origin to a point on the surface of the sphere indicating the rotation axis of the spin. How does this translate to the spin states in quantum mechanics? We immediately notice a problem: the state vectors of electron spin are two-dimensional (the column vectors have two entries, corresponding to two distinguishable measurement outcomes), while the classical states for spin are vectors in three-dimensions. However, we have also seen that the state vectors can have complex components: it turns out that two-dimensional complex state vectors can be represented in terms of three-dimensional real vectors. We will now see how this is done.

The general spin state can be written as

$$|\psi(\theta, \phi)\rangle = \cos\left(\frac{\theta}{2}\right) |\uparrow_z\rangle + e^{i\phi} \sin\left(\frac{\theta}{2}\right) |\downarrow_z\rangle \quad (345)$$

We now identify a unit vector with this state:

$$\vec{n} = \begin{pmatrix} \cos\phi \sin\theta \\ \sin\phi \sin\theta \\ \cos\theta \end{pmatrix}. \quad (346)$$

The vectors \vec{n} form a sphere of radius one in three dimensions. Since each state $|\psi(\theta, \phi)\rangle$ corresponds to a unique vector \vec{n} , the entire state space of the electron spin is *also* a sphere in three dimensions. This is called the **Bloch sphere** and it is the state space of the electron spin. If the angles θ and ϕ depend on time, then the state of the electron spin evolves over time, tracing out a path on the Bloch sphere.

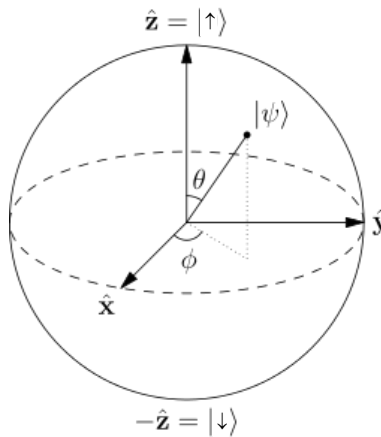


Figure 6: Bloch Sphere

We'll look here at one property of the Bloch sphere. If the spin vector is pointing in the direction given by \vec{n} , then we obtain the spin in the opposite direction by making the substitution $\theta \rightarrow \pi - \theta$ and $\phi \rightarrow \phi + \pi$. This will take \vec{n} to $-\vec{n}$. Since these are opposite spin directions, a measurement in a Stern-Gerlach experiment along this direction can perfectly distinguish between these two spins, and the corresponding quantum states should be orthogonal. We define $|\psi_1\rangle = |\psi(\theta, \phi)\rangle$ and $|\psi_2\rangle = |\psi(\pi - \theta, \phi + \pi)\rangle$, and the inner product $\langle\psi_1|\psi_2\rangle$ should be equal to zero. First, we calculate $\langle\psi_1|$:

$$\langle\psi_1| = \cos\left(\frac{\theta}{2}\right) \langle\uparrow_z| + e^{-i\phi} \sin\left(\frac{\theta}{2}\right) \langle\downarrow_z| \quad (347)$$

Notice the sign change in ϕ due to the complex conjugate. The state $|\psi_2\rangle$ with $\theta \rightarrow \pi - \theta$ and $\phi \rightarrow \phi + \pi$ becomes

$$|\psi_2\rangle = \cos\left(\frac{\pi - \theta}{2}\right) |\uparrow_z\rangle + e^{i(\phi + \pi)} \sin\left(\frac{\pi - \theta}{2}\right) |\downarrow_z\rangle \quad (348)$$

$$|\psi_2\rangle = \sin\left(\frac{\theta}{2}\right) |\uparrow_z\rangle + e^{i\phi} e^{i\pi} \cos\left(\frac{\theta}{2}\right) |\downarrow_z\rangle \quad (349)$$

$$|\psi_2\rangle = \sin\left(\frac{\theta}{2}\right) |\uparrow_z\rangle - e^{i\phi} \cos\left(\frac{\theta}{2}\right) |\downarrow_z\rangle. \quad (350)$$

The inner product then becomes

$$\langle\psi_1|\psi_2\rangle = \cos\left(\frac{\theta}{2}\right) \sin\left(\frac{\theta}{2}\right) - e^{-i\phi} e^{i\phi} \cos\left(\frac{\theta}{2}\right) \sin\left(\frac{\theta}{2}\right) = 0 \quad (351)$$

as expected.

This means that antipodal (at opposite sides) states in the Bloch sphere are orthogonal spin states. Moreover, each axis passing through the origin in the Bloch sphere connects two antipodal points, and therefore each axis through the origin can be seen as a spin component observable.

✓ From the expression for $|\uparrow_n\rangle$, what are the values of θ and ϕ for the state $|\uparrow_x\rangle$? Where on the Bloch sphere is this state located? ⁶⁵

12.5 Picturing time evolution on the Bloch sphere

Let's now imagine we want to track the time development of a quantum state on the Bloch sphere.

Imagine that our Hamiltonian \mathbf{H} is an operator proportional to the Pauli matrix σ_z :

$$\mathbf{H} = \frac{\hbar\epsilon}{2} \sigma_z \quad (352)$$

⁶⁵ $|\uparrow_x\rangle = \frac{1}{\sqrt{2}}(|\uparrow_z\rangle + |\downarrow_z\rangle) = \cos\left(\frac{\theta}{2}\right) |\uparrow_z\rangle + e^{i\phi} \sin\left(\frac{\theta}{2}\right) |\downarrow_z\rangle$. Thus, $\cos\left(\frac{\theta}{2}\right) = \sin\left(\frac{\theta}{2}\right) = \frac{1}{\sqrt{2}}$ so that $\theta = \frac{\pi}{2}$ and $e^{i\phi} = 1$ so that $\phi = 0$. Thus, the state is located along the $+x$ direction on the Bloch sphere.

where the eigenstates of the Hamiltonian, which are obviously $|\uparrow_z\rangle$ and $|\downarrow_z\rangle$ have eigenvalues of $\pm \frac{\hbar\epsilon}{2}$. Now consider an initial and general state on the Bloch sphere

$$|\psi(0)\rangle = \cos\left(\frac{\theta_0}{2}\right) |\uparrow_z\rangle + \sin\left(\frac{\theta_0}{2}\right) e^{i\phi_0} |\downarrow_z\rangle. \quad (353)$$

The time evolution operator is $\hat{T}(t) = \exp(-i\hat{H}t/\hbar)$ which for the Hamiltonian above (Eq. 352) becomes

$$\hat{T}(t) = \exp\left(-\frac{i\epsilon t}{2}\right) |\uparrow_z\rangle\langle\uparrow_z| + \exp\left(\frac{i\epsilon t}{2}\right) |\downarrow_z\rangle\langle\downarrow_z|. \quad (354)$$

Then

$$|\psi(t)\rangle = \exp\left(-\frac{i\epsilon t}{2}\right) \cos\left(\frac{\theta_0}{2}\right) |\uparrow_z\rangle + \exp\left(\frac{i\epsilon t}{2} + \phi_0\right) \sin\left(\frac{\theta_0}{2}\right) |\downarrow_z\rangle \quad (355)$$

$$|\psi(t)\rangle = \exp\left(-\frac{i\epsilon t}{2}\right) \left[\cos\left(\frac{\theta_0}{2}\right) |\uparrow_z\rangle + \exp(i\epsilon t + \phi_0) \sin\left(\frac{\theta_0}{2}\right) |\downarrow_z\rangle \right]. \quad (356)$$

Dropping the irrelevant global phase yields

$$|\psi(t)\rangle = \cos\left(\frac{\theta_0}{2}\right) |\uparrow_z\rangle + \exp[i\phi(t)] \sin\left(\frac{\theta_0}{2}\right) |\downarrow_z\rangle \quad (357)$$

with $\phi(t) = \phi_0 + \epsilon t$. Thus we find that the Bloch vector has an azimuthal angle that rotates with angular frequency ϵ , whereas the polar angle remains fixed. This motion corresponds to a rotating about the z axis of the Bloch sphere - which is the axis defined by connecting the Hamiltonian eigenstates.

We chose a Hamiltonian proportional to σ_z since it makes the analysis simple. However, there is nothing special about having the eigenstates aligned along this axis - by rotational symmetry if we had chosen the a Hamiltonian proportional to σ_x then the time evolution would have been to precess around the line eigenstates of σ_x , which is nothing more than the x axis in Fig. 6.

So we can conclude: *the time evolution of a quantum state on the Bloch sphere is a rotation about the axis connecting the Hamiltonian eigenstates, at a frequency given by the difference in their energies divided by \hbar .*

13 Multi-spin problems: the Heisenberg Hamiltonian

In this lecture, we shall make our first foray into a topic that you have not previously touched in quantum mechanics, viz. how to find the eigenkets and eigenvalues when the system consists of more than one particle. To make things easy, we'll start off by considering an insulator: this is easier because in an insulator the electrons are fixed on their lattice positions, and so the only thing that can change is the orientation of their spins.

13.1 The two-site Heisenberg Hamiltonian

Let's start with the case in which there are only two electrons, each fixed in space. The first question is: what Hamiltonian describes the interaction between their spins? This can be derived from a microscopic basis⁶⁶, but we can in fact work it out just by symmetry.

As we saw last time, each spin is represented by a vector of operators. Let's call these vectors $\hat{\mathbf{S}}_1$ and $\hat{\mathbf{S}}_2$, referring respectively to the spin of the first and second electrons in the problem. Their components are⁶⁷

$$\hat{\mathbf{S}}_1 = \begin{pmatrix} \hat{S}_1^x \\ \hat{S}_1^y \\ \hat{S}_1^z \end{pmatrix}; \quad \hat{\mathbf{S}}_2 = \begin{pmatrix} \hat{S}_2^x \\ \hat{S}_2^y \\ \hat{S}_2^z \end{pmatrix}. \quad (358)$$

The spin operators on a given site obey the usual anticommutation relations; those on different sites commute with each other.

The Hamiltonian must be written in terms of these vectors of spin operators. Furthermore, it should not pick out a special direction, and it should of course be a scalar (as it's an energy). There's only one Hamiltonian that satisfies this requirement: the dot product of $\hat{\mathbf{S}}_1$ and $\hat{\mathbf{S}}_2$. Hence the Hamiltonian must be

$$\hat{H} = J \hat{\mathbf{S}}_1 \cdot \hat{\mathbf{S}}_2. \quad (359)$$

This is called the **two-site Heisenberg Hamiltonian**⁶⁸. The sign of the constant J affects the physics, since if $J < 0$ the energy is minimised by aligning the spins (**ferromagnetic** behaviour), while if $J > 0$ it is minimised by anti-aligning the spins (**antiferromagnetic** behaviour). Thus if the spins were classical vectors, the ground state would be:

- if $J < 0$, \mathbf{S}_1 oriented arbitrarily, and \mathbf{S}_2 parallel to \mathbf{S}_1 ;
- if $J > 0$, \mathbf{S}_1 oriented arbitrarily, and \mathbf{S}_2 antiparallel⁶⁹ to \mathbf{S}_1 .

⁶⁶See the PH5014 'Interacting Electron Problem in Solids' course for that.

⁶⁷Note that we are now writing the component (x , y , or z) as a superscript, to make way for the particle index (1 or 2) as a subscript.

⁶⁸Because it was first written down by Heitler and London...

⁶⁹The somewhat upsetting term 'antiparallel' means 'oriented along the same axis but in the opposite direction'.

13.2 Quantum-mechanical solution

13.2.1 Basis kets

But the spins are not classical vectors — they are quantum-mechanical operators. Let us therefore try to find the eigenkets of the Hamiltonian and their eigenvalues: this will tell us the energy spectrum of the problem.

As for the single spin, let's use a matrix representation of the states. There are four of them:

- $|\uparrow\uparrow\rangle$, in which both spins have z -projection $+\hbar/2$;
- $|\uparrow\downarrow\rangle$, in which the first spin has z -projection $+\hbar/2$ while the second has z -projection $-\hbar/2$;
- $|\downarrow\uparrow\rangle$, in which the first spin has z -projection $-\hbar/2$ while the second has z -projection $+\hbar/2$;
- $|\downarrow\downarrow\rangle$, in which both spins have z -projection $-\hbar/2$.

✓What would be the dimension of the matrix representing the two-site Heisenberg Hamiltonian $\hat{H} = J\hat{\mathbf{S}}_1 \cdot \hat{\mathbf{S}}_2$? ⁷⁰

✓How could you write the four basis states $|\uparrow\uparrow\rangle, |\uparrow\downarrow\rangle, |\downarrow\uparrow\rangle, |\downarrow\downarrow\rangle$ as vectors? ⁷¹

13.2.2 Rewriting the Hamiltonian

We need to find out how the Hamiltonian acts on these basis states. To do this, let's rewrite it⁷² in terms of the ladder operators S^+ and S^- . Noting that

$$S^x = \frac{1}{2}(S^+ + S^-); \quad (360)$$

$$S^y = \frac{1}{2i}(S^+ - S^-), \quad (361)$$

we may rewrite the two-site Heisenberg Hamiltonian as follows:

$$H = J\mathbf{S}_1 \cdot \mathbf{S}_2 \quad (362)$$

$$= J[S_1^z S_2^z + S_1^x S_2^x + S_1^y S_2^y] \quad (363)$$

$$= J\left[S_1^z S_2^z + \frac{1}{4}(S_1^+ + S_1^-)(S_2^+ + S_2^-) - \frac{1}{4}(S_1^+ - S_1^-)(S_2^+ - S_2^-)\right] \quad (364)$$

⁷⁰ \mathbf{H} would be a 4×4 matrix, as there are four basis states so that the vector space for the two spins is 4-dimensional.

⁷¹Each basis state will be represented by a 4×1 column vector. The assignment is arbitrary, but we could choose $|\uparrow\uparrow\rangle = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}; |\uparrow\downarrow\rangle = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}; |\downarrow\uparrow\rangle = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}; |\downarrow\downarrow\rangle = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}$

⁷²Notice that we've again dropped the hats for brevity — remember, though, that all these capital S 's are operators.

$$= J \left[S_1^z S_2^z + \frac{1}{2} (S_1^+ S_2^- + S_1^- S_2^+) \right]. \quad (365)$$

The effect of these terms is clearer now. The first ($S^z S^z$) term just measures the z -projection of each spin without changing it — hence it is sometimes referred to as the classical (or **Ising**⁷³) term. The other terms alter the state, by flipping neighbouring pairs of spins: $\uparrow\downarrow \leftrightarrow \downarrow\uparrow$. These terms are sometimes said to create **quantum fluctuations**.

13.2.3 Effects of H on the basis states

✓ Recall that $S^+|\downarrow\rangle = \hbar|\uparrow\rangle$ and $S^-|\uparrow\rangle = \hbar|\downarrow\rangle$. What would be $S_1^+|\downarrow\uparrow\rangle$? What would be $S_2^+|\downarrow\uparrow\rangle$?⁷⁴

Let us now consider the effect of the Hamiltonian (365) on each of the basis states given above.

- When H acts on $|\uparrow\uparrow\rangle$, each operator in the Ising term returns $+\hbar/2$, resulting in $\hbar^2 J/4$ overall. The quantum fluctuation terms give zero, since the S^+ term in each will kill the state. Hence

$$H|\uparrow\uparrow\rangle = \frac{\hbar^2 J}{4} |\uparrow\uparrow\rangle, \quad (366)$$

i.e. the classical state $|\uparrow\uparrow\rangle$ is already an eigenstate of H , with eigenvalue $\hbar^2 J/4$.

- When H acts on $|\uparrow\downarrow\rangle$, S_1^z returns $+\hbar/2$ but S_2^z returns $-\hbar/2$; the Ising term as a whole therefore returns $-\hbar^2 J/4$. This time the quantum fluctuation terms can also act, producing a term in the result that is proportional to $|\downarrow\uparrow\rangle$ (the result of doing a nearest-neighbour spin exchange on $|\uparrow\downarrow\rangle$). Adding this in with the correct prefactor, we obtain

$$H|\uparrow\downarrow\rangle = -\frac{\hbar^2 J}{4} |\uparrow\downarrow\rangle + \frac{\hbar^2 J}{2} |\downarrow\uparrow\rangle. \quad (367)$$

- Similar reasoning yields the action of H on the two remaining basis vectors⁷⁵:

$$H|\downarrow\uparrow\rangle = -\frac{\hbar^2 J}{4} |\downarrow\uparrow\rangle + \frac{\hbar^2 J}{2} |\uparrow\downarrow\rangle; \quad (368)$$

$$H|\downarrow\downarrow\rangle = \frac{\hbar^2 J}{4} |\downarrow\downarrow\rangle. \quad (369)$$

We conclude that the kets $|\uparrow\uparrow\rangle$ and $|\downarrow\downarrow\rangle$ are already eigenkets of the two-site Heisenberg Hamiltonian, with eigenenergies $\hbar^2 J/4$. The other two basis kets are mixed by H .

✓ Show equation (369) explicitly.⁷⁶

⁷³After the physicist/mathematician who first studied the model with only this term in its Hamiltonian.

⁷⁴The operator S_1^+ only acts on spin 1 and leaves spin 2 unchanged. Thus, $S_1^+|\downarrow\uparrow\rangle = \hbar|\uparrow\uparrow\rangle$. The operator S_2^+ only acts on spin 2 and leaves spin 1 unchanged. Thus, $S_2^+|\downarrow\uparrow\rangle = 0$ (ladder termination condition).

⁷⁵These can also be obtained just by changing z to $-z$, which changes \uparrow to \downarrow and vice versa. The Hamiltonian, being a dot product, is invariant under this change, so its action on the transformed states must be the same as on the original ones.

⁷⁶ $H|\downarrow\downarrow\rangle = J [S_1^z S_2^z + \frac{1}{2} (S_1^+ S_2^- + S_1^- S_2^+)] |\downarrow\downarrow\rangle = J S_1^z S_2^z |\downarrow\downarrow\rangle$ as $S_2^-|\downarrow\downarrow\rangle = 0$ and $S_1^-|\downarrow\downarrow\rangle = 0$. Thus, $H|\downarrow\downarrow\rangle = J (-\frac{\hbar}{2}) (-\frac{\hbar}{2}) |\downarrow\downarrow\rangle = \frac{\hbar^2 J}{4} |\downarrow\downarrow\rangle$

13.2.4 Matrix representation

Finding the linear combinations of $|\uparrow\downarrow\rangle$ that are not mixed by H is again a problem in linear algebra. We therefore introduce a vector representation of the states, in which

$$|\uparrow\downarrow\rangle : \begin{pmatrix} 1 \\ 0 \end{pmatrix}; \quad |\downarrow\uparrow\rangle : \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \quad (370)$$

The Hamiltonian (more precisely the part of it that acts on these two basis kets) is accordingly written as a matrix \mathbf{H} . The equations (367) and (368) are accordingly written

$$\mathbf{H} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = -\frac{\hbar^2 J}{4} \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \frac{\hbar^2 J}{2} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -\alpha \\ 2\alpha \end{pmatrix}, \quad (371)$$

$$\mathbf{H} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = -\frac{\hbar^2 J}{4} \begin{pmatrix} 0 \\ 1 \end{pmatrix} + \frac{\hbar^2 J}{2} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 2\alpha \\ -\alpha \end{pmatrix}, \quad (372)$$

$$(373)$$

where we have defined the energy

$$\alpha \equiv \frac{\hbar^2 J}{4} \quad (374)$$

just to keep things neat. Since we know the action of \mathbf{H} on both basis vectors, we know all of its entries, and can simply write

$$\mathbf{H} = \begin{pmatrix} -\alpha & 2\alpha \\ 2\alpha & -\alpha \end{pmatrix}. \quad (375)$$

where the first column is the action of \mathbf{H} on the $(1, 0)$ basis vector and the second is the action of \mathbf{H} on the $(0, 1)$ basis vector.

We have solved this kind of problem before — for example, in the degenerate perturbation theory section of this very course! The eigenvalues are given by

$$\begin{vmatrix} -\alpha - \lambda & 2\alpha \\ 2\alpha & -\alpha - \lambda \end{vmatrix} = 0; \quad (376)$$

expanding the determinant, this becomes

$$(\alpha + \lambda)^2 = 4\alpha^2, \quad (377)$$

and hence

$$\lambda + \alpha = \pm 2\alpha \quad \rightarrow \quad \lambda = -3\alpha, \alpha. \quad (378)$$

The eigenvectors associated with (respectively) the positive and negative eigenvalues are

$$\mathbf{e}_+ = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \mathbf{e}_- = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix} \quad (379)$$

as usual.

13.2.5 Eigenkets and eigenvalues: summary

Copying the eigenvectors (379) back into the ket language, and combining them with the two eigenkets we already know, we find the following table of eigenkets and eigenenergies for the two-site Heisenberg model:

Eigenket	Eigenenergy
$ \uparrow\uparrow\rangle$	
$\frac{1}{\sqrt{2}}(\uparrow\downarrow\rangle + \downarrow\uparrow\rangle)$	$\frac{\hbar^2 J}{4}$
$ \downarrow\downarrow\rangle$	
$\frac{1}{\sqrt{2}}(\uparrow\downarrow\rangle - \downarrow\uparrow\rangle)$	$-\frac{3\hbar^2 J}{4}$

Surely it can't be accident that we have ended up with a triplet of degenerate states at $\hbar^2 J/4$? We'll take up this question in the next lecture.

✓ From your knowledge of two-particle entangled states from PH2012, which if any of these eigenkets are entangled spin states? ⁷⁷

⁷⁷The eigenkets $\frac{1}{\sqrt{2}}(|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle)$ and $\frac{1}{\sqrt{2}}(|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle)$ are entangled, as they can not be written as the product of the state of spin 1 and the state of spin 2. If the z -components of the two spins were measured for many pairs, the outcomes for spin 1 and spin 2 individually would be completely random (on average as many outcomes $+\frac{\hbar}{2}$ as $-\frac{\hbar}{2}$), but the outcomes for each pair would be perfectly anticorrelated (the two outcomes would always be opposite).

14 Adding angular momenta in quantum mechanics

Last time we solved, by brute force, the two-site Heisenberg model. We found that the eigenstates were grouped together in energy: a triplet of states at $\hbar^2 J/4$, and a singlet state at $-3\hbar^2 J/4$. The chances of a randomly selected 4×4 matrix having three eigenvalues the same is practically zero — there must be some reason why it has happened in this case. Read on...

14.1 The operator for total spin

What we are doing with the Heisenberg model is coupling two spins together. It therefore seems natural to consider their sum. The classical sum of two vectors would be written

$$\mathbf{S}_T = \mathbf{S}_1 + \mathbf{S}_2, \quad (380)$$

where the ‘T’ stands for ‘total’. Therefore (by the correspondence principle) the operator for the total spin must obey a similar equation:

$$\hat{\mathbf{S}}_T = \hat{\mathbf{S}}_1 + \hat{\mathbf{S}}_2. \quad (381)$$

Two important facts follow from this:

- The components of $\hat{\mathbf{S}}_T$ obey the angular momentum commutation relations, just as the components of $\hat{\mathbf{S}}_1$ and $\hat{\mathbf{S}}_2$ do.
- Squaring equation (381) leads to⁷⁸

$$\mathbf{S}_T^2 = (\mathbf{S}_1 + \mathbf{S}_2)^2 = \mathbf{S}_1^2 + \mathbf{S}_2^2 + 2\mathbf{S}_1 \cdot \mathbf{S}_2. \quad (382)$$

Looking at (382), we realise that we know the eigenvalues of all the terms on the right-hand side. Since the first spin has $S_1 = 1/2$, we know that the eigenvalue of \mathbf{S}_1^2 is (according to the usual rule) $S_1(S_1 + 1)\hbar^2$, which in this case gives $3\hbar^2/4$. Since the second spin also has $S_2 = 1/2$, the second term gives the same result. And the third term is proportional to the Hamiltonian of the two-site Heisenberg model! Since we’ve already worked out that the eigenvalues of $J\mathbf{S}_1 \cdot \mathbf{S}_2$ are $\hbar^2 J/4$ (for the triplet states) and $-3\hbar^2 J/4$ (for the singlet state), we may set $J = 2$ to discover that the eigenvalues of $2\mathbf{S}_1 \cdot \mathbf{S}_2$ are $\hbar^2/2$ (for the triplet states) and $-3\hbar^2/2$ (for the singlet state).

Putting these into (382), we find that the eigenvalues are:

$$\mathbf{S}_T^2 : \begin{cases} \frac{3}{4}\hbar^2 + \frac{3}{4}\hbar^2 + \frac{1}{2}\hbar^2 = 2\hbar^2 & \text{(triplets)} \\ \frac{3}{4}\hbar^2 + \frac{3}{4}\hbar^2 - \frac{3}{2}\hbar^2 = 0 & \text{(singlet)} \end{cases} \quad (383)$$

On the other hand, since the operator \mathbf{S}_T^2 is the square of an angular momentum, its eigenvalues must follow the usual rule:

$$\mathbf{S}_T^2 : S_T(S_T + 1)\hbar^2. \quad (384)$$

⁷⁸Hats dropped for brevity.

Comparing these two, we see that the triplet states correspond to $S_T = 1$, while the singlet states correspond to $S_T = 0$.

✓ For the total spin quantum number $S_T = 1$, what values of m_{S_T} are possible? What values of S_T^z are possible? ⁷⁹

✓ The operator $S_T^z = S_1^z + S_2^z$. What is $S_T^z | \uparrow\uparrow \rangle$? What is $S_T^z | \uparrow\downarrow \rangle$? What is $S_T^z | \downarrow\downarrow \rangle$? Hence, what are the values of S_T and m_{S_T} for the three triplet states? ⁸⁰

The reason that the three triplet states are degenerate in energy is that they differ only in the z -projection of the total spin; but the Hamiltonian

$$H = J\mathbf{S}_1 \cdot \mathbf{S}_2 = \frac{J}{2} \{ \mathbf{S}_T^2 - \mathbf{S}_1^2 - \mathbf{S}_2^2 \} \quad (385)$$

has no dependence on S_T^z .

14.2 Rules for coupling angular momenta

Let us reflect on what we have just discovered. We coupled together two spin-1/2 particles, and found that the total spin quantum number of the resulting states was either 1 or 0. Let us represent this as⁸¹

$$\frac{1}{2} \otimes \frac{1}{2} = 0 \oplus 1. \quad (386)$$

This, it transpires, is a particular instance of a general rule for coupling angular momenta in quantum mechanics. We state the general rule without proof:

$$S_1 \otimes S_2 = \{|S_1 - S_2|\} \oplus \{|S_1 - S_2| + 1\} \oplus \cdots \oplus \{S_1 + S_2\}. \quad (387)$$

In words, it says that the energy spectrum when we couple a spin S_1 to S_2 will contain levels corresponding to any total spin between $|S_1 - S_2|$ and $S_1 + S_2$ in integer steps. For example, if we couple a spin-5/2 particle to a spin-2 one, the rule would give

$$\frac{5}{2} \otimes 2 = \frac{1}{2} \oplus \frac{3}{2} \oplus \frac{5}{2} \oplus \frac{7}{2} \oplus \frac{9}{2} \quad (388)$$

As a check, we can make sure that the total number of states for the coupled system comes out right:

- Before the coupling, the spin-5/2 had six states, corresponding to the six possible z -projections ($-5/2, -3/2, -1/2, 1/2, 3/2$, and $5/2$), while the spin-2 had five states (with z -projections $-2, -1, 0, 1$, and 2 respectively). The total number of states was therefore the product of these, viz. 30.

⁷⁹ S_T obeys the same relations as S . Thus, the total spin magnetic quantum number m_{S_T} goes from $-S_T$ to S_T in integer steps. Thus, $m_{S_T} = -1, 0, 1$. From $S_T^z = m_{S_T}\hbar$ it follows that $S_T^z = -\hbar, 0, \hbar$.

⁸⁰ $S_T^z | \uparrow\uparrow \rangle = (S_1^z + S_2^z) | \uparrow\uparrow \rangle = \frac{\hbar}{2} | \uparrow\uparrow \rangle + \frac{\hbar}{2} | \uparrow\uparrow \rangle = \hbar | \uparrow\uparrow \rangle$. Similarly, one finds $S_T^z | \uparrow\downarrow \rangle = 0$ and $S_T^z | \downarrow\downarrow \rangle = -\hbar | \downarrow\downarrow \rangle$. Thus, the three triplet states all have total spin quantum number $S_T = 1$ but different quantum numbers $m_{S_T} = +1, 0, -1$ respectively.

⁸¹We're using here a notation of direct sums and direct products borrowed from group theory. Roughly, we are saying that the 4-dimensional product space of the two 2-dimensional spin spaces of spins 1 and 2 can be written as the direct sum of two subspaces, a 1-dimensional ($S_T = 0$) and a 3-dimensional ($S_T = 1$) space. However, there's no need to know this if you don't want to.

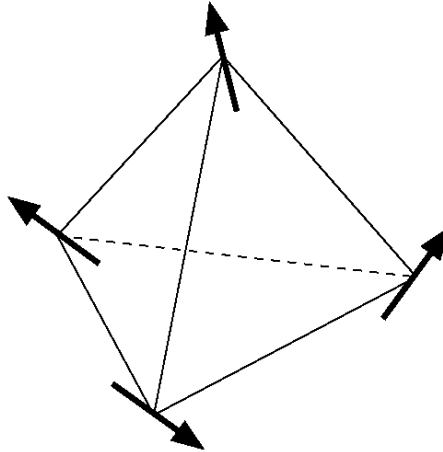


Figure 7: A tetrahedron with a spin-1/2 at each corner.

- After the coupling, how many states are there? Well, the spin-1/2 state on the right-hand side has 2 possible z -projections; the spin-3/2 has 4; the spin-5/2 has 6; the spin-7/2 has 8; and the spin-9/2 has 10. The sum of these is $2 + 4 + 6 + 8 + 10 = 30$.

✓ Explain where these numbers come from in the sum $2 + 4 + 6 + 8 + 10 = 30$.⁸²

✓ What values of S_T are possible if you couple angular momenta with spin quantum numbers $S_1 = 1$ and $S_2 = \frac{5}{2}$?⁸³

It is not too hard to prove, if you wish, that the numbers of states on the two sides will always agree, as they should — the coupling only causes the states to assume new linear combinations; it should not cause them to increase or decrease in number.

14.3 Fully coupled clusters

Armed with the rule (387), we can quickly calculate the exact energy spectra of Heisenberg models that would be a pain to solve by brute force. Take, for example, a tetrahedron with a spin-1/2 at each of its corners, as shown in Fig. 7. These structures are surprisingly common in magnetic insulators. For example, the dysprosium ions in $\text{Dy}_2\text{Ti}_2\text{O}_7$ make a three dimensional lattice of corner-sharing tetrahedra: such a lattice is termed ‘pyrochlore’.

The tetrahedron of spins is an example of a **fully coupled cluster**, i.e. a group of spins in which every spin interacts with every other. Labelling the spins by the numbers 1 to 4, and generalising the Heisenberg Hamiltonian in a reasonably obvious way, we obtain the following Hamiltonian for this cluster:

$$H = J \{ \mathbf{S}_1 \cdot \mathbf{S}_2 + \mathbf{S}_1 \cdot \mathbf{S}_3 + \mathbf{S}_1 \cdot \mathbf{S}_4 + \mathbf{S}_2 \cdot \mathbf{S}_3 + \mathbf{S}_2 \cdot \mathbf{S}_4 + \mathbf{S}_3 \cdot \mathbf{S}_4 \}. \quad (389)$$

⁸²For a given spin quantum number S , the spin magnetic quantum number m_S goes from $-S$ to $+S$ in integer steps. Thus, there are $2m_S + 1$ possible values of m_S for a given S . For example, for $S = \frac{5}{2}$, we have $m_S = -\frac{5}{2}, -\frac{3}{2}, -\frac{1}{2}, \frac{1}{2}, \frac{3}{2}, \frac{5}{2}$. Thus, this is a 6-dimensional subspace with $2m_S + 1 = 2\frac{5}{2} + 1 = 6$ basis states.

⁸³ S_T goes from $|S_1 - S_2|$ to $S_1 + S_2$ in integer steps. Thus, $S_T = \frac{3}{2}, \frac{5}{2}, \frac{7}{2}$.

Similarly to the above example, we can define a total spin for the whole cluster as

$$\mathbf{S}_T \equiv \mathbf{S}_1 + \mathbf{S}_2 + \mathbf{S}_3 + \mathbf{S}_4. \quad (390)$$

Its square is therefore

$$\begin{aligned} \mathbf{S}_T^2 = & \mathbf{S}_1^2 + \mathbf{S}_2^2 + \mathbf{S}_3^2 + \mathbf{S}_4^2 \\ & + 2\mathbf{S}_1 \cdot \mathbf{S}_2 + 2\mathbf{S}_1 \cdot \mathbf{S}_3 + 2\mathbf{S}_1 \cdot \mathbf{S}_4 + 2\mathbf{S}_2 \cdot \mathbf{S}_3 + 2\mathbf{S}_2 \cdot \mathbf{S}_4 + 2\mathbf{S}_3 \cdot \mathbf{S}_4. \end{aligned} \quad (391)$$

Notice that this has generated, on the second line of the right-hand side, precisely the same set of terms as appears in the Hamiltonian (389).⁸⁴ Hence we may write the Hamiltonian as

$$H = \frac{J}{2} (\mathbf{S}_T^2 - \mathbf{S}_1^2 - \mathbf{S}_2^2 - \mathbf{S}_3^2 - \mathbf{S}_4^2). \quad (392)$$

But now this is entirely written in terms of squares of angular momenta, and we know the forms their eigenvalues take. Therefore the eigenvalues of the Hamiltonian — i.e. the energy levels of the system — are given by

$$\begin{aligned} E &= \frac{\hbar^2 J}{2} \{S_T(S_T + 1) - S_1(S_1 + 1) - S_2(S_2 + 1) - S_3(S_3 + 1) - S_4(S_4 + 1)\} \\ &= \frac{\hbar^2 J}{2} \{S_T(S_T + 1) - 3\}, \end{aligned} \quad (393)$$

where to get to the last line we have used the fact that $S_1 = S_2 = S_3 = S_4 = 1/2$.

Hence, to determine the possible total energies of the system, all we have to do is determine the possible values of the total spin quantum number S_T . Using the rule (387) several times, we obtain

$$\left(\left(\frac{1}{2} \otimes \frac{1}{2} \right) \otimes \frac{1}{2} \right) \otimes \frac{1}{2} = \left((0 \oplus 1) \otimes \frac{1}{2} \right) \otimes \frac{1}{2} \quad (395)$$

$$= \left(\left(0 \otimes \frac{1}{2} \right) \oplus \left(1 \otimes \frac{1}{2} \right) \right) \otimes \frac{1}{2} \quad (396)$$

$$= \left(\frac{1}{2} \oplus \frac{1}{2} \oplus \frac{3}{2} \right) \otimes \frac{1}{2} \quad (397)$$

$$= \left(\frac{1}{2} \otimes \frac{1}{2} \right) \oplus \left(\frac{1}{2} \otimes \frac{1}{2} \right) \oplus \left(\frac{3}{2} \otimes \frac{1}{2} \right) \quad (398)$$

$$= 0 \oplus 1 \oplus 0 \oplus 1 \oplus 1 \oplus 2. \quad (399)$$

Hence the energy spectrum has two singlet levels (total degeneracy 2), three triplet levels (total degeneracy 9), and one quintuplet level (total degeneracy 5). Note that $2 + 9 + 5 = 16$, which is equal to the 2^4 basis states we started with.

Putting these three allowed values of S_T into (394), we find that the energies of the three levels are

$$E = \begin{cases} \frac{\hbar^2 J}{2} \{6 - 3\} = \frac{3\hbar^2 J}{2} & S_T = 2 \text{ (quintuplet)} \\ \frac{\hbar^2 J}{2} \{2 - 3\} = -\frac{\hbar^2 J}{2} & S_T = 1 \text{ (triplets)} \\ \frac{\hbar^2 J}{2} \{0 - 3\} = -\frac{3\hbar^2 J}{2} & S_T = 0 \text{ (singlets)} \end{cases} \quad (400)$$

⁸⁴This trick, of course, worked only because the cluster is fully coupled.

Notice a sort of ‘conservation of centre of mass’ here. Starting from an original energy of zero: 5 states have moved up by $3\hbar^2 J/2$; 9 states have moved down by $\hbar^2 J/2$; and 2 states have moved down by $3\hbar^2 J/2$. The total displacement in energy is thus

$$\frac{\hbar^2 J}{2} (5 \times 3 - 9 \times 1 - 2 \times 3) = 0. \quad (401)$$

This can be used as a useful check that no arithmetic error has been made in the calculation.

You are encouraged to solve the tetrahedron problem by brute force — this will certainly make you appreciate better how much time this method saves!

15 Many-body wave functions; bosons and fermions

In this lecture, we shall embark on the very final part of the course: how to formulate quantum mechanics for systems containing more than one particle. The total number of particles will be N , each of them moving in d dimensions; but for most of the time we shall take $N = 2$ and $d = 1$ for simplicity.

15.1 The many-body wave function

To extend the formalism of quantum mechanics to N particles, let us return to the wave function notation, and consider measurement. For a single particle in one dimension, the wave function is interpreted as giving the probability of various possible outcomes of position measurements. In particular, the probability of finding the particle to be between $x = a$ and $x = b$ on measurement of its position is given by⁸⁵

$$\Pr(a, b) = \int_a^b dx_1 |\psi(x_1)|^2, \quad (402)$$

the usual Born interpretation. For several particles, the natural generalisation of this question is to ask about a simultaneous position measurement of all of them — think of it as a flash photograph, instantaneously revealing (with some finite resolution) the positions of all the N particles. Then we are led to ask for the joint probability that the first particle is found in the region $a_1 < x_1 < b_1$ and the second is found in the region $a_2 < x_2 < b_2$ and so on. Presumably there should be one integral involved for each co-ordinate, which suggests that

$$\Pr(a_1, b_1, a_2, b_2, \dots, a_N, b_N) = \int_{a_1}^{b_1} dx_1 \int_{a_2}^{b_2} dx_2 \dots \int_{a_N}^{b_N} dx_N |\psi(x_1, x_2, \dots, x_N)|^2. \quad (403)$$

The entity appearing in the integrand,

$$\psi(x_1, x_2, \dots, x_N), \quad (404)$$

is referred to as the **many-body wave function** of the system.

15.2 Uncorrelated particles: the Hartree wave function

What can we say about the many-body wave function (404)? In the case of particles whose positions are not correlated, it's actually quite simple. In this case, the different position measurements are statistically independent events, and so we know that their probabilities must just multiply:

$$\Pr(a_1, b_1, a_2, b_2, \dots, a_N, b_N) = \Pr(a_1, b_1) \Pr(a_2, b_2) \dots \Pr(a_N, b_N) \quad (405)$$

⁸⁵The eagle-eyed will notice that we've called the particle's position x_1 rather than the usual x . This is to prepare the ground for the existence of other particles, whose co-ordinates will be x_2, x_3, \dots, x_N .

$$= \left(\int_{a_1}^{b_1} dx_1 |\phi_{n_1}(x_1)|^2 \right) \left(\int_{a_2}^{b_2} dx_2 |\phi_{n_2}(x_2)|^2 \right) \dots \left(\int_{a_N}^{b_N} dx_N |\phi_{n_N}(x_N)|^2 \right) \quad (406)$$

$$= \int_{a_1}^{b_1} dx_1 \int_{a_2}^{b_2} dx_2 \dots \int_{a_N}^{b_N} dx_N |\phi_{n_1}(x_1)|^2 |\phi_{n_2}(x_2)|^2 \dots |\phi_{n_N}(x_N)|^2. \quad (407)$$

Comparing the right-hand side of (407) with (403), we see that the wave function itself has factorised. Hence for particles with uncorrelated positions, the many-body wave function is just a product of single-particle wave functions,

$$\psi(x_1, x_2, \dots, x_N) = \phi_{n_1}(x_1) \phi_{n_2}(x_2) \dots \phi_{n_N}(x_N). \quad (408)$$

Such a product wave function is called a **Hartree wave function**.

✓ Consider two distinguishable particles in the same one-dimensional infinite square well of width L , given by a Hartree wave function. Particle 1 is in the 2nd excited state, particle 2 in the 3rd excited state. Write down an integral expression for the probability of finding both particles in the left half of the well. ⁸⁶

15.3 Indistinguishable particles

However, the Hartree wave function (408) cannot be used to describe a system in which the particles are indistinguishable. To illustrate why not, let us consider a system made of two indistinguishable particles (e.g. two electrons). The many-body wave function is

$$\psi(x_1, x_2), \quad (409)$$

and thus the probability density for finding the first particle at x_1 and the second at x_2 is given by

$$|\psi(x_1, x_2)|^2. \quad (410)$$

However, if the particles are indistinguishable, this must be the same as the probability density for finding the first particle at x_2 and the second at x_1 , i.e.

$$|\psi(x_2, x_1)|^2 = |\psi(x_1, x_2)|^2. \quad (411)$$

The truly remarkable fact is that, although the probability densities have to be equal whichever way we label the particles, the wave functions themselves *don't* have to be! Indeed, the equation (411) only implies that the wave functions differ by a complex phase,⁸⁷

$$\psi(x_2, x_1) = e^{i\delta} \psi(x_1, x_2). \quad (412)$$

⁸⁶ $\int_0^{L/2} dx_1 \int_0^{L/2} dx_2 |\phi_3(x_1)|^2 |\phi_4(x_2)|^2 = \left(\int_0^{L/2} dx_1 |\phi_3(x_1)|^2 \right) \left(\int_0^{L/2} dx_2 |\phi_4(x_2)|^2 \right)$

⁸⁷ Such a phase does not survive the taking of the square modulus: $|e^{i\delta} \psi(x_1, x_2)|^2 = e^{-i\delta} \psi^*(x_1, x_2) e^{i\delta} \psi(x_1, x_2) = \psi^*(x_1, x_2) \psi(x_1, x_2) = |\psi(x_1, x_2)|^2$.

But a double exchange is the identity, which implies that $(e^{i\delta})^2 = 1$. Thus, $e^{i\delta} = \pm 1$, i.e.

$$\psi(x_2, x_1) = \pm \psi(x_1, x_2). \quad (413)$$

We can derive this result in a second way by introducing an *exchange operator* \hat{X} that swaps the locations of the two particles, with

$$\hat{X}\psi(x_1, x_2) = \psi(x_2, x_1). \quad (414)$$

Not all wavefunctions will be eigenfunctions of \hat{X} , but let us assume in what follows that $\psi(x_1, x_2)$ is an eigenfunction of \hat{X} with eigenvalue p . Thus

$$\hat{X}\psi(x_1, x_2) = p\psi(x_1, x_2). \quad (415)$$

Swapping the locations of the particles twice is equivalent with doing nothing. This can be seen mathematically using equation (414) to obtain

$$\hat{X}^2\psi(x_1, x_2) = \hat{X}\hat{X}\psi(x_1, x_2) = \hat{X}\psi(x_2, x_1) = \psi(x_1, x_2). \quad (416)$$

Using equation (415), we can also write

$$\hat{X}^2\psi(x_1, x_2) = \hat{X}p\psi(x_1, x_2) = p\hat{X}\psi(x_1, x_2) = p^2\psi(x_1, x_2). \quad (417)$$

Comparing equations (416) and (417) gives $p^2 = 1$ and thus the only possible eigenvalues of the exchange operator \hat{X} are $p = \pm 1$. For the eigenvalue $+1$, we obtain $\psi(x_2, x_1) = +\psi(x_1, x_2)$, for the eigenvalue -1 , we obtain $\psi(x_2, x_1) = -\psi(x_1, x_2)$.

We call wave functions with $\psi(x_2, x_1) = +\psi(x_1, x_2)$ **symmetric under particle interchange**, and wave functions with $\psi(x_2, x_1) = -\psi(x_1, x_2)$ **antisymmetric under particle interchange**.

If the particles have spin, interchanging labels must also involve interchanging spins. In the two-particle case for spin quantum number $s = 1/2$, we already know the spin eigenkets:

Eigenket	Total spin quantum number
$ \uparrow\uparrow\rangle$	$s_T = 1$ (Triplet)
$\frac{1}{\sqrt{2}}(\uparrow\downarrow\rangle + \downarrow\uparrow\rangle)$	
$ \downarrow\downarrow\rangle$	
$\frac{1}{\sqrt{2}}(\uparrow\downarrow\rangle - \downarrow\uparrow\rangle)$	$s_T = 0$ (Singlet)

Notice that all triplet states are symmetric under interchange of labels, while the singlet state is antisymmetric. If the particles have spin, the total wave function is given by the product of the spatial and spin parts: $\psi_{tot} = \psi_{space}\psi_{spin}$. There is a theorem linking the magnitude of a particle's spin to the "statistics" (i.e. the symmetry of its wave function): the spin statistics theorem:

Particles with integer spin are bosons: the total wave function for indistinguishable bosons is symmetric under particle interchange. Particles with

half-integer spin are fermions: the total wave function for indistinguishable fermions is antisymmetric under particle interchange. There is no known non-relativistic proof of this statement. Electrons, neutrons and protons are spin 1/2 particles and thus are fermions.

For example, the total wave function for two indistinguishable electrons must be antisymmetric under particle interchange. There are two ways this can happen: either the spatial part is antisymmetric and the spin part is symmetric ($s_T = 1$), or the spatial part is symmetric and the spin part is antisymmetric ($s_T = 0$). For indistinguishable bosons with spin, the spatial part and the spin part need to be both symmetric or both antisymmetric under particle interchange. For spinless (spin zero) bosons, the wave function just has a spatial part which needs to be symmetric.

✓ Is $\psi(x_1, x_2) = \phi_1(x_1)\phi_3(x_2)$ symmetric or antisymmetric under particle interchange? ⁸⁸

✓ Could $\psi(x_1, x_2) = \frac{1}{\sqrt{2}}(\phi_1(x_1)\phi_2(x_2) - \phi_2(x_1)\phi_1(x_2))$ be a valid spatial wave function for i) two spinless bosons; ii) two electrons with total spin quantum number $s_T = 0$; iii) two electrons with $s_T = 1$; iv) two distinguishable particles? ⁸⁹

✓ Could $\phi_1(x_1)\phi_1(x_2)$ (where both particles have the quantum number $n = 1$) be a valid spatial wave function for i) two spinless bosons; ii) two electrons with total spin quantum number $s_T = 0$; iii) two electrons with $s_T = 1$; iv) two distinguishable particles? ⁹⁰

15.4 The Pauli Exclusion Principle

Consider the case of two indistinguishable particles where the spatial wave function is antisymmetric (e.g. two electrons in a triplet state). If we assume both particles are in energy eigenstates n_1 and n_2 with $n_1 \neq n_2$, then

$$\psi(x_1, x_2 = x_1) = \frac{1}{\sqrt{2}}(\phi_{n_1}(x_1)\phi_{n_2}(x_1) - \phi_{n_2}(x_1)\phi_{n_1}(x_1)) = 0. \quad (418)$$

Thus, $|\psi(x_1, x_1)|^2 = 0$, meaning we never find the two particles in the same place.

If the spatial wave function is symmetric (e.g. two spinless bosons, or two electrons in a singlet state), then

$$\psi(x_1, x_2 = x_1) = \frac{1}{\sqrt{2}}(\phi_{n_1}(x_1)\phi_{n_2}(x_1) + \phi_{n_2}(x_1)\phi_{n_1}(x_1)) = \sqrt{2}\phi_{n_1}(x_1)\phi_{n_2}(x_1). \quad (419)$$

Thus, $|\psi(x_1, x_1)|^2 = 2|\phi_{n_1}(x_1)\phi_{n_2}(x_1)|^2$, meaning the probability of finding the two particles in the same place is increased by a factor of two compared with distinguishable particles.

Now consider the case of two indistinguishable particles where the spatial wave function is antisymmetric (e.g. two electrons in a triplet state). Assume that both particles

⁸⁸No, as $\psi(x_2, x_1) = \phi_3(x_1)\phi_1(x_2) \neq \pm\phi_1(x_1)\phi_3(x_2)$.

⁸⁹No, no, yes and yes. Note that ψ is antisymmetric under particle interchange. For distinguishable particles, the wave function can be (but does not need to be) symmetric/antisymmetric under particle exchange.

⁹⁰Yes, yes, no and yes. Note that $\phi_1(x_1)\phi_1(x_2)$ is symmetric under particle interchange.

are in the same quantum state, e.g. $n_1 = n_2$. This gives

$$\psi(x_1, x_2) = \frac{1}{\sqrt{2}} (\phi_{n_1}(x_1)\phi_{n_1}(x_2) - \phi_{n_1}(x_1)\phi_{n_1}(x_2)) = 0. \quad (420)$$

Thus, **no two electrons can occupy the same quantum state**. The term “quantum state” here refers to all the quantum numbers characterizing the electrons, including the spatial and spin quantum numbers. This leads to the shell structure of electrons in atoms and the shell structure of nuclei (as protons and neutrons are also spin 1/2 particles).

15.5 The Slater determinant

Assume we have N non-interacting particles in a 1D infinite square well. Assume the spin part of the wave function is symmetric, and that each of the particles is in an energy eigenstate n_i . If the particles were distinguishable, then

$$\psi_{tot, distinguishable} = \phi_{n_1}(x_1)\phi_{n_2}(x_2) \cdots \phi_{n_N}(x_N). \quad (421)$$

For indistinguishable particles, we need to add further terms with different permutations, so that the total wave function is antisymmetric under interchange of each pair of particle labels. There are $N!$ permutations of $\{n_1, n_2, \dots, n_N\}$, and so our wave function will have $N!$ terms. For indistinguishable particles,

$$\psi(x_1, x_2, \dots, x_N) = \frac{1}{\sqrt{N!}} \sum_P (-1)^P \phi_{P(n_1)}(x_1) \phi_{P(n_2)}(x_2) \cdots \phi_{P(n_N)}(x_N) \quad (422)$$

where the sum is over all permutations of $\{n_1, n_2, \dots, n_N\}$, and the exponent in $(-1)^P$ is the number of permutation inversions in permutation P . This is very similar to the definition of the determinant! Thus, we can find the antisymmetric wave function through the following so-called Slater determinant:

$$\psi(x_1, x_2, \dots, x_N) = \frac{1}{\sqrt{N!}} \begin{vmatrix} \phi_{n_1}(x_1) & \phi_{n_1}(x_2) & \cdots & \phi_{n_1}(x_N) \\ \phi_{n_2}(x_1) & \phi_{n_2}(x_2) & \cdots & \phi_{n_2}(x_N) \\ \vdots & \vdots & \ddots & \vdots \\ \phi_{n_N}(x_1) & \phi_{n_N}(x_2) & \cdots & \phi_{n_N}(x_N) \end{vmatrix} \quad (423)$$

Example: assume we have three electrons with parallel spins in the same 1D infinite square well, one each in the ground state, first excited state and second excited state. Thus, $N = 3$, $n_1 = 1$, $n_2 = 2$ and $n_3 = 3$. This gives

$$\begin{aligned} \psi(x_1, x_2, x_3) &= \frac{1}{\sqrt{3!}} \begin{vmatrix} \phi_1(x_1) & \phi_1(x_2) & \phi_1(x_3) \\ \phi_2(x_1) & \phi_2(x_2) & \phi_2(x_3) \\ \phi_3(x_1) & \phi_3(x_2) & \phi_3(x_3) \end{vmatrix} \\ &= \frac{1}{\sqrt{6}} \left(\phi_1(x_1) \begin{vmatrix} \phi_2(x_2) & \phi_2(x_3) \\ \phi_3(x_2) & \phi_3(x_3) \end{vmatrix} - \phi_1(x_2) \begin{vmatrix} \phi_2(x_1) & \phi_2(x_3) \\ \phi_3(x_1) & \phi_3(x_3) \end{vmatrix} + \phi_1(x_3) \begin{vmatrix} \phi_2(x_1) & \phi_2(x_2) \\ \phi_3(x_1) & \phi_3(x_2) \end{vmatrix} \right) \end{aligned} \quad (424)$$

Thus, we obtain

$$\begin{aligned}\psi(x_1, x_2, x_3) = & \frac{1}{\sqrt{6}} \{ \phi_1(x_1)\phi_2(x_2)\phi_3(x_3) - \phi_1(x_1)\phi_3(x_2)\phi_2(x_3) - \phi_1(x_2)\phi_2(x_1)\phi_3(x_3) \\ & + \phi_1(x_2)\phi_3(x_1)\phi_2(x_3) + \phi_1(x_3)\phi_2(x_1)\phi_3(x_2) - \phi_1(x_3)\phi_3(x_1)\phi_2(x_2) \}\end{aligned}$$

You are encouraged to verify that

$$\psi(x_2, x_1, x_3) = -\psi(x_1, x_2, x_3),$$

$$\psi(x_3, x_2, x_1) = -\psi(x_1, x_2, x_3),$$

$$\psi(x_1, x_3, x_2) = -\psi(x_1, x_2, x_3).$$

The wave function is thus antisymmetric under interchange of each pair of particles.

✓ Consider again the state from equation (424) with $N = 3$ electrons with parallel spins in the same 1D infinite well, but now with $n_1 = n_2 = 1$ and $n_3 = 3$. Show that in this case the Slater determinant has two identical rows. Without any calculation, what is the determinant in this case? ⁹¹

✓ Consider the general Slater determinant as given by equation (423) for an antisymmetric spatial wave function. If you exchange a pair of particles in the state, this corresponds to swapping two columns in the determinant. What happens to the value of the determinant in this case? ⁹²

91

$$\psi(x_1, x_2, x_3) = \frac{1}{\sqrt{3!}} \begin{vmatrix} \phi_1(x_1) & \phi_1(x_2) & \phi_1(x_3) \\ \phi_1(x_1) & \phi_1(x_2) & \phi_1(x_3) \\ \phi_3(x_1) & \phi_3(x_2) & \phi_3(x_3) \end{vmatrix} \quad (425)$$

The determinant of a matrix with two identical rows is zero, so that $\psi = 0$. Thus this state can not exist, as expected from the Pauli Principle.

⁹²The determinant of a matrix changes sign if you swap two columns. Thus, exchanging any two particles leads to a minus sign in the expression for ψ , as it must for an antisymmetric wave function.