

Mathematical Foundations of Quantum Mechanics

Original notes by Judith McGovern (December 2018)

Updated by Mike Birse

November 22, 2019

Contents

1	The Fundamentals of Quantum Mechanics	3
1.1	Postulates of Quantum Mechanics	3
1.2	From the ket to the wave function	4
1.3	The propagator or time-evolution operator	5
1.4	Simple examples	6
1.4.1	Two-state system	6
1.4.2	Propagator in free space	8
1.5	Ehrenfest's Theorem and the Classical Limit	9
1.6	The Harmonic Oscillator Without Tears	10
1.7	Product spaces	13
1.7.1	Product spaces and entanglement	13
1.7.2	Operators in product spaces	15
1.7.3	Two or more particle wave functions	16
1.7.4	Decoupling	17
2	Angular momentum	18
2.1	General properties of angular momentum	18
2.2	Electron spin and the Stern-Gerlach experiment	21
2.3	Spin- $\frac{1}{2}$	22
2.3.1	Spin precession	24
2.3.2	Spin and measurement: Stern Gerlach revisited	24
2.4	Higher spins	25
2.5	Addition of angular momenta	26
2.5.1	Using tables of Clebsch-Gordan tables	29
2.5.2	Example: Two spin- $\frac{1}{2}$ particles	31
2.5.3	Angular Momentum of Atoms and Nuclei	32
2.6	Vector Operators	32
3	Approximate methods I: variational method and WKB	35
3.1	Approximate methods in Quantum Mechanics	35
3.2	Variational methods: ground state	36
3.2.1	Variational methods: excited states	37
3.2.2	Variational methods: the helium atom	38
3.3	WKB approximation	38
3.3.1	WKB approximation for bound states	39
3.3.2	WKB approximation for tunnelling	43

4	Approximate methods II: Time-independent perturbation theory	45
4.1	Non-degenerate perturbation theory	45
4.1.1	Connection to variational approach	47
4.1.2	Perturbed infinite square well	47
4.1.3	Perturbed Harmonic oscillator	48
4.2	Degenerate perturbation theory	49
4.2.1	Example of degenerate perturbation theory	51
4.2.2	Symmetry as a guide to the choice of basis	53
4.3	The fine structure of hydrogen	54
4.3.1	Pure Coulomb potential and nomenclature	54
4.3.2	Fine structure: the lifting of l degeneracy	54
4.4	The Zeeman effect: hydrogen in an external magnetic field	57
4.5	The Stark effect: hydrogen in an external electric field	59
5	Quantum Measurement	61
5.1	The Einstein-Poldosky-Rosen “paradox” and Bell’s inequalities	61
A	Revision and background	65
A.1	Index notation—free and dummy indices	65
A.2	Vector spaces	67
A.2.1	Vectors and operators in finite spaces	67
A.2.2	Functions as vectors	70
A.2.3	Commutators	72
A.3	Recap of 2nd year quantum mechanics	73
A.3.1	The wave function and Schrödinger’s equation	73
A.3.2	Measurement	75
A.3.3	Bound states	78
A.3.4	Circular and spherical symmetry	81
A.3.5	Tunnelling	83
A.4	Series Solution of Hermite’s equation and the Harmonic Oscillator	85
A.5	Angular Momentum in Quantum Mechanics	87
A.6	Hydrogen wave functions	91
A.7	Properties of δ -functions	92
A.8	Gaussian integrals	94
A.9	Airy functions	95
A.10	Units in EM	96

1. The Fundamentals of Quantum Mechanics

1.1 Postulates of Quantum Mechanics

Shankar 4.1; Mandl 1; Griffiths 3

Summary: All of quantum mechanics follows from a small set of assumptions, which cannot themselves be derived.

All of quantum mechanics follows from a small set of assumptions, which cannot themselves be derived. There is no unique formulation or even number of postulates, but all formulations I've seen have the same basic content. This formulation follows Shankar most closely, though he puts III and IV together. Nothing significant should be read into my separating them (as many other authors do), it just seems easier to explore the consequences bit by bit.

I: The state of a particle is given by a vector $|\psi(t)\rangle$ in a Hilbert space. The state is normalised: $\langle\psi(t)|\psi(t)\rangle = 1$.

This is as opposed to the classical case where the position and momentum can be specified at any given time.

This is a pretty abstract statement, but more informally we can say that the wave function $\psi(x, t)$ contains all possible information about the particle. How we extract that information is the subject of subsequent postulates.

The really major consequence we get from this postulate is *superposition*, which is behind most quantum weirdness such as the two-slit experiment.

II: There is a Hermitian operator corresponding to each observable property of the particle. Those corresponding to position $\hat{\mathbf{x}}$ and momentum $\hat{\mathbf{p}}$ satisfy $[\hat{x}_i, \hat{p}_j] = i\hbar\delta_{ij}$.

Other examples of observable properties are energy and angular momentum. The choice of these operators may be guided by classical physics (e.g. $\hat{\mathbf{p}} \cdot \hat{\mathbf{p}}/2m$ for kinetic energy and $\hat{\mathbf{x}} \times \hat{\mathbf{p}}$ for orbital angular momentum), but ultimately is verified by experiment (e.g. Pauli matrices for spin- $\frac{1}{2}$ particles).

The commutation relation for $\hat{\mathbf{x}}$ and $\hat{\mathbf{p}}$ is a formal expression of Heisenberg's uncertainty principle.

III: Measurement of the observable associated with the operator $\hat{\Omega}$ will result in one of the eigenvalues ω_i of $\hat{\Omega}$. Immediately after the measurement the particle will be in the corresponding eigenstate $|\omega_i\rangle$.

This postulate ensures *reproducibility* of measurements. If the particle was not initially in the state $|\omega_i\rangle$ the result of the measurement was not predictable in advance, but for the result of

a measurement to be meaningful the result of a subsequent measurement must be predictable. (“Immediately” reflects the fact that subsequent time evolution of the system will change the value of ω unless it is a constant of the motion.)

IV: The probability of obtaining the result ω_i in the above measurement (at time t_0) is $|\langle \omega_i | \psi(t_0) \rangle|^2$.

If a particle (or an ensemble of particles) is repeatedly prepared in the same initial state $|\psi(t_0)\rangle$ and the measurement is performed, the result each time will in general be different (assuming this state is not an eigenstate of $\hat{\Omega}$; if it is the result will be the corresponding ω_i each time). Only the distribution of results can be predicted. The postulate expressed this way has the same content as saying that the average value of ω is given by $\langle \psi(t_0) | \hat{\Omega} | \psi(t_0) \rangle$. (Note the distinction between repeated measurements on freshly-prepared particles, and repeated measurements on the same particle which will give the same ω_i each subsequent time.)

Note that if we expand the state in the (orthonormal) basis $\{|\omega_i\rangle\}$, $|\psi(t_0)\rangle = \sum_i c_i |\omega_i\rangle$, the probability of obtaining the result ω_i is $|c_i|^2$, and $\langle \hat{\Omega} \rangle = \sum_i |c_i|^2 \omega_i$.

V: The time evolution of the state $|\psi(t)\rangle$ is given by $i\hbar \frac{d}{dt} |\psi(t)\rangle = \hat{H} |\psi(t)\rangle$, where \hat{H} is the operator corresponding to the classical Hamiltonian.

In most cases the Hamiltonian is just the energy and is expressed as $\hat{\mathbf{p}} \cdot \hat{\mathbf{p}}/2m + V(\hat{\mathbf{x}})$. (They differ in some cases though - see texts on classical mechanics such as Kibble and Berkshire.) In the presence of velocity-dependent forces such as magnetism the Hamiltonian is still equal to the energy, but its expression in terms of $\hat{\mathbf{p}}$ is more complicated.

VI: The Hilbert space for a system of two or more subsystems (for example, several particles) is a product space.

This is true whether the subsystems interact or not, i.e. if the states $|\phi_i\rangle$ span the space for one subsystem, the states $|\phi_i\rangle \otimes |\phi_j\rangle$ will span the space for two. If they do interact, the eigenstates of the Hamiltonian will not be simple products of that form, but will be linear superpositions of such states.

1.2 From the ket to the wave function

Shankar 4.3; Mandl 1.3; Griffiths 3.1,2

Summary: The position-space representation allows us to make contact with quantum mechanics expressed in terms of wave functions

The position operator in 3-D is $\hat{\mathbf{x}}$, and has eigenkets $|\mathbf{r}\rangle$. A state of a particle can therefore be associated with a function of position, the wave function: $\psi(\mathbf{r}, t) = \langle \mathbf{r} | \psi(t) \rangle$. Note that position and time are treated quite differently in non-relativistic quantum mechanics. There is no operator corresponding to time, and t is just part of the label of the state. By the fourth postulate, the probability of finding the particle in an infinitesimal volume dV at a position \mathbf{r} , $\rho(\mathbf{r})dV$, is given by $\rho(\mathbf{r}, t) = |\langle \mathbf{r} | \psi(t) \rangle|^2 = |\psi(\mathbf{r}, t)|^2$. Thus a measurement of position can yield many answers, and as well as an average x -position $\langle \psi | \hat{x} | \psi \rangle$ there will be an uncertainty, Δx , where $\Delta x^2 = \langle \psi | \hat{x}^2 | \psi \rangle - \langle \psi | \hat{x} | \psi \rangle^2$.

Though the notation $\langle \psi | \hat{A} | \phi \rangle$ is compact, to calculate it if \hat{A} is a function of position and momentum operators we will usually immediately substitute the integral form $\int_{-\infty}^{\infty} \psi^*(\mathbf{r}) \hat{A} \phi(\mathbf{r}) d^3r$,

where—rather sloppily— \hat{A} is here understood as the position-representation of the operator.

The momentum operator $\hat{\mathbf{p}}$ has the representation in the position-basis of $-i\hbar\nabla$. Eigenstates of momentum, in the position representation, are just plane waves. In order that they satisfy the following normalisations

$$\hat{I} = \int_{-\infty}^{\infty} |\mathbf{p}\rangle\langle\mathbf{p}| d^3\mathbf{p}, \quad \text{and} \quad \langle\mathbf{p}|\mathbf{p}'\rangle = \delta(\mathbf{p} - \mathbf{p}') = \delta(p_x - p'_x)\delta(p_y - p'_y)\delta(p_z - p'_z),$$

we need

$$\phi_{\mathbf{p}}(\mathbf{r}) \equiv \langle\mathbf{r}|\mathbf{p}\rangle = \left(\frac{1}{2\pi\hbar}\right)^{3/2} e^{i\mathbf{p}\cdot\mathbf{r}/\hbar}.$$

From the time-evolution equation $i\hbar\frac{d}{dt}|\psi(t)\rangle = \hat{H}|\psi(t)\rangle$ we obtain in the position representation

$$i\hbar\frac{\partial}{\partial t}\psi(\mathbf{r}, t) = \hat{H}\psi(\mathbf{r}, t),$$

which is the Schrödinger equation. The x -representation of \hat{H} is usually $-\hbar^2\nabla^2/2m + V(\mathbf{r})$. The Hamiltonian is a Hermitian operator, and so its eigenstates $|E_n\rangle$ form a basis in the space.

Together with the probability density, $\rho(\mathbf{r}) = |\psi(\mathbf{r})|^2$, we also have a probability flux or current

$$\mathbf{j}(\mathbf{r}) = -\frac{i\hbar}{2m}(\psi^*(\mathbf{r})\nabla\psi(\mathbf{r}) - \psi(\mathbf{r})\nabla\psi^*(\mathbf{r})).$$

The continuity equation $\nabla\cdot\mathbf{j} = -\partial\rho/\partial t$ which ensures local conservation of probability density follows from the Schrödinger equation.

A two-particle state has a wave function which is a function of the two positions (6 coordinates), $\Phi(\mathbf{r}_1, \mathbf{r}_2)$. For states of non-interacting distinguishable particles where it is possible to say that the first particle is in single-particle state $|\psi\rangle$ and the second in $|\phi\rangle$, $\Psi(\mathbf{r}_1, \mathbf{r}_2) = \psi(\mathbf{r}_1)\phi(\mathbf{r}_2)$. But most two-particle states are not factorisable (separable) like that one. The underlying vector space for two particles is called a “tensor direct product space” with separable states written $|\Psi\rangle = |\psi\rangle \otimes |\phi\rangle$, where “ \otimes ” is a separator which is omitted in some texts. We will come back to this soon.

1.3 The propagator or time-evolution operator

Shankar 4.3

Summary: Construcing the time-evolution operator from the Hamiltonian allows us to find the time evolution of a general state

The Schrödinger equation tells us the rate of change of the state at a given time. From that we can deduce an operator that acts on the state at time t_0 to give that at a subsequent time t : $|\psi(t)\rangle = \hat{U}(t, t_0)|\psi(t_0)\rangle$, which is called the propagator or time-evolution operator. We need the identity

$$\lim_{N\rightarrow\infty} \left(1 + \frac{x}{N}\right)^N = e^x$$

(to prove it, take the log of the L.H.S. and use the Taylor expansion for $\ln(1+x)$ about the point $x=0$).

An infinitesimal time step $\hat{U}(t+dt, t)$ follows immediately from the Schrödinger equation:

$$\begin{aligned} i\hbar \frac{d}{dt} |\psi(t)\rangle = \hat{H} |\psi(t)\rangle &\Rightarrow |\psi(t+dt)\rangle - |\psi(t)\rangle = -\frac{i}{\hbar} \hat{H} dt |\psi(t)\rangle \\ &\Rightarrow |\psi(t+dt)\rangle = \left(1 - \frac{i}{\hbar} \hat{H} dt\right) |\psi(t)\rangle. \end{aligned}$$

For a finite time interval $t - t_0$, we break it into N small steps and take the limit $N \rightarrow \infty$, in which limit every step is infinitesimal and we can use the previous result N times:

$$|\psi(t)\rangle = \lim_{N \rightarrow \infty} \left(1 - \frac{i}{\hbar} \hat{H} \frac{(t-t_0)}{N}\right)^N |\psi(t_0)\rangle = e^{-i\hat{H}(t-t_0)/\hbar} |\psi(t_0)\rangle \equiv \hat{U}(t, t_0) |\psi(t_0)\rangle$$

We note that this is a unitary operator (the exponential of i times a Hermitial operator always is). Thus, importantly, it conserves the norm of the state; there remains a unit probability of finding the particle somewhere!

If $|\psi(t_0)\rangle$ is an eigenfunction $|n\rangle$ of the Hamiltonian with energy E_n ,

$$|\psi(t)\rangle = \hat{U}(t, t_0) |n\rangle = e^{-iE_n(t-t_0)/\hbar} |n\rangle.$$

If we are able to decompose $|\psi(t_0)\rangle$ as a sum of such terms, $|\psi(t_0)\rangle = \sum_n c_n |n\rangle$ with $c_n = \langle n | \psi \rangle$, then

$$|\psi(t)\rangle = \sum_n c_n e^{-iE_n(t-t_0)/\hbar} |n\rangle;$$

each term evolves with a different phase and non-trivial time evolution takes place. Note that this implies an alternative form for the propagator:

$$\hat{U}(t, t_0) = \sum_n e^{-iE_n(t-t_0)/\hbar} |n\rangle \langle n|.$$

(Aside: If the Hamiltonian depends explicitly on time, we have

$$\hat{U}(t, t_0) = T \exp \left(-i \int_{t_0}^t \hat{H}(t') dt' / \hbar \right),$$

where the *time-ordered* exponential denoted by $T \exp$ means that in expanding the exponential, the operators are ordered so that $\hat{H}(t_1)$ always sits to the right of $\hat{H}(t_2)$ (so that it acts first) if $t_1 < t_2$. This will come up in Advanced Quantum Mechanics.)

1.4 Simple examples

1.4.1 Two-state system

Let us introduce a toy system with which to explore some of the ideas from the postulates. Consider a quantum system in which the states belong to a two-dimensional, rather than infinite-dimensional, vector space, spanned by the two orthonormal states $\{|a+\rangle, |a-\rangle\}$ (notation to be explained shortly). We will need two operators in this space, \hat{A} and \hat{B} ; the basis states are eigenstates of \hat{A} with eigenvalues ± 1 and the action of \hat{B} on the basis states is

$$\hat{B}|a+\rangle = |a-\rangle \quad \text{and} \quad \hat{B}|a-\rangle = |a+\rangle.$$

In this basis the states and the two operators have the following representations:

$$|a+\rangle \longrightarrow \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad |a-\rangle \longrightarrow \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad \hat{A} \longrightarrow \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad \hat{B} \longrightarrow \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

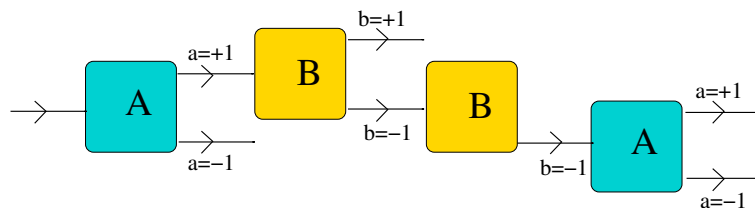
The eigenkets of \hat{B} are

$$|b\pm\rangle = \sqrt{\frac{1}{2}}(|a+\rangle \pm |a-\rangle) \longrightarrow \sqrt{\frac{1}{2}} \begin{pmatrix} 1 \\ \pm 1 \end{pmatrix}$$

with eigenvalues ± 1 .

Measurement

The flow-chart below represents an arbitrary series of measurement on a particle (or series of identically prepared particles) in an unknown initial state. We carry out consecutive measurements “immediately”, that is quickly compared with the timescale which characterises the evolution of the system in between measurements. We will talk of “measuring A ” when we strictly mean “measuring the physical quantity associated with the operator \hat{A} ”.



A priori, the possible outcomes on measuring A are the eigenvalues of \hat{A} , ± 1 . In general the particle will not start out in an eigenstate of \hat{A} , so either outcome is possible, with probabilities that depend on the initial state.

If we obtain the outcome $a = +1$ and then measure B , what can we get? We know that the state is now no longer $|\phi\rangle$ but $|a+\rangle$. The possible outcomes are $b = +1$ with probabilities $|\langle b+|a+\rangle|^2$ and $b = -1$ with probabilities $|\langle b-|a+\rangle|^2$. Both of these probabilities are $1/2$: there is a 50:50 chance of getting $b = \pm 1$. (Note that the difference between this and the previous measurement of A where we did not know the probabilities is that now we know the state before the measurement.)

If we obtain the outcome $b = -1$ and then measure B again immediately, we can only get $b = -1$ again. (This is reproducibility). The particle is in the state $|b-\rangle$ before the measurement, an eigenstate of \hat{B} .

Finally we measure A again. What are the possible outcomes and their probabilities?

Propagation

First let us consider the time-evolution of this system if the Hamiltonian is $\hat{H} = \hbar\gamma\hat{B}$. Assume we start the evolution at $t = 0$ with the system in the state $|\psi(0)\rangle$. Then $|\psi(t)\rangle = \hat{U}(t, 0)|\psi(0)\rangle$ with $\hat{U}(t, 0) = e^{-i\hat{H}t/\hbar}$. Now in general the exponentiation of an operator can't be found in

closed form, but in this case it can, because $\hat{B}^2 = \hat{I}$ and so $\hat{B}^3 = \hat{B}$. So in the power series that defines the exponential, successive terms will be alternately proportional to \hat{B} and \hat{I} :

$$\begin{aligned}\hat{U}(t, 0) &= e^{-i\gamma t \hat{B}} = \hat{I} - i\gamma t \hat{B} - \frac{1}{2}\gamma^2 t^2 \hat{B}^2 + i\frac{1}{3!}\gamma^3 t^3 \hat{B}^3 + \dots \\ &= (1 - (\gamma t)^2/2 + (\gamma t)^4/4! - \dots)\hat{I} - i(\gamma t - (\gamma t)^3/3! + (\gamma t)^5/5! - \dots)\hat{B} \\ &= \cos \gamma t \hat{I} - i \sin \gamma t \hat{B} \longrightarrow \begin{pmatrix} \cos \gamma t & -i \sin \gamma t \\ -i \sin \gamma t & \cos \gamma t \end{pmatrix}\end{aligned}$$

So if we start, say, with $|\psi(0)\rangle = |b+\rangle$, an eigenstate of \hat{B} , as expected we stay in the same state: $|\psi(t)\rangle = \hat{U}(t, 0)|b+\rangle = e^{-i\gamma t}|b+\rangle$. All that happens is a change of phase. But if we start with $|\psi(0)\rangle = |a+\rangle$,

$$|\psi(t)\rangle = \cos \gamma t |a+\rangle - i \sin \gamma t |a-\rangle.$$

Of course we can rewrite this as

$$|\psi(t)\rangle = \sqrt{\frac{1}{2}}(e^{-i\gamma t}|b+\rangle + e^{i\gamma t}|b-\rangle)$$

as expected. The expectation value of \hat{A} is not constant: $\langle\psi(t)|\hat{A}|\psi(t)\rangle = \cos 2\gamma t$. The system oscillates between $|a+\rangle$ and $|a-\rangle$ with a frequency 2γ . (This is twice as fast as you might think—but after time π/γ the state of the system is $-|a+\rangle$, which is not distinguishable from $|a+\rangle$.)

1.4.2 Propagator in free space

One case where the propagator can be calculated even in position space is the case of a free particle, in which the Hamiltonian is $\hat{H} = \hat{\mathbf{p}}^2/2m$. We want to be able to find $\psi(\mathbf{r}, t)$ given $\psi(\mathbf{r}, 0)$, using

$$\psi(\mathbf{r}, t) = \langle\mathbf{r}|\psi(t)\rangle = \langle\mathbf{r}|\hat{U}(t, 0)|\psi(0)\rangle = \int \langle\mathbf{r}|\hat{U}(t, 0)|\mathbf{r}'\rangle\psi(\mathbf{r}', 0)d^3\mathbf{r}'.$$

The object $U(\mathbf{r}, \mathbf{r}'; t, 0) \equiv \langle\mathbf{r}|\hat{U}(t, 0)|\mathbf{r}'\rangle$ is the position-space matrix element of the propagator. (Some texts call this the propagator, referring to \hat{U} only as the time-evolution operator.) This is the probability of finding the particle at position \mathbf{r}' at time t , given that at time 0 it was at \mathbf{r} . To calculate it we will use the fact that momentum eigenstates $|\mathbf{p}\rangle$ are eigenstates of \hat{H} :

$$\begin{aligned}\langle\mathbf{r}|\hat{U}(t, 0)|\mathbf{r}'\rangle &= \iint \langle\mathbf{r}|\mathbf{p}\rangle\langle\mathbf{p}|\hat{U}(t, 0)|\mathbf{p}'\rangle\langle\mathbf{p}'|\mathbf{r}'\rangle d^3\mathbf{p}d^3\mathbf{p}' \\ &= \iint \langle\mathbf{r}|\mathbf{p}\rangle\langle\mathbf{p}|\exp\left(\frac{-i\hat{\mathbf{p}}^2 t}{2m\hbar}\right)|\mathbf{p}'\rangle\langle\mathbf{p}'|\mathbf{r}'\rangle d^3\mathbf{p}d^3\mathbf{p}' \\ &= \frac{1}{(2\pi\hbar)^3} \iint \exp\left(\frac{i\mathbf{p}\cdot\mathbf{r}}{\hbar}\right) \exp\left(\frac{-i\mathbf{p}^2 t}{2m\hbar}\right) \delta(\mathbf{p} - \mathbf{p}') \exp\left(\frac{-i\mathbf{p}'\cdot\mathbf{r}'}{\hbar}\right) d^3\mathbf{p}d^3\mathbf{p}' \\ &= \frac{1}{(2\pi\hbar)^3} \int \exp\left(\frac{-i\mathbf{p}^2 t}{2m\hbar} + \frac{i\mathbf{p}\cdot(\mathbf{r} - \mathbf{r}')}{\hbar}\right) d^3\mathbf{p} \\ &= \left(\frac{m}{2i\pi\hbar t}\right)^{3/2} \exp\left(\frac{im|\mathbf{r} - \mathbf{r}'|^2}{2\hbar t}\right)\end{aligned}$$

In the last stage, to do the three Gaussian integrals ($dp_x dp_y dp_z$) we “completed the square”, shifted the variables and used the standard results $\int e^{-\alpha x^2} dx = \sqrt{\pi/\alpha}$ which is valid even if α is imaginary.

Suppose the initial wave function is a spherically symmetric Gaussian wave packet with width Δ :

$$\psi(\mathbf{r}, 0) = N \exp(-|\mathbf{r}|^2/(2\Delta^2)) \text{ with } N = (\pi\Delta^2)^{-3/4}.$$

Then the (pretty ghastly) Gaussian integrals give

$$\begin{aligned} \psi(\mathbf{r}, t) &= N \left(\frac{m}{2i\pi\hbar t} \right)^{3/2} \int \exp\left(\frac{im|\mathbf{r} - \mathbf{r}'|^2}{2\hbar t} \right) \exp\left(\frac{-|\mathbf{r}'|^2}{2\Delta^2} \right) d^3\mathbf{r}' \\ &= N' \exp\left(-\frac{|\mathbf{r}|^2}{2\Delta^2(1 + i\hbar t/m\Delta^2)} \right) \end{aligned}$$

where N' does preserve the normalisation but we do not display it. This is an odd-looking function, but the probability density is more revealing:

$$P(\mathbf{r}, t) = |\psi(\mathbf{r}, t)| = \pi^{-3/2} (\Delta^2 + (\hbar t/m\Delta)^2)^{-3/2} \exp\left(-\frac{|\mathbf{r}|^2}{\Delta^2 + (\hbar t/m\Delta)^2} \right);$$

this is a Gaussian wavepacket with width $\Delta(t) = \sqrt{\Delta^2 + (\hbar t/m\Delta)^2}$. The narrower the initial wavepacket (in position space), the faster the subsequent spread, which makes sense as the momentum-space wave function will be wide, built up of high-momentum components. On the other hand for a massive particle with Δ not too small, the spread will be slow. For $m = 1 \text{ g}$ and $\Delta(0) = 1 \mu\text{m}$, it would take longer than the age of the universe for $\Delta(t)$ to double.

1.5 Ehrenfest’s Theorem and the Classical Limit

Shankar chs 2.7, 6; Mandl 3.2; (Griffiths 3.5.3)

Summary: The form of classical mechanics which inspired Heisenberg’s formulation of Classical Mechanics allows us to see when particles should behave classically.

Using $i\hbar \frac{d}{dt} |\psi(t)\rangle = \hat{H} |\psi(t)\rangle$ and hence $-i\hbar \frac{d}{dt} \langle \psi(t) | = \langle \psi(t) | \hat{H}$, and writing $\langle \hat{\Omega} \rangle \equiv \langle \psi(t) | \hat{\Omega} | \psi(t) \rangle$, we have Ehrenfest’s Theorem

$$\frac{d}{dt} \langle \hat{\Omega} \rangle = \frac{1}{i\hbar} \langle [\hat{\Omega}, \hat{H}] \rangle + \left\langle \frac{\partial \hat{\Omega}}{\partial t} \right\rangle$$

The second term is absent if $\hat{\Omega}$ is a time-independent operator (like position, momentum, spin...). Note we are distinguishing between *intrinsic* time-dependence of an operator, and the time-dependence of its expectation value in a given state.

This is very reminiscent of a result which follows from Hamilton’s equations in classical me-

chanics, for a function $\Omega(p, x, t)$ of position, momentum (and possibly time explicitly)

$$\begin{aligned} \frac{d}{dt}\Omega(p, x, t) &= \frac{\partial\Omega}{\partial x} \frac{dx}{dt} + \frac{\partial\Omega}{\partial p} \frac{dp}{dt} + \frac{\partial\Omega}{\partial t} \\ &= \frac{\partial\Omega}{\partial x} \frac{\partial H}{\partial p} - \frac{\partial\Omega}{\partial p} \frac{\partial H}{\partial x} + \frac{\partial\Omega}{\partial t} \\ &\equiv \{\Omega, H\} + \frac{\partial\Omega}{\partial t} \end{aligned}$$

where the notation $\{\Omega, H\}$ is called the *Poisson bracket* of Ω and H , and is simply defined in terms of the expression on the line above which it replaced. (For $\Omega = x$ and $\Omega = p$ we can in fact recover Hamilton's equations for \dot{p} and \dot{x} from this more general expression.)

In fact for $\hat{H} = \frac{\hat{p}^2}{2m} + V(\hat{x})$, we can further show that

$$\frac{d}{dt}\langle\hat{x}\rangle = \left\langle\frac{\hat{p}}{m}\right\rangle \quad \text{and} \quad \frac{d}{dt}\langle\hat{p}\rangle = -\left\langle\frac{dV(\hat{x})}{d\hat{x}}\right\rangle$$

which looks very close to Newton's laws. Note though that $\langle dV(\hat{x})/d\hat{x}\rangle \neq d\langle V(\hat{x})\rangle/d\langle\hat{x}\rangle$ in general.

This correspondence is not just a coincidence, in the sense that Heisenberg was influenced by it in coming up with his formulation of quantum mechanics. It confirms that it is the expectation value of an operator, rather than the operator itself, which is closer to the classical concept of the time evolution of some quantity as a particle moves along a trajectory.

A further similarity is that in both quantum and classical mechanics, anything that commutes with the Hamiltonian (vanishing Poisson bracket in the latter case) is a constant of the motion (a conserved quantity). Examples are momentum for a free particle and angular momentum for a particle in a spherically symmetric potential.

In the QM case, we further see that even if $[\hat{\Omega}, \hat{H}] \neq 0$, if the system is in an eigenstate of \hat{H} the expectation value of $\hat{\Omega}$ will not change with time. That's why the eigenstates of the Hamiltonian are also called *stationary states*.

Similarity of formalism is not the same as identity of concepts though. Ehrenfest's Theorem does not say that the expectation value of a quantity follows a classical trajectory in general. What it does ensure is that if the uncertainty in the quantity is sufficiently small, in other words if Δx and Δp are both small (in relative terms) then the quantum motion will approximate the classical path. Of course because of the uncertainty principle, if Δx is small then Δp is large, and it can only be *relatively* small if p itself is really large—i.e. if the particle's mass is macroscopic. More specifically, we can say that we will be in the classical regime if the de Broglie wavelength is much less than the (experimental) uncertainty in x . (In the Stern-Gerlach experiment the atoms are heavy enough that (for a given component of their magnetic moment) they follow approximately classical trajectories through the inhomogeneous magnetic field.)

1.6 The Harmonic Oscillator Without Tears

Shankar 7.4,5; Mandl 12.5; Griffiths 2.3.1

Summary: Operator methods lead to a new way of viewing the harmonic oscillator in which quanta of energy are primary.

We are concerned with a particle of mass m in a harmonic oscillator potential $\frac{1}{2}kx^2 \equiv \frac{1}{2}m\omega^2x^2$ where ω is the classical frequency of oscillation. The Hamiltonian is

$$\hat{H} = \frac{\hat{p}^2}{2m} + \frac{1}{2}m\omega^2\hat{x}^2$$

and we are going to temporarily forget that we know what the energy levels and wave functions are. In order to work with dimensionless quantities, we define the length $x_0 = \sqrt{\hbar/m\omega}$.

Then if we define

$$\hat{a} = \frac{1}{\sqrt{2}} \left(\frac{\hat{x}}{x_0} + i \frac{x_0}{\hbar} \hat{p} \right) \quad \text{and} \quad \hat{a}^\dagger = \frac{1}{\sqrt{2}} \left(\frac{\hat{x}}{x_0} - i \frac{x_0}{\hbar} \hat{p} \right)$$

we can prove the following:

- $\hat{x} = (x_0/\sqrt{2})(\hat{a}^\dagger + \hat{a})$; $\hat{p} = (i\hbar/\sqrt{2}x_0)(\hat{a}^\dagger - \hat{a})$
- $[\hat{x}, \hat{p}] = i\hbar \Rightarrow [\hat{a}, \hat{a}^\dagger] = 1$
- $\hat{H} = \hbar\omega(\hat{a}^\dagger\hat{a} + \frac{1}{2})$
- $[\hat{H}, \hat{a}] = -\hbar\omega\hat{a}$ and $[\hat{H}, \hat{a}^\dagger] = \hbar\omega\hat{a}^\dagger$

Without any prior knowledge of this system, we can derive the spectrum and the wave functions of the energy eigenstates. We start by assuming we know one normalised eigenstate of \hat{H} , $|n\rangle$, with energy E_n . Since

$$E_n = \langle n|\hat{H}|n\rangle = \hbar\omega\langle n|\hat{a}^\dagger\hat{a} + \frac{1}{2}|n\rangle = \hbar\omega\langle n|\hat{a}^\dagger\hat{a}|n\rangle + \frac{1}{2}\hbar\omega$$

and also $\langle n|\hat{a}^\dagger\hat{a}|n\rangle = \langle \hat{a}n|\hat{a}n\rangle \geq 0$, we see that $E_n \geq \frac{1}{2}\hbar\omega$. There must therefore be a lowest-energy state, $|0\rangle$ (*not* the null state!).

Now consider the state $\hat{a}|n\rangle$. Using the commutator $[\hat{H}, \hat{a}]$ above we have

$$\hat{H}(\hat{a}|n\rangle) = \hat{a}\hat{H}|n\rangle - \hbar\omega\hat{a}|n\rangle = (E_n - \hbar\omega)\hat{a}|n\rangle,$$

so $\hat{a}|n\rangle$ is another eigenstate with energy $E_n - \hbar\omega$. A similar calculation shows that $\hat{a}^\dagger|n\rangle$ is another eigenstate with energy $E_n + \hbar\omega$. So starting with $|n\rangle$ it seems that we can generate an infinite tower of states with energies higher and lower by multiples of $\hbar\omega$.

However this contradicts the finding that there is a lowest energy state, $|0\rangle$. Looking more closely at the argument, though, we see there is a get-out: *either* $\hat{a}|n\rangle$ is another energy eigenstate *or it vanishes*. Hence $\hat{a}|0\rangle = 0$ (where 0 is the null state or vacuum).

The energy of this ground state is $E_0 = \langle 0|\hat{H}|0\rangle = \frac{1}{2}\hbar\omega$. The energy of the state $|n\rangle$, the n th excited state, obtained by n applications of \hat{a}^\dagger , is therefore $(n + \frac{1}{2})\hbar\omega$. Thus

$$\hat{H}|n\rangle \equiv \hbar\omega(\hat{a}^\dagger\hat{a} + \frac{1}{2})|n\rangle = (n + \frac{1}{2})\hbar\omega|n\rangle$$

and it follows that $\hat{a}^\dagger\hat{a}$ is a “number operator”, with $\hat{a}^\dagger\hat{a}|n\rangle = n|n\rangle$. The number in question is the number of the excited state ($n = 1$ —first excited state, etc) but also the number of quanta of energy in the oscillator.

Up to a phase, which we chose to be zero, the normalisations of the states $|n\rangle$ are:

$$\hat{a}|n\rangle = \sqrt{n}|n-1\rangle \quad \text{and} \quad \hat{a}^\dagger|n\rangle = \sqrt{n+1}|n+1\rangle.$$

As a result we have

$$|n\rangle = \frac{\hat{a}^{\dagger n}}{\sqrt{n!}}|0\rangle.$$

The operators \hat{a}^\dagger and \hat{a} are called “raising” and “lowering” operators, or collectively “ladder” operators.

We can also obtain the wave functions in this approach. Writing $\phi_0(x) \equiv \langle x|0\rangle$, from $\langle x|\hat{a}|0\rangle = 0$ we obtain $d\phi_0/dx = -(x/x_0^2)\phi_0$ and hence

$$\phi_0 = (\pi x_0^2)^{-1/4} e^{-x^2/2x_0^2}$$

(where the normalisation has to be determined separately). This is a much easier differential equation to solve than the one which comes direct from the Schrödinger equation!

The wave function for the n -th state is

$$\phi_n(x) = \frac{1}{\sqrt{2^n n!}} \left(\frac{x}{x_0} - x_0 \frac{d}{dx} \right)^n \phi_0(x) = \frac{1}{\sqrt{2^n n!}} H_n\left(\frac{x}{x_0}\right) \phi_0(x)$$

where here the definition of the Hermite polynomials is $H_n(z) = e^{z^2/2} (z - \frac{d}{dz})^n e^{-z^2/2}$. The equivalence of this formulation and the Schrödinger-equation-based approach means that Hermite polynomials defined this way are indeed solutions of Hermite’s equation.

This framework makes many calculations almost trivial which would be very hard in the traditional framework; in particular matrix elements of powers of \hat{x} and \hat{p} between general states can be easily found by using $\hat{x} = (\hat{a} + \hat{a}^\dagger)(x_0/\sqrt{2})$ and $\hat{p} = i(\hat{a}^\dagger - \hat{a})(\hbar/\sqrt{2}x_0)$. For example, $\langle m|\hat{x}^n|m'\rangle$ and $\langle m|\hat{p}^n|m'\rangle$ will vanish unless $|m - m'| \leq n$ and $|m - m'|$ and n are either both even or both odd (the last condition being a manifestation of parity, since $\phi_n(x)$ is odd/even if n is odd/even).

For a particle in a two-dimensional potential $\frac{1}{2}m\omega_x^2 x^2 + \frac{1}{2}m\omega_y^2 y^2$, the Hamiltonian is separable: $\hat{H} = \hat{H}_x + \hat{H}_y$. Defining $x_0 = \sqrt{\hbar/m\omega_x}$ and $y_0 = \sqrt{\hbar/m\omega_y}$, creation operators \hat{a}_x and \hat{a}_x^\dagger can be constructed from \hat{x} and \hat{p}_x as above, and we can construct a second set of operators \hat{a}_y and \hat{a}_y^\dagger from \hat{y} and \hat{p}_y (using y_0 as the scale factor) in the same way. It is clear that \hat{a}_x and \hat{a}_x^\dagger commute with \hat{a}_y and \hat{a}_y^\dagger , and each of \hat{H}_x and \hat{H}_y independently has a set of eigenstates just like the ones discussed above.

In fact the space of solutions to the two-dimensional problem can be thought of as a tensor direct product space¹ of the x and y spaces, with energy eigenstates $|n_x\rangle \otimes |n_y\rangle$, n_x and n_y being integers, and the Hamiltonian properly being written $\hat{H} = \hat{H}_x \otimes \hat{I}_y + \hat{I}_x \otimes \hat{H}_y$, and the eigenvalues being $(n_x + \frac{1}{2})\hbar\omega_x + (n_y + \frac{1}{2})\hbar\omega_y$. The ground state is $|0\rangle \otimes |0\rangle$ and it is annihilated by both \hat{a}_x ($= \hat{a}_x \otimes \hat{I}_y$) and \hat{a}_y ($= \hat{I}_x \otimes \hat{a}_y$).

The direct product notation is clumsy though, and we often write the states as just $|n_x, n_y\rangle$. Then for instance

$$\hat{a}_x|n_x, n_y\rangle = \sqrt{n_x}|n_x-1, n_y\rangle \quad \text{and} \quad \hat{a}_y^\dagger|n_x, n_y\rangle = \sqrt{n_y+1}|n_x, n_y+1\rangle.$$

¹see section 1.7

The corresponding wave functions of the particle are given by $\langle \mathbf{r} | n_x, n_y \rangle = \langle x | n_x \rangle \langle y | n_y \rangle$:

$$\begin{aligned} \phi_{0,0}(x, y) &= (\pi x_0 y_0)^{-1/2} e^{-x^2/2x_0^2} e^{-y^2/2y_0^2} \\ \phi_{n_x, n_y}(x, y) &= \frac{1}{\sqrt{2^{n_x} n_x!}} \frac{1}{\sqrt{2^{n_y} n_y!}} H_{n_x}\left(\frac{x}{x_0}\right) H_{n_y}\left(\frac{y}{y_0}\right) \phi_{0,0}(x, y) \end{aligned}$$

In many cases we are interested in a symmetric potential in which case $\omega_x = \omega_y$, $x_0 = y_0$, and $\phi_{0,0} \propto \exp(-r^2/x_0^2)$.

This formalism has remarkably little reference to the actual system in question—all the parameters are buried in x_0 . What is highlighted instead is the number of quanta of energy in the system, with \hat{a} and \hat{a}^\dagger annihilating or creating quanta (indeed they are most frequently termed “creation” and “annihilation” operators). Exactly the same formalism can be used in a quantum theory of photons, where the oscillator in question is just a mode of the EM field, and the operators create or destroy photons of the corresponding frequency.

1.7 Product spaces

Shankar pp 248-249 (chapter 10)

Summary: Product states describe multi-particle systems and systems with intrinsic (spin) angular momentum

1.7.1 Product spaces and entanglement

In quantum mechanics, the concept of the **tensor direct product** of two (or more) vector spaces occurs in various contexts. They arise when a system has two distinct aspects, described by states in two distinct vector spaces. In order to specify the state of the whole system, the states of both parts need to be given. Examples are states of two or more particles, or two contributions to the angular momentum such as orbital and spin, or the three coordinates of a single particle. The different aspects are sometimes called “degrees of freedom”.

Consider a system in which the two aspects are members of the two spaces, \mathbb{V}_a^M and \mathbb{V}_b^N . The notation just means that the first is M -dimensional, the second N -dimensional, and a and b are just labels. If $\{|a_m\rangle : m = 1 \dots M\}$ and $\{|b_n\rangle : n = 1 \dots N\}$ are basis sets for the two spaces, one possible basis for the product space is formed by picking one from each. There are $M \times N$ ways of doing this, so the product space has dimension $M \times N$. These states are written $|m, n\rangle \equiv |a_m\rangle \otimes |b_n\rangle$. The \otimes is best regarded *simply as a separator* and indeed is omitted in some texts; it doesn’t indicate any operation that is carried out. The combined space is denoted $\mathbb{V}_a^M \otimes \mathbb{V}_b^N$. There is no intrinsic difference between this and $\mathbb{V}_b^N \otimes \mathbb{V}_a^M$, but whichever order we choose, we need to stick to it in a particular problem—we can’t chop and change.

We have implied without proof so far that $\mathbb{V}_a^M \otimes \mathbb{V}_b^N$ is in fact itself a vector space. For that we need to specify the rules of combination: for linearly-independent vectors $|p\rangle, |q\rangle \in \mathbb{V}_a^M$ and $|v\rangle, |w\rangle \in \mathbb{V}_b^N$, we can take a sum of two vectors, $\alpha|p\rangle \otimes |v\rangle + \beta|q\rangle \otimes |w\rangle$, to give another vector in the product space. In addition $\alpha|p\rangle \otimes |v\rangle = (\alpha|p\rangle) \otimes |v\rangle = |p\rangle \otimes (\alpha|v\rangle)$. It can be shown that

all the other rules of a vector space hold ($\alpha = -1$ generates the additive inverse, $\alpha = 0$ the null vector), and linearity and the associative and distributive laws hold, eg

$$(\alpha|p\rangle) \otimes (\beta|v\rangle + \gamma|w\rangle) = \alpha\beta(|p\rangle \otimes |v\rangle) + \alpha\gamma(|p\rangle \otimes |w\rangle)$$

The inner product in the product space is defined as the arithmetic product of the individual inner products: $(\langle p| \otimes \langle v|)(|q\rangle \otimes |w\rangle) = \langle p|q\rangle \langle v|w\rangle$, which of course is a scalar. If there is more than one term in the bra or ket the usual linearity applies. If we choose orthonormal bases in each space, then $\{|a_m\rangle \otimes |b_n\rangle\}$ is an orthonormal basis in the product space.

Note that inner products $\langle p|q\rangle$ and outer products $|p\rangle\langle q|$ are formed between vectors in the *same* space. The tensor direct product is different, the two vectors involved are in *different* spaces, even if the two spaces are copies of one another (for two particles, for instance).

Importantly, while all vectors $|p\rangle \otimes |v\rangle$ are in the product space, as we have just seen not all vectors in the product space can be written in this way. Those that can are called **separable**, i.e. they have a specified vector in each separate space. Those that cannot, such as $\alpha|p\rangle \otimes |v\rangle + \beta|q\rangle \otimes |w\rangle$, are in the product space but are not separable. This is where the distinction between classical and quantum mechanics comes in: quantum mechanics allows for *superposition*. A non-separable state is also called an **entangled** state. In QM, entanglement is routine.

It is a straightforward extension to form direct product spaces from three or more individual spaces.

A common situation is the case of two subsystems, each of which can be in one of two states, $|0\rangle$ and $|1\rangle$. In the context of quantum information or computing, these are known as “Qbits” (or “qubits”) – the quantum analogues of classical binary bits. Choosing an order for the Qbits and sticking to it, a possible basis of states for the combined system is

$$\{|A\rangle = |0\rangle \otimes |0\rangle, \quad |B\rangle = |0\rangle \otimes |1\rangle, \quad |C\rangle = |1\rangle \otimes |0\rangle, \quad |D\rangle = |1\rangle \otimes |1\rangle\}.$$

States $|B\rangle$ and $|C\rangle$ differ by which of the two Qbits is in which state. Note that here the label $|\dots\rangle$ tells us whether a particular state is in the product space or one of the individual spaces, but in general we need to specify. These are all separable, as are the following states (unnormalised, for simplicity):

$$\begin{aligned} |A\rangle + |B\rangle &= |0\rangle \otimes (|0\rangle + |1\rangle); & |B\rangle + |D\rangle &= (|0\rangle + |1\rangle) \otimes |1\rangle; \\ |A\rangle + |B\rangle + |C\rangle + |D\rangle &= (|0\rangle + |1\rangle) \otimes (|0\rangle + |1\rangle). \end{aligned}$$

But the following are entangled:

$$\begin{aligned} |A\rangle + |D\rangle &= |0\rangle \otimes |0\rangle + |1\rangle \otimes |1\rangle; & |B\rangle - |C\rangle &= |0\rangle \otimes |1\rangle - |1\rangle \otimes |0\rangle; \\ |A\rangle + |C\rangle - |D\rangle &= (|0\rangle + |1\rangle) \otimes |0\rangle - |1\rangle \otimes |1\rangle = |0\rangle \otimes |0\rangle + |1\rangle \otimes (|0\rangle - |1\rangle) \end{aligned}$$

When a measurement of the state of one Qbit is carried out, the system will collapse into those parts of the state compatible with the result, which may or may not specify the state of the other Qbit. Restricting ourselves to states where neither Qbit’s state is known in advance (i.e. we might get 0 or 1 whichever of the two we measure) then we have two possibilities. If the initial state is separable,

$$|\psi_i\rangle = (\alpha|0\rangle + \beta|1\rangle) \otimes (\gamma|0\rangle + \delta|1\rangle),$$

(normalised, so $|\alpha|^2 + |\beta|^2 = |\gamma|^2 + |\delta|^2 = 1$), then if we measure the state of the first Qbit and get 0 (something that will happen with probability $|\alpha|^2$), then the subsequent state of the system is

$$|\psi_f\rangle = |0\rangle \otimes (\gamma|0\rangle + \delta|1\rangle),$$

and we still don't know what result we'd get if we measured the state of the second Qbit. If on the other hand the state is not separable, for instance

$$|\psi_i\rangle = \alpha|0\rangle \otimes |1\rangle - \beta|1\rangle \otimes |0\rangle,$$

then if measure the state of the first Qbit and get 0, the subsequent state of the system is

$$|\psi_f\rangle = |0\rangle \otimes |1\rangle$$

—we now know that the second Qbit is in state $|1\rangle$ and if we measured it we would be certain to get 1. This particular case where measurement of the first determines the result of the second is called maximal entanglement.

1.7.2 Operators in product spaces

Let \hat{C}_a be an operator in a vector space \mathbb{V}_a and \hat{D}_b one in \mathbb{V}_b . Then in the product space $\mathbb{V}_a \otimes \mathbb{V}_b$ we can form product operators $\hat{C}_a \otimes \hat{D}_b$, which act on the kets as follows:

$$(\hat{C}_a \otimes \hat{D}_b)(|p\rangle \otimes |v\rangle) = (\hat{C}_a|p\rangle) \otimes (\hat{D}_b|v\rangle).$$

Here it is particularly important to be clear that we are not multiplying \hat{C}_a and \hat{D}_b together; they act in different spaces. Once again \otimes should be regarded as a separator, not a multiplication.

Denoting the identity operators in each space as \hat{I}_a and \hat{I}_b respectively, in the product space the identity operator is $\hat{I}_a \otimes \hat{I}_b$. An operator in which each additive term acts in only one space, such as $\hat{C}_a \otimes \hat{I}_b + \hat{I}_a \otimes \hat{D}_b$, is called a **separable operator**. $\hat{C}_a \otimes \hat{I}_b$ and $\hat{I}_a \otimes \hat{D}_b$ commute.

The inverse of $\hat{C}_a \otimes \hat{D}_b$ is $\hat{C}_a^{-1} \otimes \hat{D}_b^{-1}$ and the adjoint, $\hat{C}_a^\dagger \otimes \hat{D}_b^\dagger$. (The order is NOT reversed, since each still has to act in the correct space.)

Matrix elements work as follows: $(\langle p| \otimes \langle v|) (\hat{C}_a \otimes \hat{D}_b) (|q\rangle \otimes |w\rangle) = \langle p|\hat{C}_a|q\rangle \langle v|\hat{D}_b|w\rangle$. (This is the arithmetic product of two scalars.)

The labels a and b are redundant since the order of the operators in the product tells us which acts in which space. The names of the operators might also tell us. Alternatively if we keep the labels, it is common to write \hat{C}_a when we mean $\hat{C}_a \otimes \hat{I}_b$ and $\hat{C}_a \hat{D}_b$ (or even, since they commute, $\hat{D}_b \hat{C}_a$) when we mean $\hat{C}_a \otimes \hat{D}_b$.

If an operator is separable, i.e. it can be written as $\hat{C}_a \otimes \hat{I}_b + \hat{I}_a \otimes \hat{D}_b$, then the eigenvectors are $|c_i\rangle \otimes |d_j\rangle$ with eigenvalue $c_i + d_j$. As already mentioned the operator is often written $\hat{C}_a + \hat{D}_b$, where the label makes clear which space each operator acts in; similarly the eigenstates are often written $|c_i, d_j\rangle$.

Examples of separable operators are the centre-of-mass position of a two-particle system, $m_1\hat{\mathbf{x}}_1 + m_2\hat{\mathbf{x}}_2$, which we would never think of writing in explicit direct-product form but which, none-the-less, acts on states of two particles such as $\psi_a(\mathbf{r}_1) \otimes \phi_b(\mathbf{r}_2)$. Another is the total angular momentum $\hat{\mathbf{J}} = \hat{\mathbf{L}} + \hat{\mathbf{S}}$. Here we will sometimes find it useful to write $\hat{J}_z = \hat{L}_z \otimes \hat{I} + \hat{I} \otimes \hat{S}_z$ to stress the fact that the operators act on different aspects of the state, see later in the course.

Vectors and operators in finite-dimensional direct-product spaces can be represented by column vectors and matrices, just as in any other vector space. In the two-Qbit space above, consider the single-Qbit operator $\hat{\Omega}$ defined by $\hat{\Omega}|0\rangle = |1\rangle$ and $\hat{\Omega}|1\rangle = |0\rangle$. Taking the states $\{|I\rangle \dots |D\rangle$ defined above as a basis, so that $|1\rangle \rightarrow (1, 0, 0, 0)^\top$ etc, the operator $\hat{\Omega} \otimes \hat{\Omega}$ has matrix elements such as

$$\langle B|\hat{\Omega} \otimes \hat{\Omega}|C\rangle = (\langle 0|\otimes\langle 1|)(\hat{\Omega} \otimes \hat{\Omega})(|1\rangle \otimes |0\rangle) = \langle 0|\hat{\Omega}|1\rangle\langle 1|\hat{\Omega}|0\rangle = 1 \times 1 = 1$$

and

$$\begin{aligned} \hat{\Omega} \otimes \hat{\Omega} &\rightarrow \begin{pmatrix} \langle A|\hat{\Omega} \otimes \hat{\Omega}|A\rangle & \langle A|\hat{\Omega} \otimes \hat{\Omega}|B\rangle & \langle A|\hat{\Omega} \otimes \hat{\Omega}|C\rangle & \langle A|\hat{\Omega} \otimes \hat{\Omega}|D\rangle \\ \langle B|\hat{\Omega} \otimes \hat{\Omega}|A\rangle & \langle B|\hat{\Omega} \otimes \hat{\Omega}|B\rangle & \langle B|\hat{\Omega} \otimes \hat{\Omega}|C\rangle & \langle B|\hat{\Omega} \otimes \hat{\Omega}|D\rangle \\ \langle C|\hat{\Omega} \otimes \hat{\Omega}|A\rangle & \langle C|\hat{\Omega} \otimes \hat{\Omega}|B\rangle & \langle C|\hat{\Omega} \otimes \hat{\Omega}|C\rangle & \langle C|\hat{\Omega} \otimes \hat{\Omega}|D\rangle \\ \langle D|\hat{\Omega} \otimes \hat{\Omega}|A\rangle & \langle D|\hat{\Omega} \otimes \hat{\Omega}|B\rangle & \langle D|\hat{\Omega} \otimes \hat{\Omega}|C\rangle & \langle D|\hat{\Omega} \otimes \hat{\Omega}|D\rangle \end{pmatrix} \\ &= \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \end{aligned}$$

Since $|0\rangle \pm |1\rangle$ are eigenvectors of $\hat{\Omega}$ with eigenvalues ± 1 , the four eigenvectors of $\hat{\Omega} \otimes \hat{\Omega}$ are $(|0\rangle \pm |1\rangle) \otimes (|0\rangle \pm |1\rangle)$ (where the two \pm are independent), i.e.

$$|A\rangle + |B\rangle + |C\rangle + |D\rangle \rightarrow \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} \quad \text{and} \quad |A\rangle - |B\rangle - |C\rangle + |D\rangle \rightarrow \begin{pmatrix} 1 \\ -1 \\ -1 \\ 1 \end{pmatrix}$$

both with eigenvalue $+1$, and

$$|A\rangle + |B\rangle - |C\rangle - |D\rangle \rightarrow \begin{pmatrix} 1 \\ 1 \\ -1 \\ -1 \end{pmatrix} \quad \text{and} \quad |A\rangle - |B\rangle + |C\rangle - |D\rangle \rightarrow \begin{pmatrix} 1 \\ -1 \\ 1 \\ -1 \end{pmatrix}$$

both with eigenvalue -1 . Of course because there is degeneracy, these are not unique; for example the simpler vectors $|A\rangle \pm |D\rangle$ and $|B\rangle \pm |C\rangle$ are also eigenvectors with eigenvalue ± 1 . It is easily checked that these are indeed eigenvectors of the corresponding matrix above. They are orthogonal as expected.

1.7.3 Two or more particle wave functions

Shankar 10.1; Mandl 1.4; Griffiths 5.1 We can now better understand the states of two particles with wave function $\Psi(\mathbf{r}_1, \mathbf{r}_2)$ obeying the Schrödinger equation, as mentioned above. The corresponding two-particle state $|\Psi\rangle$ space is a direct-product of the single-particle spaces and the basis kets of the position representation are $|\mathbf{r}_1\rangle \otimes |\mathbf{r}_2\rangle$, with $\Psi(\mathbf{r}_1, \mathbf{r}_2) = (\langle \mathbf{r}_1| \otimes \langle \mathbf{r}_2|)|\Psi\rangle$.

For states of non-interacting distinguishable particles the two-particle Hamiltonian will be separable, $\hat{H} = \hat{H}^{(1)} \otimes \hat{I}^{(2)} + \hat{I}^{(1)} \otimes \hat{H}^{(2)}$ (or just $\hat{H}^{(1)} + \hat{H}^{(2)}$). If $|\psi\rangle$ and $|\phi\rangle$ are eigenstates of the single-particle Hamiltonians respectively, separable states such as $|\psi\rangle \otimes |\phi\rangle$ will be eigenstates

of the two-particle Hamiltonian with an energy which is just $E^{(1)} + E^{(2)}$, and with wave function $\langle \mathbf{r}_1 | \psi_a \rangle \langle \mathbf{r}_2 | \phi \rangle = \psi(\mathbf{r}_1)\phi(\mathbf{r}_2)$. In general though the Hamiltonian will contain interaction terms $V(\mathbf{r}_1 - \mathbf{r}_2)$ and these separable states will not be eigenstates. But even then the set of all such states will still form a basis in the full space, and solutions can be written as superpositions of these basis states. For identical particles $\hat{H}^{(1)}$ and $\hat{H}^{(2)}$ are the same with both $|\psi\rangle$ and $|\phi\rangle$ as eigenstates. Since we can't label the particles themselves, the only allowed states are those which don't allow us to say which one is in which state, such as $\sqrt{\frac{1}{2}}(\psi(\mathbf{r}_1)\phi(\mathbf{r}_2) \pm \phi(\mathbf{r}_1)\psi(\mathbf{r}_2))$. The extension to more than two particles is (conceptually) straightforward.

1.7.4 Decoupling

One final note: we often encounter systems with many degrees of freedom but ignore all but one for simplicity. This is legitimate if the the Hamiltonian does not couple them, or if the coupling is sufficiently weak, because then the states in the space of interest will evolve independently of the other degrees of freedom. For instance we typically ignore the centre-of-mass motion of a two-particle system in the absence of external forces, because it is trivial (it will be at rest or have constant velocity). Centre-of-mass and relative motions generally **decouple**, as is the case in classical physics too. We often discuss the Schrödinger equation for spatial wave functions and ignore spin, or focus purely on spin for illustrative purposes, which again is fine if spin and spatial motion are only weakly coupled.

2. Angular momentum

We have previously encountered orbital angular momentum via the vector operator $\hat{\mathbf{L}} = \hat{\mathbf{r}} \times \hat{\mathbf{p}}$, the components of which obey the commutation relations

$$[\hat{L}_x, \hat{L}_y] = i\hbar\hat{L}_z, \quad [\hat{L}_y, \hat{L}_z] = i\hbar\hat{L}_x, \quad [\hat{L}_z, \hat{L}_x] = i\hbar\hat{L}_y, \quad [\hat{\mathbf{L}}^2, \hat{L}_i] = 0,$$

where $\hat{\mathbf{L}}^2 = \hat{L}_x^2 + \hat{L}_y^2 + \hat{L}_z^2$. We also found that $\hat{\mathbf{L}}^2$ is identical to the angular part of $-\hbar^2 r^2 \nabla^2$ and so occurs in 3D problems; it commutes with the Hamiltonian if the potential $V(r)$ is independent of angle. The commutation relations imply that we can only simultaneously know $\hat{\mathbf{L}}^2$ and one component, taken conventionally to be \hat{L}_z . The common eigenfunctions of $\hat{\mathbf{L}}^2$ and \hat{L}_z are the spherical harmonics, $Y_l^m(\theta, \phi)$ and the requirements that these be single-valued and finite everywhere restrict the eigenvalues to be $\hbar^2 l(l+1)$ and $\hbar m$ respectively, where l and m are integers which satisfy

$$l = 0, 1, 2, \dots, \quad m = -l, -l+1, \dots, l.$$

More details including the forms of the differential operators for the \hat{L}_i and of the first few spherical harmonics can be found in section A.3.4. A recap of other aspects of angular momentum from last year can be found in section A.5. In this course, we are going to re-derive the properties of angular momentum, and discover some new ones, based only on the commutation relations and nothing else.

2.1 General properties of angular momentum

Shankar 12.5; Griffiths 4.3; Mandl 5.2

Summary: Focussing purely on the commutation relations of the operators, we can derive all the properties of orbital angular momentum without a θ or ϕ in sight—and discover spin too.

In the case of the harmonic oscillator, we found that an approach which focused on operators and abstract states rather than differential equations was extremely powerful. We are going to do something similar with angular momentum, with the added incentive that we know that orbital angular momentum is not the only possible form, we will need to include spin as well—and that has no classical analogue or position-space description.

Consider three Hermitian operators, \hat{J}_1 , \hat{J}_2 and \hat{J}_3 , components of the vector operator $\hat{\mathbf{J}}$, about which we will only assume one thing, their commutation relations:

$$[\hat{J}_1, \hat{J}_2] = i\hbar\hat{J}_3, \quad [\hat{J}_2, \hat{J}_3] = i\hbar\hat{J}_1, \quad [\hat{J}_3, \hat{J}_1] = i\hbar\hat{J}_2 \quad (2.1)$$

or succinctly, $[\hat{J}_i, \hat{J}_j] = i\hbar \sum_k \epsilon_{ijk} \hat{J}_k$.¹ It can be shown that the orbital angular momentum operator defined previously satisfy these rules, but we want to be more general, hence the new name $\hat{\mathbf{J}}$, and the use of indices 1-3 rather than x, y, z . Note that $\hat{\mathbf{J}}$ has the same dimensions (units) as \hbar .

From these follow the fact that all three commute with $\hat{\mathbf{J}}^2 = \hat{J}_1^2 + \hat{J}_2^2 + \hat{J}_3^2$:

$$[\hat{\mathbf{J}}^2, \hat{J}_i] = 0$$

It follows that we will in general be able to find simultaneous eigenstates of $\hat{\mathbf{J}}^2$ and *only one* of the components \hat{J}_i . We quite arbitrarily choose \hat{J}_3 . We denote the normalised states $|\lambda, \mu\rangle$ with eigenvalue $\hbar^2\lambda$ of $\hat{\mathbf{J}}^2$ and eigenvalue $\hbar\mu$ of \hat{J}_3 . (We've written these so that λ and μ are dimensionless.) All we know about μ is that it is real, but recalling that for any state and Hermitian operator, $\langle\alpha|\hat{A}^2|\alpha\rangle = \langle\hat{A}\alpha|\hat{A}\alpha\rangle \geq 0$, we know in addition that λ must be non-negative. Furthermore

$$\hbar^2(\lambda - \mu^2) = \langle\lambda, \mu|(\hat{\mathbf{J}}^2 - \hat{J}_3^2)|\lambda, \mu\rangle = \langle\lambda, \mu|(\hat{J}_1^2 + \hat{J}_2^2)|\lambda, \mu\rangle \geq 0$$

so $|\mu| \leq \sqrt{\lambda}$. The magnitude of a component of a vector can't be bigger than the length of the vector!

Now let us define raising and lowering operators \hat{J}_\pm (appropriateness of the names still to be shown):

$$\hat{J}_+ \equiv \hat{J}_1 + i\hat{J}_2; \quad \hat{J}_- \equiv \hat{J}_1 - i\hat{J}_2.$$

Note these are not Hermitian, but $\hat{J}_- = \hat{J}_+^\dagger$. These satisfy the following commutation relations:

$$\begin{aligned} [\hat{J}_+, \hat{J}_-] &= 2\hbar\hat{J}_3, & [\hat{J}_3, \hat{J}_+] &= \hbar\hat{J}_+, & [\hat{J}_3, \hat{J}_-] &= -\hbar\hat{J}_-, \\ [\hat{\mathbf{J}}^2, \hat{J}_\pm] &= 0, \\ \hat{\mathbf{J}}^2 &= \frac{1}{2}(\hat{J}_+\hat{J}_- + \hat{J}_-\hat{J}_+) + \hat{J}_3^2 = \hat{J}_+\hat{J}_- + \hat{J}_3^2 - \hbar\hat{J}_3 = \hat{J}_-\hat{J}_+ + \hat{J}_3^2 + \hbar\hat{J}_3. \end{aligned} \quad (2.2)$$

Since \hat{J}_\pm commute with $\hat{\mathbf{J}}^2$, we see that the states $\hat{J}_\pm|\lambda, \mu\rangle$ are also eigenstates of $\hat{\mathbf{J}}^2$ with eigenvalue $\hbar^2\lambda$.

Why the names? Consider the state $\hat{J}_+|\lambda, \mu\rangle$:

$$\hat{J}_3(\hat{J}_+|\lambda, \mu\rangle) = \hat{J}_+\hat{J}_3|\lambda, \mu\rangle + \hbar\hat{J}_+|\lambda, \mu\rangle = \hbar(\mu + 1)(\hat{J}_+|\lambda, \mu\rangle)$$

So *either* $\hat{J}_+|\lambda, \mu\rangle$ is another eigenstate of \hat{J}_3 with eigenvalue $\hbar(\mu + 1)$, *or* it is the null vector. Similarly either $\hat{J}_-|\lambda, \mu\rangle$ is another eigenstate of \hat{J}_3 with eigenvalue $\hbar(\mu - 1)$, or it is the null vector. Leaving aside for a moment the case where the raising or lowering operator annihilates the state, we have $\hat{J}_+|\lambda, \mu\rangle = C_{\lambda\mu}|\lambda, \mu + 1\rangle$, where

$$|C_{\lambda\mu}|^2 = \langle\lambda, \mu|\hat{J}_+^\dagger\hat{J}_+|\lambda, \mu\rangle = \langle\lambda, \mu|\hat{J}_-\hat{J}_+|\lambda, \mu\rangle = \langle\lambda, \mu|(\hat{\mathbf{J}}^2 - \hat{J}_3^2 - \hbar\hat{J}_3)|\lambda, \mu\rangle = \hbar^2(\lambda - \mu^2 - \mu)$$

There is an undetermined phase that we can choose to be +1, so $C_{\lambda\mu} = \hbar\sqrt{\lambda - \mu^2 - \mu}$.

We can repeat the process to generate more states with quantum numbers $\mu \pm 2, \mu \pm 3 \dots$ unless we reach states that are annihilated by the raising or lowering operators. All these states are in the λ -subspace of $\hat{\mathbf{J}}^2$.

¹ ϵ_{ijk} is 1 if i, j, k is a cyclic permutation of 1, 2, 3, -1 if an anticyclic permutation such as 2, 1, 3 and 0 if any two indices are the same.

However we saw above that the magnitude of the eigenvalues μ of \hat{J}_3 must not be greater than $\sqrt{\lambda}$. So the process cannot go on indefinitely, there must be a maximum and minimum value μ_{\max} and μ_{\min} , such that $\hat{J}_+|\lambda, \mu_{\max}\rangle = 0$ and $\hat{J}_-|\lambda, \mu_{\min}\rangle = 0$. Furthermore by repeated action of \hat{J}_- , we can get from $|\lambda, \mu_{\max}\rangle$ to $|\lambda, \mu_{\min}\rangle$ in an integer number of steps: $\mu_{\max} - \mu_{\min}$ is an integer, call it N .

Now the expectation value of $\hat{J}_-\hat{J}_+$ in the state $|\lambda, \mu_{\max}\rangle$ must also be zero, but as we saw above that expectation value, for general μ , is $C_{\lambda\mu}^2 = \hbar^2(\lambda - \mu^2 - \mu)$. Thus

$$\lambda - \mu_{\max}(\mu_{\max} + 1) = 0.$$

Similarly, considering the expectation value of $\hat{J}_+\hat{J}_-$ in the state $|\lambda, \mu_{\min}\rangle$ gives

$$\lambda - \mu_{\min}(\mu_{\min} - 1) = 0.$$

Taking these two equations together with $\mu_{\min} = \mu_{\max} - N$, we find

$$\mu_{\max}(\mu_{\max} + 1) = (\mu_{\max} - N)(\mu_{\max} - N - 1) \Rightarrow (N + 1)(2\mu_{\max} - N) = 0 \Rightarrow \mu_{\max} = \frac{N}{2}.$$

Hence μ_{\max} is either an integer or a half-integer, $\mu_{\min} = -\mu_{\max}$ and there are $2\mu_{\max} + 1$ possible values of μ . Furthermore λ is restricted to values $\lambda = \frac{N}{2}(\frac{N}{2} + 1)$ for *any* integer $N \geq 0$.

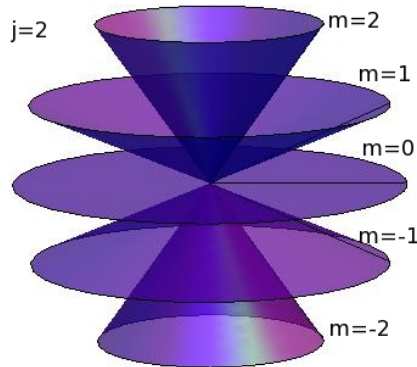
Let's compare with what we found for orbital angular momentum. There we found that what we have called λ had to have the form $l(l + 1)$ for integer l , and what we've called μ was an integer m , with $-l \leq m \leq l$. That agrees exactly with the integer case above. From now on we will use m for μ , and j for μ_{\max} ; furthermore instead of writing the state $|j(j + 1), m\rangle$ we will use $|j, m\rangle$. We refer to it as "a state with angular momentum j " but this is sloppy—if universally understood; the magnitude of the angular momentum is $\hbar\sqrt{j(j + 1)}$. The component of this along any axis, though, cannot be greater than $\hbar j$.

But there is one big difference between the abstract case and the case of orbital angular momentum, and that is that j can be half integer $\frac{1}{2}, \frac{3}{2}, \dots$. If these cases are realised in Physics, the source of the angular momentum cannot be orbital, but something without any parallel in classical Physics.

We end this section by rewriting the relations we have already found in terms of j and m , noting m can only take the one of the $2j + 1$ values $-j, -j + 1, \dots, j - 1, j$:

$$\begin{aligned} \mathbf{J}^2|j, m\rangle &= \hbar^2 j(j + 1)|j, m\rangle; & \hat{J}_3|j, m\rangle &= \hbar m|j, m\rangle; \\ \hat{J}_{\pm}|j, m\rangle &= \hbar\sqrt{j(j + 1) - m(m \pm 1)}|j, m \pm 1\rangle. \end{aligned} \tag{2.3}$$

In the diagram below, the five cones show the possible locations of the angular momentum vector with length $\sqrt{6}\hbar$ and z -component m . The x - and y -components are not fixed, but must satisfy $\langle \hat{J}_1^2 + \hat{J}_2^2 \rangle = (6 - m^2)\hbar^2 > 0$ (unless $j = 0$).



2.2 Electron spin and the Stern-Gerlach experiment

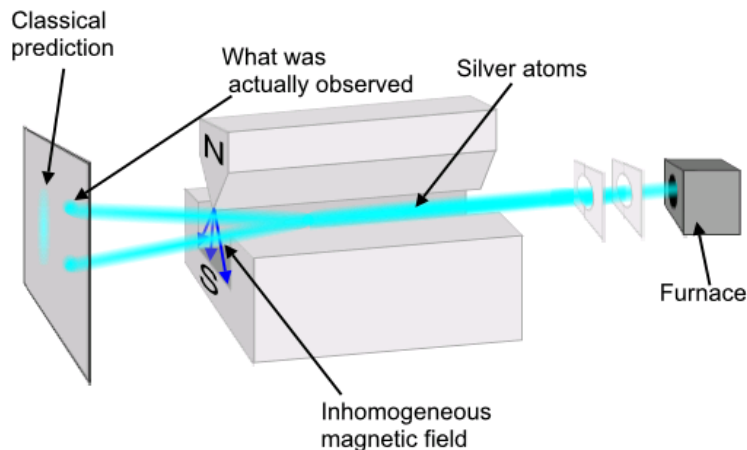
(Shankar 14.5); Griffiths 4.4; Mandl 5.2

Summary: Electrons have an intrinsic angular momentum—spin—independently of their orbital motion.

From classical physics, we know that charged systems with angular momentum have a magnetic moment $\boldsymbol{\mu}$, which means that they experience a torque $\boldsymbol{\mu} \times \mathbf{B}$ if not aligned with an external magnetic field \mathbf{B} , and their interaction energy with the magnetic field is $-\boldsymbol{\mu} \cdot \mathbf{B}$. For an electron in a circular orbit with angular momentum \mathbf{L} , the classical prediction is $\boldsymbol{\mu} = -(e/2m)\mathbf{L} = -(\mu_B/\hbar)\mathbf{L}$, where $\mu_B = |e|\hbar/2m$ is called the Bohr magneton and has dimensions of a magnetic moment.

Since the torque is perpendicular to the angular momentum the system is like a gyroscope, and (classically) the direction of the magnetic moment precesses about \mathbf{B} , with L_z being unchanged. If the field is not uniform, though, there will also be a net force causing the whole atom to move so as to reduce its energy $-\boldsymbol{\mu} \cdot \mathbf{B}$; taking the magnetic field along the +ve z axis the atom will move to regions of stronger field if $\mu_z > 0$ but to weaker field regions if $\mu_z < 0$. If a beam of atoms enters a region of inhomogeneous magnetic field one (classically) expects the beam to spread out, each atom having a random value of μ_z and so being deflected a different amount.

The Stern-Gerlach experiment, in 1922, aimed to test if silver atoms had a magnetic moment, and found that it did. The figure below (from Wikipedia) shows the apparatus; the shape of the poles of the magnet ensures that the field is stronger near the upper pole than the lower one.



The first run just showed a smearing of the beam demonstrating that there was a magnetic moment, but further running showed that atoms were actually deflected either up or down by a fixed amount, indicating that μ_z only had two possible values relative to the magnetic field. The deflection was what would be expected for $L_z = \pm\hbar$. That accorded nicely with Bohr's planetary orbits, and was taken as a confirmation of a prediction of what we now call the "old" quantum theory.

From a post-1926 perspective, however, $l = 1$ would give three spots ($m = -1, 0, 1$) not two—and anyway we now know that the electrons in silver atoms have zero net orbital magnetic

moment. By that time though other considerations, particularly the so-called anomalous Zeeman splitting of spectroscopic lines in a magnetic field, had caused first Kronig then, in 1925, Goudschmidt and Uhling, to suggest that electrons could have a further source of angular momentum that they called spin, which would have only two possible values ($m = -\frac{1}{2}, +\frac{1}{2}$) but which couples twice as strongly to a magnetic field as orbital angular momentum ($g_s = 2$)—hence the Stern-Gerlach result ($\hat{\boldsymbol{\mu}} = -(\mu_B/\hbar)2\hat{\mathbf{S}}$). We now know that the electron does indeed carry an intrinsic angular momentum, called spin but not mechanical in origin, which is an example of the $j = \frac{1}{2}$ possibility that we deduced above.

Thus the full specification of the state of an electron has two parts, spatial and spin. The vector space is a *tensor direct product space* of the space of square-integrable functions of which the spatial state is a member, states like $|\psi_{\mathbf{r}}(t)\rangle$, for which $\langle \mathbf{r} | \psi_{\mathbf{r}}(t) \rangle = \psi(\mathbf{r}, t)$, and spin space, containing states $|\psi_s(t)\rangle$, the nature of which we will explore in more detail in the next section. While in non-relativistic QM this has to be put in by hand, it emerges naturally from the Dirac equation, which also predicts $g_s = 2$.

Because this product space is itself a vector space, sums of vectors are in the space and not all states of the system are separable, (that is, they do not all have the form $|\psi_{\mathbf{r}}(t)\rangle \otimes |\psi_s(t)\rangle$). We can also have states like $|\psi_{\mathbf{r}}(t)\rangle \otimes |\psi_s(t)\rangle + |\phi_{\mathbf{r}}(t)\rangle \otimes |\phi_s(t)\rangle$. As we will see spin-space is two dimensional (call the basis $\{|+\rangle, |-\rangle\}$ just now), so including spin doubles the dimension of the state space; as a result we never need more than two terms, and can write

$$|\Psi(t)\rangle = c_1 |\psi_{\mathbf{r}}(t)\rangle \otimes |+\rangle + c_2 |\phi_{\mathbf{r}}(t)\rangle \otimes |-\rangle.$$

But this still means that the electron has two spatial wave functions, one for each spin state. In everything we've done so far the spin is assumed not to be affected by the dynamics, in which case we return to a single common spatial state. But that is not general.

2.3 Spin- $\frac{1}{2}$

Shankar 14.3; Griffiths 4.4; Mandl 5.3

Summary: A 2-D complex vector space perfectly describes spin

Whereas with orbital angular momentum we were talking about an infinite-dimension space which could be considered as a sum of subspaces with $l = 0, 1, 2, \dots$, when we talk about intrinsic angular momentum—spin—we are confined to a single subspace with fixed j . We also use $\hat{\mathbf{S}}$ in place of $\hat{\mathbf{J}}$, but the operators \hat{S}_i obey the same rules as the J_i . The simultaneous eigenstates of $\hat{\mathbf{S}}^2$ and \hat{S}_z are $|s, m\rangle$, but as ALL states in the space have the same s , we often drop it in the notation. In this case, $s = \frac{1}{2}$, $m = -\frac{1}{2}, +\frac{1}{2}$, so the space is two-dimensional with a basis variously denoted

$$\{|\frac{1}{2}, \frac{1}{2}\rangle, |\frac{1}{2}, -\frac{1}{2}\rangle\} \equiv \{|\frac{1}{2}\rangle, |-\frac{1}{2}\rangle\} \equiv \{|+\rangle, |-\rangle\} \equiv \{|\hat{\mathbf{z}}+\rangle, |\hat{\mathbf{z}}-\rangle\}$$

In the last case $\hat{\mathbf{z}}$ is a unit vector in the z -direction, so we are making it clear that these are states with spin-up (+) and spin-down (-) in the z -direction. We will also construct states with definite spin in other directions.

In this basis, the matrices representing \hat{S}_z (which is diagonal), \hat{S}_+ and $\hat{S}_- = \hat{S}_+^\dagger$, can be written down directly. Recall $\hat{J}_+ |j, m\rangle = \hbar \sqrt{j(j+1) - m(m+1)} |j, m+1\rangle$, so

$$\hat{S}_+ |\frac{1}{2}, -\frac{1}{2}\rangle = \hbar \sqrt{\frac{3}{4} + \frac{1}{4}} |\frac{1}{2}, \frac{1}{2}\rangle, \quad \hat{S}_+ |\frac{1}{2}, \frac{1}{2}\rangle = 0,$$

and so $\langle \frac{1}{2}, \frac{1}{2} | \hat{S}_+ | \frac{1}{2}, -\frac{1}{2} \rangle = \hbar$ is the only non-vanishing matrix element of \hat{S}_+ . From these $\hat{S}_x = \frac{1}{2}(\hat{S}_+ + \hat{S}_-)$ and $\hat{S}_y = -\frac{1}{2}i(\hat{S}_+ - \hat{S}_-)$ can be constructed:

$$\begin{aligned} |\hat{\mathbf{z}}+\rangle &\xrightarrow{S_z} \begin{pmatrix} 1 \\ 0 \end{pmatrix} & |\hat{\mathbf{z}}-\rangle &\xrightarrow{S_z} \begin{pmatrix} 0 \\ 1 \end{pmatrix} & \hat{S}_+ &\xrightarrow{S_z} \hbar \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} & \hat{S}_- &\xrightarrow{S_z} \hbar \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \\ \hat{S}_z &\xrightarrow{S_z} \frac{\hbar}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} & \hat{S}_x &\xrightarrow{S_z} \frac{\hbar}{2} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} & \hat{S}_y &\xrightarrow{S_z} \frac{\hbar}{2} \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}. \end{aligned}$$

The label S_z on the arrows reminds us of the particular basis we are using. It is easily shown that the matrices representing the \hat{S}_i obey the required commutation relations.

The matrices

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

are called the Pauli matrices. They obey $\sigma_i \sigma_j = \delta_{ij} \mathbf{I} + i \sum_k \epsilon_{ijk} \sigma_k$. The three together, $\boldsymbol{\sigma}$, form a vector (a vector of matrices, as $\hat{\mathbf{S}}$ is a vector of operators) while for some ‘‘ordinary’’ vector \mathbf{a} , $\mathbf{a} \cdot \boldsymbol{\sigma}$ is a single Hermitian matrix:

$$\mathbf{a} \cdot \boldsymbol{\sigma} = \begin{pmatrix} a_z & a_x - ia_y \\ a_x + ia_y & -a_z \end{pmatrix}.$$

and $\mathbf{a} \cdot \boldsymbol{\sigma} \mathbf{b} \cdot \boldsymbol{\sigma} = \mathbf{a} \cdot \mathbf{b} \mathbf{I} + i(\mathbf{a} \times \mathbf{b}) \cdot \boldsymbol{\sigma}$. Together with the identity matrix they form a basis (with real coefficients) for all Hermitian 2×2 matrices.

The component of $\hat{\mathbf{S}}$ in an arbitrary direction defined by the unit vector \mathbf{n} is $\hat{\mathbf{S}} \cdot \mathbf{n}$. We can parametrise the direction of \mathbf{n} by the polar angles θ, ϕ , so $\mathbf{n} = \sin \theta \cos \phi \mathbf{e}_x + \sin \theta \sin \phi \mathbf{e}_y + \cos \theta \mathbf{e}_z$. Then in the basis of eigenstates of \hat{S}_z , $\hat{\mathbf{S}} \cdot \mathbf{n}$ and its eigenstates are given by

$$\hat{\mathbf{S}} \cdot \mathbf{n} \xrightarrow{S_z} \frac{\hbar}{2} \begin{pmatrix} \cos \theta & \sin \theta e^{-i\phi} \\ \sin \theta e^{i\phi} & -\cos \theta \end{pmatrix} \quad |\mathbf{n}+\rangle \xrightarrow{S_z} \begin{pmatrix} \cos \frac{\theta}{2} e^{-i\phi/2} \\ \sin \frac{\theta}{2} e^{i\phi/2} \end{pmatrix} \quad |\mathbf{n}-\rangle \xrightarrow{S_z} \begin{pmatrix} -\sin \frac{\theta}{2} e^{-i\phi/2} \\ \cos \frac{\theta}{2} e^{i\phi/2} \end{pmatrix}$$

The eigenvalues, of course, are $\pm \hbar/2$. (There is nothing special about the z -direction.)

The most general normalised state of a spin-half vector has the form

$$a|\hat{\mathbf{z}}+\rangle + b|\hat{\mathbf{z}}-\rangle \xrightarrow{S_z} \begin{pmatrix} a \\ b \end{pmatrix} \quad \text{where } |a|^2 + |b|^2 = 1.$$

In fact any such state is an eigenstate of $\hat{\mathbf{S}} \cdot \mathbf{n}$ for some direction \mathbf{n} given by $\tan \frac{1}{2}\theta = |b/a|$ and $\phi = \text{Arg}(b) - \text{Arg}(a)$.

Note that (from the matrix representation, or by checking explicitly for the two states of a basis) $(2\hat{\mathbf{S}} \cdot \mathbf{n}/\hbar)^2 = \hat{I}$ for any \mathbf{n} . So

$$\exp(i(\alpha/\hbar)\hat{\mathbf{S}} \cdot \mathbf{n}) = \cos \frac{\alpha}{2} \hat{I} + i \sin \frac{\alpha}{2} \frac{2}{\hbar} \hat{\mathbf{S}} \cdot \mathbf{n}.$$

The lack of higher powers of \hat{S}_i follows from the point about Hermitian operators noted above.

Some calculation in the matrix basis reveals the useful fact that $\langle \mathbf{n} \pm | \hat{\mathbf{S}} | \mathbf{n} \pm \rangle = \pm \frac{\hbar}{2} \mathbf{n}$; that is, the *expectation value* of the vector operator $\hat{\mathbf{S}}$ is parallel or antiparallel to \mathbf{n} .

2.3.1 Spin precession

The Hamiltonian of a spin- $\frac{1}{2}$ electron in a uniform magnetic field is, with $g_s = 2$ and charge $-|e|$,

$$\hat{H} = -\boldsymbol{\mu} \cdot \mathbf{B} = (g_s \mu_B / \hbar) \hat{\mathbf{S}} \cdot \mathbf{B} \xrightarrow{S_z} \mu_B \boldsymbol{\sigma} \cdot \mathbf{B}.$$

Consider the case of a field in the x -direction, so that $\hat{H} \xrightarrow{S_z} \mu_B B \sigma_x$, and a particle initially in the state $|\hat{\mathbf{z}}+\rangle$. It turns out that we have already done this problem in Section 1.4.1, with the γ of that problem being equal to $\mu_B B / \hbar$. Giving the result in terms of $\omega = 2\gamma = 2\mu_B B / \hbar$, which is the frequency corresponding to the energy splitting of the eigenstates of \hat{H} , we obtained

$$|\psi(t)\rangle = \cos(\omega t/2) |\hat{\mathbf{z}}+\rangle - i \sin(\omega t/2) |\hat{\mathbf{z}}-\rangle. \quad \langle \psi(t) | \hat{S}_z | \psi(t) \rangle = (\hbar/2) \cos \omega t.$$

To this we can now add $\langle \psi(t) | \hat{S}_y | \psi(t) \rangle = -(\hbar/2) \sin \omega t$, and $\langle \psi(t) | \hat{S}_x | \psi(t) \rangle = 0$. So the *expectation value* of $\hat{\mathbf{S}}$ is a vector of length $\hbar/2$ in the yz plane which rotates with frequency $\omega = 2\mu_B B / \hbar$. This is exactly what we would get from Ehrenfest's theorem.

Alternatively, we can take the magnetic field along $\hat{\mathbf{z}}$ so that the energy eigenstates are $|\hat{\mathbf{z}}\pm\rangle$ with energies $\pm\mu_B B \equiv \pm\hbar\omega/2$. If the initial state is spin-up in an arbitrary direction \mathbf{n} , that is $|\mathbf{n}+\rangle$, we can decompose this in terms of the energy eigenstates, each with its own energy dependence, and obtain

$$|\psi(t)\rangle = \cos \frac{\theta}{2} e^{-i(\omega t + \phi)/2} |\hat{\mathbf{z}}+\rangle + \sin \frac{\theta}{2} e^{i(\omega t + \phi)/2} |\hat{\mathbf{z}}-\rangle = |\mathbf{n}(t)\rangle$$

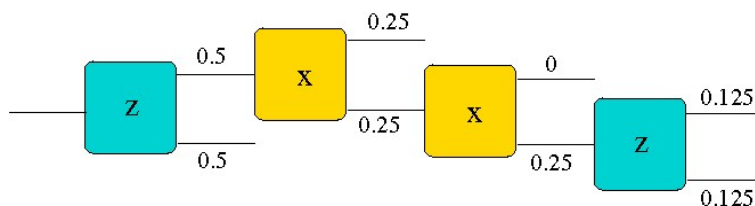
where $\mathbf{n}(t)$ is a vector which, like the original \mathbf{n} , is oriented at an angle θ to the $\hat{\mathbf{z}}$ (i.e. \mathbf{B}) axis, but which rotates about that axis so that the azimuthal angle changes with time: $\phi(t) = \phi(0) + \omega t$. The expectation value $\langle \hat{\mathbf{S}} \rangle$ precesses likewise, following the same behaviour as a classical magnetic moment of $-g_s \mu_B$.

2.3.2 Spin and measurement: Stern Gerlach revisited

We now understand, in a way that the original experimenters did not, what the Stern-Gerlach experiment does: for each atom that passes through, it measures the component of the spin angular momentum along the direction \mathbf{n} of the magnetic field, $\mathbf{S} \cdot \mathbf{n}$. Each time, the answer is either up or down, $\pm\hbar/2$. With the initial beam being unpolarised, the numbers of up and down will be equal.

The apparatus also gives us access to a beam of particles which are all spin-up in a particular direction; say the z direction. We can then run that beam through a second copy of the apparatus rotated through some angle θ relative to the first. The particles exiting from this copy will be either spin-up or down along the new magnetic field axis, and the probability of getting each is $|\langle \mathbf{n} \pm \hat{\mathbf{z}}+\rangle|^2$, that is $\cos^2 \frac{\theta}{2}$ and $\sin^2 \frac{\theta}{2}$ respectively. If $\theta = \pi/2$ (new field along the x axis, assuming a beam in the y -direction), the probabilities are both 50%.

Successive measurements can be schematically represented below, each block being labelled by the direction of the magnetic field. It should look very familiar.



2.4 Higher spins

Summary: Angular momentum s is represented by a complex $(2s + 1)$ -D vector space

The Particle Data Group lists spin-1 particles and spin- $\frac{3}{2}$ particles; gravitons if they exist are spin-2, and nuclei can have much higher spins (at least $\frac{9}{2}$ for known ground states of stable nuclei).

Furthermore, since in many situations total angular momentum commutes with the Hamiltonian (see later), even when orbital angular momentum is involved we are often only concerned with a subspace of fixed j (or l or s). All such sub-spaces are finite dimensional, of dimension $N = 2j + 1$, and spanned by the basis $\{|j, m\rangle\}$ with $m = j, j - 1 \dots -j + 1, -j$. It is most usual (though of course not obligatory) to order the states by descending m .

In this subspace, with this basis, the operators $\hat{J}_x, \hat{J}_y, \hat{J}_z$ are represented by three $N \times N$ matrices with matrix elements eg $(J_x)_{m'm} = \langle j, m' | \hat{J}_x | j, m \rangle$. (Because states with different j are orthogonal, and because the \hat{J}_i only change m , not j , $\langle j', m' | \hat{J}_x | j, m \rangle = 0$ if $j \neq j'$: that's why we can talk about non-overlapping subspaces in the first place.) The matrix representation of \hat{J}_z of course is diagonal with diagonal elements $j, j - 1 \dots -j + 1, -j$. As with spin- $\frac{1}{2}$, it is easiest to construct \hat{J}_+ first, then \hat{J}_- as its transpose (the elements of the former having been chosen to be real), then $\hat{J}_x = \frac{1}{2}(\hat{J}_+ + \hat{J}_-)$ and $\hat{J}_y = -\frac{1}{2}i(\hat{J}_+ - \hat{J}_-)$.

As an example we construct the matrix representation of the operators for spin-1. The three basis states $|s, m\rangle$ are $|1, 1\rangle, |1, 0\rangle$ and $|1, -1\rangle$. Recall $\hat{J}_+ |j, m\rangle = \hbar\sqrt{j(j+1)-m(m+1)} |j, m+1\rangle$, so $\hat{S}_+ |1, -1\rangle = \hbar\sqrt{2-0} |1, 0\rangle$, $\hat{S}_+ |1, 0\rangle = \hbar\sqrt{2-0} |1, 1\rangle$ and $\hat{S}_+ |1, 1\rangle = 0$, and the only non-zero matrix elements of \hat{S}_+ are

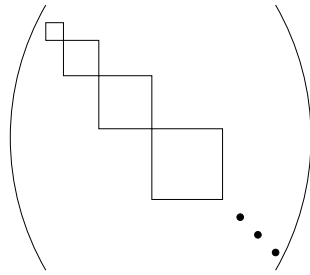
$$\langle 1, 1 | \hat{S}_+ | 1, 0 \rangle = \langle 1, 0 | \hat{S}_+ | 1, -1 \rangle = \sqrt{2}\hbar.$$

So:

$$\begin{aligned} |1, 1\rangle &\xrightarrow{s_z} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} & |1, 0\rangle &\xrightarrow{s_z} \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} & |1, -1\rangle &\xrightarrow{s_z} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \\ \hat{S}_z &\xrightarrow{s_z} \hbar \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix} & \hat{S}_+ &\xrightarrow{s_z} \sqrt{2}\hbar \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} & \hat{S}_- &\xrightarrow{s_z} \sqrt{2}\hbar \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \\ \hat{S}_x &\xrightarrow{s_z} \frac{\hbar}{\sqrt{2}} \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} & \hat{S}_y &\xrightarrow{s_z} \frac{\hbar}{\sqrt{2}} \begin{pmatrix} 0 & -i & 0 \\ i & 0 & -i \\ 0 & i & 0 \end{pmatrix} \end{aligned}$$

Of course this is equally applicable to any system with $j = 1$, including the $l = 1$ spherical harmonics.

Once all possible values of j and m are allowed, any angular momentum operator is represented in the $\{|j, m\rangle\} = \{|0, 0\rangle, |\frac{1}{2}, \frac{1}{2}\rangle, |\frac{1}{2}, -\frac{1}{2}\rangle, |1, 1\rangle, |1, 0\rangle, |1, -1\rangle \dots\}$ basis by a block-diagonal matrix



The first block is a single element, zero in fact, since all components of $\hat{\mathbf{J}}$ in the one-dimensional space of states of $j = 0$ are zero. The next block is the appropriate 2×2 spin- $\frac{1}{2}$ matrix, the next a 3×3 spin-1 matrix, and so on. This block-diagonal structure reflects the fact that the vector space can be written as a *direct sum* of spaces with $j = 0, j = \frac{1}{2}, j = 1, \dots$: $\mathbb{V} = \mathbb{V}^1 \oplus \mathbb{V}^2 \oplus \mathbb{V}^3 \oplus \dots$ (where the superscripts of course are $2j + 1$).

In fact, any given physical system can *only* have integer or half-integer angular momentum. So the picture would be similar, but with only odd- or even-dimensional blocks. For orbital angular momentum, for instance, the blocks would be $1 \times 1, 3 \times 3, 5 \times 5, \dots$

2.5 Addition of angular momenta

Shankar 15.1,2; Griffiths 4.4; Mandl 4.4

Summary: Two sources of angular momentum are described by a direct-product vector space. States with definite total angular momentum J are not separable, but form an alternative basis in the space.

Up till now, we have in general spoken rather loosely as if an electron has either orbital or spin angular momentum—or more precisely, we've considered cases where only one affects the dynamics, so we can ignore the other. But many cases are not like that. The electron in an excited hydrogen atom may have orbital as well as spin angular momentum, and if is placed in a magnetic field, both will affect how the energy levels shift, and hence how the spectral lines split. Or the deuteron (heavy hydrogen nucleus) consists of both a proton and a neutron, and both have spin; heavier atoms and nuclei have many components all with spin and angular momentum. Only the total angular momentum of the whole system is guaranteed by rotational symmetry to be conserved in the absence of external fields. So we need to address the question of the addition of angular momentum.

Because the notation is clearest, we will start with the spin and orbital angular momentum of a particle. We consider the case where l as well as s is fixed: electrons in a p -wave orbital, for instance. These two types of angular momentum are independent and live in different vector spaces, so this is an example of a tensor direct product space, spanned by the basis $\{|l, m_l\rangle \otimes |s, m_s\rangle\}$ and hence $(2l + 1) \times (2s + 1)$ dimensional.

Now angular momentum is a vector, and we expect the total angular momentum to be the vector sum of the orbital and spin angular momenta. We can form a new vector operator in the product space

$$\hat{\mathbf{J}} = \hat{\mathbf{L}} \otimes \hat{I} + \hat{I} \otimes \hat{\mathbf{S}} \quad \hat{\mathbf{J}}^2 = \hat{\mathbf{L}}^2 \otimes \hat{I} + \hat{I} \otimes \hat{\mathbf{S}}^2 + 2\hat{\mathbf{L}} \otimes \hat{\mathbf{S}}$$

where the last term represents a scalar product as well as a tensor product and would more clearly be written $2(\hat{L}_x \otimes \hat{S}_x + \hat{L}_y \otimes \hat{S}_y + \hat{L}_z \otimes \hat{S}_z)$.

In practice, the tensor product notation for operators proves cumbersome, and we always just write

$$\hat{\mathbf{J}} = \hat{\mathbf{L}} + \hat{\mathbf{S}} \quad \hat{\mathbf{J}}^2 = \hat{\mathbf{L}}^2 + \hat{\mathbf{S}}^2 + 2\hat{\mathbf{L}} \cdot \hat{\mathbf{S}}$$

We know that the \hat{L}_i and \hat{S}_i act on different parts of the state, and we don't need to stress that when we act with \hat{S}_i alone we are not changing the orbital state, etc. An alternative form, in which the tensor product notation is again suppressed, is

$$\hat{\mathbf{J}}^2 = \hat{\mathbf{L}}^2 + \hat{\mathbf{S}}^2 + \hat{L}_+ \hat{S}_- + \hat{L}_- \hat{S}_+ + 2\hat{L}_z \hat{S}_z.$$

Now in calling the sum of angular momenta $\hat{\mathbf{J}}$, which we previously used for a generic angular momentum, we are assuming that the \hat{J}_i do indeed obey the defining commutation rules for angular momentum, and this can easily be demonstrated. For instance

$$[\hat{J}_x, \hat{J}_y] = [\hat{L}_x + \hat{S}_x, \hat{L}_y + \hat{S}_y] = [\hat{L}_x, \hat{L}_y] + [\hat{S}_x, \hat{S}_y] = i\hbar\hat{L}_z + i\hbar\hat{S}_z = i\hbar\hat{J}_z,$$

where we have used the fact that $[\hat{L}_i, \hat{S}_j] = 0$, since they act in different spaces. Hence we expect that an alternative basis in the product space will be $\{|j, m_j\rangle\}$, with allowed values of j not yet determined. The question we want to answer, then, is the connection between the $\{|l, m_l\rangle \otimes |s, m_s\rangle\}$ and $\{|j, m_j\rangle\}$ bases. Both, we note, must have dimension $(2l+1) \times (2s+1)$.

We note some other points about the commutator: \hat{L}_z , \hat{S}_z and \hat{J}_z all commute; \hat{J}_z commutes with $\hat{\mathbf{J}}^2$ (of course) and with $\hat{\mathbf{L}}^2$ and with $\hat{\mathbf{S}}^2$ (because both \hat{L}_z and \hat{S}_z do), but \hat{L}_z and \hat{S}_z do not commute with $\hat{\mathbf{J}}^2$. Thus we can, as implied when we wrote down the two bases, always specify l and s , but then *either* m_l and m_s (with $m_j = m_l + m_s$) *or* j and m_j . (We will sometimes write $|l, s; j, m_j\rangle$ instead of just $|j, m_j\rangle$, if we need a reminder of l and s in the problem.) What this boils down to is that the state of a given j and m_j will be linear superpositions of the states of given m_s and m_l that add up to that m_j . If there is more than one such state, there must be more than one allowed value of j for that m_j .

Let's introduce a useful piece of jargon: the state of maximal m in a multiplet, $|j, j\rangle$, is called the **stretched state**.

We start with the state of maximal m_l and m_s $|l, l\rangle \otimes |s, s\rangle$, which has $m_j = l + s$. This is clearly the maximal value of m_j , and hence of j : $j_{\max} = l + s$, and since the state is unique, it must be an eigenstate of $\hat{\mathbf{J}}^2$ ². If we act on this with $\hat{J}_- = \hat{L}_- + \hat{S}_-$, we get a new state with two terms in it; recalling the general rule $\hat{J}_- |j, m\rangle = \hbar\sqrt{j(j+1) - m(m-1)} |j, m-1\rangle$ where j can stand for j or l or s , we have (using \bar{j} as a shorthand for $j_{\max} = l + s$)

$$\begin{aligned} |\bar{j}, \bar{j}\rangle = |l, l\rangle \otimes |s, s\rangle &\Rightarrow \hat{J}_- |\bar{j}, \bar{j}\rangle = (\hat{L}_- |l, l\rangle) \otimes |s, s\rangle + |l, l\rangle \otimes (\hat{S}_- |s, s\rangle) \\ &\Rightarrow \sqrt{2\bar{j}} |\bar{j}, \bar{j}-1\rangle = \sqrt{2l} |l, l-1\rangle \otimes |s, s\rangle + \sqrt{2s} |l, l\rangle \otimes |s, s-1\rangle \end{aligned}$$

From this state we can continue operating with \hat{J}_- ; at the next step there will be three terms on the R.H.S. with $\{m_l, m_s\}$ equal to $\{l-2, s\}$, $\{l-1, s-1\}$ and $\{l, s-2\}$, then four, but eventually we will reach states which are annihilated by \hat{L}_- or \hat{S}_- and the number of terms will start to shrink again, till we finally reach $|\bar{j}, -\bar{j}\rangle = |l, -l\rangle \otimes |s, -s\rangle$ after $2\bar{j}$ steps ($2\bar{j}+1$ states in all).

²This can also be seen directly by acting with $\hat{\mathbf{J}}^2 = \hat{J}_- \hat{J}_+ + \hat{J}_z^2 + \hbar\hat{J}_z$, since $|l, l\rangle \otimes |s, s\rangle$ is an eigenstate of \hat{J}_z with eigenvalue $\hbar(l+s)$, and is annihilated by both \hat{L}_+ and \hat{S}_+ , and hence by \hat{J}_+ .

Whichever is the smaller of l or s will govern the maximum number of $\{m_l, m_s\}$ that can equal any m_j ; for example if s is smaller, the maximum number is $2s + 1$.

Now the state we found with $m_j = l+s-1$ is not unique, there must be another orthogonal combination of the two states with $\{m_l, m_s\}$ equal to $\{l-1, s\}$ and $\{l, s-1\}$. This cannot be part of a multiplet with $j = \bar{j}$ because we've "used up" the only state with $m_j = \bar{j}$. So it must be the highest m_j state (the stretched state) of a multiplet with $j = \bar{j} - 1$ (ie $l + s - 1$):

$$|\bar{j} - 1, \bar{j} - 1\rangle = -\sqrt{\frac{s}{l+s}} |l, l-1\rangle \otimes |s, s\rangle + \sqrt{\frac{l}{l+s}} |l, l\rangle \otimes |s, s-1\rangle$$

Successive operations with \hat{J}_- will generate the rest of the multiplet ($2\bar{j} - 1$ in all); all the states will be orthogonal to the states of the same m_j but higher j already found.

However there will be a third linear combination of the states with $\{m_l, m_s\}$ equal to $\{l-2, s\}$, $\{l-1, s-1\}$ and $\{l, s-2\}$, which cannot have $j = \bar{j}$ or $\bar{j}-1$. So it must be the stretched state of a multiplet with $j = \bar{j}-2$, ($2\bar{j} - 3$ states in all).

And so it continues, generating multiplets with successively smaller values of j . However the process will come to an end. As we saw, the maximum number of terms in any sum is whichever is smaller of $2l + 1$ or $2s + 1$, so this is also the maximum number of mutually orthogonal states of the same m_j , and hence the number of different values of j . So j can be between $l + s$ and the larger of $l + s - 2s$ and $l + s - 2l$; that is, $l + s \geq j \geq |l - s|$. The size of the $\{|j, m_j\rangle\}$ basis is then $\sum_{j=|l-s|}^{l+s} 2j + 1$, which is equal to $(2l + 1)(2s + 1)$.

The table below illustrates the process for $l = 2, s = 1$; we go down a column by applying \hat{J}_- , and start a new column by constructing a state orthogonal to those in the previous columns. The three columns correspond to $j = 3, j = 2$ and $j = 1$, and there are $7 + 5 + 3 = 5 \times 3$ states in total.

$ 3, 3\rangle = 2, 2\rangle \otimes 1, 1\rangle$	$ 2, 2\rangle = -\sqrt{\frac{1}{3}} 2, 1\rangle \otimes 1, 1\rangle + \sqrt{\frac{2}{3}} 2, 2\rangle \otimes 1, 0\rangle$	$ 1, 1\rangle = \sqrt{\frac{1}{10}} 2, 0\rangle \otimes 1, 1\rangle - \sqrt{\frac{3}{10}} 2, 1\rangle \otimes 1, 0\rangle$
$ 3, 2\rangle = \sqrt{\frac{2}{3}} 2, 1\rangle \otimes 1, 1\rangle + \sqrt{\frac{1}{3}} 2, 2\rangle \otimes 1, 0\rangle$	$ 2, 1\rangle = -\sqrt{\frac{1}{2}} 2, 0\rangle \otimes 1, 1\rangle + \sqrt{\frac{1}{6}} 2, 1\rangle \otimes 1, 0\rangle$	$ 1, 0\rangle = \sqrt{\frac{3}{10}} 2, -1\rangle \otimes 1, 1\rangle - \sqrt{\frac{2}{5}} 2, 0\rangle \otimes 1, 0\rangle$
$ 3, 1\rangle = \sqrt{\frac{2}{5}} 2, 0\rangle \otimes 1, 1\rangle + \sqrt{\frac{8}{15}} 2, 1\rangle \otimes 1, 0\rangle$	$ 2, 0\rangle = -\sqrt{\frac{1}{2}} 2, -1\rangle \otimes 1, 1\rangle + 0 2, 0\rangle \otimes 1, 0\rangle$	$ 1, -1\rangle = \sqrt{\frac{3}{5}} 2, -2\rangle \otimes 1, 1\rangle - \sqrt{\frac{3}{10}} 2, -1\rangle \otimes 1, 0\rangle$
$ 3, 0\rangle = \sqrt{\frac{1}{5}} 2, -1\rangle \otimes 1, 1\rangle + \sqrt{\frac{8}{15}} 2, 0\rangle \otimes 1, 0\rangle$	$ 2, -1\rangle = -\sqrt{\frac{1}{3}} 2, -2\rangle \otimes 1, 1\rangle - \sqrt{\frac{1}{6}} 2, -1\rangle \otimes 1, 0\rangle$	$ 1, -1\rangle = \sqrt{\frac{3}{5}} 2, -2\rangle \otimes 1, 1\rangle - \sqrt{\frac{3}{10}} 2, -1\rangle \otimes 1, 0\rangle$
$ 3, -1\rangle = \sqrt{\frac{1}{15}} 2, -2\rangle \otimes 1, 1\rangle + \sqrt{\frac{8}{15}} 2, -1\rangle \otimes 1, 0\rangle$	$ 2, -2\rangle = -\sqrt{\frac{2}{3}} 2, -2\rangle \otimes 1, 0\rangle + \sqrt{\frac{1}{3}} 2, -1\rangle \otimes 1, -1\rangle$	$ 1, -1\rangle = \sqrt{\frac{3}{5}} 2, -2\rangle \otimes 1, 1\rangle - \sqrt{\frac{3}{10}} 2, -1\rangle \otimes 1, 0\rangle$
$ 3, -2\rangle = \sqrt{\frac{1}{3}} 2, -2\rangle \otimes 1, 0\rangle + \sqrt{\frac{2}{3}} 2, -1\rangle \otimes 1, -1\rangle$	$ 3, -3\rangle = 2, -2\rangle \otimes 1, -1\rangle$	$ 1, -1\rangle = \sqrt{\frac{3}{5}} 2, -2\rangle \otimes 1, 1\rangle - \sqrt{\frac{3}{10}} 2, -1\rangle \otimes 1, 0\rangle$

The coefficients in the table are called **Clebsch-Gordan coefficients**. They are the inner products $(\langle l, m_l | \otimes \langle s, m_s |) | j, m_j \rangle$ but that is too cumbersome a notation; with a minimum modification Shankar uses $\langle l, m_l; s, m_s | j, m_j \rangle$; Mandl uses $C(l, m_l, s, m_s; j, m_j)$, but $\langle l, s, m_l, m_s | j, m_j \rangle$ and other minor modifications, including dropping the commas, are common. They are all totally clear when symbols are being used, but easily confused when numerical values are substituted! We use the "Condon-Shortley" phase convention, which is the most common; in this convention Clebsch-Gordan coefficients are real, which is why we won't write $\langle l, m_l; s, m_s | j, m_j \rangle^*$ in the second equation of Eq. (2.4) below.

All of this has been written for the addition of orbital and spin angular momenta. But we did not actually assume at any point that l was integer. So in fact the same formulae apply for the addition of any two angular momenta of any origin: a very common example is two spin- $\frac{1}{2}$ particles. The more general form for adding two angular momenta j_1 and j_2 , with J and M

being the quantum numbers corresponding to the total angular momentum of the system, is

$$\begin{aligned}
 |J, M\rangle &= \sum_{m_1, m_2} \langle j_1, m_1; j_2, m_2 | J, M \rangle |j_1, m_1\rangle \otimes |j_2, m_2\rangle, \\
 |j_1, m_1\rangle \otimes |j_2, m_2\rangle &= \sum_{J, M} \langle j_1, m_1; j_2, m_2 | J, M \rangle |J, M\rangle.
 \end{aligned}
 \tag{2.4}$$

For the common case of $s = \frac{1}{2}$, $j = l \pm \frac{1}{2}$, we have

$$|l \pm \frac{1}{2}, m_j\rangle = \sqrt{\frac{l \mp m_j + \frac{1}{2}}{2l+1}} |l, m_j + \frac{1}{2}\rangle \otimes |\frac{1}{2}, -\frac{1}{2}\rangle \pm \sqrt{\frac{l \pm m_j + \frac{1}{2}}{2l+1}} |l, m_j - \frac{1}{2}\rangle \otimes |\frac{1}{2}, \frac{1}{2}\rangle.$$

To summarise, the states of a system with two contributions to the angular momentum, j_1 and j_2 , written in a basis in which the total angular momentum J and z -component M are specified; the values of J range from $|j_1 - j_2|$ to $j_1 + j_2$ in unit steps. In this basis the total angular momentum operators \hat{J}_i and $\hat{\mathbf{J}}^2$ are cast in block-diagonal form, one $(2J+1)$ -square block for each value of J . The vector space, which we started by writing as a product, $\mathbb{V}^{2j_1+1} \otimes \mathbb{V}^{2j_2+1}$, can instead be written as a direct sum: $\mathbb{V}^{2(j_1+j_2)+1} \oplus \dots \oplus \mathbb{V}^{2|j_1-j_2|+1}$. In particular for some orbital angular momentum l and $s = \frac{1}{2}$, $\mathbb{V}^{2l+1} \otimes \mathbb{V}^2 = \mathbb{V}^{2l+2} \oplus \mathbb{V}^{2l}$. The overall dimension of the space is of course unchanged.

2.5.1 Using tables of Clebsch-Gordan tables

Griffiths 4.4

General formulae for the Clebsch-Gordan coefficients are not used (the already-met case of $s = \frac{1}{2}$ is an exception). One may use the Mathematica function `ClebschGordan` $[\{l, m_l\}, \{s, m_s\}, \{j, m_j\}]$, and on-line calculator at [Wolfram Alpha](#).

Frequently though one consults tables and this section gives instructions on their use.

In a system with two contributions to angular momentum j_1 and j_2 , Clebsch-Gordan coefficients are used to write states good of total angular momentum J and z -component M , $|j_1, j_2; J, M\rangle$ or just $|J, M\rangle$, in terms of the basis $\{m_1, m_2\}$, $|j_1, m_1\rangle \otimes |j_2, m_2\rangle$:

$$\begin{aligned}
 |j_1, j_2; J, M\rangle &= \sum_{m_1 m_2} \langle j_1, m_1; j_2, m_2 | J, M \rangle \left(|j_1, m_1\rangle \otimes |j_2, m_2\rangle \right) \quad \text{and} \\
 |j_1, m_1\rangle \otimes |j_2, m_2\rangle &= \sum_{J, M} \langle j_1, m_1; j_2, m_2 | J, M \rangle |j_1, j_2; J, M\rangle
 \end{aligned}$$

where the numbers denoted by $\langle j_1, m_1; j_2, m_2 | J, M \rangle$ are the Clebsch-Gordan coefficients; they vanish unless $j_1 + j_2 \geq J \geq |j_1 - j_2|$, and $m_1 + m_2 = M$. There is a conventional tabulation which can be found in various places including the [Particle Data Group](#) site, but the notation takes some explanation.

There is one table for each j_1, j_2 pair. The table consists of a series of blocks, one for each value of M . Along the top are possible values of $\frac{J}{M}$ and at the left are possible values of $m_1 m_2$.

Each block stands for something which could be written like this one for $j_1 = 1$, $j_2 = \frac{1}{2}$ and

$$M = m_1 + m_2 = \frac{1}{2}$$

		J	$3/2$	$1/2$
		M	$+1/2$	$+1/2$
m_1	m_2			
$+1$	$-1/2$		$1/3$	$2/3$
0	$+1/2$		$2/3$	$-1/3$

For compactness the numbers in the blocks are the coefficients squared times their sign; thus $-\frac{1}{2}$ stands for $-\sqrt{\frac{1}{2}}$.

As an example consider the table for coupling $j_1 = 1$ and $j_2 = \frac{1}{2}$ to get $J = \frac{3}{2}$ or $\frac{1}{2}$. For clarity we will use the notation $|J, M\rangle$ in place of $|j_1, j_2; J, M\rangle$.

In red the coefficients of $|1, 1\rangle \otimes |\frac{1}{2}, -\frac{1}{2}\rangle$ and $|1, 0\rangle \otimes |\frac{1}{2}, \frac{1}{2}\rangle$ in $|\frac{1}{2}, \frac{1}{2}\rangle$ are highlighted

Notation:

		J	J	...
		M	M	...
m_1	m_2	Coefficients		
m_1	m_2			
.	.			
.	.			

$1 \times 1/2$

		$3/2$			
		$+3/2$	$3/2$	$1/2$	
$+1$	$+1/2$	1	$+1/2$	$+1/2$	
$+1$	$-1/2$	$1/3$	$2/3$	$3/2$	$1/2$
0	$+1/2$	$2/3$	$-1/3$	$1/2$	$-1/2$
		0	$-1/3$	$3/3$	$-1/3$
		-1	$+1/2$	$1/3$	$-2/3$
				-1	$-1/2$
					1

$$|\frac{1}{2}, \frac{1}{2}\rangle = \sqrt{\frac{2}{3}}|1, 1\rangle \otimes |\frac{1}{2}, -\frac{1}{2}\rangle - \sqrt{\frac{1}{3}}|1, 0\rangle \otimes |\frac{1}{2}, \frac{1}{2}\rangle.$$

In green are the components for the decomposition

$$|1, -1\rangle \otimes |\frac{1}{2}, \frac{1}{2}\rangle = \sqrt{\frac{1}{3}}|\frac{3}{2}, -\frac{1}{2}\rangle - \sqrt{\frac{2}{3}}|\frac{1}{2}, -\frac{1}{2}\rangle.$$

Or for coupling $j_1 = \frac{3}{2}$ and $j_2 = 1$:

$3/2 \times 1$

		$5/2$			
		$+5/2$	$5/2$	$3/2$	
$+3/2$	$+1$	1	$+3/2$	$+3/2$	
$+3/2$	0	$2/5$	$3/5$	$5/2$	$3/2$
$+1/2$	$+1$	$3/5$	$-2/5$	$+1/2$	$+1/2$
$+3/2$	-1	$1/10$	$2/5$	$1/2$	
$+1/2$	0	$3/5$	$1/15$	$-1/3$	$5/2$
$-1/2$	$+1$	$3/10$	$-8/15$	$1/6$	$-1/2$
		$3/10$	$8/15$	$1/6$	$3/2$
		$-1/2$	0	$3/5$	$-1/15$
		$-3/2$	$+1$	$1/10$	$-2/5$
				$1/2$	$1/2$
				$1/2$	-1
				$3/5$	$2/5$
				$2/5$	$-3/5$
				$-3/2$	-1
					1

giving for example

$$|\frac{3}{2}, -\frac{1}{2}\rangle = \sqrt{\frac{8}{15}}|\frac{3}{2}, \frac{1}{2}\rangle \otimes |1, -1\rangle - \sqrt{\frac{1}{15}}|\frac{3}{2}, -\frac{1}{2}\rangle \otimes |1, 0\rangle - \sqrt{\frac{2}{5}}|\frac{3}{2}, -\frac{3}{2}\rangle \otimes |1, 1\rangle$$

If instead one wants $j_1 = 1$ and $j_2 = \frac{3}{2}$, we use the relation

$$\langle j_2, m_2; j_1, m_1 | J, M \rangle = (-1)^{J-j_1-j_2} \langle j_1, m_1; j_2, m_2 | J, M \rangle$$

Note that table of Clebsch-Gordan coefficients are given for states of j_1 and j_2 coupling up to total J . But as j is a generic angular momentum, that covers s and l coupling to j , or s_1 and s_2 coupling to S etc.

2.5.2 Example: Two spin- $\frac{1}{2}$ particles

Here we will call the operators $\hat{\mathbf{S}}^{(1)}$, $\hat{\mathbf{S}}^{(2)}$ and $\hat{\mathbf{S}} = \hat{\mathbf{S}}^{(1)} + \hat{\mathbf{S}}^{(2)}$ for the individual and total spin operators, and S and M for the total spin quantum numbers. (The use of capitals is standard in a many-particle system.) Because both systems are spin- $\frac{1}{2}$, we will omit the label from our states, which we will write in the $\{m_1, m_2\}$ basis as

$$|1\rangle = |+\rangle \otimes |+\rangle, \quad |2\rangle = |+\rangle \otimes |-\rangle, \quad |3\rangle = |-\rangle \otimes |+\rangle, \quad |4\rangle = |-\rangle \otimes |-\rangle.$$

(The 1...4 are just labels here.) In this basis

$$\hat{S}_+ \longrightarrow \hbar \begin{pmatrix} 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \quad \hat{S}_z \longrightarrow \hbar \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \quad \hat{\mathbf{S}}^2 \longrightarrow \hbar^2 \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix}$$

where we use explicit calculation for the matrix elements, eg

$$\langle 1 | (\hat{S}_+^{(1)} + \hat{S}_+^{(2)}) | 2 \rangle = \langle + | \hat{S}_+^{(1)} | + \rangle \langle + | \hat{I}^{(2)} | - \rangle + \langle + | \hat{I}^{(1)} | + \rangle \langle + | \hat{S}_+^{(2)} | - \rangle = 0 + \hbar,$$

then $\hat{S}_- = (\hat{S}_+)^{\dagger}$ and $\hat{\mathbf{S}}^2 = \hat{S}_+ \hat{S}_- + \hat{S}_z^2 - \hbar \hat{S}_z$.

It is clear that $|1\rangle$ and $|4\rangle$ are eigenstates of $\hat{\mathbf{S}}^2$ with eigenvalue $2\hbar^2$ and hence $S = 1$. They are also eigenstates of \hat{S}_z with eigenvalues $\pm\hbar$. In the $\{|2\rangle, |3\rangle\}$ subspace, which has $M = 0$, $\hat{\mathbf{S}}^2$ is represented by the matrix $\hbar^2 \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$ which has eigenvalues $2\hbar^2$ and 0 corresponding to states $\sqrt{\frac{1}{2}}(|2\rangle \pm |3\rangle)$. We label these four simultaneous eigenstates of $\hat{\mathbf{S}}^2$ and \hat{S}_z as $|S, M\rangle$, and take the ordering for the new basis as $\{|0, 0\rangle, |1, 1\rangle, |1, 0\rangle, |1, -1\rangle\}$. Then the matrix of eigenvectors, \mathbf{U} , is

$$\mathbf{U} = \hbar \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1/\sqrt{2} & 0 & 1/\sqrt{2} & 0 \\ -1/\sqrt{2} & 0 & 1/\sqrt{2} & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

and the transformed matrices $\mathbf{U}^{\dagger} \mathbf{S}_i \mathbf{U}$ are

$$\hat{S}_x \longrightarrow \frac{\hbar}{\sqrt{2}} \begin{pmatrix} 0 & | & 0 & 0 & 0 \\ 0 & | & 0 & 1 & 0 \\ 0 & | & 1 & 0 & 1 \\ 0 & | & 0 & 1 & 0 \end{pmatrix} \quad \hat{S}_y \longrightarrow \frac{\hbar}{\sqrt{2}} \begin{pmatrix} 0 & | & 0 & 0 & 0 \\ 0 & | & 0 & -i & 0 \\ 0 & | & i & 0 & -i \\ 0 & | & 0 & i & 0 \end{pmatrix} \quad \hat{S}_z \longrightarrow \hbar \begin{pmatrix} 0 & | & 0 & 0 & 0 \\ 0 & | & 1 & 0 & 0 \\ 0 & | & 0 & 0 & 0 \\ 0 & | & 0 & 0 & -1 \end{pmatrix}$$

where the 1×1 plus 3×3 block-diagonal structure has been emphasised and the 3×3 blocks are just the spin-1 matrices we found previously.

2.5.3 Angular Momentum of Atoms and Nuclei

Both atoms and nuclei consist of many spin- $\frac{1}{2}$ fermions, each of which has both spin and orbital angular momentum. In the independent-particle model we think of each fermion occupying a well-defined single-particle orbital which is an eigenstate of a central potential and hence has well defined orbital angular momentum l . The notation s, p, d, f, g, \dots is used for orbitals of $l = 0, 1, 2, 3, 4, \dots$. For each fermion there is also a total angular momentum j . All the angular momenta of all the fermions can be added in a variety of ways, and the following quantum numbers are defined: L for the sum of all the orbital angular momenta (that is, the eigenvalues of $\hat{\mathbf{L}}_{\text{tot}}^2$ are $\hbar^2 L(L+1)$); S for the sum of all the spin angular momenta, and J for the total angular momentum of the atom or nucleus from all sources. The use of capitals for the quantum numbers shouldn't be confused with the operators themselves.

In reality the independent-particle model is only an approximation, and only the total angular momentum J is a conserved quantum number (only $\hat{\mathbf{J}}_{\text{tot}}^2$ commutes with the Hamiltonian of the whole system). For light atoms, it is a good starting point to treat L and S as if they were conserved too, and the notation $^{2S+1}L_J$ is used, with L being denoted by S, P, D, F, G, \dots . This is termed LS coupling. So 3S_1 has $L = 0, S = J = 1$. Energy levels of light atoms are split according to J by the spin-orbit splitting $\propto \hat{\mathbf{L}} \cdot \hat{\mathbf{S}}$ (of which more later). However for heavy atoms and for nuclei, it is a better approximation to sum the individual total angular momenta j to give J as the only good quantum number. ($j-j$ coupling.)

Somewhat confusingly, J is often called the spin of the atom or nucleus, even though its origin is both spin and angular momentum. This composite origin shows up in a magnetic coupling g which is neither 1 (pure orbital) or 2 (pure spin). For light atoms g can be calculated from L, S and J (the Landé g -factor). For nuclei things are further complicated by the fact that protons and neutrons are not elementary particles, and their "spin" is likewise of composite origin, something which shows up through their g values of $g_p = 5.59$ and $g_n = -3.83$ rather than 2 and 0 respectively. Using these the equivalent of the Landé g -factor can be calculated for individual nucleon orbitals, and hence for those odd-even nuclei for which the single-particle model works (that is, assuming that only the last unpaired nucleon contributes to the total angular momentum). Beyond that it gets complicated.

2.6 Vector Operators

Shankar 15.3

*This section is not examinable. The take-home message is that **vector operators** such as $\hat{\mathbf{x}}$ and $\hat{\mathbf{p}}$ can change the angular momentum of the state they act on in the same way as coupling in another source of angular momentum with $l = 1$. If the components of the vector operator are written in a spherical basis analogously to $\hat{\mathbf{J}}_{\pm}$, the dependence of the matrix elements on the m quantum numbers is given by Clebsch Gordan coefficients, with the non-trivial dependence residing only in a single "reduced matrix element" for each pair j and j' of the angular momenta of the initial and final states. This is the **Wigner-Eckart theorem** of Eq. (2.6).*

We have now met a number of vector operators: $\hat{\mathbf{x}} = (\hat{x}, \hat{y}, \hat{z})$, $\hat{\mathbf{p}} = (\hat{p}_x, \hat{p}_y, \hat{p}_z)$, and of course $\hat{\mathbf{L}}, \hat{\mathbf{S}}$ and $\hat{\mathbf{J}}$. We have seen, either in lectures or examples, that they all satisfy the following

relation: if $\hat{\mathbf{V}}$ stands for the vector operator

$$[\hat{J}_i, \hat{V}_j] = \sum_k i\hbar \epsilon_{ijk} \hat{V}_k$$

for example, $[\hat{J}_x, \hat{y}] = i\hbar \hat{z}$. (We could have substituted \hat{L}_x for \hat{J}_x here as spin and space operators commute.)

We can take this to be the definition of a vector operator: a triplet of operators makes up a vector operator if it satisfies these commutation relations.

Just as it was useful to define \hat{J}_+ and \hat{J}_- , so it is useful to define

$$\hat{V}_{+1} = -\sqrt{\frac{1}{2}}(\hat{V}_1 + i\hat{V}_2) \quad \hat{V}_{-1} = \sqrt{\frac{1}{2}}(\hat{V}_1 - i\hat{V}_2) \quad \hat{V}_0 = \hat{V}_3$$

where the subscripts are no longer Cartesian coordinates ($1 \equiv x$ etc) but analogous to the m of the spherical harmonics—and indeed

$$\mp \sqrt{\frac{1}{2}}(x \pm iy) = \sqrt{\frac{4\pi}{3}} r Y_1^{\pm 1}(\theta, \phi) \quad z = \sqrt{\frac{4\pi}{3}} r Y_1^0(\theta, \phi).$$

Note a slight change of normalisation and sign: $\hat{J}_{\pm 1} = \mp \sqrt{\frac{1}{2}} \hat{J}_{\pm}$. In terms of these spherical components \hat{V}_m ,

$$[\hat{J}_0, \hat{V}_m] = m\hbar \hat{V}_m \quad [\hat{J}_{\pm}, \hat{V}_m] = \hbar \sqrt{(1 \mp m)(2 \pm m)} \hat{V}_{m \pm 1}$$

If we compare these to the effects on states

$$\hat{J}_3 |j, m\rangle = \hbar m |j, m\rangle \quad \hat{J}_{\pm} |j, m\rangle = \hbar \sqrt{(j \mp m)(j \pm m + 1)} |j, m \pm 1\rangle$$

we see a close parallel, so long as we take $j = 1$ for the vector operators.³

Consider the following two calculations. First, we consider matrix elements of the commutator of the components of a tensor operator \hat{V}_m with \hat{J}_{\pm} , in which $l = 1$, and p and q are magnetic quantum numbers like m ; in the second line we note that $\langle j', m | \hat{J}_{\pm}$ is the bra associated with $\hat{J}_{\mp} |j, m\rangle$

$$\begin{aligned} \langle j', p | [\hat{J}_{\pm}, \hat{V}_m] | j, q \rangle &= \hbar \sqrt{(l \mp m)(l \pm m + 1)} \langle j', p | \hat{V}_{m \pm 1} | j, q \rangle \\ \text{and} \quad \langle j', p | \hat{J}_{\pm} \hat{V}_m - \hat{V}_m \hat{J}_{\pm} | j, q \rangle &= \hbar \sqrt{(j' \pm p)(j' \mp p + 1)} \langle j', p \mp 1 | \hat{V}_m | j, q \rangle \\ &\quad - \hbar \sqrt{(j \mp q)(j \pm q + 1)} \langle j', p | \hat{V}_m | j, q \pm 1 \rangle \\ \Rightarrow \sqrt{(l \mp m)(l \pm m + 1)} \langle j', p | \hat{V}_{m \pm 1} | j, q \rangle &= \sqrt{(j' \pm p)(j' \mp p + 1)} \langle j', p \mp 1 | \hat{V}_m | j, q \rangle \\ &\quad - \sqrt{(j \mp q)(j \pm q + 1)} \langle j', p | \hat{V}_m | j, q \pm 1 \rangle \end{aligned}$$

Secondly we take matrix elements of $\hat{J}_{\pm} = \hat{J}_{\pm}^{(1)} + \hat{J}_{\pm}^{(2)}$, giving us a relation between the Clebsch-Gordan coefficients for l and j coupling up to j'

$$\begin{aligned} \langle j', p | \hat{J}_{\pm} (|l, m\rangle \otimes |j, q\rangle) &= \hbar \sqrt{(j' \pm p)(j' \mp p + 1)} \langle j', p \mp 1 | l, m; j, q \rangle \\ \text{and} \quad \langle j', p | \hat{J}_{\pm}^{(1)} + \hat{J}_{\pm}^{(2)} (|l, m\rangle \otimes |j, q\rangle) &= \hbar \sqrt{(l \mp m)(l \pm m + 1)} \langle j', p | l, m \pm 1; j, q \rangle \\ &\quad + \hbar \sqrt{(j \mp q)(j \pm q + 1)} \langle j', p | l, m; j, q \pm 1 \rangle \\ \Rightarrow \sqrt{(l \mp m)(l \pm m + 1)} \langle j', p | l, m \pm 1; j, q \rangle &= \sqrt{(j' \pm p)(j' \mp p + 1)} \langle j', p \mp 1 | l, m; j, q \rangle \\ &\quad - \sqrt{(j \mp q)(j \pm q + 1)} \langle j', p | l, m; j, q \pm 1 \rangle \end{aligned}$$

³Note that in this section, we use the algebraically equivalent $(j \mp m)(j \pm m + 1)$ for $j(j+1) - m(m \pm 1)$ in the normalisation of $\hat{J}_{\pm} |j, m\rangle$.

Comparing the two, we see that the coefficients are identical, but in the first they multiply matrix elements of $\hat{\mathbf{V}}$ whereas in the second, they multiply Clebsch-Gordan coefficients. This can only be true if the matrix elements are proportional to the Clebsch-Gordan coefficients, with a constant of proportionality which must be independent of magnetic quantum numbers, and which we will write as $\langle j' || \hat{\mathbf{V}} || j \rangle$, the **reduced matrix element**:

$$\langle j', p | \hat{V}_m | j, q \rangle = \langle j' || \hat{\mathbf{V}} || j \rangle \langle j', p | j, q; l, m \rangle_{l=1}$$

This is a specific instance of the **Wigner-Eckart theorem**. It says that acting on a state with a vector operator is like coupling in one unit of angular momentum; only states with $|j' - 1| \leq j \leq j' + 1$ and with $p = m + q$ will have non-vanishing matrix elements. It also means that if one calculates one matrix element, which ever is the simplest (so long as it is non-vanishing), then the others can be written down directly.

Since $\hat{\mathbf{J}}$ is a vector operator, it follows that matrix elements of $\hat{\mathbf{J}}_q$ can also be written in terms of a reduced matrix element $\langle j' || \hat{\mathbf{J}} || j \rangle$, but of course this vanishes unless $j' = j$.

Writing $|j_1, j_2; J, M\rangle = \sum_{m_1 m_2} \langle J, M | j_1, m_1; j_2, m_2 \rangle |j_1, m_1\rangle \otimes |j_2, m_2\rangle$, and using orthonormality of the states $\{|J, M\rangle\}$, allows us to show that

$$\sum_{m_1 m_2} \langle J, M | j_1, m_1; j_2, m_2 \rangle \langle J', M' | j_1, m_1; j_2, m_2 \rangle = \delta_{JJ'} \delta_{MM'} \quad (2.5)$$

Noting too that a scalar product of vector operators $\hat{\mathbf{P}} \cdot \hat{\mathbf{Q}}$ can be written in spherical components as $\sum_q (-1)^q \hat{P}_{-q} \hat{Q}_q$, we can show that

$$\begin{aligned} \langle j, m | \hat{\mathbf{P}} \cdot \hat{\mathbf{J}} | j, m \rangle &= \sum_{q, j', m'} (-1)^q \langle j, m | \hat{P}_{-q} | j', m' \rangle \langle j', m' | \hat{J}_q | j, m \rangle \\ &= \sum_{q, m'} \langle j', m' | \hat{P}_q | j, m \rangle \langle j', m' | \hat{J}_q | j, m \rangle = \langle j || \hat{\mathbf{P}} || j \rangle \langle j || \hat{\mathbf{J}} || j \rangle; \end{aligned}$$

(we insert a complete set of states at the first step, then use the Wigner-Eckart theorem and Eq. (2.5)).

Replacing $\hat{\mathbf{P}}$ with $\hat{\mathbf{J}}$ gives us $\langle j || \hat{\mathbf{J}} || j \rangle = \sqrt{j(j+1)}$. Hence we have the extremely useful relation

$$\langle j, m | \hat{\mathbf{P}} | j, m \rangle = \langle j, m | \hat{\mathbf{J}} | j, m \rangle \frac{\langle j, m | \hat{\mathbf{P}} \cdot \hat{\mathbf{J}} | j, m \rangle}{j(j+1)}. \quad (2.6)$$

which we will use in calculating the Landé g factor in the next section.

Finally, we might guess from the way that we used a general symbol l instead of 1, that there are operators which couple in 2 or more units of angular momentum. Simple examples are obtained by writing $r^l Y_l^m$ in terms of x, y and z , then setting $x \rightarrow \hat{x}$ etc; so $(\hat{x} \pm i\hat{y})^2$, $(\hat{x} \pm i\hat{y})\hat{z}$, and $2\hat{z}^2 - \hat{x}^2 - \hat{y}^2$ are the $m = \pm 2$, $m = \pm 1$ and $m = 0$ components of an operator with $l = 2$ (a **rank-two tensor operator**, in the jargon). There are six components of $\hat{x}_i \hat{x}_j$, but $(\hat{z}^2 + \hat{x}^2 + \hat{y}^2)$ is a scalar ($l = 0$). This is an example of the tensor product of two $l = 1$ operators giving $l = 2$ and $l = 0$ operators.

3. Approximate methods I: variational method and WKB

3.1 Approximate methods in Quantum Mechanics

It is often (almost always!) the case that we cannot solve real problems analytically. Only a very few potentials have analytic solutions, by which I mean we can write down the energy levels and wave functions in closed form, as we can for the familiar examples of the square well, harmonic oscillator and Coulomb potential. In fact those are really the only useful ones, as they do crop up in various physical contexts. In the last century, a number of approximate methods have been developed to obtain information about systems which can't be solved exactly.

These days, this might not seem very relevant. Computers can solve differential equations very efficiently. But:

- It is always useful to have a check on numerical methods.
- They don't provide much insight on which aspects of the physics are more important.
- Even supercomputers can't solve the equations for many interacting particles exactly in a reasonable time (where "many" may be as low as four, depending on the complexity of the interaction) — ask a nuclear physicist or quantum chemist.
- Quantum field theories are systems with infinitely many degrees of freedom. All approaches to QFT must be approximate.
- If the system we are interested in is close to a soluble one, we might obtain more insight from approximate methods than from numerical ones. This is the realm of perturbation theory. The most accurate prediction ever made, for the anomalous magnetic moment of the electron, which is good to one part in 10^{12} , is a 4th order perturbative calculation.

In the next chapter we will consider perturbation theory, which is the probably the most widely used approximate method, and in this chapter we consider two other very useful approaches: variational methods and WKB.

3.2 Variational methods: ground state

Shankar 16.1; Mandl 8.1; Griffiths 7.1; Gasiorowicz 14.4

Summary: Whatever potential we are considering, we can always obtain an *upper bound* on the ground-state energy.

Suppose we know the Hamiltonian of a bound system but don't have any idea what the energy of the ground state is, or the wave function. The variational principle states that if we simply guess the wave function, the expectation value of the Hamiltonian in that wave function will be greater than the true ground-state energy:

$$\frac{\langle \Psi | \hat{H} | \Psi \rangle}{\langle \Psi | \Psi \rangle} \geq E_0$$

This initially surprising result is more obvious if we consider expanding the (normalised) $|\Psi\rangle$ in the true energy eigenstates $|n\rangle$, which gives $\langle \hat{H} \rangle = \sum_n P_n E_n$. Since all the probabilities P_n are non-negative, and all the E_n greater than or equal to E_0 , this is obviously not less than E_0 .

It is also clear that the better the guess (in the sense of maximising the overlap with the true ground state) the lower the energy bound, till successively better guesses converge on the true result.

As a very simple example, consider the infinite square well with $V = 0$, $0 < x < a$ and $V = \infty$ otherwise. As a trial function, we use $\Psi(x) = x(a - x)$, $0 < x < a$ and $\Psi(x) = 0$ otherwise. Then

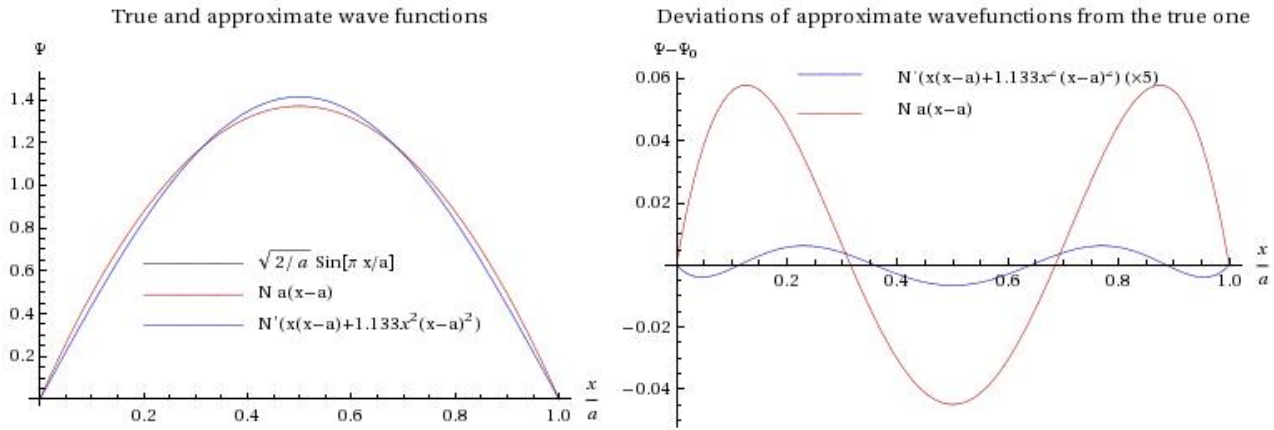
$$\frac{\langle \Psi | \hat{H} | \Psi \rangle}{\langle \Psi | \Psi \rangle} = \frac{10 \hbar^2}{2ma^2} = 1.013 \frac{\pi^2 \hbar^2}{2ma^2}$$

This is spectacularly good! Obviously it helped that our trial wave function looked a lot like what we'd expect of the true solution - symmetric about the midpoint, obeying the boundary conditions, no nodes...

In general, we will do better if we have an adjustable parameter, because then we can find the value which minimises our upper bound. So we could try $\Psi(x) = x(a - x) + bx^2(a - x)^2$ (with our previous guess corresponding to $b = 0$). Letting Mathematica do the dirty work, we get an energy bound which is a function of b , which takes its minimum value of $1.00001E_0$ at $b = 1.133/a^2$. Not much room for further improvement here!

Above we have plotted, on the left, the true and approximate wave functions (except that the true one is hidden under the second approximation, given in blue) and on the right, the deviations of the approximate wave functions from the true one (except that for the second approximation the deviation has been multiplied by 5 to render it visible!) This illustrates a general principle though: the wave function does have deviations from the true one on the part-per-mil scale, while the energy is good to 1 part in 10^5 . This is because the error in the energy is proportional to the coefficients squared of the admixture of "wrong" states, whereas the error in the wave function is linear in them.

Another example, for which the analytic solution is not known, is the quartic potential, $V(x) = \beta x^4$. Here a Gaussian trial wave function $\Psi(x) = e^{-ax^2/2}$ gives an upper bound for the ground energy state of $\frac{3}{8}6^{1/3} = 0.68$ in units of $(\hbar^4\beta/m^2)^{1/3}$. (The value obtained from numerical solution of the equation is 0.668).



3.2.1 Variational methods: excited states

Shankar 16.1; Mandl 8.3

Summary: Symmetry considerations may allow us to extend the variational method to certain excited states.

Looking again at the expression $\langle \hat{H} \rangle = \sum_n P_n E_n$, and recalling that the P_n are the squares of the overlap between the trial function and the actual eigenstates of the system, we see that we can only find bounds on excited states if we can arrange for the overlap of the trial wave function with all lower states to be zero. Usually this is not possible.

However an exception occurs where the states of the system can be separated into sets with different symmetry properties or other quantum numbers. Examples include parity and (in 3 dimensions) angular momentum. For example the lowest state with odd parity will automatically have zero overlap with the (even-parity) ground state, and so an upper bound can be found for it as well.

For the square well, the relevant symmetry is reflection about the midpoint of the well. If we choose a trial function which is antisymmetric about the midpoint, it must have zero overlap with the true ground state. So we can get a good bound on the first excited state, since $\langle \hat{H} \rangle = \sum_{n>0} P_n E_n > E_1$. Using $\Psi_1(x) = x(a-x)(2x-a)$, $0 < x < a$ we get $E_1 \leq 42 \hbar^2 / 2ma^2 = 1.064 E_1$.

If we wanted a bound on E_2 , we'd need a wave function which was orthogonal to both the ground state and the first excited state. The latter is easy by symmetry, but as we don't know the exact ground state (or so we are pretending!) we can't ensure the first. We can instead form a trial wave function which is orthogonal to the best *trial* ground state, but we will no longer have a strict upper bound on the energy E_2 , just a guess as to its value.

In this case we can choose $\Psi(x) = x(a-x) + bx^2(a-x)^2$ with a new value of b which gives orthogonality to the previous state, and then we get $E_2 \sim 10.3 E_0$ (as opposed to 9 for the actual value).

3.2.2 Variational methods: the helium atom

Griffiths 7.2; Gasiorowicz 14.2,4; Mandl 7.2, 8.8.2; Shankar 16.1

Summary: The most famous example of the variational principle is the ground state of the two-electron helium atom.

If we could switch off the interactions between the electrons, we would know what the ground state of the helium atom would be: $\Psi(\mathbf{r}_1, \mathbf{r}_2) = \phi_{100}^{Z=2}(\mathbf{r}_1)\phi_{100}^{Z=2}(\mathbf{r}_2)$, where ϕ_{nlm}^Z is a single-particle wave function of the hydrogenic atom with nuclear charge Z . For the ground state $n = 1$ and $l = m = 0$ (spherical symmetry). The energy of the two electrons would be $2Z^2 E_{\text{Ry}} = -108.8$ eV. But the experimental energy is only -78.6 eV (ie it takes 78.6 eV to fully ionise neutral helium). The difference is obviously due to the fact that the electrons repel one another.

The full Hamiltonian (ignoring the motion of the proton - a good approximation for the accuracy to which we will be working) is

$$-\frac{\hbar^2}{2m}(\nabla_1^2 + \nabla_2^2) - 2\hbar c\alpha \left(\frac{1}{|\mathbf{r}_1|} + \frac{1}{|\mathbf{r}_2|} \right) + \hbar c\alpha \frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|}$$

where ∇_1^2 involves differentiation with respect to the components of \mathbf{r}_1 , and $\alpha = e^2/(4\pi\epsilon_0 \hbar c) \approx 1/137$. (See [here](#) for a note on units in EM.)

A really simple guess at a trial wave function for this problem would just be $\Psi(\mathbf{r}_1, \mathbf{r}_2)$ as written above. The expectation value of the repulsive interaction term is $(5Z/4)E_{\text{Ry}}$ giving a total energy of -74.8 eV. (Gasiorowicz demonstrates the integral, as do [Fitzpatrick](#) and [Branson](#).)

It turns out we can do even better if we use the atomic number Z in the wave function Ψ as a variational parameter (that in the Hamiltonian, of course, must be left at 2). The best value turns out to be $Z = 27/16$ and that gives a better upper bound of -77.5 eV – just slightly higher than the experimental value. (Watch the sign – we get an lower bound for the ionization energy.) This effective nuclear charge of less than 2 presumably reflects the fact that to some extent each electron screens the nuclear charge from the other.

3.3 WKB approximation

Shankar 16.2; Gasiorowicz [Supplement 4A](#), Griffiths 8.1

Summary: The WKB approximation works for potentials which are slowly-varying on the scale of the wavelength of the particle and is particularly useful for describing tunnelling.

The WKB approximation is named for G. Wentzel, H.A. Kramers, and L. Brillouin, who independently developed the method in 1926. There are pre-quantum antecedents due to Jeffreys and Raleigh, though.

We can always write the one-dimensional time-independent Schrödinger equation as

$$\frac{d^2\psi}{dx^2} = -k(x)^2\psi(x)$$

where $k(x) \equiv \sqrt{2m(E - V(x))}/\hbar$. We could think of the quantity $k(x)$ as a spatially-varying wavenumber ($k = 2\pi/\lambda$), though we anticipate that this can only make sense if it doesn't change too quickly with position - else we can't identify a wavelength at all.

If the potential, and hence k , were constant, the plain waves $e^{\pm ikx}$ would be solutions. Let's see under what conditions a solution of the form

$$\psi(x) = A \exp\left(\pm i \int^x k(x') dx'\right)$$

might be a good approximate solution when the potential varies with position. Plugging this into the SE above, the LHS reads $-(k^2 \mp ik')\psi$. (Here and hereafter, primes denote differentiation wrt x — except when they indicate an integration variable.) So provided $|k'/k^2| \ll 1$, or equivalently $|\lambda'| \ll 1$, this is indeed a good solution as the second term can be ignored. And $|\lambda'| \ll 1$ does indeed mean that the wavelength is slowly varying. (One sometimes reads that what is needed is that the potential is slowly varying. But that is not a well defined statement, because dV/dx is not dimensionless. For any smooth potential, at high-enough energy we will have $|\lambda'| \ll 1$. What is required is that the *length-scale* of variation of λ , or k , or V (the scales are all approximately equal) is large compared with the de Broglie wavelength of the particle.

An obvious problem with this form is that the **probability current** isn't constant: if we calculate it we get $|A|^2 \hbar k(x)/m$. A better approximation is

$$\psi(x) = \frac{A}{\sqrt{k(x)}} \exp\left(\pm i \int^x k(x') dx'\right)$$

which gives a constant flux. (Classically, the probability of finding a particle in a small region is inversely proportional to the speed with which it passes through that region.) Furthermore one can show that if the error in the first approximation is $\mathcal{O}|\lambda'|$, the residual error with the second approximation is $\mathcal{O}|\lambda'|^2$. At first glance there is a problem with the second form when $k(x) = 0$, ie when $E = V(x)$. But near these points—the classical turning points—the whole approximation scheme is invalid anyway, because $\lambda \rightarrow \infty$ and so the potential cannot be “slowly varying” on the scale of λ .

For a region of constant potential, of course, there is no difference between the two approximations and both reduce to a plain wave, since $\int^x k(x') dx' = kx$.

For regions where $E < V(x)$, $k(x)$ will be imaginary and there is no wavelength as such. Instead we get real exponentials, which describe tunnelling. But defining $\lambda = 2\pi/|k|$ still, the WKB approximation will continue to be valid if $|\lambda'| \ll 1$ (where λ should now be interpreted as a decay length).

Tunnelling and bound-state problems inevitably include regions where $E \approx V(x)$ and the WKB approximation isn't valid. This would seem to be a major problem. However if such regions are short the requirement that the wave function and its derivative be continuous can help us to “bridge the gap”.

3.3.1 WKB approximation for bound states

Shankar 16.2; Gasiorowicz Supplement 4A; Griffiths 8.3

In a bound state problem with potential $V(x)$, for a given energy E , we can divide space into classically allowed regions, for which $E > V(x)$, and classically forbidden regions for which

$E < V(x)$. For simplicity we will assume that there are only three regions in total, classically forbidden for $x < a$ and $x > b$, and classically allowed for $a < x < b$.

In the classically allowed region $a < x < b$ the wave function will be oscillatory and we can write it either as an equal superposition of right- and left-moving complex exponentials or as

$$\psi(x) = \frac{A}{\sqrt{k(x)}} \cos \left(\int_{x_0}^x k(x') dx' + \phi \right)$$

where x_0 is some reference point; if it is changed, ϕ will also change.

For the particular case of a well with infinite sides the solution must vanish at the boundaries. Hence, taking the $x_0 = a$, we can choose $\phi = -\frac{1}{2}\pi$. We then require $\int_a^b k(x') dx' + \phi = (n + \frac{1}{2})\pi$ so that $\psi(b)$ vanishes; in other words

$$\int_a^b k(x') dx' = (n + 1)\pi, \quad \text{for } n = 0, 1, 2, \dots$$

(The integral is positive-definite so n cannot be negative.) Now $k(x)$ depends on E , and this condition will not hold for a general E , any more than the boundary conditions can be satisfied for an arbitrary E in the infinite square well. Instead we obtain an expression for E in terms of n , giving a spectrum of levels E_0, E_1, \dots

Of course for the infinite square well $k = \sqrt{2mE}/\hbar$ is constant and the condition gives $k = (n + 1)\pi/(b - a)$, which is exact. (Using $n + 1$ rather than n allows us to start with $n = 0$; starting at $n = 1$ is more usual for an infinite well.)

For a more general potential, outside the classically allowed region we will have decaying exponentials. In the vicinity of the turning points these solutions will not be valid, but if we approximate the potential as linear we can solve the Schrödinger equation exactly (in terms of Airy functions). Matching these to our WKB solutions in the vicinity of $x = a$ and $x = b$ gives the surprisingly simple result

$$\int_a^b k(x') dx' = (n + \frac{1}{2})\pi.$$

This is the quantisation condition for a finite well; it is different from the infinite well because the solution can leak into the forbidden region. (For a semi-infinite well, the condition is that the integral equal $(n + \frac{3}{4})\pi$. This is the appropriate form for the $l = 0$ solutions of a spherically symmetric well.) Unfortunately we can't check this against the finite *square* well though, because there the potential is definitely not slowly varying at the edges, nor can it be approximated as linear. But we can try the harmonic oscillator, for which the integral gives $E\pi/\hbar\omega$ and hence the quantisation condition gives $E = (n + \frac{1}{2})\hbar\omega$! The approximation was only valid for large n (small wavelength) but in fact we've obtained the exact answer for all levels.

Matching with Airy Functions

Griffiths 8.3

This subsection gives details of the matching process. The material in it is not examinable. More about Airy functions can be found in section [A.9](#).

If we can treat the potential as linear over a wide-enough region around the turning points that, at the edges of the region, the WKB approximation is valid, then we can match the WKB and exact solutions.

Consider a linear potential $V = \beta x$ as an approximation to the potential near the right-hand turning point b . We will scale $x = (\hbar^2/(2m\beta))^{1/3}z$ and $E = (\hbar^2\beta^2/2m)^{1/3}\mu$, so the turning point is at $z = \mu$. Then the differential equation is $y''(z) - zy(z) + \mu y(z) = 0$ and the solution which decays as $z \rightarrow \infty$ is $y(z) = AAi(z - \mu)$. This has to be matched, for z not too close to μ , to the WKB solution. In these units, $k(x) = (\mu - z)^{1/2}$ and

$$\int_b^x k(x')dx' = \int_\mu^z (\mu - z')^{1/2}dz' = -\frac{2}{3}(\mu - z)^{3/2},$$

so

$$\psi_{x < b}^{\text{WKB}}(z) = \frac{B}{(\mu - z)^{1/4}} \cos\left(-\frac{2}{3}(\mu - z)^{3/2} + \phi\right) \quad \text{and} \quad \psi_{x > b}^{\text{WKB}}(z) = \frac{C}{(z - \mu)^{1/4}} \exp\left(-\frac{2}{3}(z - \mu)^{3/2}\right).$$

(We chose the lower limit of integration to be b in order that the constant of integration vanished; any other choice would just shift ϕ .) Now the asymptotic forms of the Airy function are known:

$$Ai(z) \xrightarrow{z \rightarrow -\infty} \frac{\cos\left(\frac{2}{3}|z|^{3/2} - \frac{\pi}{4}\right)}{\sqrt{\pi}|z|^{1/4}} \quad \text{and} \quad Ai(z) \xrightarrow{z \rightarrow \infty} \frac{e^{-\frac{2}{3}z^{3/2}}}{2\sqrt{\pi}z^{1/4}}$$

so

$$Ai(z - \mu) \xrightarrow{z \rightarrow -\infty} \frac{\cos\left(\frac{2}{3}(\mu - z)^{3/2} - \frac{\pi}{4}\right)}{\sqrt{\pi}(\mu - z)^{1/4}} \quad \text{and} \quad Ai(z - \mu) \xrightarrow{z \rightarrow \infty} \frac{e^{-\frac{2}{3}(z - \mu)^{3/2}}}{2\sqrt{\pi}(z - \mu)^{1/4}}$$

and these will match the WKB expressions exactly provided $C = 2B$ and $\phi = \pi/4$.

At the left-hand turning point a , the potential is $V = -\beta'x$ (with $\beta \neq \beta'$ in general) and the solution $y(z) = AAi(-z + \nu)$. On the other hand the WKB integral is

$$\int_a^x k(x')dx' = \int_\nu^z (-\nu + z')dz' = \frac{2}{3}(-\nu + z)^{3/2}.$$

So in the classically allowed region we are comparing

$$Ai(-z + \nu) \xrightarrow{z \rightarrow \infty} \frac{\cos\left(\frac{2}{3}(z - \nu)^{3/2} - \frac{\pi}{4}\right)}{\sqrt{\pi}(z - \nu)^{1/4}} \quad \text{with} \quad \psi_{x > a}^{\text{WKB}}(z) = \frac{D}{(z - \nu)^{1/4}} \cos\left(\frac{2}{3}(z - \nu)^{3/2} + \phi'\right)$$

which requires $\phi' = -\pi/4$. (Note that $\phi' \neq \phi$ because we have taken a different lower limit of the integral.)

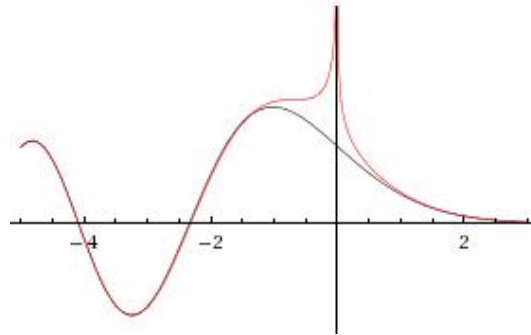
So now we have two expressions for the solution, valid everywhere except in the vicinity of the boundaries,

$$\psi(x) = \frac{D}{\sqrt{k(x)}} \cos\left(\int_a^x k(x')dx' - \pi/4\right) \quad \text{and} \quad \psi(x) = \frac{B}{\sqrt{k(x)}} \cos\left(\int_b^x k(x')dx' + \pi/4\right)$$

which can be satisfied only if $D = \pm B$ and $\int_a^b k(x')dx' = (n + \frac{1}{2})\pi$, as required.

It is worth stressing that although—for a linear potential—the exact (Airy function) and WKB solutions match “far away” from the turning point, they do not do so close in. The $(z - \mu)^{-1/4}$

Exact vs WKB solutions for a linear potential



terms in the latter mean that they blow up, but the former are perfectly smooth. They are shown (for $\mu = 0$) in the figure, in red for the WKB and black for the exact functions. We can see they match very well so long as $|z - \mu| > 1$; in fact $z \rightarrow \infty$ is overkill!

So now we can be more precise about the conditions under which the matching is possible: we need the potential to be linear over the region $\Delta x \sim (\hbar^2/(2m\beta))^{1/3}$ where $\beta = dV/dx$. Linearity means that $\Delta V/\Delta x \approx dV/dx$ at the turning point, or $|d^2V/dx^2| \Delta x \ll dV/dx$ (assuming the curvature is the dominant non-linearity, as is likely if V is smooth). For the harmonic oscillator, $|d^2V/dx^2| \Delta x/(dV/dx) = 2(\hbar\omega/E)^{2/3}$ which is only much less than 1 for very large n , making the exact result even more surprising!

3.3.2 WKB approximation for tunnelling

Shankar 16.2; Gasiorowicz Supplement 4B; Griffiths 8.2

For the WKB approximation to be applicable to tunnelling through a barrier, we need as always $|\lambda| \ll 1$. In practice that means that the barrier function is reasonably smooth and that $E \ll V(x)$. Now it would of course be possible to do a careful calculation, writing down the WKB wave function in the three regions (left of the barrier, under the barrier and right of the barrier), linearising in the vicinity of the turning points in order to match the wave function and its derivatives at both sides. This however is a tiresomely lengthy task, and we will not attempt it. Instead, recall the result for a high, wide square barrier; the transmission coefficient in the limit $e^{-2\kappa L} \ll 1$ is given by

$$T = \frac{16k_1 k_2 \kappa^2}{(\kappa^2 + k_1^2)(\kappa^2 + k_2^2)} e^{-2\kappa L},$$

where k_1 and k_2 are the wavenumbers on either side of the barrier (width L , height V) and $\kappa = \sqrt{2m(V - E)}$. (See the **revision notes**.) All the prefactors are not negligible, but they are weakly energy-dependent, whereas the $e^{-2\kappa L}$ term is very strongly energy dependent. If we plot $\log T$ against energy, the form will be essentially $\text{const} - 2\kappa(E)L$, and so we can still make predictions without worrying about the constant.

For a barrier which is not constant, the WKB approximation will yield a similar expression for the tunnelling probability:

$$T = [\text{prefactor}] \times \exp\left(-2 \int_a^b \kappa(x') dx'\right),$$

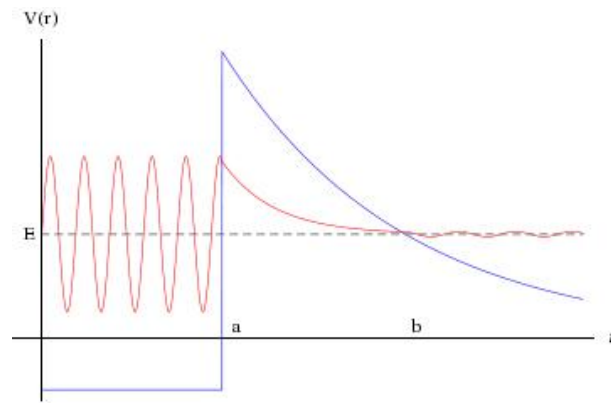
where $\kappa(x) \equiv \sqrt{2m(V(x) - E)}/\hbar$. The WKB approximation is like treating a non-square barrier like a succession of square barriers of different heights. The need for $V(x)$ to be slowly varying is then due to the fact that we are slicing the barrier sufficiently thickly that $e^{-2\kappa\Delta L} \ll 1$ for each slice.

The classic application of the WKB approach to tunnelling is alpha decay. The potential here is a combination of an attractive short-range nuclear force and the repulsive Coulomb interaction between the alpha particle and the daughter nucleus. Unstable states have energies greater than zero, but they are long-lived because they are classically confined by the barrier. (It takes some thought to see that a repulsive force can cause quasi-bound states to exist!) The semi-classical model is of a pre-formed alpha particle of energy E bouncing back and forth many times (f) per second, with a probability of escape each time given by the tunnelling probability, so the decay rate is given by $1/\tau = fT$. Since we can't calculate f with any reliability we would be silly to worry about the prefactor in T , but the primary dependence of the decay rate on the energy of the emitted particle will come from the easily-calculated exponential.

In the figure above the value of a is roughly the nuclear radius R , and b is given by $V_c(b) = E$, with the Coulomb potential $V_c(r) = zZ\hbar c\alpha/r$. (Z is the atomic number of the daughter nucleus and $z = 2$ that of the alpha.) The integral in the exponent can be done (see Gasiorowicz Supplement 4 B for details¹; the substitution $r = b \cos^2 \theta$ is used), giving in the limit $b \gg a$

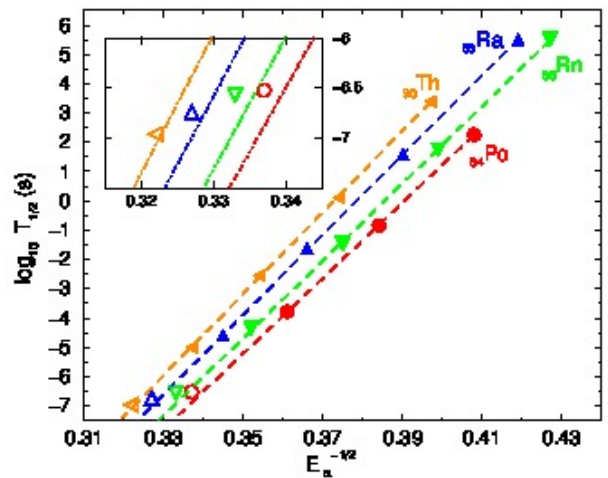
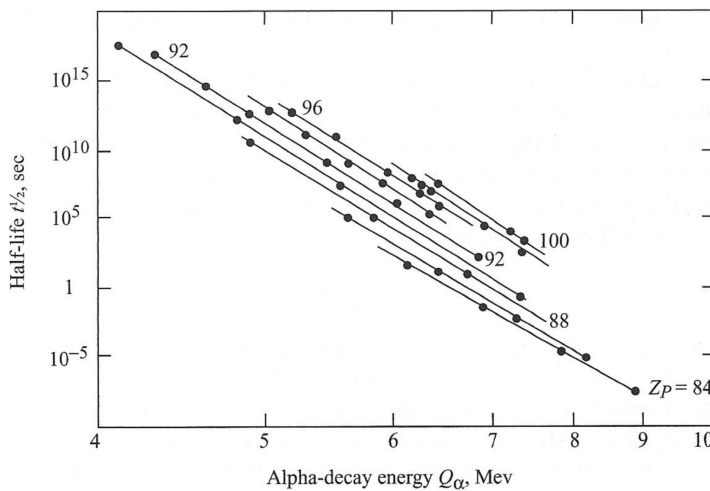
$$2 \int_R^b \kappa(r) dr = 2\pi z Z \alpha \sqrt{\frac{mc^2}{2E}} = 39 \frac{Z}{\sqrt{E(\text{MeV})}} \Rightarrow \log_{10} \tau = \text{const} + 1.72 \frac{Z}{\sqrt{E(\text{MeV})}}.$$

¹But note that there is a missing minus sign between the two terms in square brackets in eq. 4B-4.



This is the Geiger-Nuttall law. Data for the lifetimes of long-lived isotopes (those with low-energy alphas) fit such a functional form well, but with 1.61 rather than 1.72. In view of the fairly crude approximations made, this is a pretty good result. Note it is independent of the nuclear radius because we used $b \gg R$; we could have kept the first correction, proportional to $\sqrt{b/R}$, to improve the result. Indeed the first estimates of nuclear radii came from exactly such studies.

A version of the classic figure of the results is given below on the left.² The marked energy scale is non-linear; the linear variable on the x -axis is actually $-E^{-\frac{1}{2}}$, so that the slope is negative. Straight lines join isotopes of the same element. On the right is a more recent plot³ of $\log \tau$ against $E^{-\frac{1}{2}}$, showing deviations from linearity at the high-energy (left-hand) end.



²found at <https://web-docs.gsi.de/~wolle/TELEKOLLEG/KERN/LECTURE/Fraser/L14.pdf>, which may not be the original source.

³Astier *et al*, Eur. Phys. J. A46 (2010) 165

4. Approximate methods II: Time-independent perturbation theory

Perturbation theory is the most widely used approximate method. It requires we have a set of exact solutions to a Hamiltonian which is close to the realistic one.

4.1 Non-degenerate perturbation theory

Shankar 17.1; Mandl 7.1; Griffiths 6.1

Summary: First assume the exactly-solvable Hamiltonian has no degeneracies (repeated eigenvalues); then the rules are straightforward.

Perturbation theory is applicable when the Hamiltonian \hat{H} can be split into two parts, with the first part being exactly solvable and the second part being small in comparison. The first part is always written $\hat{H}^{(0)}$, and we will denote its eigenstates by $|n^{(0)}\rangle$ and energies by $E_n^{(0)}$ (with wave functions $\phi_n^{(0)}$). These we know, *and for now assume to be non-degenerate*. The eigenstates and energies of the full Hamiltonian are denoted $|n\rangle$ and E_n , and the aim is to find successively better approximations to these. The zeroth-order approximation is simply $|n\rangle = |n^{(0)}\rangle$ and $E_n = E_n^{(0)}$, which is just another way of saying that the perturbation is small and at a crude enough level of approximation we can ignore it entirely.

Nomenclature for the perturbing Hamiltonian $\hat{H} - \hat{H}^{(0)}$ varies. δV , $\hat{H}^{(1)}$ and $\lambda\hat{H}^{(1)}$ are all common. It usually is a perturbing potential but we won't assume so here, so we won't use the first. The second and third differ in that the third has explicitly identified a small, dimensionless parameter (eg α in EM), so that the residual $\hat{H}^{(1)}$ isn't itself small. With the last choice, our expressions for the eigenstates and energies of the full Hamiltonian will be explicitly power series in λ , so $E_n = E_n^{(0)} + \lambda E_n^{(1)} + \lambda^2 E_n^{(2)} + \dots$ etc. With the second choice the small factor is hidden in $\hat{H}^{(1)}$, and is implicit in the expansion which then reads $E_n = E_n^{(0)} + E_n^{(1)} + E_n^{(2)} + \dots$. In this case one has to remember that anything with a superscript (1) is first order in this implicit small factor, or more generally the superscript (m) denotes something which is m th order. For the derivation of the equations we will retain an explicit λ , but thereafter we will set it equal to one to revert to the other formulation. We will take λ to be real so that \hat{H}_1 is Hermitian.

We start with the master equation

$$(\hat{H}^{(0)} + \lambda\hat{H}^{(1)})|n\rangle = E_n|n\rangle.$$

Then we substitute in $E_n = E_n^{(0)} + \lambda E_n^{(1)} + \lambda^2 E_n^{(2)} + \dots$ and $|n\rangle = |n^{(0)}\rangle + \lambda|n^{(1)}\rangle + \lambda^2|n^{(2)}\rangle + \dots$ and expand. Then since λ is a free parameter, we have to match terms on each side with the

same powers of λ , to get

$$\begin{aligned}\hat{H}^{(0)}|n^{(0)}\rangle &= E_n^{(0)}|n^{(0)}\rangle \\ \hat{H}^{(0)}|n^{(1)}\rangle + \hat{H}^{(1)}|n^{(0)}\rangle &= E_n^{(0)}|n^{(1)}\rangle + E_n^{(1)}|n^{(0)}\rangle \\ \hat{H}^{(0)}|n^{(2)}\rangle + \hat{H}^{(1)}|n^{(1)}\rangle &= E_n^{(0)}|n^{(2)}\rangle + E_n^{(1)}|n^{(1)}\rangle + E_n^{(2)}|n^{(0)}\rangle\end{aligned}$$

We have to solve these sequentially. The first we assume we have already done. The second will yield $E_n^{(1)}$ and $|n^{(1)}\rangle$. Once we know these, we can use the third equation to yield $E_n^{(2)}$ and $|n^{(2)}\rangle$, and so on. The expressions for the changes in the states, $|n^{(1)}\rangle$ etc, will make use of the fact that the unperturbed states $\{|n^{(0)}\rangle\}$ form a basis, so we can write

$$|n^{(1)}\rangle = \sum_m c_m |m^{(0)}\rangle = \sum_m \langle m^{(0)}|n^{(1)}\rangle |m^{(0)}\rangle.$$

In each case, to solve for the energy we take the inner product with $\langle n^{(0)}|$ (i.e. the same state) whereas for the wave function, we use $\langle m^{(0)}|$ (another state). We use, of course, $\langle m^{(0)}|\hat{H}^{(0)} = E_m^{(0)}\langle m^{(0)}|$ and $\langle m^{(0)}|n^{(0)}\rangle = \delta_{mn}$.

At first order we get

$$E_n^{(1)} = \langle n^{(0)}|\hat{H}^{(1)}|n^{(0)}\rangle \quad (4.1)$$

$$\langle m^{(0)}|n^{(1)}\rangle = \frac{\langle m^{(0)}|\hat{H}^{(1)}|n^{(0)}\rangle}{E_n^{(0)} - E_m^{(0)}} \quad \forall m \neq n. \quad (4.2)$$

The second equation tells us the overlap of $|n^{(1)}\rangle$ with all the other $|m^{(0)}\rangle$, but not with $|n^{(0)}\rangle$. This is obviously not constrained by the eigenvalue equation, because we can add any amount of $|n^{(0)}\rangle$ and the equations will still be satisfied. However we need the state to continue to be normalised, and when we expand $\langle n|n\rangle = 1$ in powers of λ we find that $\langle n^{(0)}|n^{(1)}\rangle$ is required to be imaginary. This is just like a phase rotation of the original state and we can ignore it. (Recall that an infinitesimal change in a unit vector has to be at right angles to the original.) Hence

$$|n^{(1)}\rangle = \sum_{m \neq n} \frac{\langle m^{(0)}|\hat{H}^{(1)}|n^{(0)}\rangle}{E_n^{(0)} - E_m^{(0)}} |m^{(0)}\rangle. \quad (4.3)$$

If the spectrum of $\hat{H}^{(0)}$ is degenerate, there may be a problem with this expression because the denominator can vanish, making the corresponding term infinite. In fact nothing that we have done so far is directly valid in that case, and we have to use “degenerate perturbation theory” instead. For now we assume that for any two states $|m^{(0)}\rangle$ and $|n^{(0)}\rangle$, either $E_n^{(0)} - E_m^{(0)} \neq 0$ (non degenerate) or $\langle m^{(0)}|\hat{H}^{(1)}|n^{(0)}\rangle = 0$ (the states are not mixed by the perturbation.)

Then at second order

$$E_n^{(2)} = \langle n^{(0)}|\hat{H}^{(1)}|n^{(1)}\rangle = \sum_{m \neq n} \frac{|\langle m^{(0)}|\hat{H}^{(1)}|n^{(0)}\rangle|^2}{E_n^{(0)} - E_m^{(0)}}. \quad (4.4)$$

The expression for the second-order shift in the wave function $|n^{(2)}\rangle$ can also be found but it is tedious. The main reason we wanted $|n^{(1)}\rangle$ was to find $E_n^{(2)}$ anyway, and we’re not planning to find $E_n^{(3)}$! Note that though the expression for $E_n^{(1)}$ is generally applicable, those for $|n^{(1)}\rangle$ and $E_n^{(2)}$ would need some modification if the Hamiltonian had continuum eigenstates as well

as bound states (eg hydrogen atom). Provided the state $|n\rangle$ is bound, that is just a matter of integrating rather than summing. This restriction to bound states is why Mandl calls chapter 7 “bound-state perturbation theory”. The perturbation of continuum states (eg scattering states) is usually dealt with separately.

Note that the equations above hold whether we have identified an explicit small parameter λ or not. So from now on we will set λ to one, assume that $\hat{H}^{(1)}$ has an implicit small parameter within it, and write $E_n = E_n^{(0)} + E_n^{(1)} + E_n^{(2)} + \dots$; the expressions above for $E^{(1,2)}$ and $|n^{(1)}\rangle$ are still valid.

4.1.1 Connection to variational approach

We have shown that $\langle\psi|\hat{H}|\psi\rangle \geq E_0$ for all normalised states $|\psi\rangle$ (with equality implying $|\psi\rangle = |0\rangle$). Thus for the (non-degenerate) ground state, $E_0^{(0)} + E_0^{(1)}$ is an upper bound on the exact energy E_0 , since it is obtained by using the unperturbed ground state as a trial state for the full Hamiltonian. It follows that the sum of all higher corrections $E_0^{(2)} + \dots$ must be negative. We can see indeed that $E_0^{(2)}$ will always be negative, since for every term in the sum the numerator is positive and the denominator negative. (The fact that $\langle\psi|\hat{H}|\psi\rangle \geq E_0$ is the basis of the variational approach to finding the ground state energy, where we vary a trial state $|\psi\rangle$ to optimise an upper bound on E_0 .)

4.1.2 Perturbed infinite square well

Probably the simplest example we can think of is an infinite square well with a low step half way across, so that

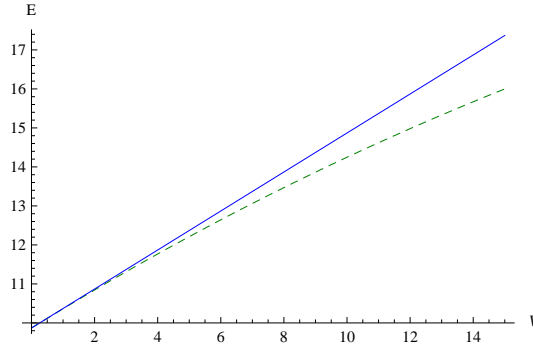
$$V(x) = \begin{cases} 0 & \text{for } 0 < x < a/2, \\ V_0 & \text{for } a/2 < x < a \\ \infty & \text{elsewhere} \end{cases}$$

We treat this as a perturbation on the flat-bottomed well, so $\hat{H}^{(1)} \xrightarrow{x} V_0$ for $a/2 < x < a$ and zero elsewhere.

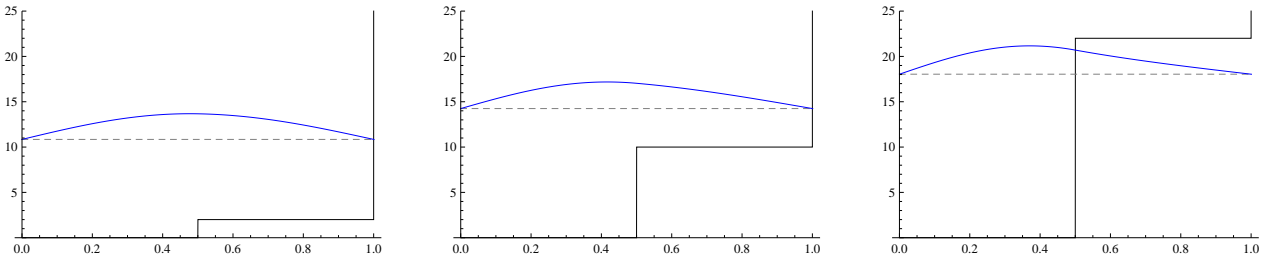
The ground-state unperturbed wave function is $\psi_0^{(0)} = \sqrt{\frac{2}{a}} \sin \frac{\pi x}{a}$, with unperturbed energy $E_0^{(0)} = \pi^2 \hbar^2 / (2ma^2)$. A “low” step will mean $V_0 \ll E_0^{(0)}$. Then we have

$$E_0^{(1)} = \langle\psi_0^{(0)}|\hat{H}^{(1)}|\psi_0^{(0)}\rangle = \frac{2}{a} \int_{a/2}^a V_0 \sin^2 \frac{\pi x}{a} dx = \frac{V_0}{2}$$

This problem can be solved semi-analytically; in both regions the solutions are sinusoids, but with wavenumbers $k = \sqrt{2mE}/\hbar$ and $k' = \sqrt{2m(E - V_0)}/\hbar$ respectively; satisfying the boundary conditions and matching the wave functions and derivatives at $x = a/2$ gives the condition $k \cot(ka/2) = k' \cot(k'a/2)$ which can be solved numerically for E . Below the exact solution (green, dotted) and $E_0^{(0)} + E_0^{(1)}$ (blue) are plotted; we can see that they start to diverge when V_0 is about 5, which is higher than we might have expected (everything is in units of $\hbar^2/(2ma^2) \approx 0.1E_0$).



We can also plot the exact wave functions for different step size, and see that for $V_0 = 10$ (the middle picture, well beyond the validity of first-order perturbation theory) it is significantly different from a simple sinusoid.



4.1.3 Perturbed Harmonic oscillator

Another example is the harmonic oscillator, $\hat{H}^{(0)} = \frac{\hat{p}^2}{2m} + \frac{1}{2}m\omega^2\hat{x}^2$, with a perturbing potential $H^{(1)} = \lambda\hat{x}^2$. The states of the unperturbed oscillator are denoted $|n^{(0)}\rangle$ with energies $E_n^{(0)} = (n + \frac{1}{2})\hbar\omega$.

Recalling that in terms of creation and annihilation operators (see section 1.6), $\hat{x} = (x_0/\sqrt{2})(\hat{a} + \hat{a}^\dagger)$, with $[\hat{a}, \hat{a}^\dagger] = 1$, and $x_0 = \sqrt{\hbar/(m\omega)}$, and so

$$E_n^{(1)} = \langle n^{(0)} | \hat{H}^{(1)} | n^{(0)} \rangle = \frac{x_0^2 \lambda}{2} \langle n^{(0)} | (\hat{a}^\dagger)^2 + \hat{a}^2 + 2\hat{a}^\dagger \hat{a} + 1 | n^{(0)} \rangle = \frac{\lambda}{m\omega^2} \hbar\omega (n + \frac{1}{2}).$$

The first-order change in the wave function is also easy to compute, as $\langle m^{(0)} | \hat{H}^{(1)} | n^{(0)} \rangle = 0$ unless $m = n \pm 2$. Thus

$$\begin{aligned} |n^{(1)}\rangle &= \sum_{m \neq n} \frac{\langle m^{(0)} | \hat{H}^{(1)} | n^{(0)} \rangle}{E_n^{(0)} - E_m^{(0)}} |m^{(0)}\rangle \\ &= \frac{\hbar\lambda}{2m\omega} \left(\frac{\sqrt{(n+1)(n+2)}}{-2\hbar\omega} |(n+2)^{(0)}\rangle + \frac{\sqrt{n(n-1)}}{2\hbar\omega} |(n-2)^{(0)}\rangle \right). \end{aligned}$$

We can now also calculate the second order shift in the energy:

$$\begin{aligned} E_n^{(2)} &= \langle n^{(0)} | \hat{H}^{(1)} | n^{(1)} \rangle = \sum_{m \neq n} \frac{|\langle m^{(0)} | \hat{H}^{(1)} | n^{(0)} \rangle|^2}{E_n^{(0)} - E_m^{(0)}} \\ &= \left(\frac{\hbar\lambda}{2m\omega} \right)^2 \left(\frac{(n+1)(n+2)}{-2\hbar\omega} + \frac{n(n-1)}{2\hbar\omega} \right) = -\frac{1}{2} \left(\frac{\lambda}{m\omega^2} \right)^2 \hbar\omega (n + \frac{1}{2}) \end{aligned}$$

We can see a pattern emerging, and of course this is actually a soluble problem, as all that the perturbation has done is change the frequency. Defining $\omega' = \omega\sqrt{1 + 2\lambda/(m\omega^2)}$, we see that the exact solution is

$$E_n = (n + \frac{1}{2})\hbar\omega' = (n + \frac{1}{2})\hbar\omega \left(1 + \frac{\lambda}{m\omega^2} - \frac{1}{2} \left(\frac{\lambda}{m\omega^2}\right)^2 + \dots\right)$$

in agreement with the perturbative calculation.

4.2 Degenerate perturbation theory

Shankar 17.3; Mandl 7.3; Griffiths 6.6

Summary: If there are degeneracies, we need to first find a set of eigenstates of the unperturbed Hamiltonian which are not mixed by the perturbation. Then we can proceed as above.

None of the formalism that we have developed so far works “out of the box” if $\hat{H}^{(0)}$ has degenerate eigenstates. To be precise, it is still fine for the non-degenerate states, but it fails to work in a subspace of degenerate states if $\hat{H}^{(1)}$ is not also diagonal in this subspace. We can see that in Eq. (4.2), where the vanishing energy denominator clearly signals a problem, but even Eq. (4.1) is wrong. The reason is simple: we assumed from the start that the shifts in the states due to the perturbation would be small. But suppose $|1^{(0)}\rangle$ and $|2^{(0)}\rangle$ are degenerate eigenstates of $\hat{H}^{(0)}$; then so are $\sqrt{\frac{1}{2}}(|1^{(0)}\rangle \pm |2^{(0)}\rangle)$. Now the eigenstates of the full Hamiltonian $|1\rangle$ and $|2\rangle$ are not degenerate—but which of the possible choices for the eigenstate of $\hat{H}^{(0)}$ are they close to? If for example it is the latter (as is often the case in simple examples) then even a tiny perturbation $\hat{H}^{(1)}$ will induce a big change in the eigenstates.

Consider a potential in two dimensions which is symmetric: $V(x, y) = V_0(r)$. The ground state will be non-degenerate, but all higher states, of energy E_n , will be doubly degenerate; a possible but not unique choice has wave functions $\psi_n(r) \sin(n\phi)$ and $\psi_n(r) \cos(n\phi)$ (where ϕ is the angle in the plane from the x -axis); the form of $\psi_n(r)$ will of course depend on $V_0(r)$. For $n = 1$, the probability-density maps of these two states looks like the last two figures of the top line of figure 4.1. Now imagine a perturbation which is not circularly symmetric, say λx^2 . The potential now rises more steeply along the x -axis than along the y -axis (the top line of figure 4.1 shows a contour map of the—somewhat exaggeratedly—deformed potential). This will clearly lift the degeneracy, because the first state, which vanishes along the x -axis, will “feel” the perturbation less and have a lower energy than the second. The new energy eigenstates will be similar to the original ones, and the first order energy shifts will indeed be given by the usual formula: a naïve application of perturbation theory is fine.

But what if the perturbation is $\frac{1}{2}\lambda(x + y)^2$? (See the bottom line of figure 4.1 for the new contour map.) This is symmetric between x and y , $\cos\phi$ and $\sin\phi$: do we conclude that the states remain degenerate? No: this is really the same problem as before with a more and a less steep direction, the only difference being that the orientation is rotated by 45° . And so there will again be a pair of solutions split in energy according to whether the probability of being found along the 45° line is larger or smaller. We expect them to look like those shown in the bottom line of figure 4.1. But these are not close to our original pairs of solutions; we can’t get from one to the other by perturbation theory. Does that mean that perturbation theory is

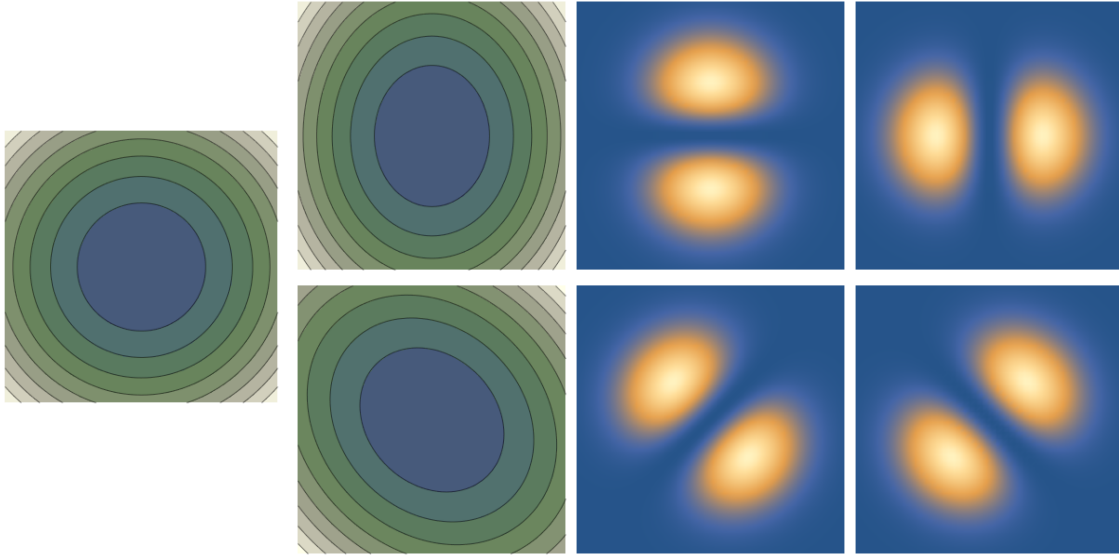


Figure 4.1: Far left: contour plots of unperturbed potential. Left: contour plots of perturbed potentials. Right and far-right: density plots of $|\psi|^2$ for lower- and higher-energy eigenstates with $n = 1$, for each perturbed potential. See text for explanation.

useless in this case? Surely not: effectively all we have done between the two cases is rotate the axes by 45° ! The resolution comes from recognising that our initial choice of eigenstates of V_0 was not unique, and with hindsight was not the appropriate starting point for this perturbation.

In fact the circular symmetry of the original potential means that for any angle α , the alternative pair of wave functions $\psi_n(r) \cos(n(\phi - \alpha))$ and $\psi_n(r) \sin(n(\phi - \alpha))$ is an equally good choice of orthogonal unperturbed eigenstates. (We can rotate the x -axis at will.) For the perturbation $\frac{1}{2}\lambda(x + y)^2$, the choice of $\alpha = 45^\circ$ is the right one to ensure that the perturbed solutions are close to the unperturbed ones, and the usual formulae can be used if we chose that basis.

To sum up: the perturbations broke the symmetry of the original problem and hence took away the original freedom we had to choose our basis. If we chose inappropriately initially, we need to choose again once we know the form of the perturbation we are dealing with.

Sometimes the right choice of unperturbed eigenstates is immediately obvious, as in the case above, or can be deduced by considerations of symmetry. If not, then starting with our initial choice $\{|n^{(0)}\rangle\}$, we need to find a new set of states in the degenerate space, linear combinations of our initial choice, which we will call $\{|n'^{(0)}\rangle\}$, and which crucially are *not mixed* by the perturbation, in other words,

$$\langle m'^{(0)} | \hat{H}^{(1)} | n'^{(0)} \rangle = 0 \quad \text{if } E_m^{(0)} = E_n^{(0)} \text{ for } m \neq n.$$

In the new basis $\hat{H}^{(1)}$ is *diagonal* in the degenerate subspace. That will ensure that all state and energy shifts are small, as assumed in the set-up of perturbation theory.

Thus we write down the matrix which is the representation of $\hat{H}^{(1)}$ in the degenerate subspace of the originally-chosen basis, and find the eigenvectors. These, being linear combinations of the old basis states, are still eigenstates of $\hat{H}^{(0)}$ and so are equally good new basis states. But because they are eigenstates of $\hat{H}^{(1)}$ in this restricted space, they are not mixed by $\hat{H}^{(1)}$. In the process we will probably find the eigenvalues too, and these are the corresponding first-order energy shifts.

We then proceed almost exactly as in the non-degenerate case having replaced (say) $|1^{(0)}\rangle$ and $|2^{(0)}\rangle$ with the new linear combinations which we can call $|1'^{(0)}\rangle$ and $|2'^{(0)}\rangle$. The expressions for the energy and state shifts, using the new basis, are as before, Eqs. (4.1,4.2,4.3,4.4), except instead of summing over all states $m \neq n$, we sum over all states for which $E_m^{(0)} \neq E_n^{(0)}$. As already mentioned the first-order energy-shifts of the originally-degenerate states, $\langle n'^{(0)} | \hat{H}^{(1)} | n'^{(0)} \rangle$, are just the eigenvalues of the representation of $\hat{H}^{(1)}$ in the degenerate subspace.

For example suppose $\hat{H}^{(0)}$ has many eigenstates but two, $|1^{(0)}\rangle$ and $|2^{(0)}\rangle$, are degenerate, and that $\hat{H}^{(1)}|1^{(0)}\rangle = \beta|2^{(0)}\rangle + \dots$ and $\hat{H}^{(1)}|2^{(0)}\rangle = \beta|1^{(0)}\rangle + \dots$, with β real and \dots referring to $|3^{(0)}\rangle$ and higher. Then *in the degenerate subspace*

$$\hat{H}^{(1)} \longrightarrow \begin{pmatrix} \langle 1^{(0)} | \hat{H}^{(1)} | 1^{(0)} \rangle & \langle 1^{(0)} | \hat{H}^{(1)} | 2^{(0)} \rangle \\ \langle 2^{(0)} | \hat{H}^{(1)} | 1^{(0)} \rangle & \langle 2^{(0)} | \hat{H}^{(1)} | 2^{(0)} \rangle \end{pmatrix} = \beta \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

whose eigenvectors are $\sqrt{\frac{1}{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}$ and $\sqrt{\frac{1}{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}$, with eigenvalues $\mp\beta$. With these, we define new eigenstates:

$$|1'^{(0)}\rangle = \sqrt{\frac{1}{2}}(|1^{(0)}\rangle - |2^{(0)}\rangle), \quad |2'^{(0)}\rangle = \sqrt{\frac{1}{2}}(|1^{(0)}\rangle + |2^{(0)}\rangle).$$

Now that we have the appropriate basis, we can apply Eq. (4.1) to obtain

$$E_{1'}^{(1)} = \langle 1'^{(0)} | \hat{H}^{(1)} | 1'^{(0)} \rangle = -\beta, \quad E_{2'}^{(1)} = \langle 2'^{(0)} | \hat{H}^{(1)} | 2'^{(0)} \rangle = \beta.$$

The expressions for the first order state changes $|n'^{(1)}\rangle$ and second-order energy changes $E_{n'}^{(2)}$ are just given by Eq. (4.3,4.4) but with primed states where appropriate; in these the higher states will of course enter. However since $\langle 2'^{(0)} | \hat{H}^{(1)} | 1'^{(0)} \rangle = 0$ by construction, $|2'^{(0)}\rangle$ does not appear in the sum over states and there is no problem with vanishing denominators.

We should note that, unless the perturbation does not mix the degenerate states and the higher-order states, we have not solved the problem exactly. At this stage we have just found the correct approximate eigenstates and the first-order energy shifts. Of course that is often all we want.

4.2.1 Example of degenerate perturbation theory

If we have a two-state problem, then the work done to diagonalise $\hat{H}^{(1)}$ is equivalent to finding the exact solution of the problem. So the simplest non-trivial example requires three states of which two are degenerate. (If all three are degenerate and mixed by the perturbation, we end up solving the problem exactly.)

Suppose we have a three-state basis and an $\hat{H}^{(0)}$ whose eigenstates, $|1^{(0)}\rangle$, $|2^{(0)}\rangle$ and $|3^{(0)}\rangle$, have energies $E_1^{(0)}$, $E_2^{(0)}$ and $E_3^{(0)}$ (all initially assumed to be different). A representation of this system is

$$|1^{(0)}\rangle \longrightarrow \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad |2^{(0)}\rangle \longrightarrow \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad |3^{(0)}\rangle \longrightarrow \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \quad \hat{H}^{(0)} \longrightarrow \begin{pmatrix} E_1^{(0)} & 0 & 0 \\ 0 & E_2^{(0)} & 0 \\ 0 & 0 & E_3^{(0)} \end{pmatrix}.$$

To this, we add the perturbation

$$\hat{H}^{(1)} \longrightarrow a \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}.$$

First, let's consider a non-degenerate case, with $E_1^{(0)} = E_0$, $E_2^{(0)} = 2E_0$ and $E_3^{(0)} = 3E_0$. Then we can show that, to first order in a

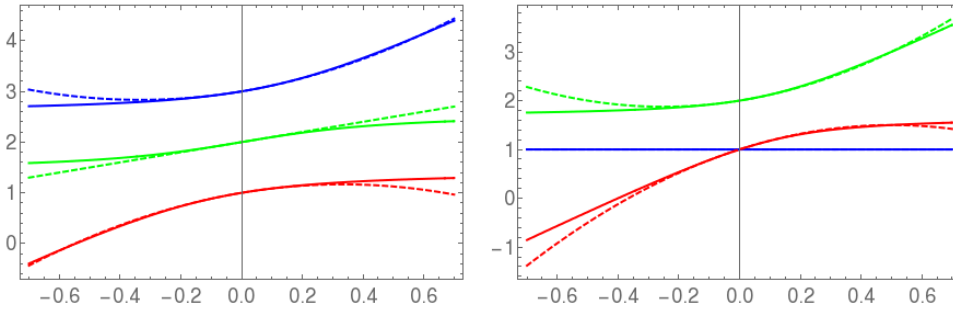
$$E_1^{(1)} = E_2^{(1)} = E_3^{(1)} = a,$$

$$|1^{(1)}\rangle = -\frac{a}{E_0}|2^{(0)}\rangle - \frac{a}{2E_0}|3^{(0)}\rangle \longrightarrow \frac{a}{2E_0} \begin{pmatrix} 0 \\ -2 \\ -1 \end{pmatrix}, \quad |2^{(1)}\rangle = \frac{a}{E_0}|1^{(0)}\rangle - \frac{a}{E_0}|3^{(0)}\rangle \longrightarrow \frac{a}{E_0} \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix},$$

$$|3^{(1)}\rangle = \frac{a}{2E_0}|1^{(0)}\rangle + \frac{a}{E_0}|2^{(0)}\rangle \longrightarrow \frac{a}{2E_0} \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}, \quad E_1^{(2)} = -\frac{3a^2}{2E_0}, \quad E_2^{(2)} = 0, \quad E_3^{(2)} = \frac{3a^2}{2E_0}.$$

Note that all of these terms are just the *changes* in the energies and states, which have to be added to the zeroth-order ones to get expressions which are complete to the given order.

In this case the exact eigenvalues of $\hat{H}^{(0)} + \hat{H}^{(1)}$ can only be found numerically. The left-hand plot below shows the energies as a function of a , both in units of E_0 , with the dashed lines being the expansion to second order:



The right-hand plot above shows the partially degenerate case discussed below.

Now we consider a case with two degenerate states, with $E_1^{(0)} = E_2^{(0)} = E_0$, and $E_3^{(0)} = 2E_0$. We note that $|1^{(0)}\rangle$ and $|2^{(0)}\rangle$ are just two of an infinite set of eigenstates with the same energy $E_1^{(0)}$, since any linear combination of them is another eigenstate. We have to make the choice which diagonalises $\hat{H}^{(1)}$ in this subspace: in this subspace

$$\hat{H}^{(1)} \longrightarrow a \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$$

whose eigenstates are $\frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}$ and $\frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}$, with eigenvalues 0 and $2a$. So

$$|1'^{(0)}\rangle = \frac{1}{\sqrt{2}} \left(|1^{(0)}\rangle - |2^{(0)}\rangle \right) \quad \text{and} \quad |2'^{(0)}\rangle = \frac{1}{\sqrt{2}} \left(|1^{(0)}\rangle + |2^{(0)}\rangle \right).$$

These new states don't diagonalise the full $\hat{H}^{(1)}$, of course. To go further we need the matrix elements $\langle 3^{(0)} | \hat{H}^{(1)} | 1'^{(0)} \rangle = 0$ and $\langle 3^{(0)} | \hat{H}^{(1)} | 2'^{(0)} \rangle = \sqrt{2}a$. Then

$$E_{1'}^{(1)} = 0 \quad E_{2'}^{(1)} = 2a, \quad E_3^{(1)} = a,$$

$$|1'^{(1)}\rangle = 0, \quad |2'^{(1)}\rangle = -\frac{a\sqrt{2}}{E_0}|3^{(0)}\rangle \longrightarrow -\frac{\sqrt{2}a}{E_0} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \quad |3^{(1)}\rangle = \frac{\sqrt{2}a}{E_0}|2'^{(0)}\rangle \longrightarrow \frac{a}{E_0} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$$

$$E_{1'}^{(2)} = 0, \quad E_{2'}^{(2)} = -\frac{2a^2}{E_0}, \quad E_3^{(2)} = \frac{2a^2}{E_0}$$

In this particular case it is easy to show that $|1^{(0)}\rangle$ is actually an eigenstate of $\hat{H}^{(1)}$, so there will be no change to any order. We can check our results against the exact eigenvalues and see that they are correct, which is left as an exercise for the reader; for that purpose it is useful to write $\hat{H}^{(1)}$ in the new basis ($\hat{H}^{(0)}$ of course being unchanged) as:

$$\hat{H}^{(1)} \longrightarrow a \begin{pmatrix} 0 & 0 & 0 \\ 0 & 2 & \sqrt{2} \\ 0 & \sqrt{2} & 1 \end{pmatrix}.$$

One final comment: we calculated

$$|3^{(1)}\rangle = \frac{\langle 1^{(0)} | \hat{H}^{(1)} | 3^{(0)} \rangle}{E_0} |1^{(0)}\rangle + \frac{\langle 2^{(0)} | \hat{H}^{(1)} | 3^{(0)} \rangle}{E_0} |2^{(0)}\rangle$$

But we could equally have used the un-diagonalised states $|1^{(0)}\rangle$ and $|2^{(0)}\rangle$. This can be seen if we write

$$|3^{(1)}\rangle = \frac{1}{E_0} \left(|1^{(0)}\rangle \langle 1^{(0)}| + |2^{(0)}\rangle \langle 2^{(0)}| \right) \hat{H}^{(1)} |3^{(0)}\rangle$$

and spot that the term in brackets is the identity operator in the degenerate subspace, which can equally well be written $\left(|1^{(0)}\rangle \langle 1^{(0)}| + |2^{(0)}\rangle \langle 2^{(0)}| \right)$. Of course for a problem in higher dimensions, there would be other terms coming from the non-degenerate states $|m^{(0)}\rangle$ as well.

4.2.2 Symmetry as a guide to the choice of basis

Recall that in the majority of cases, degeneracy arises from one or more symmetries, which show themselves as operators which commute with the Hamiltonian. In the case considered in the discussion of figure 4.1, the original Hamiltonian commuted with \hat{L}_z while the perturbations did not. In a more complicated case the perturbation may break some but not all of the symmetries: in 3D for instance full rotational invariance may be broken, but symmetry about the z -axis might remain, and so the full Hamiltonian will commute with \hat{L}_z . Then the perturbation will not mix initially-degenerate states of the same l but different m , and the usual choice of spherical harmonics for the wave functions will enable normal, non-degenerate perturbation theory to be used for the first-order energy shift. We say that m remains a good quantum number, since \hat{L}_z commutes with the full Hamiltonian.

More generally, if eigenstates of $H^{(0)}$ are also eigenstates of some operator $\hat{\Omega}$, we distinguish the degenerate states with energy $E_n^{(0)}$ by the eigenvalues of that operator: $\{|n, \omega_i\rangle\}$. If $[\hat{H}^{(1)}, \hat{\Omega}] = 0$, then from $\langle n, \omega_j | [\hat{H}^{(1)}, \hat{\Omega}] | n, \omega_i \rangle = 0$ we immediately have $\langle n, \omega_j | \hat{H}^{(1)} | n, \omega_i \rangle = 0$ if $\omega_j \neq \omega_i$. $\hat{H}^{(1)}$ is diagonal in the $E_n^{(0)}$ -subspace with this choice of basis; it does not mix the degenerate states.

We are about to move on to perturbations in the hydrogen atom, where $\hat{H}^{(0)}$ commutes with both $\hat{\mathbf{L}}$ and $\hat{\mathbf{S}}$, and we have a choice of quantum numbers to classify the state: our basis can be $\{|l, m_l; s, m_s\rangle\}$ or $\{|l, s; j, m_j\rangle\}$. If $\hat{H}^{(1)}$ fails to commute with $\hat{\mathbf{L}}$ or $\hat{\mathbf{S}}$, while still commuting with $\hat{\mathbf{L}}^2$, and $\hat{\mathbf{S}}^2$ and $\hat{\mathbf{J}}$, then we avoid all problems by simply choosing the second basis from the start.

4.3 The fine structure of hydrogen

Shankar 17.3; Mandl 7.4; Griffiths 6.3; Gasiorowicz 12.1,2,4

Summary: The Schrödinger equation with a pure Coulomb potential does an excellent job of describing hydrogen energy levels, but even 100 years ago spectroscopy could detect deviations from that picture. But first-order perturbation theory is good enough for almost all purposes, since the small parameter is $\alpha_{\text{EM}}^2 \sim 10^{-4}$.

4.3.1 Pure Coulomb potential and nomenclature

Although the Schrödinger equation with a Coulomb potential reproduces the Bohr model and gives an excellent approximation to the energy levels of hydrogen, the true spectrum was known to be more complicated right from the start. The small deviations are termed “fine structure” and they are of order 10^{-4} compared with the ground-state energy (though the equivalent terms for many-electron atoms can be sizable). Hence perturbation theory is an excellent framework in which to consider them.

First a reminder of the results of the unperturbed calculation. The Coulomb potential is $V(r) = -\hbar c \alpha / r$ (written in terms of the dimensionless $\alpha = e^2 / 4\pi\epsilon_0 \hbar c \approx 1/137$); the energies turn out to depend only on the principal quantum number n and not on l : $E_n = -\frac{1}{n^2} E_{\text{Ry}}$, where $E_{\text{Ry}} = \frac{1}{2} \alpha^2 m c^2 = 13.6 \text{ eV}$ (with m the reduced mass of the electron-proton system). For a given n all values of $l < n$ are allowed, giving $\sum_l^{n-1} (2l+1) = n^2$ degenerate states. The wave function of the ground state is proportional to e^{-r/a_0} , where a_0 is the Bohr radius, $a_0 = \hbar c / (m c^2 \alpha)$. Results for other hydrogen-like single electron atoms (with nuclear charge Z) can be obtained by replacing α with $Z\alpha$ and m with the appropriate reduced mass. Lists of wave functions are given in [A.6](#).

The states of the system are tensor-direct-product states of spatial and spin states, written $|n, l, m_l\rangle \otimes |\frac{1}{2}, m_s\rangle$ or $|n, l, m_l, m_s\rangle$ ($s = \frac{1}{2}$ is usually suppressed). Since the Hamiltonian has no spin dependence, the spatial state is just the one discussed previously. The alternative basis $|n, l, j, m_j\rangle$ is often used for reasons which will become clear; subshells of states of a given $\{l, s, j\}$ are referred to using the notation $^{(2s+1)}l_j$, with $l = s, p, d, f \dots$, so for example $^2f_{5/2}$ (or just $f_{5/2}$) has $s = \frac{1}{2}$ of course, $l = 3$ and $j = l - s = \frac{5}{2}$. An example of such a state is

$$|n, 1; \frac{1}{2}, \frac{1}{2}\rangle = \sqrt{\frac{2}{3}} |n, 1, 1\rangle \otimes |\frac{1}{2}, -\frac{1}{2}\rangle - \sqrt{\frac{1}{3}} |n, 1, 0\rangle \otimes |\frac{1}{2}, \frac{1}{2}\rangle.$$

(Compared to the section on addition of orbital and spin angular momenta, the only difference is that in calculating matrix elements of the states $|n, l, m_l\rangle$, there is a radial integral to be done as well as the angular one.) With the pure Coulomb potential, we can use whichever basis we like.

4.3.2 Fine structure: the lifting of l degeneracy

There are two effects to be considered. One arises from the use of the non-relativistic expression $p^2/2m$ for the kinetic energy, which is only the first term in an expansion of $\sqrt{(mc^2)^2 + (pc)^2} -$

mc^2 . The first correction term is $-p^4/(8m^3c^2)$, and its matrix elements are most easily calculated using the trick of writing it as $-1/(2mc^2)(\hat{H}^{(0)} - V_C(r))^2$, where $\hat{H}^{(0)}$ is the usual Hamiltonian with a Coulomb potential. Now in principle we need to be careful here, because $\hat{H}^{(0)}$ is highly degenerate (energies depend only on n and not on l or m). However we have $\langle n, l', m_l' | (\hat{H}^{(0)} - V_C(r))^2 | n, l, m_l \rangle = \langle n, l', m_l' | (E_n^{(0)} - V_C(r))^2 | n, l, m_l \rangle$, and since in this form the operator is spherically symmetric and spin-independent, it can't link states of different l , m_l or m_s . So the basis $\{|n, l, m_l, s_s\rangle\}$ already diagonalises $\hat{H}^{(1)}$ in each subspace of states with the same n , and we have no extra work to do here. (We are omitting the superscript (0) on the hydrogenic states, here and below.)

The final result for the kinetic energy effect is

$$\begin{aligned} \langle n, l, m_l | \hat{H}_{\text{KE}}^{(1)} | n, l, m_l \rangle &= -\frac{1}{2mc^2} \left((E_n^{(0)})^2 + 2E_n^{(0)} \hbar c \alpha \langle n, l | \frac{1}{r} | n, l \rangle + (\hbar c \alpha)^2 \langle n, l | \frac{1}{r^2} | n, l \rangle \right) \\ &= -\frac{\alpha^2 |E_n^{(0)}|}{n} \left(\frac{2}{2l+1} - \frac{3}{4n} \right) \end{aligned}$$

In calculating this the relation $E_{\text{Ry}} = \hbar c \alpha / (2a_0)$ is useful. The matrix elements involve radial integrals only; tricks for calculating these are explained in Shankar qu. 17.3.4; they are tabulated in A.6. Details of the algebra for this and the following calculation are given [here](#).

The second correction is the spin-orbit interaction:

$$\hat{H}_{\text{SO}}^{(1)} = \frac{1}{2m^2c^2r} \frac{dV_C}{dr} \hat{\mathbf{L}} \cdot \hat{\mathbf{S}}$$

In this expression $\hat{\mathbf{L}}$ and $\hat{\mathbf{S}}$ are the vector operators for orbital and spin angular momentum respectively. The usual (somewhat hand-waving) derivation talks of the electron seeing a magnetic field from the proton which appears to orbit it; the magnetic moment of the electron then prefers to be aligned with this field. A slightly more respectable derivation uses the fact that an EM field which is purely electric in one frame (the lab frame) has a magnetic component when seen from a moving frame; this at least makes it clear that the spin-orbit interaction is also a relativistic effect. Both “derivations” give an expression which is too large by a factor of 2; an exact result requires the use of the Dirac equation.

This time we will run into trouble with the degeneracy of $\hat{H}^{(0)}$ unless we do some work first. Since the Coulomb potential is spherically symmetric, there is no mixing of states of the same n but different l . However states of different $\{m_l, m_s\}$ will mix, since $\hat{\mathbf{L}} \cdot \hat{\mathbf{S}}$ does not commute with \hat{L}_z and \hat{S}_z . The trick of writing $2\hat{\mathbf{L}} \cdot \hat{\mathbf{S}} = \hat{\mathbf{J}}^2 - \hat{\mathbf{L}}^2 - \hat{\mathbf{S}}^2$ where $\hat{\mathbf{J}} = \hat{\mathbf{L}} + \hat{\mathbf{S}}$ tells us that $\hat{\mathbf{L}} \cdot \hat{\mathbf{S}}$ *does* commute with $\hat{\mathbf{J}}^2$ and \hat{J}_z , so we should work in the basis $|n, l, j, m_j\rangle$, instead. (The label $s = \frac{1}{2}$ is suppressed.) This basis diagonalises the spin-orbit perturbation (and is an equally acceptable basis, giving the same result, for the relativistic correction term above).

Then

$$\begin{aligned} \langle n, l, j, m_j | \hat{H}_{\text{SO}}^{(1)} | n, l, j, m_j \rangle &= \frac{1}{2m^2c^2} \langle n, l | \frac{1}{r} \frac{dV_C}{dr} | n, l \rangle \langle l, j, m_j | \frac{1}{2} (\hat{\mathbf{J}}^2 - \hat{\mathbf{L}}^2 - \hat{\mathbf{S}}^2) | l, j, m_j \rangle \\ &= \frac{\hbar c \alpha}{4m^2c^2} \langle n, l | \frac{1}{r^3} | n, l \rangle \hbar^2 (j(j+1) - l(l+1) - \frac{3}{4}) \\ &= \frac{\alpha^2 |E_n^{(0)}|}{n} \left(\frac{2}{2l+1} - \frac{2}{2j+1} \right), \end{aligned}$$

where in the first line we have separated the radial integral from the angular momentum matrix elements, and where a fair amount of algebra links the last two lines. (This expression is only

correct for $l \neq 0$. However there is another separate effect, the Darwin term, which only affects s -waves and whose expectation value is just the same as above (with $l = 0$ and $j = \frac{1}{2}$), so we can use this for all l . The Darwin term can only be understood in the context of the Dirac equation.)

So finally

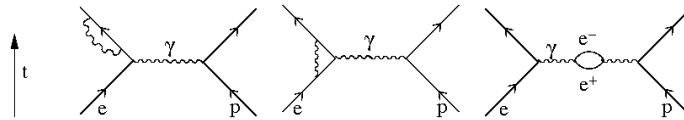
$$E_{nj}^{(1)} = \frac{\alpha^2 |E_n^{(0)}|}{n} \left(\frac{3}{4n} - \frac{2}{2j+1} \right).$$

The degeneracy of all states with the same n has been broken. States of given j with $l = j \pm \frac{1}{2}$ are still degenerate, a result that persists to all orders in the Dirac equation (where in any case orbital angular momentum is no longer a good quantum number.) So the eight $n = 2$ states are split by 4.5×10^{-5} eV, with the ${}^2p_{3/2}$ state lying higher than the degenerate ${}^2p_{1/2}$ and ${}^2s_{1/2}$ states.

For a brief comment on spin-orbit splitting in light atoms other than hydrogen, see [2.5.3](#).

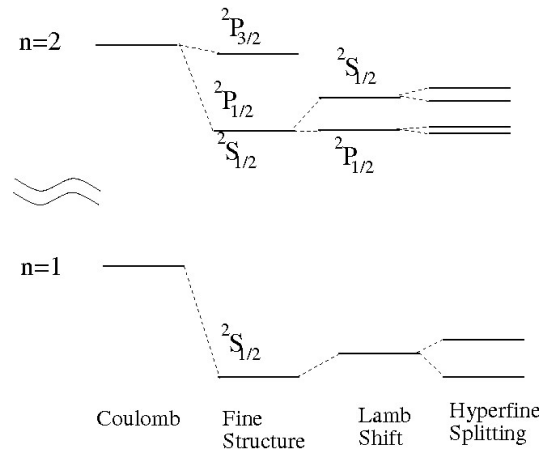
Two other effects should be mentioned here. One is the *hyperfine splitting*. The proton has a magnetic moment, and the energy of the atom depends on whether the electron spin is aligned with it or not— more precisely, whether the total spin of the electron and proton is 0 or 1. The anti-aligned case has lower energy (since the charges are opposite), and the splitting for the $1s$ state is 5.9×10^{-6} eV. (It is around a factor of 10 smaller for any of the $n = 2$ states.) Transitions between the two hyperfine states of $1s$ hydrogen give rise to the 21 cm microwave radiation which is a signal of cold hydrogen gas in the galaxy and beyond.

The final effect is called the *Lamb shift*. It cannot be accounted for in quantum mechanics, but only in quantum field theory.



The diagrams above show corrections to the simple Coulomb force which would be represented by the exchange of a single photon between the proton and the electron. The most notable effect on the spectrum of hydrogen is to lift the remaining degeneracy between the ${}^2p_{1/2}$ and ${}^2s_{1/2}$ states, so that the latter is higher by 4.4×10^{-6} eV.

Below the various corrections to the energy levels of hydrogen are shown schematically. The gap between the $n = 1$ and $n = 2$ shells is suppressed, and the Lamb and hyperfine shifts are exaggerated in comparison with the fine-structure. The effect of the last two on the ${}^2p_{3/2}$ level is not shown.



4.4 The Zeeman effect: hydrogen in an external magnetic field

(Shankar 14.5); Mandl 7.5; Griffiths 6.4; Gasiorowicz 12.3

Summary: This is an interesting application of degenerate perturbation theory because the appropriate unperturbed basis depends on the strength of the external field compared to the effective internal one which contributes to fine structure.

(Since we will not ignore spin, this whole section is about the so-called anomalous Zeeman effect. The so-called normal Zeeman effect cannot occur for hydrogen, but is a special case which pertains in certain multi-electron atoms for which the total spin is zero.)

With an external magnetic field along the z -axis, the perturbing Hamiltonian is $\hat{H}^{(1)} = -\boldsymbol{\mu} \cdot \mathbf{B} = (\mu_B B / \hbar)(\hat{L}_z + 2\hat{S}_z)$. The factor of 2 multiplying the spin is of course the famous g -factor for spin, as predicted by the Dirac equation. Clearly this is diagonalised in the $\{|n, l, m_l, m_s\rangle\}$ basis ($s = \frac{1}{2}$ suppressed in the labelling as usual). Then $E_{nlm_l m_s}^{(1)} = \mu_B B (m_l + 2m_s)$. If, for example, $l = 2$ there are 7 possible values of $m_l + 2m_s$ between -3 and 3 , with $-1, 0$ and 1 being degenerate ($5 \times 2 = 10$ states in all).

This is the “strong-field Zeeman” or “Paschen-Back” effect. It is applicable if the magnetic field is strong enough that we can ignore the fine structure discussed in the last section. In hydrogen, that means $B \gg 10^{-4} \text{ eV} / \mu_B \sim 2 \text{ T}$. But for a weak field, $B \ll 2 \text{ T}$, the fine structure effects will be stronger, so we will consider them part of $\hat{H}^{(0)}$ for the Zeeman problem; our basis is then $\{|n, l, j, m_j\rangle\}$ and states of the same j but different l and m_j are degenerate. This degeneracy however is not a problem, because the operator $(\hat{L}_z + 2\hat{S}_z)$ does not connect states of different l or m_j (since it commutes with $\hat{\mathbf{L}}^2$ and with \hat{J}_z). So we can use non-degenerate perturbation theory, with

$$E_{nljm_j}^{(1)} = \frac{\mu_B B}{\hbar} \langle n, l, j, m_j | \hat{L}_z + 2\hat{S}_z | n, l, j, m_j \rangle = \mu_B B m_j + \frac{\mu_B B}{\hbar} \langle n, l, j, m_j | \hat{S}_z | n, l, j, m_j \rangle.$$

If \hat{J}_z is conserved but \hat{L}_z and \hat{S}_z are not, the expectation values of the latter two might be expected to be proportional to the first, modified by the average degree of alignment: $\langle \hat{S}_z \rangle = \hbar m_j \langle \hat{\mathbf{S}} \cdot \hat{\mathbf{J}} \rangle / \langle \hat{\mathbf{J}}^2 \rangle$. (This falls short of a proof but is in fact correct, and follows from the Wigner

Eckart theorem as explained in section 2.6. A similar expression holds for \hat{L}_z . A semi-classical derivation may be found [here](#).) Using $2\hat{\mathbf{S}} \cdot \hat{\mathbf{J}} = \hat{\mathbf{S}}^2 + \hat{\mathbf{J}}^2 - \hat{\mathbf{L}}^2$ gives

$$E_{nljmj}^{(1)} = \mu_B B m_j \left(1 + \frac{j(j+1) - l(l+1) + s(s+1)}{2j(j+1)} \right) = \mu_B B m_j g_{jls}.$$

Of course for hydrogen $s(s+1) = \frac{3}{4}$, but the expression above, which defines the *Landé g factor*, is actually more general and hence I've left it with an explicit s . For hydrogen, $j = l \pm \frac{1}{2}$ and so $g = (1 \pm \frac{1}{2l+1})$.

Thus states of a given j (already no longer degenerate due to fine-structure effects) are further split into $(2j + 1)$ equally-spaced levels. Since spectroscopy involves observing transitions between two states, both split but by different amounts, the number of spectral lines can be quite large.

For heavier (but not too heavy) atoms the fine structure splitting of levels is much greater than in hydrogen, and the weak-field Zeeman effect is the norm. The expression above can be used for the Landé g -factor, using the quantum numbers J , L and S for the atom as a whole.

For protons and neutrons the complex internal dynamics of the quarks leads to g -factors, and hence magnetic moments, which are quite different from 2 (for the proton) and 0 (for the uncharged neutron); see comments in 2.5.3. The magnetic moments of nuclei with one unpaired nucleon in an orbital of specified l and j can then be estimated analogously to the procedure above. (These estimates give the “Schmidt limits” which will be discussed in PHYS40302.)

4.5 The Stark effect: hydrogen in an external electric field

Shankar 17.2,3; Gasiorowicz 11.3; (Griffiths problems 6.35,36)

Summary: In this case we have to get to grips with degenerate and beyond-first-order perturbation theory

In this section we consider the energy shifts of the levels of hydrogen in an external electric field, taken to be along the z -axis: $\mathbf{E} = \mathcal{E}\mathbf{e}_z$ (we use \mathcal{E} for the electric field strength to distinguish it from the energy). We will work in the strong-field limit and ignore fine structure; furthermore the dynamics are then independent of the spin so we can ignore m_s ; the unperturbed eigenstates can be taken to be $|n, l, m_l\rangle$.

The perturbing Hamiltonian is $\hat{H}^{(1)} = |e|\mathcal{E}z$. (In this section we will not write \hat{z} for fear of confusion with a unit vector.) Now it is immediately obvious that, for any state, $\langle n, l, m_l | z | n, l, m_l \rangle = 0$. The probability density is symmetric on reflection in the xy -plane, but z is antisymmetric. So for the ground state, the first order energy shift vanishes. (We will return to excited states, but think now about why we can't conclude the same for them.) This is not surprising, because an atom of hydrogen in its ground state has no electric dipole moment: there is no $\mathbf{p} \cdot \mathbf{E}$ term to match the $\boldsymbol{\mu} \cdot \mathbf{B}$ one.

To calculate the second-order energy shift we need $\langle n, l, m_l | z | 1, 0, 0 \rangle$. We can write z as $r \cos \theta$ or $\sqrt{4\pi/3} r Y_1^0(\theta, \phi)$. The lack of dependence on ϕ means that m_l can't change, and in addition l can only change by one unit, so $\langle n, l, m_l | z | 1, 0, 0 \rangle = \delta_{l1} \delta_{m_l 0} \langle n, 1, 0 | z | 1, 0, 0 \rangle$. However this isn't the whole story: there are also states in the continuum, which we will denote $|\mathbf{k}\rangle$ (though these are not plane waves, since they see the Coulomb potential). So we have

$$E_{100}^{(2)} = (e\mathcal{E})^2 \sum_{n>1} \frac{|\langle n, 1, 0 | z | 1, 0, 0 \rangle|^2}{E_1^{(0)} - E_n^{(0)}} + (e\mathcal{E})^2 \int d^3k \frac{|\langle \mathbf{k} | z | 100 \rangle|^2}{E_1^{(0)} - E_k^{(0)}}$$

(We use E_1 for E_{100}). This is a compact expression, but it would be very hard to evaluate directly. We can get a crude estimate of the size of the effect by simply replacing all the denominators by $E_1^{(0)} - E_2^{(0)}$; this overestimates the magnitude of every term but the first, for which it is exact, so it will give an upper bound on the magnitude of the shift. Then (recalling $E_1^{(2)} < 0$),

$$\begin{aligned} E_1^{(2)} &> \frac{(e\mathcal{E})^2}{E_1^{(0)} - E_2^{(0)}} \left(\sum_{n \geq 1} \sum_{l m_l} \langle 1, 0, 0 | z | n, l, m_l \rangle \langle n, l, m_l | z | 1, 0, 0 \rangle + \int d^3k \langle 1, 0, 0 | z | \mathbf{k} \rangle \langle \mathbf{k} | z | 1, 0, 0 \rangle \right) \\ &= \frac{(e\mathcal{E})^2}{E_1^{(0)} - E_2^{(0)}} \langle 1, 0, 0 | z \left(\sum_{n \geq 1} \sum_{l m_l} |n, l, m_l\rangle \langle n, l, m_l| + \int d^3k |\mathbf{k}\rangle \langle \mathbf{k}| \right) z | 1, 0, 0 \rangle \\ &= \frac{(e\mathcal{E})^2}{E_1^{(0)} - E_2^{(0)}} \langle 1, 0, 0 | z^2 | 1, 0, 0 \rangle = -\frac{4(e\mathcal{E}a_0)^2}{3E_{\text{Ry}}} = -\frac{8(e\mathcal{E})^2 a_0^3}{3\hbar c \alpha} \end{aligned}$$

where we have included $n = 1$ and other values of l and m in the sum because the matrix elements vanish anyway, and then used the completeness relation involving all the states, bound and unbound, of the hydrogen atom. For details of this integral and the one needed for the next part, see [here](#), p2.

There is a trick for evaluating the exact result, which gives $9/4$ rather than $8/3$ as the constant (See Shankar.) So our estimate of the magnitude is fairly good. (For comparison with other ways of writing the shift, note that $(e\mathcal{E})^2/\hbar c\alpha = 4\pi\epsilon_0\mathcal{E}^2$ —or, in **Gaussian units**, just \mathcal{E}^2 .)

Having argued above that the hydrogen atom has no electric dipole, how come we are getting a finite effect at all? The answer of course is that the field polarises the atom, and the induced dipole can then interact with the field. We have in fact calculated the polarisability of the hydrogen atom.

Now for the first excited state. We can't conclude that the first-order shift vanishes here, of course, because of degeneracy: there are four states and $\hat{H}^{(1)}$ is not diagonal in the usual basis $|2, l, m_l\rangle$ (with $l = 0, 1$). In fact as we argued above it only connects $|2, 0, 0\rangle$ and $|2, 1, 0\rangle$, so the states $|2, 1, \pm 1\rangle$ decouple and their first order shifts do vanish. Using $\langle 2, 1, 0|z|2, 0, 0\rangle = -3a_0$, we have in this subspace (with $|2, 0, 0\rangle = (1, 0)^\top$ and $|2, 1, 0\rangle = (0, 1)^\top$)

$$\hat{H}^{(1)} = -3a_0|e|\mathcal{E} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},$$

and the eigenstates are $\sqrt{1/2}(|2, 0, 0\rangle \pm |2, 1, 0\rangle)$ with eigenvalues $\mp 3a_0|e|\mathcal{E}$. So the degenerate quartet is split into a triplet of levels (with the unshifted one doubly degenerate).

In reality the degeneracy of the $n = 2$ states is lifted by the fine-structure splitting; are these results then actually relevant? They will be approximately true if the field is large; at an intermediate strength both fine-structure and Stark effects should be treated together as a perturbation on the pure Coulomb states. For very weak fields degenerate perturbation theory can be applied in the space of $j = \frac{1}{2}$ states ($^2s_{\frac{1}{2}}$ and $^2p_{\frac{1}{2}}$), which are shifted by $\pm\sqrt{3}a_0|e|\mathcal{E}$. The $j = \frac{3}{2}$ states though have no first-order shift.

5. Quantum Measurement

5.1 The Einstein-Poldosky-Rosen “paradox” and Bell’s inequalities

Mandl 6.3; Griffiths 12.2; Gasiorowicz 20.3,4

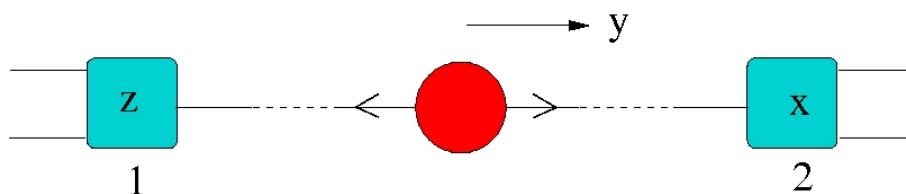
Summary: This section is about nothing less important than “the nature of reality”!

In 1935 Einstein, along with Boris Poldosky and Nathan Rosen, published a paper entitled “Can quantum-mechanical description of physical reality be considered complete?” By this stage Einstein had accepted that the uncertainty principle did place fundamental restrictions on what one could discover about a particle through measurements conducted on it. The question however was whether the measuring process actually somehow brought the properties into being, or whether they existed all along but without our being able to determine what they were. If the latter was the case there would be “hidden variables” (hidden from the experimenter) and the quantum description—the wave function—would not be a complete description of reality. Till the EPR paper came out many people dismissed the question as undecidable, but the EPR paper put it into much sharper focus. Then in 1964 John Bell presented an [analysis](#) of a variant of the EPR paper which showed that the question actually was decidable. Many experiments have been done subsequently, and they have come down firmly in favour of a positive answer to the question posed in EPR’s title.

The original EPR paper used position and momentum as the two properties which couldn’t be simultaneously known (but might still have hidden definite values), but subsequent discussions have used components of spin instead, and we will do the same. But I will be quite lax about continuing to refer to “the EPR experiment”.

There is nothing counter-intuitive or unclassical about the fact that we can produce a pair of particles whose total spin is zero, so that if we find one to be spin-up along some axis, the other must be spin down. All the variants of the experiment to which we will refer can be considered like this: such a pair of electrons is created travelling back-to-back at one point, and travel to distant measuring stations where each passes through a Stern-Gerlach apparatus (an “SG”) of a certain orientation in the plane perpendicular to the electrons’ momentum.

As I say there is nothing odd about the fact that when the two SGs have the same orientation the two sequences recorded at the two stations are perfectly anti-correlated (up to measurement errors). But consider the case where they are orientated at 90° with respect to each other as below: Suppose for a particular pair of electrons, we measure number 1 to be spin up in the z -direction and number 2 to be spin down in the x -direction. Now let’s think about what *would have happened* if we had instead measured the spin in the x -direction of particle 1. Surely, say EPR, we know the answer. Since particle 2 is spin down in the x -direction, particle 1 would



have been spin up. So now we know that before it reached the detector, particle 1 was spin up in the z -direction (because that's what we got when we measured it) and also spin up in the x -direction (because it is anti-correlated with particle 2 which was spin down). We have beaten the uncertainty principle, if only retrospectively.

But of course we know we can't construct a wave function with these properties. So is there more to reality than the wave function? Bell's contribution was to show that the assumption that the electron really has definite values for different spin components—if you like, it has an instruction set which tells it which way to go through any conceivable SG that it might encounter—leads to testable predictions.

For Bell's purposes, we imagine that the two measuring stations have agreed that they will set their SG to one of 3 possible settings. Setting A is along the z -direction, setting C is along the x direction, and setting B is at 45° to both. In the ideal set-up, the setting is chosen just before the electron arrives, sufficiently late that no possible causal influence (travelling at not more than the speed of light) can reach the other lab before the measurements are made. The labs record their results for a stream of electrons, and then get together to classify each pair as, for instance, $(A \uparrow, B \downarrow)$ or $(A \uparrow, C \uparrow)$ or $(B \uparrow, B \downarrow)$ (the state of electron 1 being given first). Then they look at the number of pairs with three particular classifications: $(A \uparrow, B \uparrow)$, $(B \uparrow, C \uparrow)$ and $(A \uparrow, C \uparrow)$. Bell's inequality says that, *if the way the electrons will go through any given orientation is set in advance*,

$$N(A \uparrow, B \uparrow) + N(B \uparrow, C \uparrow) \geq N(A \uparrow, C \uparrow)$$

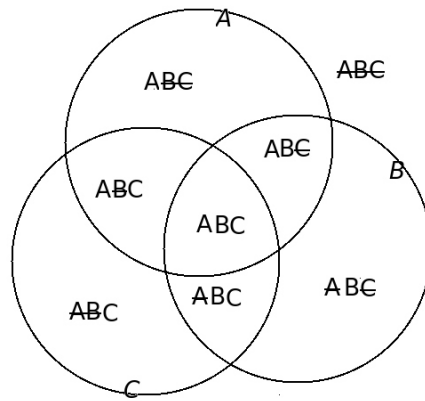
where $N(A \uparrow, B \uparrow)$ is the number of $(A \uparrow, B \uparrow)$ pairs etc.

Now let's prove that.

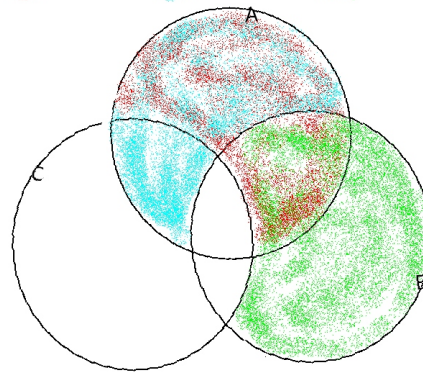
Imagine any set of objects (or people!) with three distinct binary properties a , b and c —say blue or brown eyes, right or left handed, and male or female (ignoring messy reality in which there are some people not so easily classified). In each case, let us denote the two possible values as A and \bar{A} etc (\bar{A} being “not A ” in the sense it is used in logic, so if A is blue-eyed, \bar{A} is brown-eyed). Then every object is classified by its values for the three properties as, for instance, ABC or $\bar{A}\bar{B}\bar{C}$ or $\bar{A}BC \dots$. The various possibilities are shown on a Venn diagram below (sorry that the bars are through rather than over the letters...) In any given collection of objects, there will be no fewer than zero objects in each subset, obviously. All the N s are greater than or equal to zero. Now we want to prove that the number of objects which are $\bar{A}\bar{B}$ (irrespective of c) plus those that are $B\bar{C}$ (irrespective of a) is greater than or equal to the number which are $A\bar{C}$ (irrespective of b):

$$N(\bar{A}\bar{B}) + N(B\bar{C}) \geq N(A\bar{C})$$

This is obvious from the diagram below, in which the union of the blue and green sets fully contains the red set.



$$\text{A Not C} < \text{A Not B} + \text{B Not C}$$



A logical proof is as follows:

$$\begin{aligned} N(\overline{AB}) + N(\overline{BC}) &= N(\overline{ABC}) + N(\overline{A\overline{B}\overline{C}}) + N(\overline{A\overline{B}C}) + N(\overline{\overline{A}B\overline{C}}) \\ &= N(\overline{ABC}) + N(\overline{A\overline{C}}) + N(\overline{\overline{A}B\overline{C}}) \geq N(\overline{A\overline{C}}) \end{aligned}$$

To apply to the spins we started with, we identify A with $A \uparrow$ and \overline{A} with $A \downarrow$. Now if an electron is $A \uparrow B \downarrow$ (whatever C might be) then its partner must be $A \downarrow B \uparrow$, and so the result of a measurement A on the first and B on the second will be $(A \uparrow, B \uparrow)$. Hence the inequality for the spin case is a special case of the general one. We have proved Bell's inequality assuming, remember, that the electrons really do have these three defined properties even if, for a single electron, we can only measure one of them.

Now let's consider what quantum mechanics would say. A spin-0 state of two identical particles is

$$|S = 0\rangle = \sqrt{\frac{1}{2}} \left(|\uparrow\rangle \otimes |\downarrow\rangle - |\downarrow\rangle \otimes |\uparrow\rangle \right)$$

and this is true whatever the axis we have chosen to define "up" and "down". As expected, if we choose the same measurement direction at the two stations (eg both A), the first measurement selects one of the two terms and so the second measurement, on the other particle, always gives the opposite result. (Recall this is the meaning of the 2-particle wave function being non-seperable or *entangled*.)

What about different measurement directions at the two stations (eg A and B)? Recall the relation between the spin-up and spin-down states for two directions in the xz -plane, where θ

is the angle between the two directions:

$$\begin{aligned} |\theta, \uparrow\rangle &= \cos \frac{\theta}{2} |0, \uparrow\rangle + \sin \frac{\theta}{2} |0, \downarrow\rangle & |0, \uparrow\rangle &= \cos \frac{\theta}{2} |\theta, \uparrow\rangle - \sin \frac{\theta}{2} |\theta, \downarrow\rangle \\ |\theta, \downarrow\rangle &= -\sin \frac{\theta}{2} |0, \uparrow\rangle + \cos \frac{\theta}{2} |0, \downarrow\rangle & |0, \downarrow\rangle &= \sin \frac{\theta}{2} |\theta, \uparrow\rangle + \cos \frac{\theta}{2} |\theta, \downarrow\rangle. \end{aligned}$$

(We previously showed this for the first axis being the z -axis, but, up to overall phases, it is true for any pair). For A and B or for B and C , $\theta = 45^\circ$; for A and C it is 90° .

Consider randomly-oriented spin-zero pairs and settings A , B and C equally likely. If the first SG is set to A and the second to B (which happens 1 time in 9), there is a probability of $1/2$ of getting $A \uparrow$ at the first station. But then we know that the state of the second electron is $|A \downarrow\rangle$ and the probability that we will measure spin-up in the B direction is $|\langle B \uparrow | A \downarrow \rangle|^2 = \sin^2 \frac{\pi}{8}$. (This probability doesn't change if we start with the second measurement.) Thus the fraction of pairs which are $(A \uparrow, B \uparrow)$ is $\frac{1}{2} \sin^2 22.5^\circ = 0.073$, and similarly for $(B \uparrow, C \uparrow)$. But the fraction which are $(A \uparrow, C \uparrow)$ is $\frac{1}{2} \sin^2 45^\circ = 0.25$. So the prediction of quantum mechanics for $9N_0$ measurements is

$$N(A \uparrow, B \uparrow) + N(B \uparrow, C \uparrow) = 0.146N_0 < N(A \uparrow, C \uparrow) = 0.25N_0$$

So in quantum mechanics, Bell's inequality is violated. The experiment has been done many times, starting with the pioneering work of Alain Aspect, and every time the predictions of quantum mechanics are upheld and Bell's inequality is violated. (Photons rather than electrons are often used. Early experiments fell short of the ideal in many ways, but as loopholes have been successively closed the result has become more and more robust.)

It seems pretty inescapable that the electrons have not "decided in advance" how they will pass through any given SG. Do we therefore have to conclude that the measurement made at station 1 is responsible for collapsing the wave function at station 2, even if there is no time for light to pass between the two? It is worth noting that no-one has shown any way to use this set-up to send signals between the stations; on their own they both see a totally random succession of results. It is only in the statistical correlation that the weirdness shows up...

In writing this section I found [this document](#) by David Harrison of the University of Toronto very useful.

As well as the textbook references given at the start, further discussions can be found in N. David Mermin's book *Boojums all the way through* (CUP 1990) and in John S. Bell's *Speakable and unspeakable in quantum mechanics* (CUP 1987).

A. Revision and background

A.1 Index notation—free and dummy indices

This is a brief reminder about index notation. When we have a set of N objects, we often label them with the subscript $1, 2 \dots N$. Components of a vector $|v\rangle$ in some basis, for instance, are $(v_1, v_2, \dots v_N)$. When we write v_3 , the index 3 is definite, we are talking about the third component. But often we want to make more general statements, for instance

$$|u\rangle + |v\rangle = |w\rangle \Rightarrow u_1 + v_1 = w_1 \text{ and } u_2 + v_2 = w_2 \text{ and } u_3 + v_3 = w_3 \text{ and } \dots$$

or

$$u_i + v_i = w_i.$$

Here i is a *free* index; it stands for 1 or 2 or 3 or $\dots N$. It allows us to write N equations in one, and is completely equivalent in content to the vector equation. It has to occur in each additive term on both sides of the equation. There is nothing special about the choice of i : provided we make the substitution *everywhere* in the equation we can use j or k or m or n or \dots . Often more than one free index is needed, e.g. to label the i th row and j th column of a matrix, and they must be chosen to be different (unless one actually means the diagonal elements). Free indices can also be used to label basis vectors, as in $\{|i\rangle\}$. If these basis vectors are orthonormal, they satisfy

$$\langle i|j\rangle = \delta_{ij} \equiv \begin{cases} 1 & \text{if } i = j, \\ 0 & \text{if } i \neq j. \end{cases}$$

This is an example of an equation with two free indices. δ_{ij} is the Kronecker delta. (In a different context it might also be used for the general element of the unit matrix or identity operator.)

Another use of indices is to represent summation:

$$\langle u|v\rangle = u_1^*v_1 + u_2^*v_2 + u_3^*v_3 + \dots + u_N^*v_N = \sum_{i=1}^N u_i^*v_i.$$

Here i is called a *dummy* index and represents *all* indices. This is a completely different use from free indices; we only have one equation not a set of N , and we can't replace i with a definite index of our choice. Again, though, we can use another symbol, say j , for the dummy index. (In this course we will never omit the \sum —we won't use the Einstein equation convention, because it causes problems in eigenvalue equations. But the limits 1 to N will often be implied and omitted.)

The crucial rule is that if dummy and free indices, or more than one dummy index, occur in the same equation, they must all be different. So for instance suppose we have an orthonormal basis $\{|i\rangle\}$ in terms of which we write $|v\rangle = \sum_i v_i|i\rangle$. If we want to find the value of $\langle i|a\rangle$, where

i is a free index, we have to choose something different for the dummy index in the sum, say j :

$$\langle i|v\rangle = \langle i|\left(\sum_j v_j|j\rangle\right) = \sum_j v_j\langle i|j\rangle = \sum_j v_j\delta_{ij} = v_i.$$

In the penultimate step we still had a sum over j , but only the i th term was non-zero. Suppose instead we'd used i as the dummy index, then we would end up with something like

$$\sum_i v_i\langle i|i\rangle = \sum_i v_i = v_1 + v_2 + \dots + v_N$$

which is not right!

Similarly if we also have $|u\rangle = \sum_i u_i|i\rangle$ and we want $\langle u|v\rangle$, we need to use an index other than i for at least one of the two sums (it doesn't matter which):

$$\langle u|v\rangle = \left(\sum_i u_i^*\langle i|\right)\left(\sum_j v_j|j\rangle\right) = \sum_{ij} u_i^*v_j\delta_{ij} = \sum_i u_i^*v_i, \text{ or equivalently } = \sum_j u_j^*v_j.$$

The Kronecker delta collapses the double sum to a single sum, but it doesn't matter which of the two dummy indices we use in that single sum.

A.2 Vector spaces

This section is intended as revision for those who did PHYS20672 last semester. Those who did not should consult the longer guides [here](#) and [here](#), as well as chapter 1 of Shankar.

A.2.1 Vectors and operators in finite spaces

Vectors in a **vector space** are members of a set for which addition and scalar multiplication both yield new members of the set. The familiar displacement and velocity vectors in real 3-D space are only some examples of vectors, and many more abstract instances occur in physics. In particular, the state of a quantum system is a vector in an infinite-dimensional vector space, and the possibility of *superposition*, which is one of the main ways in which classical and quantum descriptions of objects differ, follows. We denote vectors as $|\dots\rangle$, eg $|v\rangle$, $|3\rangle$, $|\psi\rangle$, $|+\rangle$, $|\heartsuit\rangle$, where the text between the “|” and the “)” is just a name or label for the ket, which can take many forms. The **dimension** of the space is the size of the largest set of vectors which can be linearly independent, and such a set is called a **basis**. Any vector in the space can be written as a sum over basis vectors

$$|v\rangle = \sum_{n=1}^N v_n |n\rangle$$

and the numbers v_n are called the **coefficients** or **components** of the vector in that basis. For a given basis, specifying the components specifies the vector.

Multiplying any vector by zero gives the null vector, which properly should be written $|0\rangle$ but is often written simply as 0. Indeed in QM, $|0\rangle$ may denote the ground state of a system, and in quantum field theory it may denote the vacuum.

We are usually concerned with spaces in which two vectors can be combined to give a complex number; this is the **inner product** which is written $\langle w|v\rangle = \langle v|w\rangle^*$. Note that if $|w\rangle = \alpha|a\rangle + \beta|b\rangle$, α and β being complex numbers, then $\langle w|v\rangle = \alpha^*\langle a|v\rangle + \beta^*\langle b|v\rangle$. This is called conjugate- or skew-linearity.

We may write vectors of a basis as $\{|v_1\rangle, |v_1\rangle \dots |v_N\rangle\}$ or simply as $\{|1\rangle, |2\rangle \dots |N\rangle\}$. It is very useful to work with **orthonormal** bases for which $\langle m|n\rangle = \delta_{mn}$.

When the vectors we are talking about are ordinary vectors in real 3-D space, we will tend not to use Dirac notation. Cartesian unit vectors forming an orthonormal basis in that space will be written $\{\mathbf{e}_x, \mathbf{e}_y, \mathbf{e}_z\}$. Here the inner product is the familiar scalar product.

Operators act on vectors to produce new vectors: $\hat{Q}|v\rangle = |w\rangle$. The **matrix element** of \hat{Q} between two vectors is defined as $\langle u|\hat{Q}|v\rangle = \langle u|w\rangle$. The **identity operator** \hat{I} leaves vectors unchanged.

The object $\hat{A} = |u\rangle\langle v|$ is an operator, since it can act on a vector to give another (which will always be proportional to $|u\rangle$): $\hat{A}|w\rangle = (\langle v|w\rangle)|u\rangle$. If the vectors $\{|n\rangle\}$ form an orthonormal basis, then

$$\sum_{n=1}^N |n\rangle\langle n| = \hat{I} \quad \text{since} \quad \left(\sum_{n=1}^N |n\rangle\langle n| \right) |v\rangle = \sum_{n=1}^N |n\rangle\langle n|v\rangle = \sum_{n=1}^N v_n |n\rangle = |v\rangle.$$

This is called the **completeness relation**.

An operator is fully defined by what it does to the vectors of a basis, since then we can find what it does to any other vector. For each basis vector $|n\rangle$, $\hat{Q}|n\rangle$ is a new vector which can itself be expanded in the basis: $\hat{Q}|n\rangle = \sum_m^N Q_{mn}|m\rangle$. These N^2 numbers Q_{mn} fully define the operator, in the same way that the components of vector fully define it (always with respect to a given basis of course). With an orthonormal basis, we have

$$v_n = \langle n|v\rangle, \quad Q_{mn} = \langle m|\hat{Q}|n\rangle \quad \text{and} \quad w_m = \sum_n^N Q_{mn}v_n.$$

The final equation is reminiscent of matrix multiplication. We can write the components of a vector as a vertical list (or column vector), and of an operator as a matrix, to give:

$$|v\rangle \longrightarrow \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_N \end{pmatrix} = \begin{pmatrix} \langle 1|v\rangle \\ \langle 2|v\rangle \\ \vdots \\ \langle N|v\rangle \end{pmatrix} \equiv \mathbf{v},$$

$$\langle v| \longrightarrow (v_1^*, v_2^*, \dots, v_N^*) = (\langle 1|v\rangle, \langle 2|v\rangle, \dots, \langle N|v\rangle) \equiv \mathbf{v}^\dagger,$$

$$\hat{Q} \longrightarrow \begin{pmatrix} Q_{11} & Q_{12} & \dots & Q_{1N} \\ Q_{21} & Q_{22} & \dots & Q_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ Q_{N1} & Q_{N2} & \dots & Q_{NN} \end{pmatrix} = \begin{pmatrix} \langle 1|\hat{Q}|1\rangle & \langle 1|\hat{Q}|2\rangle & \dots & \langle 1|\hat{Q}|N\rangle \\ \langle 2|\hat{Q}|1\rangle & \langle 2|\hat{Q}|2\rangle & \dots & \langle 2|\hat{Q}|N\rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle N|\hat{Q}|1\rangle & \langle N|\hat{Q}|2\rangle & \dots & \langle N|\hat{Q}|N\rangle \end{pmatrix} \equiv \mathbf{Q}.$$

The Q_{mn} are called the matrix elements of \hat{Q} in this basis. So

$$\langle u|\hat{Q}|v\rangle = (u_1^*, u_2^*, \dots, u_N^*) \begin{pmatrix} Q_{11} & Q_{12} & \dots & Q_{1N} \\ Q_{21} & Q_{22} & \dots & Q_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ Q_{N1} & Q_{N2} & \dots & Q_{NN} \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_N \end{pmatrix} = \mathbf{u}^\dagger \mathbf{Q} \mathbf{v}$$

The symbol $\xrightarrow{\text{name}}$ means “**is represented by**”, with *name* being a name or label for the basis, which will be omitted if the basis is obvious. In different bases, the components and matrix elements will be different. The corresponding column vectors and matrices are *different representations of the same vector/operator*. (Note though that $\langle u|\hat{Q}|v\rangle$ is a just number and independent of the representation.)

Note that in their own basis, the basis vectors themselves have extremely simple representations: in a 3-D space, if we use the symbol $\xrightarrow{\{1,2,3\}}$ to mean “is represented in the $\{|1\rangle, |2\rangle, |3\rangle\}$ basis by”, then

$$|1\rangle \xrightarrow{\{1,2,3\}} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \quad |2\rangle \xrightarrow{\{1,2,3\}} \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \quad |3\rangle \xrightarrow{\{1,2,3\}} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

If we choose a new orthonormal basis $\{|n'\rangle\}$, vectors and operators will have new coefficients. With $v_n = \langle n|v\rangle$, $v'_n = \langle n'|v\rangle$, $Q_{mn} = \langle m|\hat{Q}|n\rangle$ and $Q'_{mn} = \langle m'|\hat{Q}|n'\rangle$, and where \mathbf{S} is a unitary matrix (not an representing an operator) defined as $S_{mn} = \langle m|n'\rangle$, we have the following

relations between the two representations:

$$v'_n = \sum_j S_{jn}^* v_j \Rightarrow \mathbf{v}' = \mathbf{S}^\dagger \mathbf{v}; \quad Q'_{ij} = S_{ki}^* Q_{kl} S_{lj} \Rightarrow \mathbf{Q}' = \mathbf{S}^\dagger \mathbf{Q} \mathbf{S}.$$

For instance the vectors

$$|1'\rangle = \frac{1}{2}|1\rangle + \frac{i}{\sqrt{2}}|2\rangle - \frac{1}{2}|3\rangle, \quad |2'\rangle = \sqrt{\frac{1}{2}}(|1\rangle + |3\rangle), \quad |3'\rangle = \frac{1}{2}|1\rangle - \frac{i}{\sqrt{2}}|2\rangle - \frac{1}{2}|3\rangle$$

are orthonormal and so also form a basis. But in this new basis, the column vectors and matrices which represent states and operators will be different. For instance if $|v\rangle = |1\rangle - |3\rangle = |1'\rangle + |3'\rangle$ we write

$$|v\rangle \xrightarrow{\{1,2,3\}} \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} \equiv \mathbf{v} \quad |v\rangle \xrightarrow{\{1',2',3'\}} \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \equiv \mathbf{v}',$$

and

$$\begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & \sqrt{\frac{1}{2}} & \frac{1}{2} \\ \frac{i}{\sqrt{2}} & 0 & -\frac{i}{\sqrt{2}} \\ -\frac{1}{2} & \sqrt{\frac{1}{2}} & -\frac{1}{2} \end{pmatrix} \begin{pmatrix} v'_1 \\ v'_2 \\ v'_3 \end{pmatrix}.$$

The matrix is \mathbf{S} as defined above. We observe that its columns are just the representations of the new states $\{|1'\rangle, |2'\rangle, |3'\rangle\}$ in the old basis $\{|1\rangle, |2\rangle, |3\rangle\}$: $S_{23} = \langle 2|3'\rangle$ etc.

The **adjoint** of an operator is defined by $\langle u|\hat{Q}|v\rangle = \langle v|\hat{Q}^\dagger|u\rangle^*$. A **unitary** operator satisfies $\hat{Q}^\dagger\hat{Q} = \hat{I}$.

A **Hermitian operator** is its own adjoint: $\langle u|\hat{Q}|v\rangle = \langle v|\hat{Q}|u\rangle^*$. In practice that means that \hat{Q} can act backwards on $\langle u|$ or forward on $|v\rangle$, whichever is more convenient. In an orthonormal basis, \hat{Q} will be represented by a matrix which equals its adjoint (transposed complex-conjugate): $Q_{mn} = Q_{nm}^*$.

Hermitian operators have real eigenvalues and orthogonal eigenvectors which span the space. (If eigenvalues are repeated, all linear combinations of the corresponding eigenvectors are also eigenvectors—they form a **degenerate subspace**—but an orthogonal subset can always be chosen.) Thus the normalised eigenvectors of Hermitian operators are often chosen as a basis, typically labelled by the eigenvalues: $|\lambda_n\rangle$. Two Hermitian operators which commute will have a common set of eigenvectors with might be labelled by both eigenvalues: $|\mu_m, \lambda_n\rangle$.

In its own eigenbasis, a Hermitian operator will be diagonal, with the eigenvalues as the diagonal elements. Hence the process of finding the eigenvalues and eigenvectors is often called **diagonalisation**. The unitary matrix \mathbf{S} whose columns are the normalised eigenvectors can be used to transform other vectors and operators to this basis.

Since we can add and multiply operators and multiply them by scalars, we can form power series of an operator and hence define more general functions via their power-series expansion. The most important **function of an operator** is the exponential:

$$e^{\hat{Q}} \equiv \sum_{n=0}^{\infty} \frac{\hat{Q}^n}{n!}.$$

Since the corresponding power series for e^λ converges for all finite numbers, this is defined for all Hermitian operators, and its eigenvalues are e^{λ_i} . (In the eigenbasis of a Hermitian operator, any function of the operator is also represented by a diagonal matrix whose elements are the function of the eigenvalues.)

The exponential of a Hermitian operator is a unitary operator.

A.2.2 Functions as vectors

p th-order polynomials in the real variable x (with complex coefficients) form an $(p+1)$ -D vector space. For $p = 3$, one examples of a base in this space would be $\{1, x, x^2, x^3\}$, and the representation of $|v\rangle = v_0 + v_1x + v_2x^2 + v_3x^3$ in that basis is just the column vector $(v_0, v_1, v_2, v_3)^\top$. Another possible basis would be the first four Hermite polynomials

$$\{H_0(x) = 1, \quad H_1(x) = 2x, \quad H_2(x) = 4x^2 - 2, \quad H_3(x) = 8x^3 - 12x\}.$$

in which basis $|v\rangle \longrightarrow (v_0 + v_2/2, v_1/2 + 3v_3/4, v_2/4, v_3/8)^\top$.

More general sets of functions can also form vector spaces, but typically infinite-dimensional ones, with basis sets involving infinitely many functions. An example would be the set of all smooth functions $f(x)$ for which $\int_{-\infty}^{\infty} |f(x)|^2 dx$ is finite. We take this as the definition of $\langle f|f\rangle$, with

$$\langle f|g\rangle = \int_{-\infty}^{\infty} f^*(x)g(x) dx.$$

An example of an orthonormal basis for these functions is the set

$$\{|n\rangle = N_n H_n(x) e^{-x^2/2}\} \quad \text{for } n = 0, 1, 2, \dots$$

where N_n is a normalisation constant. Then any such function can be represented by an (infinitely long) list of numbers $|f\rangle \longrightarrow (\langle 0|f\rangle, \langle 1|f\rangle, \dots)^\top$.

If we shift our perspective, we can consider the vectors in a infinite-dimensional space as primary, and the functions as just another representation—the **position-space representation**, in which $|f\rangle \xrightarrow{x} f(x)$. With that viewpoint, the value $f(x_0)$ of the function at some value $x = x_0$ is like a component of the vector, and can be found by taking the inner product with a vector that picks out just that value, $|x_0\rangle$: $f(x_0) = \langle x_0|f\rangle$. If we don't want to specify a particular value, we have $f(x) = \langle x|f\rangle$ for the variable x .

With an eye on QM, we will often refer to vectors in a general vector space as **states**, which also helps to distinguish them from position and momentum vectors (of which more later).

Operators act on functions to turn one function into another; two simple examples are multiplication by x , and differentiation with respect to x . For their action on the abstract states, we use \hat{x} and \hat{D} , and we need¹

$$\langle x|\hat{x}|f\rangle = xf(x), \quad \langle x|\hat{D}|f\rangle = \frac{df}{dx}.$$

Since $\int f^* x g dx = (\int g^* x f dx)^*$, \hat{x} is Hermitian. So we see that $|x\rangle$ is an eigenstate of \hat{x} :

$$\hat{x}|x_0\rangle = x_0|x_0\rangle \quad \text{and} \quad \hat{x}|x\rangle = x|x\rangle.$$

¹Shankar uses \hat{X} for a dimensionless position variable, and $\hat{K} = -i\hat{D}$ as a dimensionless version of \hat{p} , but we stick with the QM notation.

These position eigenstates satisfy (where x and x' are both values of the position variable)

$$\langle x'|x\rangle = \delta(x' - x), \quad \int_{-\infty}^{\infty} |x\rangle\langle x| dx = \hat{I}, \quad \langle f|g\rangle = \langle f|\hat{I}|g\rangle = \int_{-\infty}^{\infty} f^*(x)g(x) dx.$$

Also since $\int f^* \frac{dg}{dx} dx = -(\int g^* \frac{df}{dx} dx)^*$, \hat{D} is anti-hermitian, so $i\hat{D}$ is Hermitian. In QM we work with $\hat{p} = -i\hbar\hat{D}$, and we can see that $[\hat{x}, \hat{p}] = i\hbar$. In the abstract vector space, this commutation relation defines \hat{p} . In position space, these operators are represented by²³

$$\hat{x} \xrightarrow{x} x, \quad \hat{p} \xrightarrow{x} -i\hbar \frac{d}{dx}.$$

We can define eigenstates of \hat{p} , $|p\rangle$, which have the following representation in position space:⁴

$$|p\rangle \xrightarrow{x} \langle x|p\rangle = \frac{1}{\sqrt{2\pi\hbar}} e^{ipx/\hbar},$$

and which satisfy

$$\langle p'|p\rangle = \delta(p' - p), \quad \int_{-\infty}^{\infty} |p\rangle\langle p| dp = \hat{I} \quad \langle f|g\rangle = \langle f|\hat{I}|g\rangle = \int_{-\infty}^{\infty} \tilde{f}^*(p)\tilde{g}(p) dp.$$

Up to factors of \hbar , $\langle p|f\rangle = \tilde{f}(p)$ is the Fourier transform of $f(x)$, and is an equally valid representation—in what we call **momentum space**—of the abstract state $|f\rangle$. The numerical equality of $\langle f|g\rangle$ calculated in the position and momentum representations is a reflection of Parseval's theorem.

We note that the states $|n\rangle$ defined above whose position-space representation is a Hermite polynomial times a Gaussian are actually eigenstates of $\hat{x}^2 - \hat{D}^2$, with eigenvalues $\lambda_n = 2n + 1$. In this basis \hat{x} and \hat{p} are represented by infinite-dimensional matrices, and it can be shown that for both, only matrix elements where m and n differ by ± 1 are non-zero.

We can extend the discussion to functions of three coordinates x , y and z . (Our notation for a point in space in a particular Cartesian coordinate system is $\mathbf{r} = x \mathbf{e}_x + y \mathbf{e}_y + z \mathbf{e}_z$.) There are operators associated with each coordinate, \hat{x} , \hat{y} and \hat{z} , which commute, and corresponding momentum operators \hat{p}_x , \hat{p}_y and \hat{p}_z , which also commute. Between the two sets the only non-vanishing commutators are $[\hat{x}, \hat{p}_x] = [\hat{y}, \hat{p}_y] = [\hat{z}, \hat{p}_z] = i\hbar$.

The position operator, $\hat{\mathbf{x}}$, is⁵ $\hat{x} \mathbf{e}_x + \hat{y} \mathbf{e}_y + \hat{z} \mathbf{e}_z$, and similarly $\hat{\mathbf{p}}$. Boldface-and-hat now indicates a **vector operator**, i.e. a triplet of operators. The eigenstate of position is $|x, y, z\rangle \equiv |\mathbf{r}\rangle$:

$$\begin{aligned} \hat{\mathbf{x}}|\mathbf{r}\rangle &= (\hat{x} \mathbf{e}_x + \hat{y} \mathbf{e}_y + \hat{z} \mathbf{e}_z)|\mathbf{r}\rangle = (x \mathbf{e}_x + y \mathbf{e}_y + z \mathbf{e}_z)|\mathbf{r}\rangle = \mathbf{r}|\mathbf{r}\rangle, \\ \hat{\mathbf{p}}|\mathbf{p}\rangle &= (\hat{p}_x \mathbf{e}_x + \hat{p}_y \mathbf{e}_y + \hat{p}_z \mathbf{e}_z)|\mathbf{p}\rangle = (p_x \mathbf{e}_x + p_y \mathbf{e}_y + p_z \mathbf{e}_z)|\mathbf{p}\rangle = \mathbf{p}|\mathbf{p}\rangle. \end{aligned}$$

²Actually as these are operators, it is more accurate to give the matrix elements $\hat{x} \xrightarrow{x} \langle x'|\hat{x}|x\rangle = x\delta(x' - x)$ and $\hat{p} \xrightarrow{x} \langle x'|\hat{p}|x\rangle = -i\hbar \frac{d\delta(x-x')}{dx'}$, which then are integrated over x' in any expression, but as this has just the net effect of setting x' to x we never bother with this more correct version.

³The lecturer of PHYS20672 used $|e_k\rangle \xrightarrow{x} e_k(x)$ for eigenkets of $\hat{K} = \hat{p}/\hbar$. The labelling of the kets used here is more common and is more symmetric between the x and p representations. I don't find it causes problems in practice. There is no common convention for the name of the function $\langle x|p\rangle$ which represents a plane wave.

⁴This can be derived from the differential equation obtained by combining $\langle x|\hat{p}|p\rangle = p\langle x|p\rangle$ together with the position-space representation $\langle x|\hat{p}|p\rangle = -i\hbar \frac{d}{dx} \langle x|p\rangle$.

⁵We do not use $\hat{\mathbf{r}}$ since that is reserved for the unit vector \mathbf{r}/r !

In position space, $\hat{\mathbf{x}} \xrightarrow{\mathbf{r}}$ and $\hat{\mathbf{p}} \xrightarrow{\mathbf{r}} -i\hbar\nabla$. Momentum eigenstates are

$$|\mathbf{p}\rangle \xrightarrow{\mathbf{r}} \langle \mathbf{r}|\mathbf{p}\rangle = \left(\frac{1}{2\pi\hbar}\right)^{3/2} e^{i\mathbf{p}\cdot\mathbf{r}/\hbar},$$

which is a plane wave travelling in the direction of \mathbf{p} . Also

$$\langle f|g\rangle = \int_{-\infty}^{\infty} f^*(\mathbf{r})g(\mathbf{r}) d^3r, \quad \langle \mathbf{r}|\mathbf{r}'\rangle = \delta(\mathbf{r} - \mathbf{r}') = \delta(x - x')\delta(y - y')\delta(z - z').$$

A.2.3 Commutators

Let \hat{A} , \hat{B} and \hat{C} be arbitrary operators in some space. Then the following relations are very useful:

$$\begin{aligned} \hat{A}\hat{B} &= \hat{B}\hat{A} + [\hat{A}, \hat{B}], \\ [\hat{A}, \hat{B}\hat{C}] &= [\hat{A}, \hat{B}]\hat{C} + \hat{B}[\hat{A}, \hat{C}], \\ [\hat{A}\hat{B}, \hat{C}] &= [\hat{A}, \hat{C}]\hat{B} + \hat{A}[\hat{B}, \hat{C}], \\ [\hat{A}, \hat{B}^n] &= n[\hat{A}, \hat{B}]\hat{B}^{n-1} \quad \text{provided } [\hat{A}, \hat{B}] \text{ commutes with } \hat{B}. \\ e^{\hat{A}}e^{\hat{B}} &= e^{\hat{A}+\hat{B}+[\hat{A},\hat{B}]/2} \quad \text{provided } [\hat{A}, \hat{B}] \text{ commutes with } \hat{A} \text{ and } \hat{B}. \end{aligned}$$

Let $Q(x)$ be a polynomial with derivative $R(x)$. Then

$$[\hat{p}_x, Q(\hat{x})] = -i\hbar R(\hat{x}) \xrightarrow{x} -i\hbar \frac{dQ(x)}{dx}.$$

Similarly if $V(\mathbf{r})$ is a function of position in 3-D,

$$[\hat{\mathbf{p}}, V(\hat{\mathbf{x}})] \xrightarrow{\mathbf{r}} -i\hbar \nabla V(\mathbf{r}).$$

A.3 Recap of 2nd year quantum mechanics

A.3.1 The wave function and Schrödinger's equation

The insight of de Broglie was that Bohr's atomic model with its quantised energy levels could be understood if the electron acted like a wave, with wave vector and frequency related to the momentum and energy by $\mathbf{p} = \hbar\mathbf{k}$ and $E = \hbar\omega$. Schrödinger found the appropriate wave equation to describe this wave, and quantum mechanics (also called wave mechanics in this formulation) was born. A brief recap of the theory follows.⁶ We will focus on the problem of a single spinless particle of mass M moving in an external potential $V(\mathbf{r})$.⁷ Our notation is $\mathbf{r} = x\mathbf{e}_x + y\mathbf{e}_y + z\mathbf{e}_z$.

The maximum knowledge that we can have about a quantum particle is the *wave function* $\Psi(\mathbf{r}, t)$, a complex function of position and time. The particle in general has no definite position or momentum; instead a range of possible values may be obtained if these or other aspects of the system such as the energy are measured. In particular the probability of finding a particle in an infinitesimal volume dV around position \mathbf{r} is

$$P(\mathbf{r}, dV) = |\Psi(\mathbf{r}, t)|^2 dV$$

and as a result the wave function must be normalised (it must be somewhere in space):⁸

$$\int |\Psi(\mathbf{r}, t)|^2 dV = 1$$

where the integral here and everywhere below is over all space unless otherwise specified. The wave function is a solution of the time-dependent Schrödinger equation (TDSE)

$$-\frac{\hbar^2}{2M}\nabla^2\Psi + V(\mathbf{r})\Psi = i\hbar\frac{\partial\Psi}{\partial t} \quad \text{or equivalently} \quad \hat{H}\Psi = i\hbar\frac{\partial\Psi}{\partial t},$$

where \hat{H} is the Hamiltonian or energy operator. As we often do when faced with a partial differential equation, we can look for solutions which are separable in space and time,

$$\Psi(\mathbf{r}, t) = \psi(\mathbf{r})T(t),$$

(this will not work if the potential depends on time, but for many interesting cases it does not). Substituting in the the Schrödinger equation and dividing by $\Psi(\mathbf{r}, t)$ gives

$$\frac{\hbar^2}{2M} \frac{\nabla^2\psi}{\psi} + V(\mathbf{r}) = \frac{i\hbar}{T} \frac{dT}{dt} = E;$$

using the usual argument that if two functions of different variables are equal for all values of those variables, they both must be equal to a constant which we denote as E . The time equation is trivial and independent of the potential, and gives

$$T(t) = e^{-iE/\hbar};$$

⁶These notes are designed to be read at the start of the current course, and so do not use vector space terminology. They also do not distinguish between operators and their position-space representations.

⁷ m is usually used for the mass, but that may be confused with the azimuthal quantum number.

⁸This means that the wave function is dimensioned, with units of inverse length to the power $D/2$, D being the number of spatial dimensions.

this is just a phase factor with a constant frequency satisfying de Broglie's relation $\hbar\omega = E$. Allowed values of E though are not determined at this stage.

The spatial equation is the time-independent Schrödinger equation (TISE)

$$-\frac{\hbar^2}{2M}\nabla^2\psi + V(\mathbf{r})\psi = E\psi \quad \text{or equivalently} \quad \hat{H}\psi = E\psi.$$

Though the *form* of the equation is universal, the solutions depend on the potential. In solving the TISE, at a minimum we need to ensure the solutions are finite everywhere—we don't want the probability of finding the particle in some region to be infinite. If the potential is constant, $V(\mathbf{r}) = V_0$, the solutions are plane waves characterised by a wave vector \mathbf{k}

$$\psi(\mathbf{r}) \propto e^{i\mathbf{k}\cdot\mathbf{r}}, \quad \text{where} \quad E = V_0 + \frac{\hbar^2|\mathbf{k}|^2}{2M}.$$

This is very suggestive, it looks like E equals the potential energy plus the kinetic energy, if the momentum is given by de Broglie's relation $\mathbf{p} = \hbar\mathbf{k}$. And the only allowable solutions are those with $E > V_0$, which makes sense classically. Indeed we identify E with the energy of the system in all cases.

If the potential varies with position, we have two possible types of solution. There are those for which the energy is greater than the potential over most of space, in which case the solutions are not localised and the particle may be found anywhere; these are called scattering solutions. They will look like plane waves in regions where the potential is constant. The other type of solution may exist if the potential has a well, a region of space where the potential is lower than its value at large distances. These solutions are called *bound states*; they have energies that are not large enough for the particle to climb out of the well, and the wave function is concentrated within the well—the probability of finding the particle large distances away from the well vanishes.

Elementary QM is almost exclusively concerned with bound states. The extra requirement that the wave function must vanish at infinity means that, for arbitrary values of E , no solution to the TISE exists; only for certain discrete values can acceptable solutions be found. The energy is *quantised*; there is a single lowest (ground-state) energy state and a number of higher-energy or excited states. If the potential grows without bound at large distances (an infinite well) there will be infinitely-many states, but more realistically the potential will level off eventually (a finite well); there will be a maximum energy of the bound states, and scattering states will exist as well. (In this case it is usual to set $V(|\mathbf{r}| \rightarrow \infty) = 0$, with $V < 0$ in the well.) For simplicity, we often concentrate on infinite wells.

In this case we have infinitely many states with energies E_n , $n = 1, 2, \dots$ and solutions satisfying

$$\Psi_n(\mathbf{r}, t) = \psi_n(\mathbf{r})e^{-iE_nt/\hbar} \quad \text{where} \quad -\frac{\hbar^2}{2M}\nabla^2\psi_n(\mathbf{r}) + V(\mathbf{r})\psi_n(\mathbf{r}) = E_n\psi_n(\mathbf{r}).$$

Note that while the *form* of the time-dependence is the same for all solutions, the frequency $\omega_n = E_n/\hbar$ is different for each. The TISE is an *eigenfunction equation*, with the energies E_n being the eigenvalues. The states are taken to be normalised, and it can be shown that they are also *orthogonal*:⁹

$$\int \psi_m^*(\mathbf{r})\psi_n(\mathbf{r}) dV = \delta_{mn}.$$

⁹In more than one dimension there may be distinguishable states with the same energy which are termed degenerate; in that case any superposition of degenerate solutions is also a solution but we are always able to choose an orthogonal set.

The general solution of the TDSE is a superposition of these states:

$$\Psi(\mathbf{r}, t) = \sum_{n=1}^{\infty} c_n \psi_n(\mathbf{r}) e^{-iE_n t/\hbar}, \quad \text{where} \quad \sum_{n=1}^{\infty} |c_n|^2 = 1,$$

the restriction on the sum of the magnitudes of the coefficients being the normalisation condition. In any specific case a particular wave function is determined by the coefficients, which (since the functions $\psi_n(\mathbf{r})$ are orthogonal) can be obtained from the initial state of the system by

$$c_n = \int \psi_n^*(\mathbf{r}) \Psi(\mathbf{r}, 0) dV.$$

Thus in a very simple case, if the initial state is $\Psi(\mathbf{r}, 0) = \sqrt{\frac{1}{3}}\psi_1(\mathbf{r}) + \sqrt{\frac{2}{3}}\psi_3(\mathbf{r})$, the subsequent state would be

$$\Psi(\mathbf{r}, t) = \sqrt{\frac{1}{3}}\psi_1(\mathbf{r})e^{-iE_1 t/\hbar} + \sqrt{\frac{2}{3}}\psi_3(\mathbf{r})e^{-iE_3 t/\hbar}.$$

In practical situations (such as atomic physics) the quantisation of the energy levels shows up primarily in the existence of discrete excitation or de-excitation energies—spectral lines.

A.3.2 Measurement

It is a postulate of QM that for every “observable” or measurable property of a quantum system there is a corresponding Hermitian operator. Denoting such an operator \hat{Q} , “Hermitian” means that if $\psi(\mathbf{r})$ and $\phi(\mathbf{r})$ are (complex) normalisable functions of position,

$$\int \psi^*(\mathbf{r}) \hat{Q} \phi(\mathbf{r}) dV = \left(\int \phi^*(\mathbf{r}) \hat{Q} \psi(\mathbf{r}) dV \right)^*.$$

Similarly to the case with Hermitian matrices, Hermitian operators have real eigenvalues and orthogonal eigenfunctions which form a complete set, that is any other normalisable function can be expressed as a superposition of them.

At the start we noted that a measurement of position (for which the operator is just the position vector, denoted $\hat{\mathbf{x}}$ but simply equal to \mathbf{r} in an integral) will give an answer which is not known in advance, with the probability of the possible results being governed by the modulus-squared of the wave function. So, in general, is the case with other possible measurements. If some observable is associated with an operator \hat{Q} , the average value of a measurement of that observable for a particle with wave function $\Psi(\mathbf{r}, t)$ is given by¹⁰

$$\langle \hat{Q} \rangle = \int \Psi^*(\mathbf{r}, t) \hat{Q} \Psi(\mathbf{r}, t) dV$$

Since, as is well known, measurement in quantum mechanics changes the system, we can only talk about average results if we make repeated measurements on identically-prepared copies of the system. So this average is an *ensemble average*, also called the *expectation value*.

As well as an average value, these measurements will have a spread or *uncertainty* ΔQ which is given through the usual definition of the standard deviation by

$$\Delta Q^2 = \langle \hat{Q}^2 \rangle - \langle \hat{Q} \rangle^2.$$

¹⁰ $\langle \hat{Q} \rangle$ may be written $\langle \Psi | \hat{Q} | \Psi \rangle$.

If many measurements are done (each on a fresh copy), but a single result q_i is obtained every time, then Ψ is an eigenfunction of \hat{Q} with eigenvalue q_i : $\Psi(\mathbf{r}, t) \propto \phi_i(\mathbf{r})$ where $\hat{Q}\phi_i(\mathbf{r}) = q_i\phi_i(\mathbf{r})$. This ensures that $\Delta Q = 0$.

After a measurement is done, and a particular result q is obtained, subsequent measurements carried out quickly enough *on the same system* will again yield the same value q ; this is reproducibility. It is only possible if, after a measurement is made, the system is no longer in its original state but is in an eigenstate of \hat{Q} with q as the corresponding eigenvalue. And so the only possible results of a measurement of \hat{Q} are its eigenvalues.

Any position is an eigenvalue of the position operator, likewise with momentum, and likewise (above some minimum) for energy in scattering states. Some operators though, like the energy for bound states, have discrete values.

There is a common misconception among students that making a measurement on a system is somehow related to operating with the corresponding operator on the wave function. This is not true! In general an operator turns a normalised wave function into another, quite different, unnormalised function that doesn't even have the dimensions of a wave function.

What we can say is that the probability of getting a particular eigenvalue of \hat{Q} when we make a measurement of the corresponding observable is given by¹¹

$$P(q_i) = \left| \int \phi_i^*(\mathbf{r})\Psi(\mathbf{r}, t) dV \right|^2.$$

And immediately after the measurement, the system is in state $\phi_i^*(\mathbf{r})$.

Some frequently-met operators are momentum, $\hat{\mathbf{p}} = -i\hbar\nabla$, kinetic energy $\hat{\mathbf{p}} \cdot \hat{\mathbf{p}}/2M$, energy \hat{H} as given above, and angular momentum $\hat{\mathbf{L}} = \mathbf{r} \times \hat{\mathbf{p}}$.

For energy, once we have found the eigenfunctions and eigenvalues for the relevant particular potential, as we have already seen any wave function can be expanded

$$\Psi(\mathbf{r}, t) = \sum_{n=1}^{\infty} c_n \psi_n(\mathbf{r}) e^{-iE_n t/\hbar}, \quad \text{where} \quad \sum_{n=1}^{\infty} |c_n|^2 = 1 \quad \text{and} \quad c_n = \int \psi_n^*(\mathbf{r})\Psi(\mathbf{r}, 0) dV.$$

We now see that $|c_n|^2$ is the probability that, if we measure the energy, we will get the value E_n . If only one c_n is non-zero, $\Psi(\mathbf{r}, t)$ is an energy eigenstate. The following terms are synonymous: separable solution of the TDSE; solution of the TISE; eigenstate of the Hamiltonian; state of definite energy; stationary state. The last requires some explanation: for $\Psi(\mathbf{r}, t) = \psi_n(\mathbf{r})e^{-iE_n t/\hbar}$, the expectation value $\langle \hat{Q} \rangle$ is independent of time for any operator \hat{Q} that does not itself depend on time.

For the case we considered above,

$$\Psi(\mathbf{r}, t) = \sqrt{\frac{1}{3}}\psi_1(\mathbf{r})e^{-iE_1 t/\hbar} + \sqrt{\frac{2}{3}}\psi_3(\mathbf{r})e^{-iE_3 t/\hbar},$$

a measurement of energy would yield the result E_1 one third of the time and E_3 two-thirds of the time.

Given any wave function except a plane wave, the uncertainty in any component of the momentum Δp_i will be non-zero. This is related to the Fourier transform: to make up a wave

¹¹if the eigenvalues are continuous, this should be interpreted as a probability density instead

packet of finite spatial extent Δx requires a superposition of waves of a spread of wavelengths. (Similarly, a signal of finite duration consists of a spread of frequencies). The bandwidth theorem relates the two; the narrower the wave packet, the wider the spread of wave numbers. In QM language, this is Heisenberg's uncertainty relation:

$$\Delta x \Delta p_x \geq \frac{1}{2} \hbar.$$

If we make a measurement that narrows down the region in which the particle exists, a subsequent measurement of the momentum is much more unpredictable, and vice versa. It is obvious therefore that we cannot find a state of well-defined position and momentum; position and momentum can't have common eigenfunctions. In fact if operators have common eigenfunctions, they must commute, and \hat{x} and \hat{p}_x do not:

$$(\hat{x}\hat{p}_x - \hat{p}_x\hat{x})\psi(x) = i\hbar\psi(x) \quad \Rightarrow [\hat{x}, \hat{p}_x] = i\hbar.$$

Heisenberg, with Born and Jordan, derived his formulation of QM, called matrix mechanics, starting from this relation; Schrödinger subsequently showed that it was in fact equivalent to his wave mechanics.

Other operators which do not commute are the three components of angular momentum, so that in QM we can only know one of them exactly (usually but not necessarily taken to be L_z). In fact the commutation relations are rather interesting:

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

However all three separately commute with $\hat{\mathbf{L}}^2 = \hat{L}_x^2 + \hat{L}_y^2 + \hat{L}_z^2$, so we *can* know the magnitude of the angular momentum and one of its components at the same time. For more on angular momentum operators see section [A.5](#). For more on commutators see section [A.2.3](#).

[continued overleaf]

A.3.3 Bound states

We now turn to the problem of finding the energy levels for a variety of potentials.

Infinite square well

An unrealistic but useful model potential is one which has a constant value, taken to be zero, over some region, but which is infinite elsewhere. Where a potential is discontinuous like this we have to solve the TISE in each region separately, and we match them by requiring that ψ is continuous at the boundaries (the probability density has to be unambiguous). Since the TISE is a second-order differential equation we usually need the derivatives of ψ to match as well. However in the unique case of an infinite potential step there is no solution possible in the classically-forbidden region, so in this case the condition is just that ψ vanishes at the boundaries.

In 1D, with $V(x) = 0$ for $0 < x < a$, the general solution is $Ae^{ikx} + Be^{-ikx}$ or equivalently $C \sin(kx) + D \cos(kx)$, with $\hbar k = \sqrt{2ME}$, but requiring $\psi(0) = \psi(a) = 0$ restricts the (normalised) solution to $\sqrt{\frac{2}{a}} \sin(k_n x)$, with $k_n = n\pi/a$ and $E_n = \hbar^2 \pi^2 / (2Ma^2)$. The quantisation of the energies is completely analogous to the “quantisation” of the vibrational frequencies of a string fixed at both ends.

If we’d chosen the well to stretch from $-a/2$ to $a/2$ the values k_n and hence the energies would (of course) be unchanged but the wave functions would alternate cosines and sines. (I only mention this because the finite square well is much neater with that choice, see below.)

In 2D and 3D, for a rectangular or cuboidal well, the spatial wave function is separable. The solution in 3D, with sides L_x etc, is just

$$\psi_{n_x n_y n_z} = \sqrt{\frac{8}{L_x L_y L_z}} \sin\left(\frac{n_x \pi x}{L_x}\right) \sin\left(\frac{n_y \pi y}{L_y}\right) \sin\left(\frac{n_z \pi z}{L_z}\right)$$

and $E_{n_x n_y n_z} = \frac{\hbar^2 \pi^2}{2M} \left(\frac{n_x^2}{L_x^2} + \frac{n_y^2}{L_y^2} + \frac{n_z^2}{L_z^2} \right).$

A cuboidal well (albeit better modelled as a finite than infinite well) is also called a quantum dot. If all the lengths are the same, some energy levels will correspond to more than one different wave functions—denoting the quantum numbers of the state by (n_x, n_y, n_z) , then the first excited energy level is in fact three states, with quantum numbers $(2, 1, 1)$, $(1, 2, 1)$ and $(1, 1, 2)$. These are said to be three-fold degenerate.

Quantum mechanics allows us to make systems which act as if they have fewer than 3 dimensions in a way that classical physics would not. If we have a situation where one length is much shorter than the others, say $L_z \ll L_x, L_y$, then all the low-lying states will have $n_z = 1$. If we probe such a system with energies less than the energy needed to excite the first state with $n_z = 2$, it will be effectively two-dimensional; we say the third degree of freedom is *frozen out*. Modern so-called 2D systems are often of this nature. Similarly if $L_x \gg L_y, L_z$ we have an effectively 1D system or quantum wire.

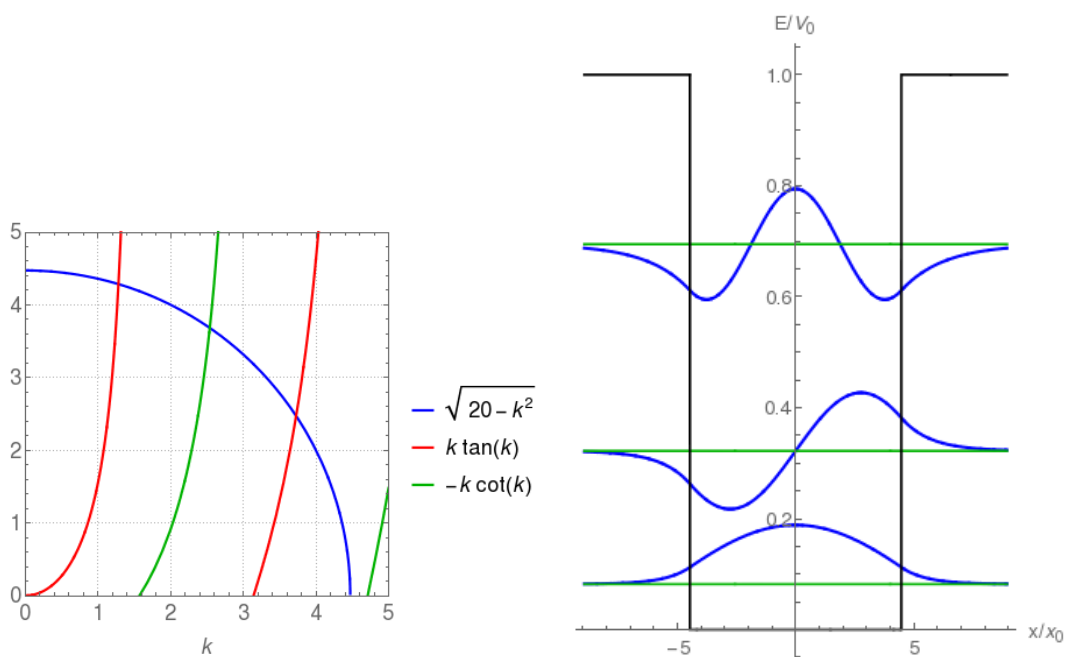
For a 2D circular well, the radial solutions are Bessel functions with argument kr . Again k is chosen so that the wave function vanishes at the boundary $r = R$, ie kR is a zero of the Bessel function. For a 3-D spherical well we get spherical Bessel functions. More on this below.

Finite square well

A somewhat more realistic well has one region of constant low potential with a constant higher potential elsewhere. We can either take these to be $-V_0$ and 0, or 0 and $+V_0$; the former allows the potential at infinity to be zero but the latter is closer to the infinite well so we will go with that.

In 1D, with $V(x) = 0$ for $-a/2 < x < a/2$ and $V(x) = V_0 > 0$ elsewhere, the solutions are again $Ae^{ikx} + Be^{-ikx}$ inside the well. But now there solutions to the equation outside the well as well: for $x > a/2$, $\psi = Ce^{-\kappa x}$ and for $x < -a/2$, $\psi = De^{\kappa x}$, with $\hbar\kappa(k) = \sqrt{2M(V_0 - E)}$. (These were not allowable for a constant potential over all space as they blow up, but they are OK here as they hold only in restricted regions.) Matching both ψ and ψ' at $x = \pm a/2$ again restricts the values of k : either ψ is symmetric (cosine inside the well) and k has to satisfy $\kappa(k) = k \tan(ka/2)$, or antisymmetric (sine inside the well) with $\kappa(k) = -k \cot(ka/2)$. The graphical solution for the allowed values of k , and the first few solutions, are shown below. A 1D well, no matter how shallow, always has one bound state.

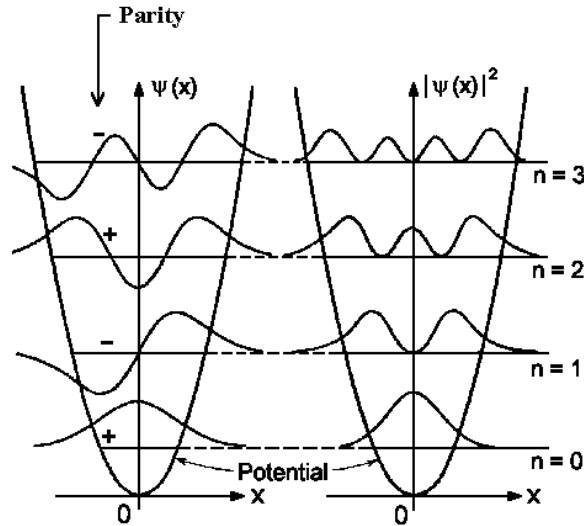
The most important part of this analysis (apart again from the quantisation of the energy) is that there is a finite probability of finding the particle in the classically-forbidden regions. The higher the energy, the higher this probability is. On the other hand we note that as V_0 increases, the tails of the low-lying states get smaller and shorter, until we reach the point where simply setting $\psi = 0$ at the boundary, as we did in the infinite well, is a good approximation.



There is a standard piece of iconography used in the picture on the right. We draw a well and the energy levels. Then we take the lines representing the energy levels as the x -axis for a graph of the corresponding wave function. These therefore have off-set y -axes, and the two y scales, for the potential and for the wave functions, are not related.

Harmonic Oscillator

A particle in a 1D harmonic potential $V(x) = \frac{1}{2}k_s x^2$, where k_s plays the role of a spring constant, is a good model for many practical situations. (Essentially all wells are harmonic at their base!). The classical oscillation frequency is $\omega = \sqrt{k_s/M}$ and the potential is often written $V(x) = \frac{1}{2}M\omega^2 x^2$. It is useful to note that $x_0 = \sqrt{\hbar/M\omega}$ has dimensions of length, and of course $\hbar\omega$ has dimensions of energy. For any given energy there will be a classically allowed region for which $E > V(x)$, between $x = \pm 2x_0(E/\hbar\omega)$, and the rest is classically forbidden in which the wave function falls to zero as $x \rightarrow \pm\infty$. We expect the solutions to look somewhat like the finite well, oscillatory (though not specifically sinusoidal) within the well, with decaying tails outside, as indeed they do; again the requirement of finite, normalisable wave functions restricts the possible energies, this time to $(n + \frac{1}{2})\hbar\omega$, for $n = 0, 1, 2, \dots$. (Hence we often use E_0 rather than E_1 for the ground state.) The ground state is just a Gaussian, $\phi_0(x) \propto e^{-x^2/2x_0^2}$, and higher states multiply this by Hermite polynomials; see section A.4 for details of the solution.



A potential which is harmonic in three directions, $\frac{1}{2}M(\omega_x^2 x^2 + \omega_y^2 y^2 + \omega_z^2 z^2)$ will have solutions which (like the square well) are just products of the 1D states, and energies which are the sum of the corresponding 1D energies. If all the “spring constants” are equal there will be a high degree of degeneracy among the excited states, for instance $(7/2)\hbar\omega$ is the energy of the states with $(n_x, n_y, n_z) = (2, 0, 0), (0, 2, 0), (0, 0, 2), (1, 1, 0), (1, 0, 1)$ and $(0, 1, 1)$.

In the symmetric case, though, the potential can be written $V(\mathbf{r}) = V(r) = \frac{1}{2}m\omega^2 r^2$, and we can work instead in spherical polars: see the next section. The energies, degeneracies, and non-degenerate wave functions such as the ground state must turn out the same in both coordinate systems, but the degenerate ones need only be linear combinations of one another.

Linear potential

In a region of space where the potential is linear, $V(x) \propto x$, the solutions are Airy functions (see section A.9). To form a well, this potential would have to have a hard wall somewhere, say $V = \infty$ for $x < 0$, or it might be part of a V-shaped potential $V(x) \propto |x|$. The energy levels have to be found numerically.

A.3.4 Circular and spherical symmetry

In 2D we will use $r^2 = x^2 + y^2$ and $y/x = \tan \phi$; in 3D, $r^2 = x^2 + y^2 + z^2$, $y/x = \tan \phi$ and $z/r = \cos \theta$. The double meaning of r will not cause problems so long as we don't consider 3D problems with cylindrical geometry.

In a 2D problem with a symmetric potential $V(\mathbf{r}) = V(r)$, we can write the wave function in a form which is separable in plane polar coordinates: $\psi(x, y) = R(r)\Phi(\phi)$. Skipping some detail we find that the angular dependence is just of the form $e^{im\phi}$ where, in order for the wave function to be single valued, we need m to be an integer (not to be confused with the mass!). Then the radial equation is

$$-\frac{\hbar^2}{2M} \left(\frac{d^2 R}{dr^2} + \frac{1}{r} \frac{dR}{dr} \right) + \frac{\hbar^2 m^2}{2Mr^2} R + V(r)R = ER.$$

In a 3D spherically-symmetric potential $V(\mathbf{r}) = V(r)$, we can write the wave function in a form which is separable in spherical polar coordinates: $\psi(\mathbf{r}) = R(r)Y(\theta, \phi)$. Then

$$\begin{aligned} & -\frac{\hbar^2}{2M} \nabla^2 \psi(\mathbf{r}) + V(r)\psi(\mathbf{r}) = E\psi(\mathbf{r}) \\ \Rightarrow & -\hbar^2 \left(\frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial Y}{\partial \theta} \right) + \frac{1}{\sin^2 \theta} \frac{\partial^2 Y}{\partial \phi^2} \right) = \hbar^2 l(l+1)Y \quad (\text{A.1}) \\ \text{and} & -\frac{\hbar^2}{2Mr} \frac{d^2(rR)}{dr^2} + \frac{\hbar^2 l(l+1)}{2Mr^2} R + V(r)R = ER \end{aligned}$$

where $l(l+1)$ is the constant of separation. The radial equation depends on the potential, and so differs from problem to problem. However the angular equation is *universal*: its solutions do not depend on the potential. It is further separable into an equation in θ and one in ϕ with separation constant m^2 ; the latter is the same as in 2D with solution $e^{im\phi}$ for integer m . Finally the allowable solutions of the θ equation are restricted to those which are finite for all θ , which is only possible if l is an integer greater than or equal to $|m|$; the solutions are associated Legendre polynomials $P_l^m(\cos \theta)$. The combined angular solutions are called spherical harmonics $Y_l^m(\theta, \phi)$:

$$\begin{aligned} Y_0^0(\theta, \phi) &= \sqrt{\frac{1}{4\pi}} & Y_1^{\pm 1}(\theta, \phi) &= \mp \sqrt{\frac{3}{8\pi}} \sin \theta e^{\pm i\phi} \\ Y_1^0(\theta, \phi) &= \sqrt{\frac{3}{4\pi}} \cos \theta & Y_2^{\pm 2}(\theta, \phi) &= \sqrt{\frac{15}{32\pi}} \sin^2 \theta e^{\pm 2i\phi} \\ Y_2^{\pm 1}(\theta, \phi) &= \mp \sqrt{\frac{15}{8\pi}} \sin \theta \cos \theta e^{\pm i\phi} & Y_2^0(\theta, \phi) &= \sqrt{\frac{5}{16\pi}} (3 \cos^2 \theta - 1) \end{aligned}$$

These are normalised and orthogonal:

$$\int (Y_{l'}^{m'})^* Y_l^m d\Omega = \delta_{l'l} \delta_{m'm} \quad \text{where} \quad d\Omega = \sin \theta d\theta d\phi$$

The physical significance of the quantum numbers l and m is not clear from this approach. However if we look at the radial equation, we see that the potential has been effectively modified by an extra term $\hbar^2 l(l+1)/(2Mr^2)$. Recalling classical mechanics, this is reminiscent of the

centrifugal potential which enters the equation for the radial motion of an orbiting particle, where $\hbar^2 l(l+1)$ is taking the place of the (conserved) square of the angular momentum. Now we have already defined the angular momentum operator $\hat{\mathbf{L}} = -i\hbar \mathbf{r} \times \nabla$. If we cast this in spherical polar coordinates, we find

$$\hat{L}_z = -i\hbar \frac{\partial}{\partial \phi}, \quad \hat{\mathbf{L}}^2 = -\hbar^2 \left(\frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial}{\partial \theta} \right) + \frac{1}{\sin^2 \theta} \frac{\partial^2}{\partial \phi^2} \right).$$

But $\hat{\mathbf{L}}^2$ is just the angular part of $-\hbar^2 r^2 \nabla^2$. So its eigenfunctions, the spherical harmonics, are in fact states of definite squared angular momentum, which is quantised and takes the values $\hbar^2 l(l+1)$ for integer l . They are also states of definite L_z , which takes the values $\hbar m$. It is good to see that as $|m| \leq l$, $L_z^2 \leq \mathbf{L}^2$, with equality only if both are zero.

We have already noted that we can simultaneously know the total angular momentum and its z component, since the corresponding operators commute. (This should be obvious as $\hat{\mathbf{L}}^2$ is independent of ϕ .) Since $\hat{\mathbf{L}}^2$ commutes with the Hamiltonian for a system with a symmetric potential, states may be fully classified by their energy, the square of their angular momentum, and the z -component of angular momentum.

For completeness we note that the formulae for \hat{L}_x and \hat{L}_y are rather lengthy (the z -axis has a special role in spherical polars), but they can be expressed more succinctly as

$$\hat{L}_x = \frac{1}{2}(\hat{L}_+ + \hat{L}_-) \quad \text{and} \quad \hat{L}_y = \frac{1}{2i}(\hat{L}_+ - \hat{L}_-),$$

where

$$\hat{L}_+ = \hbar e^{i\phi} \left(\frac{\partial}{\partial \theta} + i \cot \theta \frac{\partial}{\partial \phi} \right), \quad \hat{L}_- = \hat{L}_+^\dagger = \hbar e^{-i\phi} \left(-\frac{\partial}{\partial \theta} + i \cot \theta \frac{\partial}{\partial \phi} \right).$$

The operators \hat{L}_\pm both commute with $\hat{\mathbf{L}}^2$, though it is somewhat easier to show it for \hat{L}_x and \hat{L}_y using Cartesian coordinates.

There is more on angular momentum in section [A.5](#).

Square wells again

For a 2D circular infinite square well with $V = 0$ for $r < R$ the solutions are Bessel functions as mentioned above; actually there is a different Bessel function for each value of m so

$$\psi(r, \phi) \propto J_m(kr) e^{im\phi}.$$

The ground state is circularly symmetric, $m = 0$, and there is a tower of $m = 0$ states with k , and hence E , determined by $J_0(kR) = 0$. However the (doubly-degenerate) first excited states have $m = \pm 1$ and k fixed by requiring kR to be the first zero of J_1 .

Similarly for a 3D spherical well, the solutions are spherical Bessel functions,

$$\psi(r, \phi) \propto j_l(kr) Y_l^m(\theta, \phi).$$

This time the energy levels depend on l but not m (as expected since the z -axis is arbitrary). Similar remarks as in the previous paragraph can be made about the ordering of levels, first the $l = 0$ ground state, then an $l = 1$ excited states, etc.

Interestingly for 3D the radial part of $\nabla^2 R$ simplifies if we set $R(r) = u(r)/r$, becoming u''/r . Even in the absence of an external potential there is the centrifugal potential to stop the

equation for the radial wave function being trivial, but for the special case of $l = 0$ we simply have $u(r) = \sin(kr)$. And indeed $j_0(z) = \sin z/z$. So the $l = 0$ states of a spherical square well have energies $E_n^{l=0} = \hbar^2 n^2 \pi^2 / 2MR^2$,

Finite circular and spherical square wells can also be solved numerically. A 2D circular well, like a 1D well, no matter how shallow, always has one bound state, but in 3D a sufficiently-shallow spherical well will have no bound states. (The $l = 0$ states, having the form within the well of $\sin(kr)/r$ and outside of $e^{-\kappa r}/r$, have the same energies as the antisymmetric states of a 1D well of the same diameter; as we saw when we considered 1D finite wells above, such a state may not exist if the well is too shallow.)

Symmetric harmonic oscillator

If we solve the symmetric harmonic oscillator $V(r) = \frac{1}{2}M\omega^2 r^2$ in 2 or 3D, we find the lowest energy solution to be $e^{-r^2/2x_0^2}$ with energy $\hbar\omega$ or $\frac{3}{2}\hbar\omega$ respectively. Higher-energy solutions take the form of the same Gaussian multiplied by polynomials in r and the appropriate angular form $e^{im\phi}$ or $Y_l^m(\theta, \phi)$. The first excited states in the Cartesian basis are just x or y (or z) times the Gaussian; the relation between the two forms comes from noting that $x \pm iy = re^{\pm i\phi}$ (2D) or, up to a constant, $rY_1^{\pm 1}(\theta, \phi)$ in 3D, while $z \propto rY_1^0$. Though it becomes increasingly tiresome to verify, the degeneracies and energies do all match as they must. The three Cartesian-basis states with $E = (5/2)\hbar\omega$ all have $l = 1$. From the 6 states with $E = (7/2)\hbar\omega$ discussed above, we now find that from linear combinations we can make one state with $l = 0$ and five with $l = 2$ (and $m = 2, 1, 0, -1, -2$).

Hydrogen atom

The details of finding the radial solutions for the Coulomb potential are extremely lengthy, but are covered in all the course textbooks. The results are quite simple though, and are given in section [A.6](#).

A.3.5 Tunnelling

As we saw with the finite square well, there is a finite probability of finding a particle in a classically-forbidden region where $V > E$. This means that a particle incident on a barrier of finite height and thickness has a finite chance of tunnelling through and continuing on the other side. (See picture below.) To find how likely this is, we need to solve the Schrödinger equation and look at the probability density beyond the barrier. For an individual particle we would need to set up a localised incoming wave packet with, necessarily, a spread of momenta and energies, which would be a complicated, time-dependent problem. Instead we consider the steady-state problem of a beam of mono-energetic particles incident from the left on the barrier; there will also be a reflected beam travelling in the opposite direction to the left of the barrier and a transmitted beam beyond the barrier. (This is an example of a scattering problem). Without loss of generality we can take the potential to be zero on the left, and the wave function for the incident beam, with momentum $\hbar k$ and energy $E = \hbar\omega = \hbar^2 k^2 / (2M)$, is $Ie^{i(kx - \omega t)}$, the probability density is $|I|^2$ and the flux of particles in the beam is $(p/M)|I|^2$ particles per second. We are looking for a single solution to the TISE, albeit one with different functional

forms in different regions, so the energy is the same everywhere and the time-dependence $e^{-i\omega t}$ is likewise universal, so we will drop it.

Beyond the barrier the potential need not be the same as on the left, so we will let it be V_f ; the momentum of particles that tunnel through will be $p' = \sqrt{2M(E - V_f)}$; the wave function is $T e^{ik'x}$ and the transmitted flux $(p'/M)|T|^2$. Finally the reflected wave is $R e^{-ikx}$ and the reflected flux $(p/M)|R|^2$. The transmission and reflection probabilities are

$$t = \frac{p'}{p} \left| \frac{T}{I} \right|^2, \quad r = \left| \frac{R}{I} \right|^2; \quad t + r = 1.$$

We will focus on a single square barrier of height $V_b > E$ between $x = 0$ and $x = L$. In that region the wave function is $A e^{-\kappa x} + B e^{\kappa x}$ where $\hbar\kappa = \sqrt{2M(V_b - E)}$, that is there will be exponentially decaying and exponentially growing parts. As in the finite square well we require the wave function and its derivative to be continuous at the boundaries, which gives four equations to determine A , B , T , and R . The rather unenlightening answer is

$$t = \frac{4\kappa^2 k k'}{(\kappa^2 - k k')^2 \sinh^2(\kappa L) + \kappa^2 (k' + k)^2 \cosh^2(\kappa L)}$$

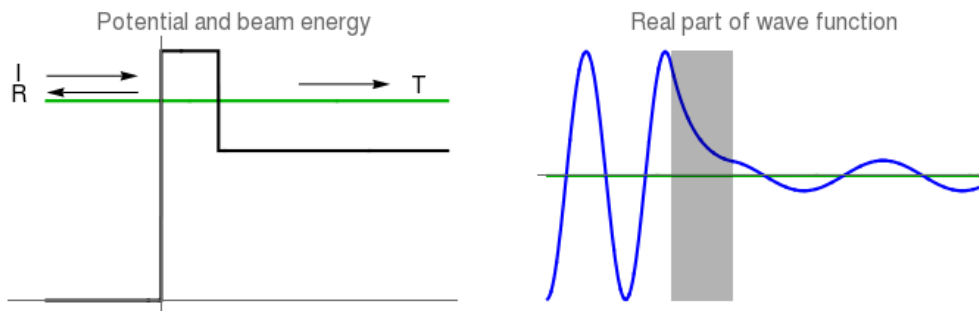
But if $\kappa L \gg 1$, $\cosh(\kappa L)$ and $\sinh(\kappa L)$ both tend to $\frac{1}{2}e^{\kappa L}$ and the expression simplifies to

$$t = \frac{16\kappa^2 k k'}{(\kappa^2 + k^2)(\kappa^2 + k'^2)} e^{-2\kappa L} = \frac{4\hbar^4 \kappa^2 k k'}{M^2 V_b (V_b - V_f)} e^{-2\kappa L}.$$

This is equivalent to ignoring the exponentially-growing wave within the barrier, so that the ratio of the wave function at either end of the barrier is just $e^{-\kappa L}$.

Recall that k , k' and κ are all functions of energy. By far the most rapidly-varying part of the expression is the exponential, and a rough estimate of the dependence of the tunnelling probability on energy for a high, wide barrier is given simply by $t \sim \exp(-2L\sqrt{2M(V_b - E)}/\hbar)$.

In the figure below we can note the approximately exponential decay within the barrier, the continuity and smoothness of the wave function, and the much smaller amplitude and (since V_f is higher than the initial potential) the longer wavelength on the right.



It may be noted that there will also be some reflection even if $E > V_b$, or indeed from a well rather than a barrier.

A.4 Series Solution of Hermite's equation and the Harmonic Oscillator

Shankar 7.3

Griffiths 2.3.2

We consider a particle moving in a 1D quadratic potential $V(x) = \frac{1}{2}m\omega^2x^2$, like a mass on a spring. The Hamiltonian operator is

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

We will work in rescaled dimensionless coordinates, defining the length scale $x_0 = \sqrt{\hbar/m\omega}$, so $\hat{x} \rightarrow x_0y$ and $\hat{p} \rightarrow (-i\hbar/x_0)d/dy$. The energy scale is $\frac{1}{2}m\omega^2x_0^2 = \frac{1}{2}\hbar\omega$. We are looking for wave functions $\phi(y)$, of energy $E = \frac{1}{2}\hbar\omega\mathcal{E}$, which satisfy

$$-\frac{d^2\phi}{dy^2} + y^2\phi = \mathcal{E}\phi. \quad (\text{A.3})$$

If we write $\phi(y) \equiv f(y)e^{-y^2/2}$, this can be rewritten as

$$\frac{d^2f}{dy^2} - 2y\frac{df}{dy} + (\mathcal{E} - 1)f = 0. \quad (\text{A.4})$$

This is Hermite's differential equation. If we look for a series solution of the form $f(y) = \sum_{j=0}^{\infty} c_j y^j$, we get

$$\begin{aligned} \sum_{j=2}^{\infty} j(j-1)c_j y^{j-2} - 2 \sum_{j=1}^{\infty} j c_j y^j + (\mathcal{E} - 1) \sum_{j=0}^{\infty} c_j y^j &= 0 \\ \Rightarrow \sum_{j=0}^{\infty} \left((j+1)(j+2)c_{j+2} + (\mathcal{E} - 1 - 2j)c_j \right) y^j &= 0 \end{aligned} \quad (\text{A.5})$$

where we have changed the summation index in the first sum before relabelling it j . The only way a polynomial can vanish for all y is if all the coefficients vanish, so we have a *recurrence relation*:

$$(j+1)(j+2)c_{j+2} + (\mathcal{E} - 1 - 2j)c_j = 0. \quad (\text{A.6})$$

Given c_0 and c_1 , we can construct all other coefficients from this equation, for any \mathcal{E} . We can obtain two independent solutions, as expected for a second order differential equation: even solutions with $c_1 = 0$ and odd ones with $c_0 = 0$.

However, we need the wave function to be *normalisable* (square integrable), which means that it tends to 0 as $x \rightarrow \pm\infty$. In general an infinite polynomial times a Gaussian will not satisfy this, and these solutions are not physically acceptable. If we look again at equation (A.6), though, we see that if $\mathcal{E} = 1 + 2n$ for some integer $n \geq 0$, then $c_{n+2}, c_{n+4}, c_{n+6} \dots$ are all zero. Thus for $\mathcal{E} = 1, 5, 9 \dots$ we have finite even polynomials, and for $\mathcal{E} = 3, 7, 11 \dots$ we have finite odd polynomials. These are called the Hermite polynomials.

Rewriting (A.6) with $\mathcal{E} = 1 + 2n$ as

$$c_{j+2} = \frac{2(j-n)}{(j+1)(j+2)}c_j, \quad (\text{A.7})$$

we have for instance, for $n = 5$,

$$c_3 = 2(1-5)c_1/(2.3) = -4c_1/3 \quad c_5 = 2(3-5)c_3/(4.5) = -c_3/5 = 4c_1/15, \quad c_7 = c_9 = \dots = 0, \quad (\text{A.8})$$

and $H_5(y) = c_1(4y^5 - 20y^3 + 15y)/15$. The conventional normalisation uses 2^n for the coefficient of the highest power of y , which would require $c_1 = 120$, and $H_5(y) = 32y^5 - 160y^3 + 120y$. The first four are:

$$H_0(y) = 1; \quad H_1(y) = 2y; \quad H_2(y) = 4y^2 - 2; \quad H_3(y) = 8y^3 - 12y; \quad H_4(y) = 16y^4 - 48y^2 + 12. \quad (\text{A.9})$$

The corresponding solutions of the original Hamiltonian, returning to unscaled coordinates, are

$$\phi_n(x) = (2^n n!)^{-1/2} H_n\left(\frac{x}{x_0}\right) \times (\pi x_0^2)^{-1/4} \exp(-x^2/(2x_0^2));$$

with energies $E_n = (n + \frac{1}{2})\hbar\omega$.

Just as in the square well, the restriction to solutions which satisfy the boundary conditions has resulted in quantised energy levels.

The wave functions and probability densities are illustrated in section [A.3.3](#).

A.5 Angular Momentum in Quantum Mechanics

The following section was prepared in another context and so does not use Dirac notation; I would use $|\uparrow\rangle$ etc for spin states. See section A.3.4 for angular momentum in the context of the Schrödinger equation in spherical polar coordinates, including the forms of the differential operators for the \hat{L}_i .

Orbital angular momentum

We start with the classical definition of orbital angular momentum. In quantum mechanics the position and momentum vectors become operators, so

$$\mathbf{L} = \mathbf{r} \times \mathbf{p} \Rightarrow \hat{L}_z = -i\hbar \left(x \frac{\partial}{\partial y} - y \frac{\partial}{\partial x} \right) = -i\hbar \frac{\partial}{\partial \phi} \quad \text{etc}$$

$$[\hat{L}_x, \hat{L}_y] = i\hbar \hat{L}_z \quad \text{etc}; \quad [\hat{\mathbf{L}}^2, \hat{L}_i] = 0;$$

The commutation relations imply that we can only simultaneously know \mathbf{L}^2 and one component, taken conventionally to be L_z . The common eigenfunctions of $\hat{\mathbf{L}}^2$ and \hat{L}_z are the spherical harmonics, $Y_l^m(\theta, \phi)$:

$$\hat{\mathbf{L}}^2 Y_l^m(\theta, \phi) = \hbar^2 l(l+1) Y_l^m(\theta, \phi) \quad \hat{L}_z Y_l^m(\theta, \phi) = \hbar m Y_l^m(\theta, \phi)$$

From requirements that the wave function must be finite everywhere, and single-valued under $\phi \rightarrow \phi + 2\pi$, it emerges that l and m are integers and must satisfy

$$l = 0, 1, 2, \dots, \quad m = -l, -l+1, \dots, l.$$

These have definite parity of $(-1)^l$, since under $\mathbf{r} \rightarrow -\mathbf{r}$,

$$Y_l^m(\theta, \phi) \rightarrow Y_l^m(\pi - \theta, \phi + \pi) = (-1)^l Y_l^m(\theta, \phi).$$

See the end of these notes for some explicit forms of spherical harmonics.

Intrinsic and total angular momentum

Orbital angular momentum is not the only source of angular momentum, particles may have *intrinsic angular momentum or spin*. The corresponding operator is $\hat{\mathbf{S}}$. The eigenvalues of $\hat{\mathbf{S}}^2$ have the same form as in the orbital case, $\hbar^2 s(s+1)$ but now s can be integer or half integer; similarly the eigenvalues of \hat{S}_z are $\hbar m_s$, with

$$s = 0, \frac{1}{2}, 1, \frac{3}{2}, \dots, \quad m_s = -s, -s+1, \dots, s.$$

$s = \frac{1}{2}$ for an electron, $s = 1$ for a photon or W boson. This means that the magnitude of the spin vector of an electron is $(\sqrt{3}/2)\hbar$, but we always just say “spin- $\frac{1}{2}$ ”.

If a particle has both orbital and spin angular momentum, we talk about its *total angular momentum*, with operator

$$\hat{\mathbf{J}} = \hat{\mathbf{L}} + \hat{\mathbf{S}}.$$

As with spin, the eigenvalues of $\hat{\mathbf{J}}^2$ are $\hbar^2 j(j+1)$

$$j = 0, \frac{1}{2}, 1, \frac{3}{2}, \dots, \quad m_j = -j, -j+1, \dots, j.$$

Systems composed of more than one particle (hadrons, nuclei, atoms) will have many contributions to their total angular momentum. It is sometimes useful to add up all the spins to give a total spin, and now, confusingly, we denote the quantum numbers by S and M_S , so it is really important to distinguish operators and the corresponding quantum numbers. Then

$$\hat{\mathbf{S}}^{\text{tot}} = \hat{\mathbf{S}}^{(1)} + \hat{\mathbf{S}}^{(2)} + \dots,$$

where the superscripts (1), (2) refer to the individual particles.

Similarly we use $\hat{\mathbf{L}}^{\text{tot}}$ with quantum numbers L and M_L , and $\hat{\mathbf{J}}^{\text{tot}}$ with quantum numbers J and M_J . When talking about angular momentum generally, we often use $\hat{\mathbf{J}}$ to refer to *any* angular momentum, whether single or multiple particle, pure spin, pure orbital or a combination.

The following rules are obeyed by *any* angular momentum (eg $\hat{\mathbf{J}}$ can be replaced by $\hat{\mathbf{L}}$ or $\hat{\mathbf{S}}$, for a single particle of composite system):

$$[\hat{J}_x, \hat{J}_y] = i\hbar \hat{J}_z \quad \text{etc}; \quad [\hat{\mathbf{J}}^2, \hat{J}_i] = 0;$$

It follows that the eigenvalues of $(\hat{\mathbf{L}}^{\text{tot}})^2$, $(\hat{\mathbf{S}}^{\text{tot}})^2$ and $(\hat{\mathbf{J}}^{\text{tot}})^2$ have exactly the same form, with the same restrictions on the quantum numbers, as those for a single particle. So for instance the eigenstates of $(\hat{\mathbf{S}}^{\text{tot}})^2$ are $\hbar^2 S(S+1)$, and of \hat{S}_z^{tot} are $\hbar M_S$, and

$$L = 0, 1, 2, \dots, \quad S = 0, \frac{1}{2}, 1, \frac{3}{2}, \dots, \quad J = 0, \frac{1}{2}, 1, \frac{3}{2}, \dots,$$

$$M_L = -L, -L+1, \dots, L, \quad M_S = -S, -S+1, \dots, S, \quad M_J = -J, -J+1, \dots, J.$$

Addition of angular momentum

The rules for the addition of angular momentum are as follows: we start with adding orbital angular momentum and spin for a composite system with quantum numbers L and S . Angular momentum is a vector, and so the total can be smaller as well as greater than the parts; however the z -components just add. The allowed values of the total angular momentum quantum numbers are

$$J = |L - S|, |L - S| + 1, \dots, L + S, \quad M_J = M_L + M_S.$$

However since \hat{L}_z and \hat{S}_z do not commute with $\hat{\mathbf{J}}^2$, we cannot know J , M_L and M_S simultaneously. For a single-particle system, replace J , L , and S with j , l , and s .

More generally, for the addition of any two angular momenta with quantum numbers J_1, M_1 and J_2, M_2 , the rules are

$$J = |J_1 - J_2|, |J_1 - J_2| + 1, \dots, J_1 + J_2, \quad M_J = M_1 + M_2$$

and again we cannot know J , M_1 and M_2 simultaneously.

Confusingly, when referring to a composite particle (eg a hadron or nucleus), the total angular momentum is often called its “spin” but given the quantum number J . Sometimes this usage even extends to elementary particles. For the electron and proton, s is more common though.

For the case of a spin- $\frac{1}{2}$ particle, the eigenvalues of \hat{S}_z are $\pm\frac{1}{2}\hbar$, and here we will just denote these states by \uparrow and \downarrow (α_z and β_z are also often used); hence

$$\begin{aligned}\hat{\mathbf{S}}^2\uparrow &= \frac{3}{4}\hbar^2\uparrow & \hat{\mathbf{S}}^2\downarrow &= \frac{3}{4}\hbar^2\downarrow \\ \hat{S}_z\uparrow &= \frac{1}{2}\hbar\uparrow & \hat{S}_z\downarrow &= -\frac{1}{2}\hbar\downarrow\end{aligned}$$

For two such particles there are four states $\uparrow\uparrow$, $\downarrow\downarrow$, $\uparrow\downarrow$ and $\downarrow\uparrow$. The first two states have $M_S = 1$ and -1 respectively, and we can show, using $\hat{\mathbf{S}}^{\text{tot}} = \hat{\mathbf{S}}^{(1)} + \hat{\mathbf{S}}^{(2)}$, that they are also eigenstates of $(\hat{\mathbf{S}}^{\text{tot}})^2$ with $S = 1$. However the second two, though they have $M_S = 0$, are not eigenstates of $(\hat{\mathbf{S}}^{\text{tot}})^2$. To make those, we need linear combinations, tabulated below:

	$S = 1$	$S = 0$
$M = 1$	$\uparrow\uparrow$	
$M = 0$	$\frac{1}{\sqrt{2}}(\uparrow\downarrow + \downarrow\uparrow)$	$\frac{1}{\sqrt{2}}(\uparrow\downarrow - \downarrow\uparrow)$
$M = -1$	$\downarrow\downarrow$	

The $S = 1$ states are symmetric under exchange of particles; the $S = 0$ states are antisymmetric. For a system of N spin- $\frac{1}{2}$ particles, S will be integer if N is even and half-integer if N is odd.

Bosons and Fermions

Particles with half-integer spin (electrons, baryons) are called fermions, those with integer spin, including $J = 0$, (mesons, photons, Higgs) are called bosons. The ‘‘Pauli exclusion principle’’ applies to fermions, but it is a special case of the ‘‘spin-statistics theorem’’ which says that the overall quantum state of a system of identical fermions must be antisymmetric under exchange of any pair, while that of a system of identical bosons must be symmetric. There may be several components to the state (spatial wave function, spin state...).

Examples of the consequences of the spin-statistics theorem are:

- If two electrons in an atom are in the same orbital (thus their spatial wave function is symmetric under exchange of the two), they must be in an $S = 0$ state.
- Thus the ground state of helium has $S = 0$, but the excited states can have $S = 0$ (parahelium) or $S = 1$ (orthohelium).
- Two π^0 mesons must have even relative orbital angular momentum L (they are spinless, so this is the only contribution to their wave function).
- Two ρ^0 mesons (spin-1 particles) can have odd or even relative orbital angular momentum L , but their spin state must have the same symmetry as their spatial state. (In this case, $S = 2$ and 0 are even, $S = 1$ is odd.)

Note that in the last two, in the centre-of-momentum frame the spatial state only depends on the relative coordinate \mathbf{r} . So interchanging the particles is equivalent to $\mathbf{r} \rightarrow -\mathbf{r}$, ie the parity operation.

Spherical Harmonics

In spherical polar coordinates the orbital angular momentum operators are

$$\begin{aligned}\hat{L}_x &= \frac{1}{2}(\hat{L}_+ + \hat{L}_-) \quad \text{and} \quad \hat{L}_y = \frac{1}{2i}(\hat{L}_+ - \hat{L}_-), \quad \text{where} \\ \hat{L}_+ &= \hbar e^{i\phi} \left(\frac{\partial}{\partial \theta} + i \cot \theta \frac{\partial}{\partial \phi} \right), \quad \hat{L}_- = \hat{L}_+^\dagger = \hbar e^{-i\phi} \left(-\frac{\partial}{\partial \theta} + i \cot \theta \frac{\partial}{\partial \phi} \right); \\ \hat{L}_z &= -i\hbar \frac{\partial}{\partial \phi}, \quad \hat{\mathbf{L}}^2 = -\hbar^2 \left(\frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial}{\partial \theta} \right) + \frac{1}{\sin^2 \theta} \frac{\partial^2}{\partial \phi^2} \right). \\ \nabla^2 \psi &= \frac{1}{r} \frac{\partial^2}{\partial r^2} r \psi - \frac{1}{\hbar^2 r^2} \hat{\mathbf{L}}^2 \psi;\end{aligned}$$

The spherical harmonics, $Y_l^m(\theta, \phi)$ are eigenfunctions of $\hat{\mathbf{L}}^2$ and \hat{L}_z ; the first few are as follows

$$\begin{aligned}Y_0^0(\theta, \phi) &= \sqrt{\frac{1}{4\pi}} & Y_1^{\pm 1}(\theta, \phi) &= \mp \sqrt{\frac{3}{8\pi}} \sin \theta e^{\pm i\phi} \\ Y_1^0(\theta, \phi) &= \sqrt{\frac{3}{4\pi}} \cos \theta & Y_2^{\pm 2}(\theta, \phi) &= \sqrt{\frac{15}{32\pi}} \sin^2 \theta e^{\pm 2i\phi} \\ Y_2^{\pm 1}(\theta, \phi) &= \mp \sqrt{\frac{15}{8\pi}} \sin \theta \cos \theta e^{\pm i\phi} & Y_2^0(\theta, \phi) &= \sqrt{\frac{5}{16\pi}} (3 \cos^2 \theta - 1)\end{aligned}$$

A.6 Hydrogen wave functions

The solutions of the Schrödinger equation for the Coulomb potential $V(r) = -\hbar c\alpha/r$ have energy $E_n = -\frac{1}{n^2}E_{\text{Ry}}$, where $E_{\text{Ry}} = \frac{1}{2}\alpha^2 mc^2 = 13.6 \text{ eV}$ (with m the reduced mass of the electron-proton system). (Recall $\alpha = e^2/(4\pi\epsilon_0\hbar c) \approx 1/137$.) The spatial wavefunctions are $\psi_{nlm}(\mathbf{r}) = R_{n,l}(r)Y_l^m(\theta, \phi)$.

The radial wavefunctions are as follows, where $a_0 = \hbar c/(mc^2\alpha)$:

$$\begin{aligned} R_{1,0}(r) &= \frac{2}{a_0^{3/2}} \exp\left(-\frac{r}{a_0}\right), \\ R_{2,0}(r) &= \frac{2}{(2a_0)^{3/2}} \left(1 - \frac{r}{2a_0}\right) \exp\left(-\frac{r}{2a_0}\right), \\ R_{2,1}(r) &= \frac{1}{\sqrt{3}(2a_0)^{3/2}} \frac{r}{a_0} \exp\left(-\frac{r}{2a_0}\right), \\ R_{3,0}(r) &= \frac{2}{(3a_0)^{3/2}} \left(1 - \frac{2r}{3a_0} + \frac{2r^2}{27a_0^2}\right) \exp\left(-\frac{r}{3a_0}\right), \\ R_{3,1}(r) &= \frac{4\sqrt{2}}{9(3a_0)^{3/2}} \frac{r}{a_0} \left(1 - \frac{r}{6a_0}\right) \exp\left(-\frac{r}{3a_0}\right), \\ R_{3,2}(r) &= \frac{2\sqrt{2}}{27\sqrt{5}(3a_0)^{3/2}} \left(\frac{r}{a_0}\right)^2 \exp\left(-\frac{r}{3a_0}\right). \end{aligned}$$

They are normalised, so $\int_0^\infty (R_{n,l}(r))^2 r^2 dr = 1$. Radial wavefunctions of the same l but different n are orthogonal (the spherical harmonics take care of orthogonality for different l s).

The following radial integrals can be proved:

$$\begin{aligned} \langle r^2 \rangle &= \frac{a_0^2 n^2}{2} (5n^2 + 1 - 3l(l+1)), \\ \langle r \rangle &= \frac{a_0}{2} (3n^2 - l(l+1)), \\ \left\langle \frac{1}{r} \right\rangle &= \frac{1}{n^2 a_0}, \\ \left\langle \frac{1}{r^2} \right\rangle &= \frac{1}{(l+1/2)n^3 a_0^2}, \\ \left\langle \frac{1}{r^3} \right\rangle &= \frac{1}{l(l+1/2)(l+1)n^3 a_0^3}. \end{aligned}$$

For hydrogen-like atoms (single-electron ions with nuclear charge $|e|Z$) the results are obtained by substituting $\alpha \rightarrow Z\alpha$ (and so $a_0 \rightarrow a_0/Z$).

A.7 Properties of δ -functions

The δ -function is defined by its behaviour in integrals:

$$\int_a^b \delta(x - x_0) dx = 1; \quad \int_a^b f(x) \delta(x - x_0) dx = f(x_0)$$

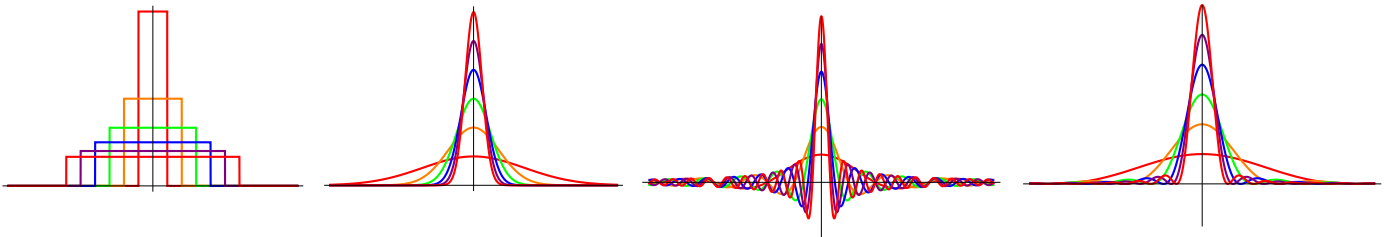
where the limits a and b satisfy $a < x_0 < b$; the integration simply has to span the point on which the δ -function is centred. The second property is called the **sifting property** because it picks out the value of f at $x = x_0$.

The following equivalences may also be proved by changing variables in the corresponding integral (an appropriate integration range is assumed for compactness of notation):

$$\begin{aligned} \delta(ax - b) &= \frac{1}{|a|} \delta\left(x - \frac{b}{a}\right) && \text{since } \int f(x) \delta(ax - b) dx = \frac{1}{a} f\left(\frac{b}{a}\right) \\ \delta(g(x)) &= \sum_i \frac{\delta(x - x_i)}{|g'(x_i)|} && \text{where the } x_i \text{ are the (simple) real roots of } g(x). \end{aligned}$$

Note that the dimensions of a δ -function are the inverse of those of its argument, as should be obvious from the first equation.

Though the δ -function is not well defined as a function (technically it is a distribution rather than a function), it can be considered as the limit of many well-defined functions. For instance the “top-hat” function which vanishes outside a range a and has height $1/a$ tends to a δ -function as $a \rightarrow \infty$. Similarly a Gaussian with width and height inversely proportional tends to a δ -function as the width tends to zero. These are shown in the first two frames below.

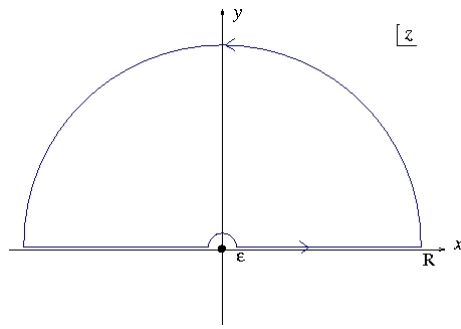


Two less obvious functions which tend to a δ -function, shown in the next two frames, are the following:

$$\begin{aligned} \frac{1}{2\pi} \int_{-L}^L e^{i(k-k')x} dx &= \frac{L}{\pi} \text{sinc}((k - k')L) \xrightarrow{L \rightarrow \infty} \delta(k - k') \\ \frac{L}{\pi} \text{sinc}^2((k - k')L) &\xrightarrow{L \rightarrow \infty} \delta(k - k') \end{aligned}$$

The first of these does not actually vanish away from the peak, but it oscillates so rapidly that there will be no contribution to any integral over k' except from the point $k' = k$. This is the integral which gives the orthogonality of two plane waves with different wavelengths: $\langle k | k' \rangle = \delta(k - k')$. It also ensures that the inverse Fourier transform of a Fourier transform recovers the original function.

That the normalisation (for integration over k) is correct follows from the following two integrals: $\int_{-\infty}^{\infty} \text{sinc}(t) dt = \pi$ and $\int_{-\infty}^{\infty} \text{sinc}^2(t) dt = \pi$. The second of these follows from the first via integration by parts. The integral $\int_{-\infty}^{\infty} \text{sinc}(t) dt = \text{Im } I$ where $I = \int_{-\infty}^{\infty} (e^{it}/t) dt$ may be done via the contour integral below:



As no poles are included by the contour, the full contour integral is zero. By Jordan's lemma the integral round the outer circle tends to zero (as $R \rightarrow \infty$, e^{iz} decays exponentially in the upper half plane). So the integral along the real axis is equal and opposite to the integral over the inner circle, namely $-\pi i$ times the residue at $x = 0$, so $I = i\pi$. So the imaginary part, the integral of $\text{sinc}(x)$, is π .

A.8 Gaussian integrals

The following integrals will be useful:

$$\int_{-\infty}^{\infty} e^{-\alpha x^2} dx = \sqrt{\frac{\pi}{\alpha}} \quad \text{and} \quad \int_{-\infty}^{\infty} x^{2n} e^{-\alpha x^2} dx = (-1)^n \frac{d^n}{d\alpha^n} \left(\sqrt{\frac{\pi}{\alpha}} \right)$$

These work even for complex α , so long as $\text{Re}[\alpha] \geq 0$

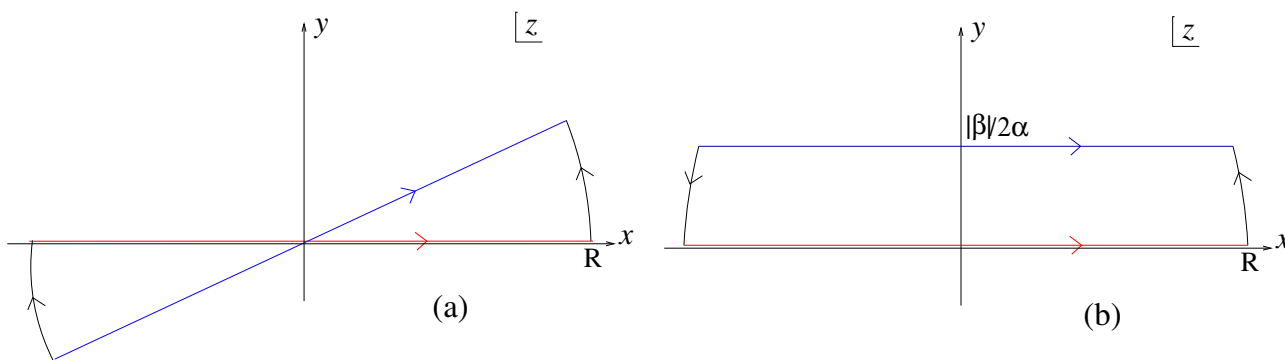
Often we are faced with a somewhat more complicated integral, which can be cast in Gaussian form by “completing the square” in the exponent and then shifting integration variable $x \rightarrow x - \beta/(2\alpha)$:

$$\int_{-\infty}^{\infty} e^{-\alpha x^2 - \beta x} dx = e^{\beta^2/(4\alpha)} \int_{-\infty}^{\infty} e^{-\alpha(x + \beta/(2\alpha))^2} dx = \sqrt{\frac{\pi}{\alpha}} e^{\beta^2/(4\alpha)}$$

This works even if β is imaginary.

The two contours below illustrate the two results for complex parameters α or β . For the first, in (a), we rewrite αx^2 as $|\alpha|z^2$ where $z = x \exp(i\text{Arg}[\alpha]/2)$, so the integral we want is along the blue line, with $R \rightarrow \infty$. Since there are no poles, by Cauchy’s theorem the integral along the blue contour must equal the sum of those along the red and black contours. As $R \rightarrow \infty$ the red one gives the known real integral. Since $e^{-|\alpha|z^2}$ tends to zero faster than $1/R$ as $R \rightarrow \infty$ providing $|x| > |y|$, the contribution from the black paths is zero as $R \rightarrow \infty$. Hence the red and blue integrals are the same, provided $\text{Arg}[\alpha] \leq \pi/2$.

For the second, in (b), the blue contour is the desired integral one after the variable change (for β imaginary). Again the red and black paths together must equal the blue and again the contribution from the black paths is zero. Hence the two integrals must be the same.

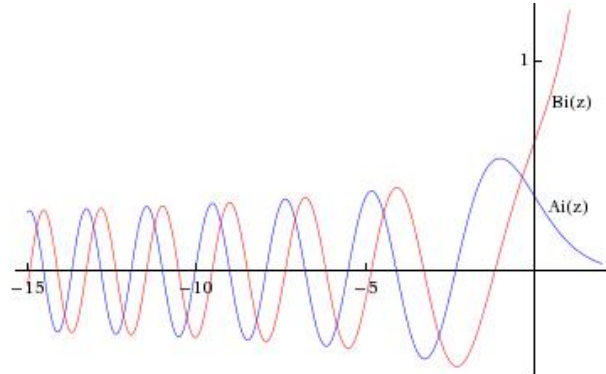


A.9 Airy functions

Airy functions are the solutions of the differential equation:

$$\frac{d^2 f}{dz^2} - zf = 0$$

There are two solutions, $\text{Ai}(z)$ and $\text{Bi}(z)$; the first tends to zero as $z \rightarrow \infty$, while the second blows up. Both are oscillatory for $z < 0$. The Mathematica functions for obtaining them are



`AiryAi[z]` and `AiryBi[z]`.

The asymptotic forms of the Airy functions are:

$$\begin{aligned} \text{Ai}(z) &\xrightarrow{z \rightarrow \infty} \frac{e^{-\frac{2}{3}z^{3/2}}}{2\sqrt{\pi}z^{1/4}} & \text{and} & \quad \text{Ai}(z) &\xrightarrow{z \rightarrow -\infty} \frac{\cos\left(\frac{2}{3}|z|^{3/2} - \frac{\pi}{4}\right)}{\sqrt{\pi}|z|^{1/4}} \\ \text{Bi}(z) &\xrightarrow{z \rightarrow \infty} \frac{e^{\frac{2}{3}z^{3/2}}}{\sqrt{\pi}z^{1/4}} & \text{and} & \quad \text{Bi}(z) &\xrightarrow{z \rightarrow -\infty} \frac{\cos\left(\frac{2}{3}|z|^{3/2} + \frac{\pi}{4}\right)}{\sqrt{\pi}|z|^{1/4}} \end{aligned}$$

The Schrödinger equation for a linear potential $V(x) = \beta x$ in one dimension can be cast in the following form

$$-\frac{\hbar^2}{2m} \frac{d^2 \psi}{dx^2} + \beta x \psi - E \psi = 0$$

Defining $z = x/x_0$, with $x_0 = (\hbar^2/(2m\beta))^{1/3}$, and $E = (\hbar^2\beta^2/(2m))^{1/3}\mu$, and with $y(z) \equiv \psi(x)$, this can be written

$$\frac{d^2 y}{dz^2} - zy + \mu y = 0.$$

The solution is

$$y(z) = C \text{Ai}(z-\mu) + D \text{Bi}(z-\mu) \quad \text{or} \quad \psi(x) = C \text{Ai}((\beta x - E)/(\beta x_0)) + D \text{Bi}((\beta x - E)/(\beta x_0))$$

where $D = 0$ if the solution has to extend to $x = \infty$. The point $z = \mu$, $x = E/\beta$ is the point at which $E = V$ and the solution changes from oscillatory to decaying / growing.

The equation for a potential with a negative slope is given by substituting $z \rightarrow -z$ in the defining equation. Hence the general solution is $\psi(x) = C \text{Ai}(-x/x_0 - \mu) + D \text{Bi}(-x/x_0 - \mu)$, with $D = 0$ if the solution has to extend to $x = -\infty$.

The first few zeros of the Airy functions are given in [Wolfram MathWorld](#).

A.10 Units in EM

There are several systems of units in electromagnetism. We are familiar with SI units, but Gaussian units are still very common and are used, for instance, in Shankar.

In SI units the force between two currents is used to define the unit of current, and hence the unit of charge. (Currents are much easier to calibrate and manipulate in the lab than charges.) The constant μ_0 is *defined* as $4\pi \times 10^{-7} \text{ N A}^{-2}$, with the magnitude chosen so that the Ampère is a “sensible” sort of size. Then Coulomb’s law reads

$$\mathbf{F} = \frac{q_1 q_2}{4\pi\epsilon_0 |\mathbf{r}_1 - \mathbf{r}_2|^2}$$

and ϵ_0 has to be obtained from experiment. (Or, these days, as the speed of light is now has a defined value, ϵ_0 is obtained from $1/(\mu_0 c^2)$.)

However one could in principle equally decide to use Coulomb’s law to define charge. This is what is done in Gaussian units, where by definition

$$\mathbf{F} = \frac{q_1 q_2}{|\mathbf{r}_1 - \mathbf{r}_2|^2}$$

Then there is no separate unit of charge; charges are measured in $\text{N}^{1/2} \text{ m}$ (or the non-SI equivalent): $e = 4.803 \times 10^{-10} \text{ g}^{1/2} \text{ cm}^{3/2} \text{ s}^{-1}$. (You should never need that!) In these units, $\mu_0 = 4\pi/c^2$. Electric and magnetic fields are also measured in different units.

The following translation table can be used:

Gauss	e	\mathbf{E}	\mathbf{B}
SI	$e/\sqrt{4\pi\epsilon_0}$	$\sqrt{4\pi\epsilon_0} \mathbf{E}$	$\sqrt{4\pi/\mu_0} \mathbf{B}$

Note that $e\mathbf{E}$ is the same in both systems of units, but $e\mathbf{B}$ in SI units is replaced by $e\mathbf{B}/c$ in Gaussian units. Thus the Bohr magneton μ_B is $e\hbar/2m$ in SI units, but $e\hbar/2mc$ in Gaussian units, and $\mu_B \mathbf{B}$ has dimensions of energy in both systems.

The fine-structure constant α is a dimensionless combination of fundamental units, and as such takes on the same value ($\approx 1/137$) in all systems. In SI it is defined as $\alpha = e^2/(4\pi\epsilon_0 \hbar c)$, in Gaussian units as $\alpha = e^2/(\hbar c)$. In all systems, therefore, Coulomb’s law between two particles of charge $z_1 e$ and $z_2 e$ can be written

$$\mathbf{F} = \frac{z_1 z_2 \hbar c \alpha}{|\mathbf{r}_1 - \mathbf{r}_2|^2}$$

and this is the form I prefer.