Further quantum theory notes

Lionel Mason, lmason@maths.ox.ac.uk

March 13, 2017

Email corrections and queries to the address above.

B7.3: Further Quantum Mechanics

Number of lectures: 16 HT2017
Course Description
Level: B/M-level Method of Assessment: Written examination.
Weight: One unit
Prerequisites: Part A Quantum Mechanics.

Overview

This course builds directly on the first course in quantum mechanics and covers a series of important topics, particularly features of systems containing several particles. The behaviour of identical particles in quantum theory is more subtle than in classical mechanics, and an understanding of these features allows one to understand the periodic table of elements and the rigidity of matter.

There are rarely neat solutions to problems involving several particles, so usually one needs some approximation methods. These are developed so as to study both energy levels of interacting Hamiltonians, and scattering.

Learning Outcomes

Students will be able to demonstrate knowledge and understanding of the mathematics of quantum mechanics of many particle systems, and atomic structure and scattering.

Synopsis

Scattering and tunnelling in one dimension. The semi-classical approximation (WKB). Symmetries in quantum mechanics. Rotations, spin and angular momentum and their addition.

Identical particles, symmetric and anti-symmetric states, Fermi-Dirac and Bose-Einstein statistics and atomic structure.

Approximation methods, Rayleigh-Schrödinger time-independent perturbation theory and variation principles. The virial theorem. Helium.

Heisenberg representation, interaction representation, time dependent perturbation theory and FeynmanDyson expansion.

Reading List

The lectures will partly follow:
S. Weinberg, Lectures on quantum mechanics, CUP, (2013). Sections 4.3-5, 5.1-7, 6.1-3, 7.1-4.
together with:
K C Hannabuss, *Introduction to quantum mechanics*, OUP (1997). Chapter 8.1-4, 8.8, 9, 16.1-4, 11.1-5, 12.1-4, 14.1-4, 15.1-3, 13.5.
But the following are also recommended:
J. Binney and D. Skinner, The physics of quantum mechanics, PUP, 2011.
Landau and Lifschitz, Quantum Mechanics, non-relativistic theory, Vol 3 of a course in theoretical physics, Pergamon press.
Gordon Screaton, Further Quantum Theory, Mathematical Institute Notes (1991). Also designed for an Oxford course, though only covering some material: This can be found online at
http://www0.maths.ox.ac.uk/system/files/attachments/Further_Quantum_Mechanics.pdf
L. I. Schiff, Quantum Mechanics (3rd edition, Mc Graw Hill, 1968).
B. J. Bransden and C. J. Joachain, Introduction to Quantum Mechanics (Longman, 1995).
A. I. M. Rae, Quantum Mechanics (4th edition, Institute of Physics, 1993).
A popular non-technical account of the subject:
A. Hey and P. Walters, The New Quantum Universe (Cambridge, 2003).

Contents

1 Introduction

Atoms are the stuff of our existence, they make up our bodies, the earth, the atmosphere, the stars and galaxies. Although it turns out that, due to so-called 'dark matter' they in fact turn out to be a relatively small proportion of the universe, they nevertheless dominate our experience of the universe and their study in various levels of aggregation make up the vast majority of the natural sciences.

Atoms cannot be explained classically; there is nothing to stop a classical system of electrons that are orbiting a nucleus from radiating away their energy and falling in. Their first consistent description requires quantum mechanics. In the earlier course(s) we saw how a stable atom could be modelled in terms of solutions to the stationary Schrödinger equation $H\psi = E\psi$ for the electron wave function $\psi(\mathbf{x})$ with energy E and Hamiltonian H incorporating a central Coulomb potential arising from the nucleus

$$H\psi := \left(\frac{\mathbf{P}^2}{2m} - \frac{Ze^2}{4\pi\epsilon_0 r}\right)\psi = E\psi\,.$$

Here e is the electron charge, Ze the nuclear charge (so that this describes a hydrogen-like ion), $\hbar$ Planck's constant, ϵ_0 the permittivity of free space, $r = |\mathbf{x}|$ and $\mathbf{P} = -i\hbar\nabla$. We saw by separation of variables that there were bound states $\psi_{nlm} = F_{nl}(r)Y_l^m(\theta,\phi)$ where $Y_l^m(\theta,\phi) := \mathrm{e}^{im\phi}P_{lm}(\theta)$ are the spherical harmonics. The energies are distributed according to

$$E_n = -\frac{1}{2n^2}\frac{Z^2e^2}{4\pi\epsilon_0 a} = -\frac{\hbar^2}{2ma^2n^2} \qquad \text{where} \qquad a = \frac{4\pi\epsilon_0\hbar^2}{me^2}\,,$$

with degeneracy n^2 where a is the Bohr radius giving a measure of the size of the atom and m the electron mass. This correctly explains Balmer's 1885 series for differences between the second and higher energy levels to account for the absorption spectra for hydrogen. However, there are still a number of ingredients required to obtain full picture of atomic structure:

- Electrons satisfy Fermi statistics.
- Electrons have intrinsic spin.
- The full Schrodinger equation with many electrons becomes too complicated to solve exactly; we need approximation methods to sufficiently calculate the spectra.

- we need to understand scattering processes.

This course will cover the basic tools required to understand the periodic table of elements, aspects of their spectra and scattering. The tools will also have wide applicability more generally in quantum theory and quantum field theory.

1.1 Review of the formalism of quantum mechanics

We will work with the algebraic formulation:

1. Physical states correspond to state vectors ψ and referred to as *kets* $|\psi\rangle$ in a Hilbert space $\mathcal{H}$ usually thought of as a space of wave functions on $\mathbb{R}^3$ although we will often think more abstractly.

2. A Hilbert space $\mathcal{H}$ is a complex vector space (often infinite dimensional) with a hermitian inner product[1].

   Notation: denote the complex conjugate of ket $|\phi\rangle$ (the wave function $\phi(\mathbf{x})$ by the *bra* $\langle\phi|$ corresponding to the conplex conjugate function $\bar{\phi}(\mathbf{x})$ and the inner product with ψ by the 'bracket' denoted $\langle\phi|\psi\rangle$ given in terms of wave functions by $\int_{\mathbb{R}^3}\bar{\phi}(\mathbf{x})\psi(\mathbf{x})d^3\mathbf{x}$.

   This is complex antilinear in ϕ as $\langle a\phi| = \bar{a}\langle\phi|$ for $a \in \mathbb{C}$.

   A normalized state ψ is one for which $||\psi||^2 := \langle\psi|\psi\rangle = 1$ although this still only fixes $\psi \to \mathrm{e}^{i\theta}\psi$ and these are understood to represent the same physical state.

3. The inner product between two normalized states $\langle\phi|\psi\rangle \in \mathbb{C}$ is an amplitude. Its modulus squared is the probability of a measurement of the state ψ to be in the state ϕ.

4. Given a linear operator $A : \mathcal{H} \to \mathcal{H}$ we define its hermitian conjugate or adjoint A^* by
$$\langle\phi|A^*\psi\rangle = \langle A\phi|\psi\rangle\,.$$

---

[1] In infinite dimensions it is also important to introduce completeness (existence of limits of Cauchy sequences) and separability (the existence of a countable basis) conditions. However, we do not emphasize these in this course. These play a major role in convergence issues and are emphasized in B4.

5. Physical observables such as position, momentum, energy etc., are given by Hermitian linear operators

   $$A : \mathcal{H} \to \mathcal{H}\,, \qquad \langle \phi | A\psi \rangle = \langle A\phi | \psi \rangle\,.$$

   Because a Hermitian A can act on either ψ or ϕ we often write

   $$\langle \phi | A | \psi \rangle := \langle \phi | A\psi \rangle = \langle A\phi | \psi \rangle\,.$$

   The allowed values of an observable A are its eigenvalues (which must be real). For a given normalized state ψ, the probability of finding the value a for A is $|\langle \phi_a | \psi \rangle|^2$ where $A|\phi_a\rangle = a|\phi_a\rangle$, i.e., ϕ_a is the corresponding eigenvector of A. (If the eigenspace is more than one dimensional, we must sum this expression over an orthonormal basis of the eigenspace).

6. The expectation value of an observable A is $E_A = \langle \psi | A | \psi \rangle / \langle \psi | \psi \rangle$.

7. If two observables A, B commute, $[A, B] := AB - BA = 0$, then it is possible to simultaneously diagonalize them, i.e., to find a basis of simultaneous eigenvectors. Such observables can be measured simultaneously.

   If they do not commute, then there will be some Heisenberg uncertainty relation between them.

A ket is often represented by its Schrodinger wave function $\psi(\mathbf{x}, t)$ a function of space, $\mathbb{R}^3$, and time t. The corresponding bra is its complex conjugate $\bar{\psi}(\mathbf{x})$ and the hermitian inner product is

$$\langle \phi | \psi \rangle = \int \bar{\phi}\psi \, \mathrm{d}^3\mathbf{x}\,.$$

This is preserved by the Schrodinger evolution equation

$$i\hbar \frac{\partial \psi}{\partial t} = H\psi$$

because the Hamiltonian H is Hermitian (exercise). Here H is taken to be the observable corresponding to the sum of kinetic and potential energy of the system, usually $H = \mathbf{P}^2/2m + V(\mathbf{x})$ where $\mathbf{P}$ is the momentum and $V(\mathbf{x})$ the potential energy.

The Hilbert space $\mathcal{H}$ = space of square integrable complex functions on $\mathbb{R}^3$ (L^2 for those who have taken the Lebesgue Integration course). Observables include the components x^i, $i = 1, 2, 3$ of the position $\mathbf{x}$ acting by multiplication, or the momenta

$$\mathbf{P} = \frac{\hbar}{i}\left(\frac{\partial}{\partial x^1}, \frac{\partial}{\partial x^2}, \frac{\partial}{\partial x^3}\right) = -i\hbar\nabla$$

acting by differentiation so $H = -\frac{\hbar^2}{2m}\nabla^2 + V(\mathbf{x})$ on the Schrödinger equation.

After the Fourier transform

$$\hat{\psi}(\mathbf{p}) = \frac{1}{\sqrt{2\pi}}\int \mathrm{e}^{i\mathbf{p}\cdot\mathbf{x}/\hbar}\psi(x)\,\mathrm{d}^3\mathbf{x}$$

$\mathbf{P}$ acts by multiplication by $\mathbf{p}$ and $\mathbf{X}$ by $\frac{\hbar}{i}\times$differentiation with respect to $\mathbf{p}$.

2 Scattering and tunnelling in one dimension

(See Hannabuss chapter 5 for this material.)

For the most part we consider stationary finite norm states. In scattering theory in one dimension, we will seek to calculate the probabilities of particles crossing and being reflected from a time independent potential $V(x)$ with Hamiltonian $H = P^2/2m + V(x)$. Although the potential is time independent, in order to work with finite norm wave functions we would need to consider *wave packets*,[2] However, this is cumbersome as it introduces time dependence and we would like to eliminate the time variable and work with stationary states. Hence we work with a momentum wave function that has fixed momentum k so with wave function $\mathrm{e}^{ikx/\hbar}$ in position space and hence non-normalizable as its magnitude is 1 all the way out to infinity. To obtain finite norm wave packets one could multiply by $\mathrm{e}^{-(k-k_0)^2/\epsilon}/\sqrt{2\pi\epsilon}$ and average over k, but we shall never need to do so. Instead we interpret the non-normalizable wave function as that associated to a continuous beam of particles. Probabilities are obtained via the *probability current*, $\mathbf{j} = \hbar(\bar{\psi}\nabla\psi - \psi\nabla\bar{\psi})/2mi = (\bar{\psi}\mathbf{P}\psi - \psi\mathbf{P}\bar{\psi})/2m$ given simply by

$$j = \frac{\hbar}{2mi}(\bar{\psi}\psi' - \psi\bar{\psi}')$$

---

[2] i.e., wave functions $\mathrm{e}^{ik_0(x-ct)/\hbar-(x-ct)^2/2\sigma^2}$ with $c = k/m$.

in this one-dimensional context. For such energy eigenstates the continuity equation implies

$$\partial_x j = -\partial_t |\psi|^2 = 0\,, \tag{1}$$

because $\psi(x,t) = \mathrm{e}^{Et/i\hbar}\psi(x)$ so the phases cancel in $|\psi|^2$ and hence has no dependence on time. This is the analogue of probability conservation in this context. j is thought of as the rate of incidence of particles.

We will assume that V vanishes for sufficiently large x, although sometimes it is merely taken to be constant for sufficiently large x, $V = V_L$ as $x \to -\infty$ and $V = V_R$ as $x \to +\infty$ (although much of what follows is true for V becoming constant to $O(1/x^2)$ as $|x| \to \infty$).

At fixed energy $E > V_L, V_R$, so that the particle has enough energy to escape to both $x = \pm\infty$ (otherwise we will be dealing with a bound state or for $V_R > E > V_L$ a barrier in which the particle cannot escape to $x = +\infty$). This gives asymptotic wave functions

$$\psi_L = A_L \mathrm{e}^{ik_L x} + B_L \mathrm{e}^{-ik_L x}\,, \qquad \frac{\hbar^2 k_L^2}{2m} = E - V_L \tag{2}$$

$$\psi_R = A_R \mathrm{e}^{ik_R x} + B_R \mathrm{e}^{-ik_R x}\,, \qquad \frac{\hbar^2 k_R^2}{2m} = E - V_R \tag{3}$$

The probability current for ψ_L is easily calculated to be

$$j_L = \frac{\hbar k_L}{m}(|A_L|^2 - |B_L|^2)\,.$$

We interpret this as an incoming beam from the left with wave function $A_L\mathrm{e}^{ik_L x}$ with momentum $\hbar k_L$ and probability current $\hbar k_L |A_L|^2/m$ moving to the right and an outgoing beam $A_L\mathrm{e}^{ik_L x}$ with probability current $-\hbar k_L |B_L|^2/m$ moving to the left. We give a similar interpretation to ψ_R as $x \to +\infty$. The current conservation (1) gives $j_L = j_R$ so that

$$k_L(|A_L|^2 - |B_L|^2) = k_R(|A_R|^2 - |B_R|^2)\,. \tag{4}$$

This encodes probability conservation. When when $k_L = k_R$, (4) gives

$$|A_L|^2 + |B_R|^2 = |A_R|^2 + |B_L|^2$$

and this is interpreted as probability conservation or unitarity for the *scattering matrix* S, $SS^* = 1$. This expresses the outgoing coefficients (B_L, A_R) in terms of the incoming (A_L, B_R)

$$\begin{pmatrix} B_L \\ A_R \end{pmatrix} = S \begin{pmatrix} A_L \\ B_R \end{pmatrix} \tag{5}$$

which exists because either (B_L, A_R) or (A_L, B_R) are linear coordinates on the two dimensional vector space of solutions to the Schrodinger equation and are therefore related by a linear transformation. The constants A_L and B_R are the amplitudes of the incoming particles and B_L and A_R are the amplitudes of the outgoing particles so the unitarity of the scattering matrix S reflects probability conservation from ingoing to outgoing. This is the scattering concept that most naturally extends to higher dimensions.

One key application is the following:

Scattering of incident beam from left: Fire a beam of particles of fixed energy E from $x = -\infty$ at an obstacle represented by a potential $V(x)$. Some particles will be transmitted through the barrier, and some reflected, so we impose the boundary condition $B_R = 0$ so that there are no incoming particles from the right.

Definition 2.1 *The reflection and transmission coefficients are respectively*

$$R = \frac{|B_L|^2}{|A_L|^2}\,, \qquad T = \frac{k_R|A_R|^2}{k_L|A_L|^2}\,.$$

Proposition 2.1 *We have probability conservation $R + T = 1$.*

Proof: This follows immediately from (4) with $B_R = 0$. □

Step function potentials: An useful class of examples follows from considering potentials that are constant on a sequence of intervals

$$V(x) = V_i \in \mathbb{R}\,, \qquad x \in (a_{i-1}, a_i]$$

with $a_0 = -\infty$ and $a_n = +\infty$ for some n, $V_L = V_1$ and $V_n = V_R$. On the ith interval $(a_{i-1}, a_i]$ we have solutions to the time independent Schrodinger equation

$$\psi_i = A_i e^{ik_i x} + B_i e^{-ik_i x}\,, \qquad \hbar^2 k_i^2/2m = E - V_i\,.$$

Since V is discontinuous, the junction conditions for ψ at the a_i are not completely obvious. We require

$$\psi_i(a_i) = \psi_{i+1}(a_i)\,, \qquad \psi_i'(a_i) = \psi_{i+1}'(a_i)\,. \tag{6}$$

Explicitly, this gives equations

$$A_i e^{ik_i a_i} + B_i e^{-ik_i a_i} = A_{i+1} e^{ik_{i+1} a_i} + B_{i+1} e^{-ik_{i+1} a_i}\,,$$

and

$$k_i A_i \mathrm{e}^{ik_i a_i} - k_i B_i \mathrm{e}^{-ik_i a_i} = k_{i+1} A_{i+1} \mathrm{e}^{ik_{i+1} a_i} - k_{i+1} B_{i+1} \mathrm{e}^{-ik_{i+1} a_i} \,.$$

The continuity of ψ and ψ' in particular implies the continuity of the probability current so that we will still have for such a potential that j is constant with $j_L = j_R$ so Proposition 2.1 will hold.

The junction conditions are a pair of equations that can be used to solve for (A_i, B_i) in terms of (A_{i-1}, B_{i-1}) and the solution can be expressed as

$$\begin{pmatrix} A_i \\ B_i \end{pmatrix} = M_i \begin{pmatrix} A_{i+1} \\ B_{i+1} \, , \end{pmatrix}$$

with

$$M_i = \frac{1}{2k_i} \begin{pmatrix} s_i \mathrm{e}^{-id_i a_i} & d_i \mathrm{e}^{-is_i a_i} \\ d_i \mathrm{e}^{is_i a_i} & s_i \mathrm{e}^{id_i a_i} \end{pmatrix} \,, \qquad s_i = k_i + k_{i+1} \,, \quad d_i = k_i - k_{i+1} \,.$$

When $E - V_i < 0$, the same formula holds, but k_i is now imaginary with $k_i = -i\kappa_i$ and $\kappa_i = \sqrt{|E - V_i|}$. With this we can systematically solve for an arbitrary sequence of potential jumps.

Single potential barrier: We consider the case of a single step with $(a_0, a_1, a_2, a_3) = (-\infty, 0, a, \infty)$. On $(-\infty, 0)$ and (a, ∞) we take $V_0 = V_2 = 0$ and in the middle $(0, a)$, we take $V_1 = V$ so that $k = k_L = k_0 = k_2 = k_R$ and $k' = k_1$.

There are three non-trivial cases (i) $E > V$, (ii) $V > E > 0$ and (iii) $0 > E > V$. In the first two we can consider scattering of an incoming beam from the left with $B_R = 0$ so no incoming beam from the right. Inn (i) the particle classically has enough energy to cross the barrier, but in the second case it does not classically. We easily solve in case (i) to find

$$\begin{pmatrix} A_L \\ B_L \end{pmatrix} = M_1 M_2 \begin{pmatrix} A_R \\ 0 \end{pmatrix} = \frac{\mathrm{e}^{ika}}{kk'} \begin{pmatrix} kk' \cos k'a - i(k^2 + k'^2) \sin k'a \\ -i(k^2 - k'^2) \sin k'a \end{pmatrix} A_R$$

and so we find the reflection and transmissions coefficients

$$R := \left| \frac{B_L}{A_L} \right|^2 = \frac{(k^2 - k'^2)^2 \sin^2 k'a}{4k^2 k'^2 \cos^2 k'a + (k^2 + k'^2)^2 \sin^2 k'a}$$

$$T := \left| \frac{A_R}{A_L} \right|^2 = \frac{4k^2 k'^2}{4k^2 k'^2 \cos^2 k'a + (k^2 + k'^2)^2 \sin^2 k'a} \,.$$

It is easily checked that they add up to one.

In case (ii) an essentially identical calculation with $k' = i\kappa'$ imaginary leads to

$$T = \frac{4k^2\kappa'^2}{4k^2\kappa'^2\cosh^2\kappa' a + (\kappa'^2 - k^2)^2\sinh^2\kappa' a}\,.$$

The remarkable feature here is that this is non zero. Although exponentially suppressed as a becomes large, quantum particles can tunnel through barriers that they do not have sufficient energy to cross classically. Such phenomena underlie nuclear reactors and such commonplace objects as the transistor.

In case (iii) $k = i\kappa$ but k', κ real and we have exponentially increasing solutions as $x \to \pm\infty$. If the coefficients of these vanish, $A_L = B_R = 0$, then we have a bound state which is possible for $k = i\kappa$, κ real, and this then gives a quantization condition on a. With k now imaginary, the scattering matrix is clearly no longer unitary and indeed must have a pole at these values, as the LHS of (5) is non-zero, but $A_L = B_R = 0$. This links scattering theory at complex k to bound states.

3 WKB approximation

(For much of this material see chapter 15 of Hannabuss.)

The importance of the WKB approximation goes beyond a calculational tool to both explain the classical correspondence and to motivate a different point of view on quantum mechanics, the Feynman path-integral. It also yields the rules of the old quantum theory and insights into the relationships between classical and quantum mechanics.

We start by decomposing the wave function into polar form

$$\psi = A\mathrm{e}^{iS/\hbar}$$

where A and S are real functions of $\mathbf{r}$ and t with $A \geq 0$.

Lemma 3.1 *In terms of A and S the Schrodinger equation with $H = P^2/2m + V(X)$ becomes the pair of equations*

$$\frac{\partial S}{\partial t} + \frac{|\nabla S|^2}{2m} + V = \frac{\hbar^2}{2m}\frac{\nabla^2 A}{A}\,, \tag{7}$$

$$\frac{\partial A^2}{\partial t} + \nabla\cdot\left(\frac{A^2}{m}\nabla S\right) = 0\,. \tag{8}$$

Proof: Direct calculation yields

$$\nabla\psi = \left(\frac{\nabla A}{A} + \frac{i}{\hbar}\nabla S\right)\psi\,, \qquad \frac{\partial\psi}{\partial t} = \left(\frac{1}{A}\frac{\partial A}{\partial t} + \frac{i}{\hbar}\frac{\partial S}{\partial t}\right)\psi\,,$$

and continuing

$$\nabla^2\psi = \left(\frac{\nabla^2 A}{A} + \frac{i}{\hbar}\nabla^2 S + 2\frac{i}{\hbar}\frac{\nabla A}{A}\cdot\nabla S - \frac{1}{\hbar^2}|\nabla S|^2\right)\psi$$

Substituting these into Schrodinger's equation and dividing by ψ yields a complex equation whose real and imaginary parts are the desired equations after a little manipulation. $\square$

The probability density $|\psi|^2 = A^2$ and the probability current is

$$\mathbf{j} := i\frac{\hbar}{2m}(\psi\nabla\bar{\psi} - \bar{\psi}\nabla\psi) = \frac{A^2}{m}\nabla S\,,$$

so we can interpret (8) as the probability conservation law. This identification of the probability current suggests $\nabla S/m$ as the velocity at $\mathbf{x}$ and we will see that ∇S indeed corresponds to the momentum.

The $\hbar \to 0$ limit follows by ignoring the righthand side of (7). This approximation will also be good for rapid phase oscillation or high energy. This yields

Definition 3.1 (The semi-classical approximation) *which uses the wave function* $\psi = A\mathrm{e}^{iS/\hbar}$ *satisfying*

$$\frac{\partial S}{\partial t} + \frac{|\nabla S|^2}{2m} + V = 0\,, \tag{9}$$

known as the Hamilton Jacobi *equation, and*

$$\frac{\partial A^2}{\partial t} + \nabla\cdot\left(\frac{A^2}{m}\nabla S\right) = 0\,, \tag{10}$$

the continuity equation.

Remark: We expect this approximation to be valid when the right hand side $\hbar^2\nabla^2 A/A$ of (7) is small, in particular, $A \neq 0$.

The Hamilton Jacobi equation arises in classical mechanics as the equation satisfied by the *action* of a classical trajectory ending at the point $\mathbf{x}$ at time t. Our classical equations of motion

$$m\ddot{\mathbf{x}} = -\nabla V \tag{11}$$

arise as the Euler-Lagrange equations that follow from extremizing the action

$$S[\mathbf{X}(t)] = \int_{t_0}^{t} L\,\mathrm{d}t'\,, \qquad \text{where } L = \frac{m}{2}|\dot{\mathbf{X}}|^2 - V(\mathbf{X})\,. \tag{12}$$

In this equation the action $S[\mathbf{X}]$ is thought of as a *functional* of the trajectory $\mathbf{X}(t')$. The solution to the Hamilton-Jacobi equation $S(t, \mathbf{x})$ arises as the value of $S[\mathbf{X}(t')]$ when $\mathbf{X}(t')$ is some family of solutions to the classical equations of motion (11) chosen so that $\mathbf{X}(t) = \mathbf{x}$ and $\mathbf{X}(0) = \mathbf{y}$ (although more generally some other boundary condition can be chosen).

This boundary condition requires us to take the integral along the classical trajectory that joins $\mathbf{y}$ to $\mathbf{x}$. Thus the initial velocity is chosen to that the classical trajectory arrives at $\mathbf{x}$ at time t (so for a free particle we set $\mathbf{v}(0) = (\mathbf{x} - \mathbf{y})/t$). This determines the classical solutions by setting

$$\mathbf{p} = \nabla S\,, \qquad \text{where} \qquad \mathbf{p} = m\dot{\mathbf{x}}$$

is the momentum[3].

The classical equations of motion follow from the Hamilton-Jacobi equations as

$$\frac{\mathrm{d}\mathbf{p}}{\mathrm{d}t} = \nabla \frac{\partial S}{\partial t} + \dot{\mathbf{x}} \cdot \nabla\nabla S = -\nabla V\,.$$

where the first equality uses the chain rule following from the dependence of $\mathbf{x}$ on t, and the second equality follows from the ∇-derivative of (9). Going in the other direction, we only obtain the ∇ derivative of the Hamilton-Jacobi equation, but this can be integrated up to an arbitrary function of t, whose integral can be absorbed into S to yield a genuine solution to the Hamilton-Jacobi equation.

The picture here is that of *geometric optics* in which we go from a wave description with wave fronts given by constant S to particle trajectories defined by the condition that they are perpendicular to the wave fronts of constant S, giving the identification of the momentum as $\mathbf{p} = \nabla S$.

---

[3] This identification follows directly from the variational definition; because the classical trajectory is an extremum for the fixed boundary problem, the variation of the integral under a variation of $\mathbf{x}$ receives only this contribution of $m\dot{\mathbf{x}}$ at the $\mathbf{x}$ endpoint.

For a free classical particle (i.e. $V = 0$) our prescription leads to

$$S = \frac{m|\mathbf{x} - \mathbf{y}|^2}{2t} = \int_0^t \frac{m}{2}\dot{\mathbf{x}}^2 \,\mathrm{d}t\,, \qquad \text{where } \dot{\mathbf{x}} = (\mathbf{x} - \mathbf{y})/t\,. \tag{13}$$

and it can be easily verified that this satisfies (9) with $V = 0$.

For a solution to the Hamilton-Jacobi equation there is always one solution for A given by

Theorem 1 *Given a solution $S(\mathbf{x}, \mathbf{y})$ to the Hamilton-Jacobi equation, a solution to the continuity equation* (8) *is given by*

$$A^2 = \det\left(\frac{\partial^2 S}{\partial x_j \partial y_k}\right)\,. \tag{14}$$

Proof: This follows by direct calculation in the 1-dimensional case given by

$$\begin{aligned} \frac{\partial A^2}{\partial t} &= \frac{\partial^3 S}{\partial x \partial y \partial t} \\ &= -\frac{\partial^2}{\partial x \partial y}\left[\frac{1}{2m}\left(\frac{\partial S}{\partial x}\right)^2 + V\right] \\ &= -\frac{\partial}{\partial x}\left(\frac{1}{m}\frac{\partial S}{\partial x}\frac{\partial^2 S}{\partial x \partial y}\right) \\ &= -\frac{1}{m}\frac{\partial}{\partial x}\left(\frac{\partial S}{\partial x}A^2\right)\,. \end{aligned}$$

which gives the continuity condition. □

Thus for our free particle, we obtain

$$A^2 = \det\left(-\frac{m}{t}I\right) = -\left(\frac{m}{t}\right)^3\,,$$

so that the approximate wave function is

$$\psi = \left(\frac{m}{t}\right)^{3/2} \mathrm{e}^{im|\mathbf{x}-\mathbf{y}|^2/2\hbar t}\,,$$

Note that this is exact as $\nabla^2 A = 0$.

For time independent potentials we can separate out the time dependence

Proposition 3.1 *For a time independent potential, the Hamilton Jacobi equation has solutions*

$$S(\mathbf{x},t) = W(\mathbf{x}) - Et$$

provided

$$\frac{|\nabla W|^2}{2m} + V = E\,.$$

The corresponding wave functions

$$\psi = A\mathrm{e}^{i(W-Et)/\hbar}$$

give approximate eigenstates of energy with eigenvalue E.

Proof: This follows by direct substitution. □

3.1 Quantization for WKB bound states

In one dimension this gives

$$\frac{1}{2m}\left(\frac{\mathrm{d}W}{\mathrm{d}x}\right)^2 + V(x) = E\,, \quad p := \frac{\mathrm{d}W}{\mathrm{d}x} = \sqrt{2m(E - V(x))}\,.$$

This can be integrated directly to give

$$W(x) = \int^x \sqrt{2m(E-V(x))}\,\mathrm{d}x$$

and the one-dimensional continuity equation becomes

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(A^2\frac{\mathrm{d}W}{\mathrm{d}x}\right) = 0\,, \quad A^2 = \text{const.}\left(\frac{\mathrm{d}W}{\mathrm{d}x}\right)^{-1}\,.$$

This leads to the time independent wave functions

$$\psi_\pm = \frac{1}{(E-V)^{1/4}}\exp \pm\frac{i}{\hbar}\left(\int_a^x \sqrt{2m(E-V(x))}\,\mathrm{d}x\right)\,. \tag{15}$$

For simple bound states, for example for infinite potential barriers at a, b, we need to impose the boundary conditions $\psi(a) = \psi(b) = 0$ so we set $\psi = \psi_+ - \psi_-$ to get $\psi(a) = 0$ and then $\psi(b) = 0$ gives the quantization rule

$$W(b) - W(a) = \int_a^b \sqrt{2m(E-V(x)}\,\mathrm{d}x = n\pi\hbar\,.$$

This gives directly, for $V =$ constant the correct energy levels

$$E - V = \frac{n^2\pi^2\hbar^2}{2m(b-a)}\,.$$

More generally, $E-V$ will become negative and since $p^2/2m = E-V < 0$, this region is classically inaccessible. Assume that $E - V \geq 0$ on the interval $[a, b]$, $b > a$, being negative outside. The solutions (15) become exponentially decreasing or increasing outside $[a, b]$. If we are working on $\mathbb{R}$, we need exponential decrease for normalizable wave functions as $x \to \pm\infty$, i.e., the solution must have the form

$$\psi = \begin{cases} \frac{C}{\sqrt{|p|}} \exp\left(-\frac{1}{\hbar}\int_b^x |p|\,\mathrm{d}x\right) & \text{for} \quad x > b\,, \\ \frac{C'}{\sqrt{|p|}} \exp\left(\frac{1}{\hbar}\int_a^x |p|\,\mathrm{d}x\right) & \text{for} \quad x < a\,. \end{cases} \tag{16}$$

We therefore need to decide how to continue these exponentially decreasing solutions (15) across a (and b) where $E - V = 0$ to some solution

$$\psi = \frac{C_1}{\sqrt{p}} \exp\left(\frac{i}{\hbar}\int_a^x p\,\mathrm{d}x\right) + \frac{C_2}{\sqrt{p}} \exp\left(-\frac{i}{\hbar}\int_a^x p\,\mathrm{d}x\right)\,,$$

in the physical region.

Proposition 3.2 *For continuation to the exponentially decreasing solution past the turning point at* b *we must have*

$$C_1 = \frac{1}{2}C\mathrm{e}^{i\pi/4}\,, \quad C_2 = \frac{1}{2}C\mathrm{e}^{-i\pi/4}\,, \qquad \psi = \frac{C}{\sqrt{p}}\cos\left(\frac{1}{\hbar}\int_b^x p\,\mathrm{d}x + \frac{\pi}{4}\right)\,,$$

where p *is the positive root. Similarly at* a *we must have*

$$\psi = \frac{C'}{\sqrt{p}}\cos\left(\frac{1}{\hbar}\int_a^x p\,\mathrm{d}x - \frac{\pi}{4}\right)\,,$$

Sketch proof: (Not examinable) At a and b the WKB approximation breaks down as the right hand side of (7), $\partial_x^2 A/A$, is no longer small at a and b. One strategy is to perform an asymptotic analysis around these points taking $E - V \approx (a - x)V'(a)$. One can find exact solution to (7) with $p^2 = (a - x)V'(a)$ in terms of Bessel's functions (or Airy functions) near

a and then match it to the semiclassical solutions on either side of a [see Hannabuss chapter 15.3 or Weinberg 5.7 for details].

An alternative more informative proof is to analytically continue in the complex around these points, i.e., set $x - a = \rho e^{i\phi}$, with ρ sufficiently large that the WKB approximation is valid along the path. This also has its subtleties as, if we continue along the path in the upper half plane we can only fix the coefficient C_2 near a as the C_1 term has an exponentially small coefficient and cannot be determined by the approximation. The C_1 term must be obtained by analytic continuation in the lower half plane. We obtain

$$C_1 = \frac{1}{2} C e^{i\pi/4}, \quad C_2 = \frac{1}{2} C e^{-i\pi/4}, \qquad \psi = \frac{C}{\sqrt{p}} \cos\left(\frac{1}{\hbar} \int_a^x p \, dx + \frac{\pi}{4} \right),$$

the important $\pi/4$ additional phases can be seen to arise from the analytic continuation of the $1/\sqrt{p} \simeq (x-a)^{-1/4}$ factors. A similar analysis follows at $x = b$. See Landau & Lifschitz §47. $\Box$

Corollary 3.1 (Bohr-Sommerfeld quantization rule) *Normalizable semiclassical solutions satisfying the conditions at turning points above are possible iff*

$$\int_{p^2/2m = E - V(x)} p \, dx = (2n+1)\pi\hbar . \tag{17}$$

The integral is understood to be taken over the full classical phase plane orbit given by $p^2 = E - V(x)$ *giving twice its total action as*

$$\int p \, dx = \int p\dot{x} \, dt = \int \frac{p^2}{m} dt = 2 \int (E - V) \, dt .$$

Proof: The arguments of the cos need to agree up to $n\pi$ (a sign can be absorbed into C), so

$$-\left(\frac{1}{\hbar} \int_b^x p \, dx + \frac{\pi}{4} \right) + \frac{1}{\hbar} \int_a^x p \, dx - \frac{\pi}{4} = \int_a^b p \, dx - \frac{\pi}{2} = n\pi$$

and this gives the above when integrated over both branches of $p = \pm\sqrt{E - V}$ for p both positive and negative in the opposite direction so doubling the answer. $\Box$

This gives surprisingly good answers. For example, it is exact for the simple harmonic oscillator (see the problem sheets).

The technique extends to three dimensions directly in the spherically symmetric case. Here $\nabla^2\psi(r) = \frac{1}{r}\partial_r^2(r\psi)$, so $r\psi$ satisfies the one dimensional Schrodinger equation so we can write, as above,

$$\psi(r) = \frac{\exp \pm i \int^r \sqrt{2m(E - V(r))}}{r(E - V(r))^{1/4}}$$

The semiclassical wave function must still satisfy proposition 3.2 at $r = a$ where $E - V(a) = 0$. The main new ingredient is that $r\psi$ should vanish at the origin so that the argument of the cosine there should be $(n + \frac{1}{2})\pi$ leading to

$$\frac{1}{\hbar}\int_0^a p\,\mathrm{d}x = n\pi + \frac{3\pi}{4}\,.$$

This yields, for example, good estimates for the energy levels for the s-orbitals[4] of hydrogen with errors that are small as $n \to \infty$.

Counting numbers of states: An application of (17) arises from turning it into an area integral to give an estimate of the number of quantum states corresponding to a given region A in the phase plane

$$n = \#\text{ states } \simeq \frac{1}{2\pi\hbar}\int_{p=\sqrt{2m(E-V(x))}} p\,\mathrm{d}x = \frac{1}{2\pi\hbar}\int\int_A \mathrm{d}p\,\mathrm{d}x\,,$$

where the second equality follows from the first as the area under the curve $p = \pm\sqrt{2m(E-V)}$ in the (p, x) phase-plane and A is the region $p^2 \leq 2m(E - V)$. (The formula works identically in the (p, r) phase plane in the spherically symmetric case.) Since wave functions decay exponentially fast outside the region, this number gives an estimate of the number of states whose wave functions are supported in A, since they can have at most n nodes.

With more degrees of freedom, or in higher dimensions, we can apply the estimate in each degree of freedom (or dimension) separately and multiply the answers together, to obtain a phase space $2d$-volume integral $\prod_{i=1}^d \mathrm{d}p_i\,\mathrm{d}x^i/(2\pi\hbar)$ when there are d degrees of freedom.

---

[4] For higher spherical harmonics the problem is more subtle as there is then an inner turning point when the centrifugal barrier kicks in as now $V(r) = l(l+1)/r^2 - e^2/4\pi\epsilon_0 r$. One then has to match to a solution that is regular at $r = 0$ going like r^{l+1} rather than the one that goes like r^{-l}.

3.2 The Feynman Path Integral

In general, this 'semi-classical solution' can be seen as the first term in an asymptotic series[5] by writing

$$\psi = \mathrm{e}^{iS/\hbar}\left(A + \sum_{n=1}^{\infty} A_n \hbar^n\right) .$$

Feynman saw that the whole series could be understood in terms of the classical action by introducing a *path-integral*

$$\psi(\mathbf{x}, T) = \int D[\mathbf{X}(t)]\mathrm{e}^{iS[\mathbf{X}(t)]/\hbar}$$

where the left hand side is the wave function for a particle that starts at the point $\mathbf{y}$ at time 0, and the integral is infinite dimensional over the space of all paths $\mathbf{X}(t)$ from $\mathbf{y}$ at time 0 to $\mathbf{x}$ at time T and S is the action (12). As $\hbar \to 0$ we obtain the asymptotic series via the *stationary phase* approximation in which the rapid oscillations cancel out all the contributions except those at the extremal for the action, i.e. the classical trajectories.

4 Particle statistics

To treat larger atoms we need to know what happens when there is more than one electron. A first approximation to this will be to consider wave-functions for n electrons $\psi(\mathbf{r}_1, \ldots, \mathbf{r}_n)$ with Hamiltonian given by the sum of the 1-particle hydrogen-like ion Hamiltonians in each variable (ignoring interactions between the electrons etc.). The modulus squared of the wave function $|\psi(\mathbf{r}_1, \ldots, \mathbf{r}_n)|^2$ can be thought of as the probability distribution (probability per unit volume) for finding particle 1 at $\mathbf{r}_1$, 2 at $\mathbf{r}_2$ and so on.

4.1 Distinguishable particles

Thus for n particles we can consider wave functions $\psi(\mathbf{r}_1, \ldots, \mathbf{r}_n)$.

---

[5] Recall that an asymptotic series in a parameter,here $\hbar$, is one in which the exact solution differs from the sum of the first n terms by an expression of the order of $\hbar^{n+1}$. Such series are not necessarily convergent, but nevertheless often give very good approximations at finite order.

More abstractly we say that if $\mathcal{H}_i$ is the Hilbert space of the ith particle for $i = 1, 2$, then the Hilbert space $\mathcal{H}$ for system consisting of both the first and second particle is the *tensor product* $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$

Definition 4.1 *Let the vector spaces* $\mathcal{H}_1$ *and* $\mathcal{H}_2$ *have bases respectively* e_i *and* f_j *with* $i = 1, \ldots . \dim \mathcal{H}_1$ *and* $j = 1, \ldots . \dim \mathcal{H}_2$ *(possibly infinite). Then the tensor product is the vector space with basis given by formal expressions* $e_i \otimes f_j$*, i.e.:*

$$\mathcal{H}_1 \otimes \mathcal{H}_2 := \operatorname{span}\{e_i \otimes f_j, i = 1, \ldots \dim \mathcal{H}_1, j = 1, \ldots \dim \mathcal{H}_2\} .$$

Physicists prefer not to use the $\otimes$ notation, and denote $e_i \otimes f_j$ by $|e_i, f_j\rangle$ or more simply $|i, j\rangle$. The inner product on $\mathcal{H}_1 \otimes \mathcal{H}_2$ is induced by that on $\mathcal{H}_1$ and $\mathcal{H}_2$ by

$$\langle i_1, j_1 | i_2, j_2\rangle = \langle i_1 | i_2\rangle\langle j_1 | j_2\rangle .$$

In the finite dimensional case its easily seen that $\dim \mathcal{H}_1 \otimes \mathcal{H}_2$ is $\dim \mathcal{H}_1 \times \dim \mathcal{H}_2$ as that is the number of basis elements $|i, j\rangle$.

This structure emerges from separation of variables in the case of wave functions. When $e_i(\mathbf{r})$ is an orthonormal basis for $\mathcal{H}_1$ and f_j one for $\mathcal{H}_2$ then separable wavefunctions of the form $e_i(\mathbf{r}_1)f_j(\mathbf{r}_2)$ span the set of all two particle wave functions $\psi(\mathbf{r}_1, \mathbf{r}_2) \in \mathcal{H}_1 \otimes \mathcal{H}_2$. Note that even when $\mathcal{H}_1 = \mathcal{H}_2$ and $e_i = f_i$, $e_i \otimes e_j \neq e_j \otimes e_i$ for $i \neq j$: the probability that particle one is in state i and two in state j is not the same as particle one being in state j and two in state i. The order matters for distinguishable particles.

We can similarly form the n-fold tensor product $\mathcal{H}_1 \otimes \mathcal{H}_2 \otimes \ldots \otimes \mathcal{H}_n$ with basis $e_{i_1} \otimes f_{i_2} \otimes \ldots \otimes h_{i_n}$ or simply $|i_1, i_2, \ldots, i_n\rangle$ with i_j indexing a basis of $\mathcal{H}_j$. This is the Hilbert space for the product of the n quantum mechanical systems $\mathcal{H}_i$, $i = 1, \ldots, n$. When the Hilbert spaces are all identical to $\mathcal{H}$ we simply write $\otimes^n \mathcal{H}$.

4.2 Indistinguishable particles

A striking feature of the quantum mechanics of fundamental particles is that particles such as the electron are not just identical but indistinguishable; we do *not* allow wave functions that would let you tell them apart.

A two particle wave function is a function of two positions $\psi(\mathbf{r}_1, \mathbf{r}_2)$. We are not allowed to distinguish particle one being at $\mathbf{r}_1$ and particle 2 at $\mathbf{r}_2$ so we require that

$$\psi(\mathbf{r}_1, \mathbf{r}_2) = \lambda\psi(\mathbf{r}_2, \mathbf{r}_1)$$

where λ is a phase. Iterating this relation we see that

$$\psi(\mathbf{r}_1, \mathbf{r}_2) = \lambda^2 \psi(\mathbf{r}_1, \mathbf{r}_2)$$

so that we see that we have just two possibilities, $\lambda = \pm 1$.

When we have many particles, given a permutation $\pi \in S_n$ the permutation group on n letters, we will require that

$$\psi_(\mathbf{r}_1, \ldots, \mathbf{r}_n) = \lambda(\pi)\psi(\mathbf{r}_{\pi(1)}, \ldots, \mathbf{r}_{\pi(n)})\,, \tag{18}$$

again with $\lambda(\pi)$ a phase. Following π by σ we find

$$\lambda(\pi \circ \sigma) = \lambda(\pi)\lambda(\sigma)\,.$$

This in particular implies that

$$\lambda(\sigma \circ \pi \circ \sigma^{-1}) = \lambda(\sigma)\lambda(\pi)\lambda(\sigma)^{-1} = \lambda(\pi)\,.$$

A transpositions consists of just swapping r and s and is denoted $(r\,s)$. Such transpositions are all conjugate:

$$(r\,s) = (1\,r)(2\,s)(1\,2)(2\,s)^{-1}(1\,r)^{-1}\,.$$

Thus from the above argument, $\lambda((r\,s)) = \pm 1$.

Permutations are all generated by transpositions, and are either even or odd according to whether they are made up of an odd or even number of transpositions. Thus we have proved

Proposition 4.1 *Let* $\psi \to \lambda(\pi)\psi$ *as in* (18) *for* $\pi \in S_n$. *Then either* $\lambda(\pi) \equiv 1$ *or* $\lambda(\pi) = \epsilon(\pi)$ *where*

$$\epsilon(\pi) := \begin{cases} 1 & \text{for } \pi \text{ even} \\ -1 & \text{for } \pi \text{ odd.} \end{cases}$$

Definition 4.2 *Indistinguishable particles satisfying* (18) *are (a) bosons if they satisfy* (18) *with* $\lambda \equiv 1$*, these are* Bose-Einstein statistics *or (b) fermions if they satisfy* (18) *with* $\lambda(\pi) = \epsilon(\pi)$*, and these are* Fermi-Dirac statistics.

Bosons include photons, gravitons and composite particles made up of even numbers of fermions such as mesons. Fermions include electrons, positrons, protons, neutrons, quarks. The general rule is

Theorem 2 (Spin-statistics theorem) *If a particle has integral spin it is a boson. Otherwise, if it is an integer plus 1/2, it is a fermion.*

We will be giving a full treatment of spin in the next section.

4.3 Explicit wave functions

Given $\lambda(\pi) = 1$ or $\epsilon(\pi)$ we can define projectors onto boson/fermion states from arbitrary wave functions by

$$Q_\lambda \psi = \frac{1}{n!} \sum_{\pi \in S_n} \lambda(\pi) \psi(\mathbf{r}_{\pi(1)}, \ldots, \mathbf{r}_{\pi(n)}) \,.$$

Proposition 4.2 *For $\sigma \in S_n$ we have*

$$(Q_\lambda \psi)(\mathbf{r}_{\sigma(1)}, \ldots, \mathbf{r}_{\sigma(n)}) = \lambda(\sigma)(Q_\lambda \psi)(\mathbf{r}_1, \ldots, \mathbf{r}_n) \,.$$

and $Q_\lambda^2 = Q_\lambda$ so that Q_λ is a projection onto boson/fermion states.

Proof: Direct calculation

$$\begin{aligned}
(Q_\lambda \psi)(\mathbf{r}_{\sigma(1)}, \ldots, \mathbf{r}_{\sigma(n)}) &= \frac{1}{n!} \sum_{\pi \in S_n} \lambda(\pi) \psi(\mathbf{r}_{\pi \circ \sigma(1)}, \ldots, \mathbf{r}_{\pi \circ \sigma(n)}) \\
&= \frac{1}{n!} \sum_{\pi \in S_n} \lambda(\pi \circ \sigma^{-1}) \psi(\mathbf{r}_{\pi(1)}, \ldots, \mathbf{r}_{\pi(n)}) \\
&= \lambda(\sigma)^{-1} (Q_\lambda \psi)(\mathbf{r}_1, \ldots, \mathbf{r}_n) \\
&= \lambda(\sigma) (Q_\lambda \psi)(\mathbf{r}_1, \ldots, \mathbf{r}_n) \,.
\end{aligned}$$

to get to the second line we observed that as π ranges over S_n, so does $\pi \circ \sigma$ and so we can replace the latter by the former. Using this we see

$$\begin{aligned}
Q_\lambda Q_\lambda \psi &= \frac{1}{n!} \sum_{\pi \in S_n} \lambda(\pi) (Q_\lambda \psi)(\mathbf{r}_{\pi(1)}, \ldots, \mathbf{r}_{\pi(n)}) \\
&= \frac{1}{n!} \sum_{\pi \in S_n} \lambda(\pi)^2 (Q_\lambda \psi)(\mathbf{r}_1, \ldots, \mathbf{r}_n) \\
&= (Q_\lambda \psi)(\mathbf{r}_1, \ldots, \mathbf{r}_n) \,.
\end{aligned}$$

Thus the result follows. $\square$

For $n = 2$ these are just the symmetric and antisymmetric combinations

$$Q_1 \psi(\mathbf{r}_1, \mathbf{r}_2) = \frac{1}{2}(\psi(\mathbf{r}_1, \mathbf{r}_2) + \psi(\mathbf{r}_2, \mathbf{r}_1)) \,, \qquad Q_\epsilon \psi(\mathbf{r}_1, \mathbf{r}_2) = \frac{1}{2}(\psi(\mathbf{r}_1, \mathbf{r}_2) - \psi(\mathbf{r}_1, \mathbf{r}_2)) \,.$$

If we start with separable wave functions

$$\psi_(\mathbf{r}_1, \ldots, \mathbf{r}_n) = \psi_1(\mathbf{r}_1) \psi_2(\mathbf{r}_2) \ldots \psi_n(\mathbf{r}_n) \,,$$

then in the antisymmetric case we obtain a *Slater determinant*

$$Q_\epsilon \psi = \begin{vmatrix} \psi_1(\mathbf{r}_1) & \psi_2(\mathbf{r}_1) & \dots & \dots & \psi_n(\mathbf{r}_1) \\ \psi_1(\mathbf{r}_2) & \psi_2(\mathbf{r}_2) & \dots & \dots & \psi_n(\mathbf{r}_2) \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \psi_1(\mathbf{r}_n) & \dots & \dots & \dots & \psi_n(\mathbf{r}_n) \end{vmatrix} .$$

Since arbitrary states can be expressed as a linear combination of separable states, all fermionic states can be obtained as a linear combination of these.

From the standard properties of determinants, this wave function vanishes identically if two of the ψ_i are proportional This is a severe restriction on the number of possible states. We can count the number of states as follows. Suppose the $\psi_i(\mathbf{r}) \in \mathcal{H}$ where $\mathcal{H}$ has dimension N (i.e., some energy eigenstate for a hydrogen-like ion).

Lemma 4.1 *The space of fermionic n-particle states built out of wave functions in an N-dimensional Hilbert space $\mathcal{H}$ is $\begin{pmatrix} N \\ n \end{pmatrix}$-dimensional.*

Proof: We can choose a basis for $\mathcal{H}$ and the ψ_i to be elements of that basis. There are N choices for ψ_1, but since ψ_2 cannot be the same as ψ_1, there are $N-1$ choices for ψ_2 and so on. The final state is independent of the ordering of $\psi_1, \dots, \psi_n$ so the answer follows. $\square$.

Corollary 4.1 *For N-dimensional $\mathcal{H}$, we can have at most N fermionic particles.*

This leads to:

The Pauli exclusion principle: *Two fermions cannot occupy the same state.*

This will become important for understanding atomic structure when there are many electrons.

Definition 4.3 *In terms of tensor products, the symmetric bosonic states, the image of Q_1, are denoted by $\odot^n \mathcal{H}$ and the antisymmetric fermionic states, the image of Q_ϵ are denoted $\wedge^n \mathcal{H}$.*

When $\mathcal{H}$ has dimension N, we can ask how many bosonic or Fermionic states are there, i.e., what is the dimension of $\odot^n \mathcal{H}$ or $\wedge^n \mathcal{H}$? (It is easily seen that $\otimes^n\mathcal{H}$ has dimensions N^n as that is the number of basis elements $|i_1,\ldots,i_N\rangle$ with each $i_j = 1,\ldots,N$.)

The bosonic case: For the bosonic case, the corresponing symmetrization of separable states leads to

$$Q_1(\psi_1(\mathbf{r}_1)\ldots\psi_n(\mathbf{r}_n)) = \frac{1}{n!}\sum_{\pi\in S_n}\psi_1(\mathbf{r}_{\pi(1)})\ldots\psi_n(\mathbf{r}_{\pi(n)})$$

Lemma 4.2 *For $\psi_i \in \mathcal{H}$ of dimension N there are $(N+n-1)!/(N-1)!n!$ linearly independent n-particle bosonic states.*

In order to prove this we introduce a generating function known as a *partition function* that has much wider applicability.

Proof: In general such a separable bosonic state can be represented as

$$Q_1\left(\psi_1^{k_1}\psi_2^{k_2}\ldots\psi_N^{k_N}\right)\,, \qquad \sum k_i = n$$

and so we need to count the number of such integer 'partitions of n', $k_i \in \mathbb{N}$ that add up to n. We can replace the ψ_i by formal variables x_i and so we are counting the number of distinct monomials of the form $x_1^{k_1}x_2^{k_2}\ldots x_N^{k_N}$. If we multiply each x_i by s, then the total power of s will be the sum of the k_i. Form the sum to find

$$\sum_{k_1,\ldots,k_N\in\mathbb{N}}(sx_1)^{k_1}(sx_2)^{k_2}\ldots(sx_N)^{k_N} = \prod_{j=1}^{N}\sum_{k_j\in\mathbb{N}}(sx_j)^{k_j} = \prod_{j=1}^{N}\frac{1}{1-sx_j}\,.$$

Now if we set all the $x_i = 1$, we obtain $1/(1-s)^N$ and the coefficient of s^n will simply count the number of terms. The binomial theorem then gives

$$\frac{1}{(1-s)^N} = \sum_{n\in\mathbb{N}}\binom{N+n-1}{n}s^n\,,$$

from which the result follows. $\Box$

This is an example of a partition function (introduced originally to count partitions). Quantum mechanically, we can think of this partition function as arising from considering the formal direct sum $\oplus_n \odot^n \mathcal{H}$ over all n-particle states for each n and taking the trace of $s^{\mathcal{N}}$ where $\mathcal{N}$ is the particle number operator that takes the value n on n-particle states $\odot^n\mathcal{H}$. On the n-particle eigenspace $\odot^n\mathcal{H}$ of $\mathcal{N}$, the trace of $s^{\mathcal{N}}$ reduces to the trace of $s^n\mathbb{1}$ and so gives $s^n \times \dim \odot^n \mathcal{H}$. The total trace is the sum of this quantity over n. The corresponding function for counting the number of fermions $\mathrm{tr} s^{\mathcal{N}}$ now on $\oplus_n \wedge^n \mathcal{H}$ is simply $(1+s)^N$ which of course terminates at s^N. This idea can be extended to count particles of different types and energies and plays a major role in statistical mechanics.

4.4 Applications: condensates & Neutron stars

Bosonic phenomena are associated with gregarious behaviour, and in particular lead to some exotic phenomena associated with 'Bose-Einstein condensates'. These are associated with very cold situations where almost all the particles collect in the ground state such as superfluidity and superconductivity.

Fermionic phenomena are on the other hand associated with exclusive behaviour. A dramatic example of this is the Neutron star which are so dense that the gravitational pull has compressed all the electrons into the protons to make neutrons. The reason it doesnt collapse further into a black hole is that it is supported simply by fermionic exclusion giving rise to 'quantum degeneracy pressure.

Fermionic exclusion underpins some of the most dramatic successes of quantum theory, in particular the qualitative theory of atomic physics when there are larger numbers of particles. However, to count the number of states correctly, we need to incorporate the intrinsic spin of the electrons (and nucleons for nucleus) so we defer the detailed discussion until after the treatment of the quantum mechanics of spin and angular momentum.

Alternative reading: See §16.1-4 of Hannabuss or §4.5 of Weinberg.

5 Rotations, angular momentum and spin

To obtain the full set of degeneracies of the hydrogen atom we need to take into account the combination of the orbital angular momentum and intrinsic spin. Here we will see that the angular momentum of a system arises from the action of the rotation group. For a particle it can be decomposed into the sum of an orbital part and an intrinsic 'spin'.

In the first course on geometry we learnt about the orthogonal group as a group of matrices acting on cartesian coordinates as $x^i = (x^1, x^2, x^3)$,

$i, j = 1, 2, 3$, by

$$x^i \to \sum_{j=1}^{3} R_{ij} x^j \qquad RR^t = \mathbb{1}\,,$$

where $\mathbb{1}$ is the identity matrix. $\mathrm{Det}(R) = \pm 1$ and if $+1$, R is a rotation.

We usually work with infinitesimal rotations. We consider a one parameter family of rotations $R = R(t)$ with $R(0) = \mathbb{1}$. For t small we can write

$$R_{ij}(t) = \delta_{ij} + t\omega_{ij} + O(t^2)\,. \tag{19}$$

Expanding $R^t(t)R(t) = \mathbb{1}$ to first order in t gives

$$\omega_{ij} + \omega_{ji} = 0$$

or in matrix notation $\omega + \omega^t = 0$ so ω is *skew symmetric* with just three nontrivial components. The components of ω_{ij} can be encoded into the 3-vector $\boldsymbol{\omega} = -(\omega_{23}, \omega_{31}, \omega_{12})$. A useful notational device is the 'epsilon symbol'

$$\varepsilon_{ijk} = \begin{cases} 1 \;\text{ for } (ijk) \text{ an even permutation of } (123) \\ -1 \;\text{ for } (ijk) \text{ an odd permutation of } (123) \\ 0 \;\text{ otherwise.} \end{cases}$$

With this we can define the vector $\boldsymbol{\omega}$ by[6]

$$\omega_i = \frac{1}{2}\sum_{j,k} \varepsilon_{ijk}\omega_{jk}\,, \qquad \omega_{ij} = \sum_k \epsilon_{ijk}\omega_k\,.$$

A rotation is determined by its axis and angle of rotation, and so the direction of $\boldsymbol{\omega}$ determines the axis and the length the (infinitesimal) angle. The epsilon symbol also defines the wedge product

$$(\mathbf{v} \wedge \mathbf{w})_i = \sum \varepsilon_{ijk} v_j w_k\,,$$

and the double wedge product identity $\mathbf{u} \wedge (\mathbf{v} \wedge \mathbf{w}) = (\mathbf{u} \cdot \mathbf{w})\mathbf{v} - (\mathbf{u} \cdot \mathbf{v})\mathbf{w}$ is expressed as

$$\sum_k \varepsilon_{ijk}\varepsilon_{klm} = \delta_{il}\delta_{jm} - \delta_{im}\delta_{jl}.$$

---

[6] We briefly mention here the Einstein summation convention that would allow us to assume that any pair of repeated indices should be summed without explicitly introducing the summation symbol, but in this course there are not enough such indices to make it worth introducing and using this convention systematically.

Thus the first order action of $R(t)$ on x^i is

$$\sum_j \omega_{ij} x_j = (\boldsymbol{\omega} \wedge \mathbf{x})_i \,. \tag{20}$$

The fact that rotations do not commute, $RR' \neq R'R$ in general, is expressed infinitesimally in the commutator: taking t, t' small in $R(t)$, $R'(t')$ we find, suppressing indices

$$R(t)R'(t') - R'(t')R(t) = (\mathbb{1} + t\omega + \ldots)(\mathbb{1} + t'\omega' + \ldots) - \leftrightarrow = tt'[\omega, \omega'] + \ldots$$

where in indices

$$[\omega, \omega']_{ik} := \sum_j \omega_{ij}\omega'_{jk} - \omega'_{ij}\omega_{jk} = \varepsilon_{ikl}(\boldsymbol{\omega} \wedge \boldsymbol{\omega}')_l \tag{21}$$

In the first instance, a rotation acts on an ordinary wave function by

$$\psi(\mathbf{x}) \to U(R)\psi := \psi(R\mathbf{x}) \,.$$

This action is obviously complex linear because

$$U(R)(a_1\psi_1 + a_2\psi_2) = a_1\psi_1(R\mathbf{x}) + a_2\psi_2(R\mathbf{x}) = a_1 U(R)\psi_1 + a_2 U(R)\psi_2 \,.$$

It is also unitary

$$\langle U(R)\psi | U(R)\psi \rangle = \int_{\mathbb{R}^3} |\psi(R\mathbf{x})|^2 \, \mathrm{d}^3\mathbf{x} = \int_{\mathbb{R}^3} |\psi|^2 \, \mathrm{d}^3\mathbf{x} \langle \psi | \psi \rangle \,,$$

because of the change of variable formula under $\mathbf{x}' = R\mathbf{x}$ for integration. This can be written[7] $U^*U = 1$. Clearly under composition

$$U(R_1R_2)\psi(\mathbf{x}) = \psi(R_1R_2\mathbf{x}) = U(R_1)\psi(R_2\mathbf{x}) = U(R_1)U(R_2)\psi(\mathbf{x})$$

so we have[8]

$$U(R_1R_2) = U(R_1)U(R_2) \,. \tag{22}$$

---

[7] this actually is equivalent to $\langle \phi | U^*U\psi \rangle = \langle \phi | \psi \rangle$ but this can be deduced from $\langle \phi + a\psi | U^*U | \phi + a\psi \rangle$ with $a = 1, i$.

[8] It is a general statement that if we have a group G of physical symmetry transformations that they must act on physical states by unitary transformations $U(g)$ for $g \in G$. They must be complex linear because they must take complex linear combinations to complex linear combinations. They must be unitary because they must preserve probability.

Infinitesimal unitary matrices are given by hermitian matrices. To see this, set $U(t) = I + itu + O(t^2)$. Then $U^*U = 1$ gives $u = u^*$ to first order. For infinitesimal rotations we introduce a vector $\mathbf{J}$ of Hermitian operators by

$$\frac{d}{dt}U(1 + t\omega + O(t^2))|_{t=0} = \frac{i}{2\hbar}\sum_{ij}\omega_{ij}J_{ij} = \frac{i}{\hbar}\boldsymbol{\omega}\cdot\mathbf{J}\,. \tag{23}$$

It follows from (22) that

$$U(R_1)U(R_2) - U(R_2)U(R_1) = U(R_1R_2) - U(R_2R_1)$$

and so infinitesimally, with (21) we must have

$$[J_i, J_j] = i\hbar\varepsilon_{ijk}J_k \tag{24}$$

When we are considering the action of rotations on a simple wave function on $\mathbb{R}^3$, we can determine this infinitesimal action directly by the chain rule for $R = R(t)$ of (19)

$$\frac{d}{dt}\psi(R(t)\mathbf{x})|_{t=0} = \frac{d}{dt}\psi(\mathbf{x} + t\omega\wedge\mathbf{x} + \ldots)|_{t=0} = \omega\wedge\mathbf{x}\cdot\nabla\psi(\mathbf{x}) = \omega\cdot(\mathbf{x}\wedge\nabla\psi)\,.$$

Thus in this case we have that $\mathbf{J}$ is given by the *orbital angular momentum* operator

$$\mathbf{L} := \mathbf{X}\wedge\mathbf{P}\,.$$

It can be checked explicitly that these satisfy

$$\begin{aligned}
[L_i, L_j] &= i\hbar\sum_k \varepsilon_{ijk}L_k\,,\\
[L_i, X_j] &= i\hbar\sum_k \varepsilon_{ijk}X_k\,,\\
[L_i, P_j] &= i\hbar\sum_k \varepsilon_{ijk}P_k\,,
\end{aligned}$$

as discussed in the Part A course last year. More generally we will say that an operator with a vector index A_i is a vector operator if the angular momentum operators induce infinitesimal rotations

$$[J_i, A_j] = i\hbar\sum_k \varepsilon_{ijk}A_k$$

and so we see that X_i, P_i, L_i and J_i are all vector operators in this sense.

When we come to study a particle with some internal structure such as a rigid body, we could attempt to model the wave function explicitly. However, the general arguments above must apply to give an action of rotations on the intrinsic spin. Thus we will have a unitary action of the rotations $U(R)$ on states as above satisfying[9] (22,23,24) generated by $\mathbf{J}$ but the quantum state will have extra degrees of freedom over and above its dependence on $\mathbf{x}$ on which the rotations act. The rotations must act on $\mathbf{X}$ and $\mathbf{P}$ via (20) and so subtracting and using the above, $\mathbf{J}-\mathbf{L}$ must commute with $\mathbf{X}$ and $\mathbf{P}$. So we can write

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

where

$$[L_i, S_j] = 0\,, \qquad \text{and} \qquad [S_i, S_j] = i\hbar \sum_k \varepsilon_{ijk} S_k\,.$$

The operator $\mathbf{S}$ is the observable associated with the *intrinsic spin* of the particle. Rather than model this action concretely, we will see in the next subsection that in fact the algebra tells us all that we need to know.

Alternative reading: See §4.1 of Weinberg or §9.1-3 of Hannabuss.

5.1 Spin and angular momentum multiplets

We remark also that for the electron to have its intrinsic spin understood mechanically would require it to be spinning much faster than the speed of light given the upper bounds on its known radius, so although mechanical analogies can provide some intuition, they are unlikely to be fundamental. So, rather than attempt to quantize a rigid body, we analyze algebraically what Hilbert spaces can arise with an action of $\mathbf{J}$ (or $\mathbf{S}$) satisfying the commutation relations (24).

Here we will not distinguish between $\mathbf{J}$, $\mathbf{L}$ or $\mathbf{S}$ and denote any of these by $\mathbf{J}$ and our discussion will be applicable to both intrinsic spin and orbital angular momentum or indeed any other infinitesimal $SO(3)$ symmetry. Define:

Definition 5.1 *A representation of the angular momentum operators is a Hilbert spaces $\mathcal{H}$ with an* action *of Hermitian operators $J_i : \mathcal{H} \to \mathcal{H}$ satisfying*

---

[9] This is only in general required to be true up to a phase, and for rotation groups will only in general be true up to signs.

the commutation relations

$$[J_i, J_j] = i\hbar\varepsilon_{ijk}J_k\,.$$

A representation is said to be irreducible if we cannot non-trivially decompose $\mathcal{H} = \mathcal{H}_1 \oplus \mathcal{H}_2$ with $\mathcal{H}_1$ orthogonal to $\mathcal{H}_2$ such that $J_i : \mathcal{H}_1 \to \mathcal{H}_1$ and $J_i : \mathcal{H}_2 \to \mathcal{H}_2$ (i.e., each is J_i-invariant).

An elementary particle is one that cannot be decomposed into smaller units and hence irreducible: if the intrinsic spin $\mathbf{S}$ is represented on $\mathcal{H}$ reducibly, then we would consider the particle to be composed of two particles, one with states in $\mathcal{H}_1$ and the other in $\mathcal{H}_2$ and hence not elementary.

In last year's course we saw that representations of the orbital angular momentum operators lead to separation of variables in terms of spherical harmonics. The same algebra applies up to considerations of half-integral spin but the ideas are the same.

We first observe that

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

so J_i preserves the different eigenspaces of $\mathbf{J}^2$. Thus if $\mathcal{H}$ is irreducible, $\mathbf{J}^2$ must act by a multiple of the identity on $\mathcal{H}$, otherwise we could decompose it into the different eigenspaces.

We can give a completely explicit description of all finite dimensional irreducible representations if we furthermore diagonalize J_3.

Theorem 3 *The irreducible representations $\mathcal{H}_j$ of the angular momentum operators are labeled by a half-integer $j = 0, 1/2, 1, \ldots \in \mathbb{N}/2$, the* spin. *The dimension of $\mathcal{H}_j$ is $2j + 1$ and $\mathbf{J}^2$ has eigenvalue $j(j + 1)\hbar^2$.*

There is an orthonormal basis of $\mathcal{H}_j$ consisting of eigenvectors ψ^m of J_3 with eigenvalues $m = -j, -j + 1, \ldots j - 1, j$.

Proof: The main idea is to introduce the ladder operators

$$J_\pm = J_1 \pm iJ_2\,.$$

They commute with $\mathbf{J}^2$ from above and we can check

$$[J_3, J_\pm] = \pm\hbar J_\pm\,.$$

This gives them the interpretation as raising and lowering operators for the eigenvectors ψ^m of J_3 with

$$J_3\psi^m = m\hbar\psi^m$$

since

$$J_3 J_\pm \psi^m = \pm\hbar J_\pm \psi^m + J_\pm J_3 \psi^m = (m \pm 1)\hbar J_\pm \psi^m \,.$$

Thus $J_\pm \psi^m$ is a multiple of an eigenvector for J_3 with eigenvalue $m \pm 1$.

Can the eigenvalues for m can be arbitrarily large? No, the following shows that $|m|$ must be bounded.

Lemma 5.1 *Let $\mathbf{J}^2\psi = \lambda\hbar^2\psi$ and $J_3\psi = m\hbar\psi$. Then for all ϕ*

$$\langle J_\pm \phi | J_\pm \psi\rangle = (\lambda - m^2 \mp m)\hbar^2 \langle \phi|\psi\rangle^2 \tag{25}$$

Thus, given that λ is fixed on an irreducible representation, $|m|$ cannot be too large as the left hand sides are positive but for large $|m|$ the right hand sides are negative.

To see this, observe from the commutation relations that

$$J_+J_- = \mathbf{J}^2 - J_3^2 + \hbar J_3 \,, \qquad J_-J_+ = \mathbf{J}^2 - J_3^2 - \hbar J_3 \,, \tag{26}$$

(see problem sheet 1) so the identities follow from

$$\langle J_-\phi|J_-\psi\rangle = \langle \phi|J_+J_-\psi\rangle = \langle\phi|(\mathbf{J}^2 - J_3^2 + \hbar J_3)\psi\rangle \tag{27}$$

and using the eigenvalue relations (and similarly for the plus version of (26)).

The only way that $|m|$ can avoid becoming arbitrarily large is that for some m_-, the smallest, $J_-\psi^{m_-} = 0$ with $\psi^{m_-} \neq 0$ *and* for some j the largest m value has $J_+\psi^j = 0$ with $\psi^j \neq 0$. Thus (25) gives

$$\lambda = j(j+1) = m_-(m_- - 1)\,.$$

Given j, the quadratic equation for m_- has solution $m_- = -j$ or $m_- = j+1$. The second solution is not compatible with j being the largest eigenvalue so $m_- = -j$. Furthermore, by construction $j - m_- = 2j$ must be an integer, hence the conclusions on the eigenvalues.

To finish the theorem, we must show that irreducibility implies that the eigenvalues are nondegenerate. Suppose that there are two linearly independent eigenvectors ψ^m and $\tilde{\psi}^m$ that, without loss of generality can be taken to be orthogonal. Then inductively we see that $J_\pm^n\psi^m$ and $J_\pm^n\tilde{\psi}^m$ are orthogonal using (25). Thus $\mathcal{H}$ has two nontrivial J_i-invariant subspaces spanned by $J_\pm^n\psi^m$ and by $J_\pm^n\tilde{\psi}^m$ contradicting irreducibility. $\square$

Remarks:

- We use the notation ψ_j^m or the kets $|j, m\rangle$ for ψ^m when we want to emphasize the role of j.

- The basis ψ_j^m of $\mathcal{H}_j$ is unique up to an overall phase if we impose the normalization conditions

$$\psi_j^{m\pm1} = \frac{J_\pm \psi_j^m}{\hbar\sqrt{j(j+1) - m(m\pm1)}} \,. \tag{28}$$

  That these provide a definition of $\psi_j^{m\pm1}$ that have unit length given that $||\psi_j^m|| = 1$ follows from (25).

  Exercise: Show using (28) and $J_3\psi_j^m = m\hbar\psi_j^m$ that J_3 and $J_\pm$ satisfy their standard commutation relations.

- Important examples of these representations were given in the first course in terms of spherical harmonics obtained by separation of variables in spherical polar coordinates (r, θ, ϕ). These are representations of the orbital angular momentum $\mathbf{L}$ given as

$$L_\pm = i\hbar \mathrm{e}^{\pm i\phi}\left(\cot\theta \frac{\partial}{\partial\phi} \pm i\frac{\partial}{\partial\theta}\right) , \qquad L_3 = -i\hbar\frac{\partial}{\partial\phi}$$

  in these coordinates. The j value is here usually denoted by j and the Ψ_l^m are then given explicitly in these coordinates as

$$Y_l^m(\theta, \phi) = P_l^m(\theta)\mathrm{e}^{im\phi}$$

  where $P_l^m(\theta)$ were the Legendre polynomials (in $\cos\theta$). Here we require both l and m to be integral so that $\mathrm{e}^{im\theta}$ is single valued. (See the second problem sheet for more on these.)

5.2 Spin 1/2

For intrinsic spin, half-integral spin is acceptable and indeed the electron, proton, quarks etc. have spin 1/2. The above gives the representation explicitly as follows:

$$\mathcal{H}_{1/2} = \mathrm{Span}\{\psi^{1/2}, \psi^{-1/2}\} = \mathbb{C}^2 \,.$$

The above action of $J_\pm$ and hence J_1 and J_2 is determined by (28) for which in this case the denominator is just $\hbar$, and the eigenvalue condition determines J_3. Thus, in this basis,

$$\mathbf{J} = \frac{\hbar}{2}\boldsymbol{\sigma}$$

where $\boldsymbol{\sigma}$ are the *Pauli spin matrices* σ_i

$$\boldsymbol{\sigma} := \frac{2}{\hbar}\langle\psi^m|\mathbf{J}|\psi^{m'}\rangle = (\sigma_1, \sigma_2, \sigma_3) := \left(\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}, \begin{pmatrix}0 & -i\\ i & 0\end{pmatrix}, \begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}\right) .$$

Here we are variously suppressing the i-index or the m, m' indices[10]. These satisfy the remarkable relation

$$\sigma_i\sigma_k = \delta_{ik} + i\sum_l \varepsilon_{ikl}\sigma_l$$

and define the imaginary quaternions. (See problem sheet again.)
Alternative reading: See §8.1-4 of Hannabuss and §4.2 of Weinberg.

5.3 $SU(2)$ and higher spins.*

The Pauli matrices are Hermitian and trace-free and therefore represent infinitesimal generators of the group SU(2). This initially rather surprising connection between a unitary matrix $U(R) \in$ SU(2) and a corresponding orthogonal $R_{ij} \in$ SO(3) can be represented using Pauli matrices by

$$U^*(R)\sigma_i U(R) = \sum_j R_{ij}\sigma_j .$$

We can see that this is a 2:1 map SU(2) → SO(3) as both U and $-U$ determine the same[11] R.

The spin half representation generates all the higher spin representations in the sense that spin $n/2$ can be obtained by considering symmetric products of spin half. The general element of the spin j Hilbert space $\mathcal{H}_j$ can also be represented by a wave function $\psi^{m_1\ldots m_{2j}} \in \mathcal{H}_j = \odot^{2j}\mathcal{H}_{1/2}$ with each $m_i = \pm 1/2$ and independent of the ordering of the m_i, i.e.,

$$\psi^{\ldots m_i \ldots m_j \ldots} = \psi^{\ldots m_j \ldots m_i \ldots} .$$

---

[10] note that we are using $m, m' = \pm 1/2$ to index the rows and columns of the 2×2 matrix rather than the more usual $1, 2$

[11] SU(2) is known as the *spin* group of $SO(3)$. Such 2:1 covers exist for all the orthogonal groups although only in low dimensions is it some other familiar group. For SO(4), the spin group is SU(2) $\times$ $SU(2)$.

Because of this ordering independence only the sum $m = \sum_i m_i$ is important so that there are only $2j + 1$ independent components and this then corresponds to Ψ_j^m. Under an SU(2) transformation $U^m_{m'}$ using $m, m' = \pm 1/2$ to index the rows and columns we have

$$\psi^{m_1 \dots m_{2j}} \to \sum_{m'_1 \dots m'_{2j}} U^{m_1}_{m'_1} \dots U^{m_{2j}}_{m'_{2j}} \psi^{m'_1 \dots m'_{2j}} .$$

A key example of this is spin 1 leading to wave functions $\psi^{m_1 m_2} = \psi^{m_2 m_1}$ and these correspond to 3-vectors via the Pauli matrices $\sum_{m_1 m_2} \psi^{m_1 m_2} \epsilon_{m_1 m_3} \sigma_i^{m_3 m_2}$ where $\epsilon_{m_1 m_2} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$. This ϵ matrix is needed because the first m-index on σ_i is a complex conjugate index from its definition and for a unitary unit determinant 2×2 matrix, $U^* = U^{-1} = \epsilon U \epsilon$. More generally, integral spin j wave functions correspond to an object with j vector indices $A_{i_1 \dots i_j}$, a so-called *tensor*.

For composite particles all spins are allowed. However, there are theorems to the effect that for fundamental particles, spin 2, describing gravity alone, is maximal.

Alternative reading: See §8.8 and 9.9 of Hannabuss or §4.2 of Weinberg.

5.4 Addition of angular momentum

If we want to understand the angular momentum of electron energy levels in a hydrogen atom, the total $\mathbf{J} = \mathbf{L} + \mathbf{S}$ is made up of the sum of the intrinsic spin-1/2 of the electron and the orbital angular momentum. For n particles we will have total angular momentum $\mathbf{J} = \mathbf{J}_1 + \dots + \mathbf{J}_n$ where $\mathbf{J}_r$ is the angular momentum of the rth particle. These can be measured separately or together[12]. The first key question is

Problem: Consider a system with total angular momentum $\mathbf{J} = \mathbf{J}' + \mathbf{J}''$ with $[J'_i, J''_j] = 0$. Consider the Hilbert space $\mathcal{H}$ of states built from those

---

[12] The Stern-Gerlach experiment measures the trajectory of the atom in a magnetic field giving the total angular momentum, or the constituent orbital and intrinsic parts; the total Hamiltonian in a constant magnetic field couples to an electron via an interaction term

$$\frac{\mu}{\hbar j} \mathbf{J} \cdot \mathbf{B}$$

where j is the spin of the total angular momentum, $\mathbf{J}^2 = \hbar^2 j(j+1)$, μ its *magnetic moment*. If $\mathbf{B}$ is constant this gives a way of measuring j and m.

of spin j' for $\mathbf{J}'$ and and j'' for $\mathbf{J}''$, i.e. states $\psi_{j'}^{m'}\psi_{j''}^{m''}$ with $m' = -j', \ldots, j'$ and $m'' = -j'', \ldots, j''$. What values of the total angular momenta j can we have?

The system could for example be a composite of two particles, one of spin j' and one of spin j'', or alternatively a particle of spin j' with orbital angular momentum j''.

Definition 5.2 *In this context we say that $\mathcal{H} = \mathcal{H}_{j'} \otimes \mathcal{H}_{j''}$ is the tensor product of $\mathcal{H}_{j'}$ and $\mathcal{H}_{j''}$. $\mathcal{H}$ has dimension $(2j'+1) \times (2j''+1)$. With this fancier notation we sometimes write the basis $\psi_{j'}^{m'}\psi_{j''}^{m''}$ as $\psi_{j'}^{m'} \otimes \psi_{j''}^{m''}$ or $|j', m'; j'', m''\rangle$.*

Proposition 5.1 *Given constituent angular momenta with spins j' and j'', $j' \geq j'' > 0$, the total angular momentum can be decomposed into single copies of multiplets with spin j taking the values from $|j' - j''|$ to $j' + j''$.*

$$\mathcal{H}_{j'} \otimes \mathcal{H}_{j''} = \oplus_{j=j'-j''}^{j'+j''} \mathcal{H}_j. \tag{29}$$

There is therefore a basis Ψ_j^m labeled by spin $j = j' - j'', \ldots, j' + j''$ and $m = -j, \ldots, j$ the eigenvalue of J_3 with both j and m increasing in unit increments.

Proof: We proceed inductively first finding the representation of highest spin $\mathcal{H}_{j'+j''}$ inside $\mathcal{H}$ and then proceeding to find the representation of next highest spin inside the orthogonal complement $\mathcal{H}_{j'+j''}^{\perp}$ inside $\mathcal{H}$ and so on.

Our basis vectors $\psi_{j'}^{m'}\psi_{j''}^{m''}$ are eigenvectors of $J_3 = J_3' + J_3''$ with eigenvalues $m = m' + m''$. This maximum m state is unique this being associated with the maximum values of $m' = j'$ and $m'' = j''$. We denote this state by

$$\Psi_{j'+j''}^{j'+j''} = \psi_{j'}^{j'}\psi_{j''}^{j''} .$$

In particular, $j = j' + j''$ must be the maximum possible spin. The other spin $j' + j''$ states are obtained by acting $j - m$ times with the lowering operator J_-. Normalizing according to (28), we obtain states $\Psi_{j'+j''}^{m}$ that still have total spin $j = j' + j''$, but now have J_3 eigenvalue m. This gives a complete spin $j = j' + j''$ representation/multiplet $\mathcal{H}_{j'+j''}$ with $m = -j' - j'', \ldots, j' + j''$ inside $\mathcal{H}$.

Now consider the states with $m = j' + j'' - 1$. There are two linearly independent such states, $\psi_{j'}^{j'-1}\psi_{j''}^{j''}$ and $\psi_{j'}^{j'}\psi_{j''}^{j''-1}$. One combination has spin

$j'+j''$ being the $\Psi^{j'+j''-1}_{j'+j''}$ defined above. A vector in the orthogonal complement must have spin $j'+j''-1$ because J_+ acting on it must vanish, otherwise the $m=j'+j''$ eigenspace would have to be two dimensional. By (28) this means it has spin $j'+j''-1$ so we can denote a normalized basis element by $\Psi^{j'+j''-1}_{j'+j''-1}$. By acting with J_- this state now generates a multiplet of spin $j'+j''-1$ again with basis elements $\Psi^m_{j'+j''-1}$ related by (28).

If $j''=1/2$ then the degeneracy for m eigenvalues with $j'+j''-1\geq m\geq -(j'+j''-1)$ is two corresponding to the $m''=\pm 1/2$ and so must be spanned by the m-eigenvalues of the two multiplets we have just found and we are done.

If $j''>1/2$ then the degeneracy for the $m=j'+j''-2$ eigenvalue is three with $(m',m'')=(j',j''-2),(j'-1,j''-1)$ or $(j'-2,j'')$. Thus, as before there is a nontrivial orthogonal normalized vector $\Psi^{j'+j"-2}_{j'+j"-2}$ orthogonal to those $m=j'+j"-2$ eigenvalues of spin $j'+j"$ and $j'+j"$, unique up to phase. This gives rise to a spin $j'+j"-2$ representation by lowering with J_-.

In general, we see that the degeneracy of the m-eigenstates of J_3 is $1+j'+j''-|m|$ for $|m|\geq J'-j''$ but is $2j''+1$ otherwise as it cannot exceed the number of choices $2j''+1$ for m''. So we can carry on by induction generating a new multiplet each time to obtain all the angular momentum multiplets with spin $|j'-j''|$ to $j'+j''$ as required. This gives a total of $2j''+1$ irreducible spin j representations of the given spins ($2j''+1$ being the maximal degeneracy of the m eigenvalue for $|m|\leq j'-j''$ being the possible number of values of m''). □

Clebsch-Gordon coefficients: The left hand side of (29) has a natural orthonormal basis $\psi^{m'}_{j'}\psi^{m''}_{j''}$ that diagonalizes $\mathbf{J}'^2$, $\mathbf{J}''^2$, J'_3 and J''_3, whereas the right hand side has a natural basis Ψ^m_j that diagonalizes $\mathbf{J}^2$ and J_3. The question arises as to how one basis is related to the other.

Definition 5.3 *The Clebsch-Gordan coefficients $C_{j'\,j''}(j,m,;m',m'')$ are defined by*

$$\Psi^m_j=\sum_{m',m'',m'+m''=m} C_{j'\,j''}(j,m;m',m'')\psi^{m'}_{j'}\psi^{m''}_{j''}\,, \tag{30}$$

where $\psi^{m'}_{j'},\psi^{m''}_{j''}$ are the standard basis for $\mathcal{H}_{j'}$ and $\mathcal{H}_{j''}$.

The coefficients are not quite unique, because the choices of the standard bases for $\mathcal{H}_j$ are only unique up to an overall phase. However, this freedom can be fixed by requiring certain of the Clebsch-Gordon coeffiecients to be real.

Because both bases are orthonormal we have the unitarity condition

$$\sum_{m',m'',m'+m''=m} |C_{j'\,j''}(j,m;m',m'')|^2 = 1\,.$$

Example: Consider $j'' = 1/2$. Then

$$\begin{aligned}
\Psi^{j'-1/2}_{j'+1/2} &= \frac{1}{\hbar\sqrt{2j'+1}} J_- \Psi^{j'+1/2}_{j'+1/2} \qquad \text{using (25)} \\
&= \frac{1}{\hbar\sqrt{2j'+1}} (J'_- + J''_-)\psi^{j'}_{j'}\psi^{1/2}_{1/2} \\
&= \frac{1}{\sqrt{2j'+1}} \left(\sqrt{2j'}\psi^{j'-1}_{j'}\psi^{1/2}_{1/2} + \psi^{j'-1}_{j'}\psi^{-1/2}_{1/2}\right)
\end{aligned}$$

Thus we can identify its orthogonal complement as

$$\Psi^{j'-1/2}_{j'-1/2} = \frac{1}{\sqrt{2j'+1}} \left(\psi^{j'-1}_{j'}\psi^{1/2}_{1/2} - \sqrt{2j'}\psi^{j'}_{j'}\psi^{-1/2}_{1/2}\right)$$

and this can now be lowered with J_- to yield the $j' - \frac{1}{2}$ multiplet. One can now read off the Clebsch-Gordon coefficients, i.e.

$$C_{j'\,\frac{1}{2}}(j' - \frac{1}{2}, j' - \frac{1}{2}; j', -\frac{1}{2}) = -\sqrt{\frac{2j'}{2j'+1}}\,.$$

Alternative reading: See §4.3 of Weinberg.

5.5 Angular momenta of hydrogen energy levels

With intrinsic spin 1/2, an electron wave function should be thought of as having two components,

$$|\psi\rangle = \psi_+(t,\mathbf{r})|\tfrac{1}{2}\rangle + \psi_-(t,\mathbf{r})|-\tfrac{1}{2}\rangle\,,$$

where $|\pm\frac{1}{2}\rangle$ are the $m = \pm\frac{1}{2}$ intrinsic spin eigenstates for S_3. Thus ψ_+ gives the wave function for the $m = \frac{1}{2}$ spin and ψ_- that for $-\frac{1}{2}$ spin. In non-relativistic quantum mechanics, the Schrodinger equation doesnt mix the two m-values and $\psi_\pm(t,\mathbf{r})$ both satisfy the same Schrodinger equation.

As far as orbital angular momentum is concerned, we can furthermore decompose $\psi_\pm$ into spherical harmonics which amounts to separating variables in spherical polars to give the basis of states

$$\Psi_{nl}^{mm_s} = f_{nl}(r)Y_l^m(\theta,\phi)|m_s\rangle \tag{31}$$

where $Y_l^m(\theta,\phi)$ are the usual spherical harmonics with eigenvalues $l(l+1)\hbar^2$ for $\mathbf{L}^2$ and $m\hbar$ for L_3, $|m_s\rangle$ are the intrinsic spin states with $m_s = \pm\frac{1}{2}$ and $f_{nl}(r)$ is a polynomial in r of degree n multiplied by $\mathrm{e}^{-Zr/na}$. These states simultaneously diagonalize $\mathbf{L}^2$, $\mathbf{S}^2$, L_3 and S_3 and H_0 with energy

$$E_n = \frac{E_0}{(n+1)^2}\,, \quad E_0 = -\frac{Z^2e^2}{8\pi\epsilon_0 a}$$

where a is the Bohr radius. The energy eigenstates are therefore only determined by n, whereas $l \leq n$ and for each l there are $2l+1$ states. When multiplied by 2 for the possible values of m_s and summed over l, we find a degeneracy of $2n^2$ for E_n.

These states do not diagonalize $\mathbf{J}^2$ where $\mathbf{J} = \mathbf{L} + \mathbf{S}$. However, according to the laws of addition of angular momentum, we can find a different basis that does do so with eigenvalue $j(j+1)\hbar^2$ with $j = l \pm \frac{1}{2}$ by taking suitable linear combinations

$$\phi_{lj}^m = \sum_{m',m_s,\ m'+m_s=j} C_{l,\frac{1}{2}}(j,m;m',m_s)Y_l^{m'}|m_s\rangle$$

where $C_{l,\frac{1}{2}}(j,m;m',m_s)$ are the Clebsch-Gordon coefficients. This basis consists of eigenvectors for $\mathbf{L}^2$, $\mathbf{J}^2$, J_3 the latter with eigenvalue m, but not for L_3, S_3. In this basis the wave functions are denoted

$$\Psi_{nlj}^m = f_{nl}(r)\phi_{lj}^m\,. \tag{32}$$

There is therefore a standard nomenclature for the electron states in this form as nl_j orbitals where n denotes the energy level, l the orbital angular momentum l, often denoted, for $l = 0,1,2,3,\ldots$, by $s,p,d,f,\ldots$ respectively then going up the alphabet. The final subscript $j = l \pm \frac{1}{2} \in \mathbb{N} + 1/2$ is the total spin incorporating both the orbital and intrinsic spin of the electron; it gives a multiplet with $2j+1$ elements as $m = -j,\ldots,j$. We will see that the energies for the different j values are separated by the *fine structure* associated to the first relativistic correction, the spin-orbit coupling.

For all n we have the $l = 0$, s-orbitals, which gives just the multiplet $ns_{1/2}$ with the two states $m = \pm\frac{1}{2}$ for $j = 1/2$. For the $l = 1$ case, the p-orbital, now for $n \geq 1$, we then have either $np_{1/2}$ with two states or $np_{3/2}$ with four states giving a total of 6 states for the np states. In general we have orbits $nl_{l\pm1/2}$ with $n \geq l$ and a similar counting gives for the nl energy level gives a total of $2(l - \frac{1}{2}) + 1 + 2(l + \frac{1}{2}) + 1 = 4l + 2$ states.

5.6 Fermionic statistics and the periodic table

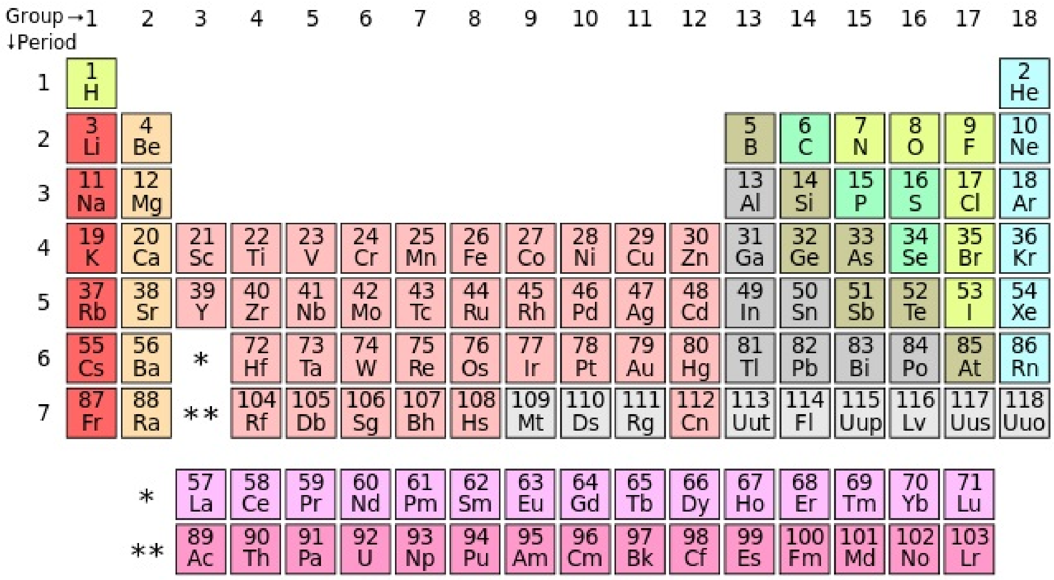
We now combine the fermionic nature of electrons with the knowledge of their intrinsic spin and angular momentum to obtain a qualitative understanding of the periodic table.

In the first approximation we take our n electrons to be non-interacting and so built from single electron (hydrogen-like ion) wave functions. Fermi statistics give the Pauli exclusion principle that two electrons cannot occupy the same state. Thus, as soon as the two spin states for the ground state orbit are filled, a third electron must be in the first excited state. It turns out that although the hydrogen-like ion calculation is good for the first few energy levels, more elaborate calculations that include shielding of nuclear charge by the inner orbits, give the following order of energy levels with levels on the same line being approximately equal:

$$\begin{array}{llllc}
 & & & & \#\text{ of states} \\
1s, & & & & 2 \\
2s, & 2p, & & & 2+6=8 \\
3s, & 3p, & & & 2+6=8 \\
4s, & 3d, & 4p & & 2+6+10=18 \\
5s, & 4d, & 5p & & 2+10+6=18 \\
6s, & 4f, & 5d, & 6p, & 2+14+10+6=32 \\
7s, & 5f, & 7p, & \ldots &
\end{array} \tag{33}$$

and so on, where the s, p, d, f are the symbols for $l =, 0, 1, 2, 3$ respectively. (We ignore the fine structure separation for different j-values at this level of approximation.) The point here is that for the larger l orbits, the electrons in the smaller l orbits are shielding the nuclear charge, reducing the attraction towards the nucleus and hence raising the energy level of the outer orbits.

This already gives a good understanding of the periodic table of the elements. This is ordered by the charge of the nucleus and hence number of electrons. The elements in the same column are grouped by their physical



Group→ ↓Period	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1	1 H																	2 He
2	3 Li	4 Be											5 B	6 C	7 N	8 O	9 F	10 Ne
3	11 Na	12 Mg											13 Al	14 Si	15 P	16 S	17 Cl	18 Ar
4	19 K	20 Ca	21 Sc	22 Ti	23 V	24 Cr	25 Mn	26 Fe	27 Co	28 Ni	29 Cu	30 Zn	31 Ga	32 Ge	33 As	34 Se	35 Br	36 Kr
5	37 Rb	38 Sr	39 Y	40 Zr	41 Nb	42 Mo	43 Tc	44 Ru	45 Rh	46 Pd	47 Ag	48 Cd	49 In	50 Sn	51 Sb	52 Te	53 I	54 Xe
6	55 Cs	56 Ba	*	72 Hf	73 Ta	74 W	75 Re	76 Os	77 Ir	78 Pt	79 Au	80 Hg	81 Tl	82 Pb	83 Bi	84 Po	85 At	86 Rn
7	87 Fr	88 Ra	**	104 Rf	105 Db	106 Sg	107 Bh	108 Hs	109 Mt	110 Ds	111 Rg	112 Cn	113 Uut	114 Fl	115 Uup	116 Lv	117 Uus	118 Uuo
		*	57 La	58 Ce	59 Pr	60 Nd	61 Pm	62 Sm	63 Eu	64 Gd	65 Tb	66 Dy	67 Ho	68 Er	69 Tm	70 Yb	71 Lu	
		**	89 Ac	90 Th	91 Pa	92 U	93 Np	94 Pu	95 Am	96 Cm	97 Bk	98 Cf	99 Es	100 Fm	101 Md	102 No	103 Lr	

Figure 1: Periodic table of the elements[13].

and chemical properties. As can be seen the lengths of the rows line up precisely with the numbers of states in the different orbitals given above. Most notable are those on the right, the noble gases. These are the most stable unreactive elements and are those for whom the orbitals of the same energy are full. This is because chemical bonds arise from sharing electrons between different atoms. However, the noble gases find it hard to accept new electrons because a new electron would have to go into a higher energy orbital than the existing one, and the existing electrons are tightly bound so they dont like to leave. On the right of the noble gases, elements have one electron sitting alone in a higher orbital. This electron can be lost easily because it has more energy. Such an element is said to have valence one. Similarly, an element on the left of a noble gas has one space left in its orbital, and so easily accepts an electron and is said to have valency -1 and so on. The qualitative theory of chemical bonding all comes down to this valency in which stable molecules have net valency zero (i.e., H_2O with $+1$ for each hydrogen and -2 for the oxygen).

---

[13] Periodic table (polyatomic) by DePiep- Own work. Licensed under Creative Commons Attribution-Share Alike 3.0 via Wikimedia Commons http://commons.wikimedia.org/wiki/File:Periodic_table_(polyatomic).svg .

Another application of Pauli's exclusion principle arises in the theory of the nucleus. It is made from protons and neutrons, again fermions of spin 1/2. These are held together by the strong interactions that dwarfs the electrostatic repulsion between the positively charged protons. The strong force is attractive and thought to vanish as particles come together and so to a first approximation can be modelled by the simple harmonic oscillator potential. As for the electrons around a central nucleus, these too form degenerate energy levels leading to shells. Like the noble gases, when these shells are filled, the nucleus is particularly stable. The magic numbers of nucleons required for a stability are $2, 8, 20, 28, 50, 82, 126, \ldots$. The protons and neutrons are considered separately since the Pauli exclusion principle only applies to like particles. Thus, for example, He^4 is stable both as regards its neutrons and protons. The first few of these numbers are explained by the naive simple harmonic oscillator model for which we have states (multiplying by 2 for the two spin states)

Energy	States	Degeneracy
0	s	2
$\hbar\omega$	p	6
$2\hbar\omega$	s, d	12
$3\hbar\omega$	p, f	20
$\ldots$	$\ldots$	$\ldots$.

This works well for the first three magic numbers, 2, $2+6 = 8$, $2+6+12 = 20$ but then fails beyond that. However, if corrections associated with deviations from the linear harmonic oscillator force law, and with the spin-orbit coupling are taken into account, the explanation can be taken further.
Alternative reading: See later parts of §4.5 of Weinberg.

5.7 Other symmetries

There are many other symmetries, both exact and approximate that play key roles in quantum mechanics that dont arise from space-time symmetries. For example, protons and neutrons that make up the nuclei of atoms are built out of triples of quarks, and quarks are labelled by both colour and flavour. There are two flavours for the quarks that make up nucleons, up (charge 2/3) and down (charge $-1/3$), and three colours chosen according to taste. There is a symmetry that rotates quarks of different flavors into each other that is approximate because although the different quarks respond in the same way to the strong interactions, the quarks have different charges and so respond differently to the weaker electromagnetic interactions. The colour symmetry is however exact. Quantum mechanically, the colour symmetry is represented on states as an SU(3) symmetry and the flavour 'isospin' symmetry by SU(2). In particular the isospin representations are the same as those for the rotation group described earlier with generators I_a, $a = 1, 2, 3$ but where now emphatically, a is not a vector index in the sense that $[\mathbf{J}, I_a] = 0$. We wont develop this further here, but the representations of such groups play a large role in the classification of elementary particles and the construction of the standard model.

6 Stationary perturbation theory

To find the exact energy levels for larger atoms we must solve

$$H\psi = E\psi \,,$$

where H contains all the interactions between the electrons and the spin with the electromagnetic field and so on. This will be too complicated to do exactly, but there are by now many good approximation methods that are sufficient to provide the corrections needed to deduce for example the deviations from the non-interacting case used in (33).

6.1 Elementary perturbation theory

We can often express $H = H_0 + H'$ where H_0 is one we can solve exactly, such as the hydrogen-like ion, and H' is small. Write $H_u = H_0 + uH'$, then we can hope to solve

$$H_u\psi_u = E_u\psi_u \tag{34}$$

in a Taylor series in u with

$$\psi_u = \psi_0 + u\psi' + u^2\psi'' + \dots \,, \qquad E_u = E_0 + uE' + u^2E'' + \dots \,.$$

Replacing ψ_u by $\psi_u/\langle\psi_0|\psi_u\rangle$, we can normalize ψ_u by

$$\langle\psi_0|\psi_u\rangle = 1 \,, \qquad \text{which gives} \quad \langle\psi_0|\psi_0\rangle = 1 \,, \quad \text{and} \quad \langle\psi_0|\psi'\rangle = 0 \,.$$

Substituting the Taylor series into (34) we obtain

$$(H_0+uH')(\psi_0+u\psi'+u^2\psi''+\ldots)=(E_0+uE'+\ldots)(\psi_0+u\psi'+u^2\psi''+\ldots)$$

and equating coefficients of different powers of u we obtain to first order

$$H_0\psi_0=E_0\psi_0\,,\qquad H'\psi_0+H_0\psi'=E'\psi_0+E_0\psi' \tag{35}$$

and this gives

$$(H_0-E_0)\psi'=-(H'-E')\psi_0\,. \tag{36}$$

By the first equation of (35), H_0-E_0 is *not* invertible as its kernel contains ψ_0 so (36) cannot simply be inverted to find ψ'. If E_0 is a nondegenerate eigenvalue, solutions can only exist if E' is chosen to make the right hand side perpendicular to ψ_0 as can be seen by taking the inner product of ψ_0 with (36).

$$0=\langle\psi_0|(H_0-E_0)|\psi'\rangle=-\langle\psi_0|(H'-E')|\psi_0\rangle\,.$$

This gives

$$E'=\langle\psi_0|H'|\psi_0\rangle\,.$$

Analytic considerations: In infinite dimensions, not all finite dimensional arguments go through. In particular quantities can diverge and some operators are only defined on certain subdomains in $\mathcal{H}$. The expansion in u will work if

$$||H'\psi||\leq c||H_0\psi||+c_1||\psi||$$

for some constants c, c_1.

Example 1: The helium atom ground state. Ignoring spin, the helium atom (assuming a fixed nucleus) has hamiltonian

$$H=H_1+H_2+H'\,,\qquad H_i=-\frac{\hbar^2}{2m}\nabla_i^2-\frac{2e^2}{4\pi\epsilon_0}\frac{1}{|\mathbf{r}_i|}\,,\ i=1,2,\quad H'=\frac{e^2}{4\pi\epsilon_0|\mathbf{r}_1-\mathbf{r}_2|}\,,$$

here the nuclear charge $Z=2$, H_1 and H_2 are the standard single electron hamiltonians and H' is the repulsion between the two electrons. The Hamiltonian H_1+H_2 is separable so we can consider eigenstates $\psi_{n_1}(\mathbf{r}_1)\psi_{n_2}(\mathbf{r}_2)$.

In particular the ground state is non-degenerate (the fermionic nature of the electrons is being absorbed into the spin degrees of freedom[14]) giving

$$\psi_0(\mathbf{r}_1,\mathbf{r}_2) = \left(\frac{Z^3}{\pi a^3}\right)^{1/2} \mathrm{e}^{-Zr_1/a} \left(\frac{Z^3}{\pi a^3}\right)^{1/2} \mathrm{e}^{-Zr_2/a} = \left(\frac{Z^3}{\pi a^3}\right) \mathrm{e}^{-Z(r_1+r_2)/a} \,.$$

The shift in the ground state is given by

$$E' = \langle\psi_0|H'|\psi_0\rangle = \left(\frac{Z^3}{\pi a^3}\right)^2 \int \frac{e^2}{4\pi\epsilon_0} \frac{\mathrm{e}^{-2Z(r_1+r_2)/a}}{|\mathbf{r}_1-\mathbf{r}_2|}\, \mathrm{d}^3\mathbf{r}_1\ \mathrm{d}^3\mathbf{r}_2\,.$$

This integral is performed using spherical polars for $\mathbf{r}_2$ set up with polar axis along $\mathbf{r}_1$ so that $\mathbf{r}_1\cdot\mathbf{r}_2 = r_1r_2\cos\theta$ and $|\mathbf{r}_1-\mathbf{r}_2| = (r_1^2+r_2^2-2r_1r_2\cos\theta)^{1/2}$. The nontrivial polar integral becomes

$$\begin{aligned}
\int_0^{2\pi}\int_0^{\pi} \frac{\sin\theta\,\mathrm{d}\theta\,\mathrm{d}\phi}{|\mathbf{r}_1-\mathbf{r}_2|} &= \frac{2\pi}{r_1r_2}(r_1^2+r_2^2-2r_1r_2\cos\theta)^{1/2}|_0^{\pi} \\
&= \frac{2\pi}{r_1r_2}\left((r_1+r_2)-|r_1+r_2|\right) \\
&= \begin{cases} \frac{4\pi}{r_1} & r_1>r_2 \\ \frac{4\pi}{r_2} & r_2>r_1\,. \end{cases}
\end{aligned}$$

The remaining polar coordinates can be done with another factor of 4π. The $r_1>r_2$ case is the same as the $r_2>r_1$ so we now have

$$E' = 2\left(\frac{Z^3}{\pi a^3}\right)^2 (4\pi^2)\frac{e^2}{4\pi\epsilon_0}\int_0^{\infty}\int_{r_2}^{\infty} \mathrm{e}^{-2Z(r_1+r_2)/a} r_1\,\mathrm{d}r_1\, r_2^2\,\mathrm{d}r_2\,.$$

This leaves radial integrations that can be performed with the help of

---

[14] In more detail, the general ground state wave function for a single electron with spin is two dimensional and can be represented as $\mathrm{e}^{-Zr/a}(a|\frac{1}{2}\rangle + b|-\frac{1}{2}\rangle)$ with a,b complex numbers. With two particles, we cannot make the wave function antisymmetric in the positions $\mathbf{r}_1$ and $\mathbf{r}_2$ because the ground state is unique, but we can make it antisymmetric instead in the spins giving the two-particle wave function

$$|\psi_0\rangle = a\mathrm{e}^{-Z(r_2+r_2)/a}(|\tfrac{1}{2},-\tfrac{1}{2}\rangle - |-\tfrac{1}{2},\tfrac{1}{2}\rangle)\,,$$

where $|\frac{1}{2},-\frac{1}{2}\rangle$ is the state in which the first particle has intrinsic spin $m=\frac{1}{2}$ and the second $m=-\frac{1}{2}$ etc.. This is a simple case of addition of angular momentum where $\mathcal{H}_{1/2}\otimes\mathcal{H}_{1/2} = \mathcal{H}_1\oplus\mathcal{H}_0$ and we take the (unique) spin 0 state.

Lemma 6.1

$$\int_R^\infty \mathrm{e}^{-kr} r^n \, \mathrm{d}r = n! k^{-(n+1)} \mathrm{e}^{-kR} \sum_{j=0}^{n} \frac{(kR)^j}{j!} \,.$$

This is proved by induction on n by differentiating with respect to k.

Putting everything together with $Z = 2$ we obtain

$$E' = \frac{5}{8} \frac{Ze^2}{4\pi\epsilon_0 a} \,,$$

and this gives a correction to the standard energy level as

$$E_1 = E_0 + E' = -\frac{2e^2}{4\pi\epsilon_0 a}\left(2 - \frac{5}{8}\right) = -0.69 \frac{e^2}{\pi\epsilon_0 a} \,.$$

This can be compared with $-0.73e^2/\pi\epsilon_0 a$ from experiment. Keeping the Z in play in the formulae, we can regard the correction as reducing the number of electron charges Z of the nucleus from 2 to 11/8 because the electrons shield the nuclear charge.

In general, if for example we are considering the higher energy eigenstates of a Hydrogen-like ion, our eigenvalue E_0 for our unperturbed H_0 will have much degeneracy and so there will be more conditions to be satisfied for (36) to have a solution, the right hand side must be orthogonal to all the eigenvectors.

Theorem 4 *Let $\phi_1, \ldots, \phi_D$ be an orthonormal basis for the E_0 eigenspace of H_0, $(H_0 - E_0)\phi_r = 0$. Then* (36) *can be solved iff E' and ψ_0 are chosen so that E' is a solution to*

$$\det\left(\langle \phi_r | H' | \phi_s \rangle - E'\delta_{rs}\right) = 0$$

and $\psi_0 = \sum_r c_r \phi_r$ is the corresponding eigenvector

$$\sum_s \langle \phi_r | H' | \phi_s \rangle c_s = E' c_r \,.$$

Proof: We need to be able to solve

$$(H_0 - E_0)\psi' = -(H' - E')\psi_0 \,, \qquad (37)$$

but $H_0 - E_0$ is not invertible and the right hand side could fail to be in the image (or range) of $H_0 - E_0$.

The image is characterized as the orthogonal complement of the kernel of $H_0 - E_0$. The orthogonality is necessary because $w = (H_0 - E_0)v$ is in the image and $\langle \phi_r | w \rangle = \langle \phi_r | (H_0 - E_0) | v \rangle = 0$ because $(H_0 - E_0)|\phi_r\rangle = 0$. It is sufficient because a vector ϕ is orthogonal to the range means that $0 = \langle \phi | (H_0 - E_0) | w \rangle = \langle (H_0 - E_0)\phi | w \rangle$ for all w so $(H_0 - E_0)\phi = 0$. Thus, taking the inner product of ϕ_r with the right hand side of (36) we must therefore have

$$0 = \langle \phi_r | (H' - E') | \psi_0 \rangle = \sum_s \langle \phi_r | H' | \phi_s \rangle c_s - E' \delta_{rs} c_s \,,$$

i.e. E' must be an eigenvalue of $\langle \phi_r | H' | \phi_s \rangle$ and c_s and hence ψ_0 the corresponding eigenvector. ☐

Since H' is hermitian

$$\overline{\langle \phi_r | H' | \phi_s \rangle} = \langle \phi_s | H' | \phi_r \rangle$$

is Hermitian and so E' is necessarily real.

Example 2: A simple nontrivial example starts with the 2-dim simple harmonic oscillator:

$$H = \frac{1}{2m}(P_x^2 + P_y^2) + \frac{1}{2} m\omega^2 (x^2 + y^2) + uxy$$

Here we can solve exactly by diagonalizing the potential for eigenvalues λ solving

$$\det \begin{pmatrix} m\omega^2 - \lambda & u \\ u & m\omega^2 - \lambda \end{pmatrix} = 0 \,,$$

which gives $\lambda =, \omega^2 \pm u$, or frequencies

$$\omega_\pm = \omega \sqrt{\left(1 \pm \frac{u}{m\omega^2}\right)} = \omega\left(1 \pm \frac{u}{2m\omega^2} + \ldots\right) ,$$

and hence true energies

$$(n_+ + \frac{1}{2})\hbar\omega_+ + (n_- + \frac{1}{2})\hbar\omega_- \,.$$

The perturbative procedure has $H' = xy$ and unperturbed energy levels

$$E_n = (n_1 + \frac{1}{2} + n_2 + \frac{1}{2})\hbar\omega\,, \qquad n = n_1 + n_2$$

with degeneracy $n{+}1$. The energy eigenstates are given by $\phi_{n_1 n_2} = \phi_{n_1}(x)\phi_{n_2}(y)$. The ground state is non degenerate so E' follows from simply computing

$$E' = \langle \phi_{00} | xy | \phi_0 0 \rangle = 0\,,$$

which follows as the integrals are odd. The first excited state is degenerate spanned by ϕ_{01} and ϕ_{10}. The corresponding matrix $\langle \phi_r | H' | \phi_s \rangle$ is now

$$\frac{a^2}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

where $a = \sqrt{\hbar/m\omega}$ and we use the relation $x\phi_0(x) = a\phi_1/\sqrt{2}$. The agreement to first order now follows by direct calculation.

Example 3: The Zeeman effect (strong field). We give a simplified description here in which we ignore the spin-orbit coupling. If we apply a constant magnetic field in the z-direction to an atom, then the energy is shifted by

$$H' = \frac{e}{2m_e c} B(L_3 + g_e S_3)\,,$$

where g_e is the *gyromagnetic ratio* of the electron. Rotate axes so that $\mathbf{B}$ is in the 3-direction. Then the matrix $\langle \phi_r | H' | \phi_s \rangle$ will be diagonalized by states that are simultaneous eigenvectors of L_3 and S_3.

If the given energy eigenspace for E_0 is associated to fixed hydrogen-like ion energy level n, then we can choose a basis by fixing orbital angular momentum l, spin $s = \frac{1}{2}$, and eigenvalues m_L and $m_S = \pm\frac{1}{2}$ for L_3 and S_3 respectively, to give states $\psi_{nl}^{m_l m_s} = f_{nl}(r) Y_l^{m_l} |m_s\rangle$ as in (31). These now diagonalize $\langle \phi_r | H' | \phi_s \rangle$ to give

$$E' = \langle \psi_l^{m_L m_S} | \delta H | \psi_{nl'}^{m'_L m'_S} \rangle = \frac{e g_e B}{2 m_e c} \delta_{m_L m'_L} \delta_{m_s m'_S} \delta_{ll'} (m_L + g_e m_S)\,.$$

and these give the corresponding shifts in the eigenvalues.

Example 4: Spin-orbit coupling. If we reduce the relativistic Dirac equation to a nonrelativistic Schrödinger equation, we find the spin-orbit coupling term

$$H' = \frac{\mu_B Z}{\hbar m_e c^2} \frac{1}{r^3} \mathbf{L} \cdot \mathbf{S}\,.$$

(Intuitively, this arises from the interaction between the spin of the electron and and the magnetic field it experiences in its rest frame due to its motion through the electric field of the nucleus.) There is a more general formula that replaces r^{-3} with $\frac{1}{r}\partial V/\partial r$ in a more general potential.

According to the theorem above, our task is to find the eigenvalues of $\langle \phi_r | H' | \phi_s \rangle$ on an eigenspace E_0 of H_0 to find the perturbations of E_0. The eigenspaces of H_0 are very degenerate being determined by the n in the nl_j parametrization for states with total angular momentum j, orbital angular momentum $l \leq n$ and intrinsic spin s ($s = 1/2$ for our example of the electron) spanned by $f_{nl}\Psi_j^m$ with $j = l \pm \frac{1}{2}$, $m = -j, \ldots, j$ where Ψ_j^m are the standard basis of $\mathcal{H}_j$ given as linear combinations of $Y_l^{m_l}|m_s\rangle$ following from the addition of orbital angular momentum and spin $\mathcal{H}_l \otimes \mathcal{H}_{1/2} = \mathcal{H}_{l+1/2} \oplus \mathcal{H}_{l-1/2}$ as in (32). In fact $\langle \phi_r | H' | \phi_s \rangle$ is already diagonal in this basis because H_0, H', $\mathbf{J}^2$, $\mathbf{L}^2$, $\mathbf{S}^2$ and J_3 all commute and we can write

$$\mathbf{L} \cdot \mathbf{S} = \frac{1}{2}(\mathbf{J}^2 - \mathbf{L}^2 - \mathbf{S}^2)\,.$$

Thus, we will have

$$\langle \psi_j^m | \mathbf{L} \cdot \mathbf{S} | \psi_{j'}^{m'} \rangle = \delta_{m\,m'}\delta_{j\,j'}\frac{\hbar^2}{2}(j(j+1) - l(l+1) - s(s+1))$$

where $s = 1/2$ using the standard normalizations of Ψ_j^m. We quote that, for the n^{th} radial eigenfunction,

$$\left\langle f_{nl} \left| \frac{1}{r^3} \right| f_{nl} \right\rangle = \int_0^\infty |f_{nl}|^2 \frac{r^2 dr}{r^3} = \frac{2}{a^3 n^3 l(l+1)(2l+1)}\,,$$

where a is the Bohr radius. Thus for a wave function $f_{nl}\psi_j^m$ in the nl_j orbital, we have

$$\langle f_{nl}\psi_j^m | H' | f_{nl}\psi_j^m \rangle = \frac{\mu_B Z \hbar}{m_e c^2} \frac{(j(j+1) - l(l+1) - 3/4)}{a^3 n^3 l(l+1)(2l+1)}\,.$$

The spin-orbit coupling separates energies with different values of both l and j. It is responsible for the energy levels of hydrogen-like ion orbitals depending both on n *and* l in the table (33). It furthermore gives the *fine structure* splitting in the energy levels for orbits with the same n and l values but different $j = l \pm 1/2$.

This is a different basis from that used for the Zeeman effect which is represented in terms of this via the Clebsch-Gordon coefficients (30). For weaker magnetic fields the spin-orbit coupling is of the same order as the Zeeman effect, a more sophisticated analysis is then required, see for example Weinberg's book.

6.2 Higher order Rayleigh-Schrodinger theory

We found $E' = \langle\psi_0|H'|\psi_0\rangle$ and so didnt need to know ψ'. Naively we need ψ^{k-1} to obtain E^k because expanding $H_u\psi_u = E_u\psi_u$ to order k gives from the coefficient of u^k

$$H_0\psi^{(k)} + H'\psi^{(k-1)} = E_0\psi^{(k)} + E'\psi^{(k-1)} + \ldots + E^k\psi_0\,,$$

and taking the inner product with ψ_0, using $\langle\psi_0|\psi^{(l)}\rangle = 0$ that follows from the normalization condition $\langle\psi_0|\psi_u\rangle = 1$, we obtain

$$E^k = \langle\psi_0|H'|\psi^{(k-1)}\rangle\,,$$

(as $\langle\psi_0|H_0|\psi^{(k)}\rangle = E_0\langle\psi_0|\psi^{(k)}\rangle = 0$). In fact there are many relations and we can do better.

Lemma 6.2

$$\langle\psi_u|H'|\psi_v\rangle = \frac{E_u - E_v}{u - v}\langle\psi_u|\psi_v\rangle\,.$$

Proof: We have $H_u - H_v = (u - v)H'$ so

$$\begin{array}{rcl} \langle\psi_u|H_u - H_v|\psi_v\rangle & = & \langle\psi_u|H_u|\psi_v\rangle - \langle\psi_u|H_v|\psi_v\rangle \\ (u - v)\langle\psi_u|H'|\psi_v\rangle & = & E_u\langle\psi_u|\psi_v\rangle - E_v\langle\psi_u|\psi_v\rangle \end{array}$$

which leads to the conclusion. □.

Different powers generate many formulae, our first naive higher order formula used the coefficient of v^{k-1} but we can do better with others:

Corollary 6.1 *E^{2k+1} can be constructed from E^r with $r < 2k + 1$ and ψ^r with $r < k$.*

Proof: From the coefficent of $u^k v^k$ we obtain the formula

$$\begin{aligned}\frac{E_u - E_v}{u - v} &= \sum_{n=1}^{\infty} \frac{u^n - v^n}{u - v} E_n \\ &= \sum_{n=1}^{\infty} E_n \left(\sum_{j=0}^{n-1} u^j v^{n-1-j} \right) \\ &= = \frac{\langle \psi_u | H' | \psi_v \rangle}{\langle \psi_u | \psi_v \rangle}\end{aligned}$$

To order $u^k v^k$ the right hand side depends only on ψ^r for $r \leq k$. □

As a simple example we have

$$E''' = \langle \psi' | H' | \psi' \rangle - E' \langle \psi' | \psi' \rangle \, .$$

However, we still need ψ'.

Theorem 5 *Let ψ_α, $\alpha = 0, 1, 2, \ldots$ be an orthonormal basis for $\mathcal{H}$ of eigenvectors of H_0*

$$H_0 \psi_\alpha = E_\alpha \psi_\alpha \, , \qquad E_\alpha \neq E_0 \text{ for } \alpha \neq 0 \, .$$

Then

$$\psi' = \sum_{\alpha \neq 0} \frac{\langle \psi_\alpha | H' | \psi_0 \rangle}{E_0 - E_\alpha} \psi_\alpha$$

and

$$E'' = \sum_{\alpha \neq 0} \frac{|\langle \psi_\alpha | H' | \psi_0 \rangle|^2}{E_0 - E_\alpha} \, .$$

Proof: We wish to solve

$$(H_0 - E_0) \psi' = -(H' - E') \psi_0 \, . \tag{38}$$

We assume that E_0 is nondegenerate and choose E' so that $\langle \psi_0 | H' - E' | \psi_0 \rangle = 0$. We can set

$$\psi' = \sum_{\alpha \neq 0} \langle \psi_\alpha | \psi' \rangle \psi_\alpha \, .$$

Take the inner product of (38) with ψ_α to find

$$\begin{aligned}\langle \psi_\alpha | H_0 - E_0 | \psi' \rangle &= -\langle \psi_\alpha | H' | \psi_0 \rangle + E' \langle \psi_\alpha | \psi_0 \rangle \, , \\ (E_\alpha - E_0) \langle \psi_\alpha | \psi' \rangle &= -\langle \psi_\alpha | H' | \psi_0 \rangle \\ \langle \psi_\alpha | \psi' \rangle &= \frac{\langle \psi_\alpha | H' | \psi_0 \rangle}{E_0 - E_\alpha}\end{aligned}$$

and this gives the first answer. Similarly

$$E'' = \langle \psi_0 | H' | \psi' \rangle = \sum_{\alpha \neq 0} \frac{|\langle \psi_0 | H' | \psi_\alpha \rangle|^2}{E_0 - E_\alpha} \qquad \Box.$$

Remarks:

- If E_0 was the ground state, we see that this second order correction is always negative.
- This formula is particularly useful if $H'\psi_0$ is an energy eigenvalue so that we only get one contribution due to the orthogonality of the energy eigenstates.
- If the energy levels are well separated, then only the first few terms dominate.

6.3 Variational methods

Rayleigh quotients:

In all physical systems, we expect the energy to be bounded below $E \geq E_0$ otherwise unstable runaway processes will occur. We can therefore characterize the ground state energy as

$$E_0 = \inf_{\psi \in \mathcal{H}} E_\psi(H) := \inf_{\psi \in \mathcal{H}} \frac{\langle \psi | H | \psi \rangle}{\langle \psi | \psi \rangle}\,.$$

Definition 6.1 *$f_H(\psi) = E_\psi(H)$ is the Rayleigh quotient.*

In fact, all stationary values are eigenvalues for $f_H(\psi)$.

Theorem 6 *Given a subspace $\mathcal{K} \subset \mathcal{H}$ then*

$$\frac{\mathrm{d}}{\mathrm{d}u} f_H(\psi + u\phi) = 0\,, \qquad \forall \phi \in \mathcal{K}$$

iff

$$\langle \phi | H - f_H(\psi) | \psi \rangle = 0 \qquad \forall \phi \in \mathcal{K}\,.$$

In particular if $\mathcal{K} = \mathcal{H}$,

$$(H - f_H(\psi))|\psi\rangle = 0\,,$$

i.e., critical points of $f_H(\psi)$ are eigenfunctions and critical values are eigenvalues of H.

Proof: For $\phi \in \mathcal{K}$, $i\phi \in \mathcal{K}$ so we require both

$$\frac{\mathrm{d}}{\mathrm{d}u} f_H(\psi + u\phi) = 0\,, \qquad \text{and} \qquad \frac{\mathrm{d}}{\mathrm{d}u} f_H(\psi + iu\phi) = 0$$

These give

$$\begin{aligned}
0 &= \left. \frac{\mathrm{d}}{\mathrm{d}u} \frac{\langle \psi + u\phi | H | \psi + u\phi \rangle}{\langle \psi + u\phi | \psi + u\phi \rangle} \right|_{u=0} \\
&= \left. \frac{\mathrm{d}}{\mathrm{d}u} \frac{\langle \psi | H | \phi \rangle + u(\langle \phi | H | \psi \rangle + \langle \psi | H | \phi \rangle) + O(u^2)}{\langle \psi | \psi \rangle + u(\langle \phi | \psi \rangle + \langle \psi | \phi \rangle) + O(u^2)} \right|_{u=0} \\
&= \frac{(\langle \phi | H | \psi \rangle + \langle \psi | H | \phi \rangle)}{\langle \psi | \psi \rangle} - \frac{\langle \psi | H | \psi \rangle (\langle \phi | \psi \rangle + \langle \psi | \phi \rangle)}{(\langle \psi | \psi \rangle^2} \\
&= \frac{\langle \phi | H - f_H(\psi) | \psi \rangle}{\langle \psi | \psi \rangle} + c.c.\,.
\end{aligned}$$

This is the real part of the desired equation. If we repeat the argument with $\phi \to i\phi$ we get the imaginary part of the desired equation so we deduce the full equation

$$\langle \phi | (H - f_H(\psi)) | \psi \rangle = 0\,.$$

Thus $(H - f_H(\psi))|\psi\rangle \in \mathcal{K}^\perp$ so if $\mathcal{K} = \mathcal{H}$, $\mathcal{K}^\perp = \{0\}$ and

$$(H - f_H(\psi))|\psi\rangle = 0$$

as desired. $\square$

Remarks: Since $f_H(\psi)$ is independent of scale of $\psi \to \lambda\psi$, we could just as well take variation over the "unit sphere" $\langle \psi | \psi \rangle = 1$ in $\mathcal{H}$ or use a Lagrange multiplier (whose value will be $f_H(\psi)$ at the extrema).

The virial theorem: We can obtain useful general information by considering the variation over just one parameter. Assume $H = T + V$ with

$$T = \frac{-\hbar^2}{2m}\nabla^2\,, \qquad V = V(\mathbf{r})\,, \quad \mathbf{r} \in \mathbb{R}^3\,.$$

Then when $H\psi = E\psi$

1. $2E_\psi(T) = E_\Psi(\mathbf{r} \cdot \nabla V)$

2. If V is homogeneous of degree N, $V(\lambda\mathbf{r}) = \lambda^N V(\mathbf{r})$, then

$$E_\psi(T) = \frac{N}{N+2}E\,, \qquad E_\psi(V) = \frac{2}{N+2}E$$

Proof: The idea is to use homogeneity, i.e., scale $\mathbf{r} \to \lambda\mathbf{r}$ and use the fact that an energy eigenstate will be an extrema for $f_H(\psi)$ with respect to λ. If $\langle\psi|\psi\rangle = 1$, then if we define $\psi_\lambda(\mathbf{r}) = \lambda^{3/2}\psi(\lambda\mathbf{r})$, this is also normalized

$$\int_{\mathbb{R}^3} |\psi_\lambda|^2 \,\mathrm{d}^3\mathbf{r} = \int \lambda^3 |\psi(\lambda\mathbf{r})|^2 \,\mathrm{d}^3\mathbf{r} = \int_{\mathbb{R}^3} |\psi(\mathbf{r}')|^2 \,\mathrm{d}^3\mathbf{r}' = 1\,,$$

substituting $\mathbf{r}' = \lambda\mathbf{r}$. Since ψ is an eigenvector of H,

$$\frac{\mathrm{d}}{\mathrm{d}\lambda} f_H(\psi_\lambda)|_{\lambda=1} = 0\,. \tag{39}$$

We have

$$\begin{aligned}
\nabla\psi_\lambda &= \lambda\lambda^{3/2}(\nabla\psi)(\lambda\mathbf{r}) \qquad \text{by the chain rule so} \\
f_H(\psi_\lambda) &= \int_{\mathbb{R}^3} \frac{\lambda^3}{2m}|\lambda\hbar(\nabla\psi)(\lambda\mathbf{r})|^2 + V(\mathbf{r})|\psi(\lambda\mathbf{r})|^2 \,\mathrm{d}^3\mathbf{r} \\
&= \lambda^2 \int_{\mathbb{R}^3} \frac{|\mathbf{P}'\psi(\mathbf{r}')|^2}{2m} + V(\lambda^{-1}\mathbf{r}')|\psi(\mathbf{r}')|^2 \,\mathrm{d}^3\mathbf{r}' \\
&= \lambda^2 E_\psi(T) + E_\psi(V(\lambda^{-1}\mathbf{r}))\,.
\end{aligned}$$

Thus (39) gives

$$0 = 2E_\psi(T) + \frac{\mathrm{d}}{\mathrm{d}\lambda}E_\psi(V(\lambda^{-1}\mathbf{r}))|_{\lambda=1} = 2E_\psi(T) - E_\psi(\mathbf{r}\cdot\nabla V(\mathbf{r}))\,.$$

For V homogeneous of degree N, $\mathbf{r}\cdot\nabla V = NV$ and

$$2E_\psi(T) = NE_\psi(V(\mathbf{r}))\,.$$

We also have

$$E = E_\psi(T) + E_\psi(V(\mathbf{r}))\,,$$

and putting these together gives

$$E_\psi(T) = \frac{N}{N+2}E\,, \qquad E_\psi(V) = \frac{2}{N+2}E$$

as required. $\square$

- We see here that for hydrogen with $N = -1$ we must have $E < 0$ as $V < 0$ and $E_\psi(V)$ is twice the size of $E_\psi(T)$.
- For the harmonic oscillator kinetic and potential energy are equal and balanced $E_\psi(V) = E_\psi(T) = E/2$. .

Approximations: We now have more generally that

Theorem 7 *If $f_H(\psi)$ is bounded below and achieves its minimum $E_{\text{ground}} = \inf_{\psi \in \mathcal{H}-\{0\}} f_H(\psi)$ then E_{ground} is the ground state energy and any state at which $f_H(\psi) = E_{\text{ground}}$ is a ground state wave function. Conversely, at a ground state wave funciton, $f_H(\psi)$ achieves its lower bound.*

Proof: $f_H(\psi)$ is automatically stationary at its minimum. $\square$

So the main approximation idea is that we can get good approximate solutions by finding the minima $\tilde{E}$ of $f_H(\psi)$ over some well-chosen finite dimensional subsets $\{\psi\}$ of $\mathcal{H}$. We will be guaranteed that $\tilde{E} \geq E_{\text{ground}}$ by the above as $\tilde{E} = f_H(\psi)$ for some ψ in our chosen subset.

Exercise: For helium again we have Hamiltonian

$$H = -\frac{\hbar^2}{2m}(\nabla_1^2 + \nabla_2^2) - \frac{2e^2}{4\pi\epsilon_0}\left(\frac{1}{r_1} + \frac{1}{r_2}\right) + \frac{e^2}{4\pi\epsilon_0|\mathbf{r}_1 - \mathbf{r}_2|} = T + V + H'$$

where

$$T = -\frac{\hbar^2}{2m}(\nabla_1^2 + \nabla_2^2), \quad V = -\frac{2e^2}{4\pi\epsilon_0}\left(\frac{1}{r_1} + \frac{1}{r_2}\right), \quad \text{and} \quad H' = \frac{e^2}{4\pi\epsilon_0|\mathbf{r}_1 - \mathbf{r}_2|}.$$

Proposition 6.1 *The Helium ground state energy E*

$$E \leq -\left(\frac{27}{32}\right)^2 \frac{e^2}{\pi\epsilon_0 a}.$$

[cf. $-\frac{11}{16}\frac{e^2}{\pi\epsilon_0 a}$ from perturbation theory.]

Proof: To account for charge screening of the nuclear charge by one electron for the field experienced by the other, use trial wave function

$$\psi_Z(\mathbf{r}_1, \mathbf{r}_2) = \frac{Z^3}{\pi a^3} \mathrm{e}^{-Z(r_1+r_2)/a},$$

and minimize over Z. This is ground state of

$$H_Z = -\frac{\hbar^2}{2m}(\nabla_1^2 + \nabla_2^2) - \frac{Ze^2}{4\pi\epsilon_0}\left(\frac{1}{r_1} + \frac{1}{r_2}\right) = T + \frac{Z}{2}V$$

satisfying $H_Z\psi_Z = E_Z\psi_Z$ with $E_Z = -Z^2e^2/4\pi\epsilon_0 a$. The virial theorem with $N = -1$ gives

$$E_{\psi_Z}(T) = -E_Z = \frac{Z^2e^2}{4\pi\epsilon_0 a}\,, \qquad E_{\psi_Z}\left(\frac{ZV}{2}\right) = 2E_Z = -\frac{2Z^2e^2}{4\pi\epsilon_0 a}$$

so $E_{\psi_Z}(V) = -Ze^2/\pi\epsilon_0 a$. We take from the perturbation theory calculation that

$$\langle\psi_Z|H'|\psi_Z\rangle = \langle\psi_Z|\frac{e^2}{4\pi\epsilon_0}\frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|}|\psi_Z\rangle = \frac{5Ze^2}{32\pi\epsilon_0 a}\,,$$

so that

$$\begin{aligned}
\langle\psi_Z|H|\psi_Z\rangle &= E_{\psi_Z}(T) + E_{\psi_Z}(V) + E_{\psi_Z}(H') \\
&= \frac{Z^2e^2}{4\pi\epsilon_0 a} - \frac{Ze^2}{\pi\epsilon_0 a} + \frac{5Ze^2}{32\pi\epsilon_0 a}\,, \\
&= \frac{e^2}{4\pi\epsilon_0 a}\left(Z^2 - \frac{27}{8}Z\right) \\
&= \frac{e^2}{4\pi\epsilon_0 a}\left[\left(Z - \frac{27}{16}\right)^2 - \left(\frac{27}{16}\right)^2\right] \\
&\geq -\frac{e^2}{4\pi\epsilon_0 a}\left(\frac{27}{16}\right)^2
\end{aligned}$$

The numerical factor here is about −0.71 as opposed to −0.69 from perturbation theory and is within 2% rather than 5% of the experimental answer -0.73.

This accuracy can be improved by considering more basis functions for the variations—accuracy to within 3 parts in 10^8 have been obtained for this calculation using 393 basis functions.

We also obtain the minimum at $Z = 27/16$ reflecting the extent to which each electron shields the charge of the nucleus to the other, i.e., it is reduced by $5e/16$ from $2e$.

The improvement over elementary perturbation theory is no coincidence:

Proposition 6.2 *Let $H = H_0 + H'$ as before and let the set of states $\{\psi_\lambda\}$ over which we will minimize be chosen so that ψ_0 the ground state of H_0 is in $\{\psi_\lambda\}$. Let $E_1 = E_0 + E'$ be the first order perturbation theoretic solution found before, then for $\tilde{E} = \inf_{\psi \in \{\psi_\lambda\}} f_H(\psi)$ we have*

$$E_1 \geq \tilde{E} \geq E_{\text{ground}} \,.$$

Proof: Recall that $E' = \langle \psi_0 | H' | \psi_0 \rangle$ so that

$$E_1 = E_0 + E' = \langle \psi_0 | H_0 + H' | \psi_0 \rangle = f_H(\psi_0) \,,$$

and since $\psi_0 \in \{\psi\}$, $E_1 \geq \tilde{E} = \inf_{\psi \in \{\psi\}} f_H(\psi)$. □

Excited states

Once we have a good approximation for a ground state, how do we obtain the higher excited states? Suppose we know the first k, $E_0, E_1, \dots, E_{k-1}$ eigenvalues with corresponding eigenvectors $\psi_0, \dots \psi_{k-1}$, let $\mathcal{H}_k = \text{Span}\{\psi_0, \dots, \psi_{k-1}\}$.

Theorem 8 *If $\inf_{\psi \in \mathcal{H}_k^\perp} f_H(\psi)$ is attained for some ψ_k then this is the $k+1$th lowest energy eigenstate, and $f_H(\psi_k) \geq E_j$, $j = 0, \dots, k-1$.*

Proof: Since $H : \mathcal{H}_k \to \mathcal{H}_k$, $H : \mathcal{H}_k^\perp \to \mathcal{H}_k^\perp$ because H is Hermitian. Applying the main variational result for $f_H(\psi)$ pn $\mathcal{H}_k^\perp$ gives that if $f_H(\psi$ achieves its infimum at ψ_k, then $(H - f_H(\psi_k))|\psi_k\rangle = 0$ gives the lowest eigenvalue/eigenstate acting on $\mathcal{H}_k^\perp$. Since the lowest k eigenvalues are in $\mathcal{H}_k$, this must be the $k+1$th. □

We can express this conversely without knowledge of the first k eigenvectors/eigenvalues as follows.

Theorem 9 (Minimax) *If the infimum*

$$\inf\{\max\{f_H(\psi); \psi \in \mathcal{K};, \dim \mathcal{K} = k\}$$

is attained where the infimum is taken over k-dimensional subspaces $\mathcal{K} \subset \mathcal{H}$, then it is the kth lowest eigenvalue and the state on which it is attained the corresponding eigenvector.

Thus the maximum value of $f_H(\psi)$ on $\mathcal{K}$ is an upper bound for E_{k-1}.
Rayleigh-Ritz theory: To use the above to obtain an approximation scheme, we need to choose trial subspaces $\mathcal{K}$ perhaps varying with respect to some parameters.

Theorem 10 *Let v_j be a basis of a finite dimensional subspace $\tilde{\mathcal{H}} \subset \mathcal{H}$. Then the kth lowest root $\tilde{E}_k$ of the* secular equation:

$$\det\left(\langle v_j|H|v_l\rangle - E\langle v_j|v_l\rangle\right) = 0\,, \tag{40}$$

provides an upper bound for the kth lowest evalue E_k of H for $k \leq \dim \tilde{\mathcal{H}}$

$$\tilde{E}_k \geq E_k\,.$$

The matrix $\langle v_j|v_l\rangle$ is known as the Grammian; it is the identity if the basis v_i for $\tilde{\mathcal{H}}$ is chosen to be orthonormal.

Proof: For $\psi \in \tilde{\mathcal{H}}$, if $f_H(\psi)$ stationary under $\psi \to \psi + uv_l$ or $\psi \to \psi + iuv_l$ we must, as before in theorem 6, have

$$\langle v_l|(H - f_H(\psi))|\psi\rangle = 0\,.$$

Now suppose $\psi \in \mathcal{K} \subset \tilde{\mathcal{H}}$, then $\psi = \sum_i c_i v_i$ so that this gives

$$\sum_j \langle v_l|H|v_j\rangle c_j - f_H(\psi)\langle v_l|v_j\rangle c_j = 0\,. \tag{41}$$

In particular $f_H(\psi)$ is a solution to the secular equation

$$\det\left(\langle v_l|H|v_j\rangle c_j - f_H(\psi)\langle v_l|v_j\rangle\right) = 0\,,$$

and ψ is the eigenvector of the projection $\langle v_l|H|v_j\rangle$ of H to $\tilde{\mathcal{H}}$ with eigenvalue $f_H(\psi)$. By minimax

$$E_k \leq \tilde{E}_k := \inf_{\mathcal{K}\subset\tilde{\mathcal{H}}}\{\max f_H(\psi), \psi \in \mathcal{K}_k\}$$

where $\mathcal{K}_k$ is a k-dimensional subspace of $\tilde{\mathcal{H}}$. To obtain desired result, take $\mathcal{K}_k$ to be the span of first k eigenvectors of (41) on which max $f_H(\psi) =$ is the kth eigenvalue of solution to (40). $\Box$

Thus, to obtain estimates, for E_k we need only consider the finite dimensional secular equations for well chosen $\tilde{\mathcal{H}}$. In the problem sheets we will see this idea used to estimate eigenvalues of the total orbital angular momentum operator $\mathbf{L}^2$.

7 Pictures in quantum mechanics

In the early days of quantum theory, Heisenberg developed 'matrix mechanics' and Schrodinger 'wave mechanics' and they were regarded as competing approaches to understanding quantum mechanics. However Dirac saw that they were equivalent. The key idea was to observe that a Hilbert space $\mathcal{H}$ has many bases and we can transform the equations into very distinct forms by different choices of basis. Here we will consider basis changes that depend on time. We will show how this leads to different 'pictures' from the Schrodinger one that we have so far been using, the Heisenberg and Dirac's interaction picture.

The Schrodinger representation

This is the representation we have been using in which the key observables X^S, P^S etc., the superscript S here emphasizing that we are in the Schrodinger picture, are given once and for all and are time independent unless some explicit dependence is introduced. All the time dependence is encoded in the Schrodinger equation for the wave function which therefore depends on time according to

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

When H is time independent, this has the formal solution

$$\Psi^S(\mathbf{r},t) = U_t\Psi^S(\mathbf{r},0)\,, \qquad U_t = \mathrm{e}^{-iHt/\hbar} := \sum_n \frac{(-iHt)^n}{\hbar^n n!}\,,$$

where U_t is a unitary operator $U_tU_t^* = I$. Here U_t is the solution to the equation

$$i\hbar\frac{\partial U_t}{\partial t} = HU_t = U_tH$$

and is unitary as H is Hermitian (so that $U_tU_t^*$ is constant and hence the identity, its value at $t = 0$). [If H has some time dependence this latter definition makes sense, but we dont have the nice exponential formula for its solution, but see later.]

The Heisenberg picture

To transform to the Heisenberg picture, we change basis of $\mathcal{H}$ using U_t so that the wave function no longer evolves. Thus we set

$$\Psi^H = U_t^* \Psi^S(\mathbf{r}, t) = \Psi^S(\mathbf{r}, 0)\,.$$

However, the operators must therefore be conjugated to give

$$A^H = U_t^* A^S U_t\,.$$

This guarantees that amplitudes and expectations are the same as

$$\langle \Phi^H | A^H | \Psi^H \rangle = \langle \Phi^S | U_t U_t^* A^S U_t U_t^* | \Psi^S \rangle = \langle \Phi^S | A^S | \Psi^S \rangle\,. \tag{42}$$

This is also compatible with commutators etc., as, if $[A^S, B^S] = C^S$, then we will have

$$[A^H, B^H] = U_t^* A^S U_t U_t^* B^S U_t - U_t^* B^S U_t U_t^* A^S U_t = U_t^* [A^S, B^S] U_t = U_t^* C U_t^S = C^H\,.$$

However the evolution equations are now different with

$$\frac{\mathrm{d}\Psi^H}{\mathrm{d}t} = 0\,, \qquad i\hbar \frac{\mathrm{d}A^H}{\mathrm{d}t} = [A^H, H]\,,$$

assuming that the operators have no explicit time dependence. In particular, the operators evolve (although, since $[H, H] = 0$, the Hamiltonian itself does not).

Example: The simple harmonic oscillator. For $H = \frac{P^2}{2m} + \frac{1}{2} m\omega^2 X^2$ we have

$$\frac{\mathrm{d}P}{\mathrm{d}t} = \frac{i}{\hbar}[H, P] = \frac{i}{\hbar}\left[\tfrac{1}{2} m\omega^2 X^2, P\right] = -m\omega^2 X$$

and

$$\frac{\mathrm{d}X}{\mathrm{d}t} = \frac{i}{\hbar}[H, X] = \frac{i}{\hbar}\left[\frac{m}{2} P^2, X\right] = \frac{P}{m}\,.$$

Thus if we define $A_t^{\pm} = P \pm i m\omega X$ we obtain

$$\frac{\mathrm{d}}{\mathrm{d}t} A_t^{\pm} = \pm i\omega A_t^{\pm}\,, \quad \text{so} \quad A_t^{\pm} = \mathrm{e}^{\pm i\omega t} A_0^{\pm}\,.$$

This gives complete soln for SHM in Heisenberg picture,although we still need to compute energy levels (see the algebraic method for example on the first problem sheet).

The interaction picture (Dirac)

Often $H = H_0 + H'$ with H_0 time independent such as the free particle such as $P^2/2m$ and H' an interaction term, perhaps a potential. When H' is small,we can use the above idea to separate the evolution of the large H_0 from the smaller H'. Essentially this picture is Heisenber for H_0 and Schrodinger for H'.

Set $W_t = \mathrm{e}^{-iH_0t/\hbar}$, $U_t = \mathrm{e}^{-iHt/\hbar}$, then the wave functions Ψ^I and operators A^I in the interaction picture will be defined by

$$\begin{aligned} \Psi^I_t &= W^*_t \Psi^S_t = W^*_t U_t \Psi^S_0 \\ A^I_t &= W^*_t A^S W_t\,. \end{aligned}$$

If instead we put $W = I$ this gives the Schrodinger picture, and $W_t = U_t$ this is the Heisenberg picture.

The same argument as (42) gives the equivalence of amplitudes in the different pictures. For the time evolution we now obtain

Theorem 11 *In the interaction picture we have*

$$i\hbar\frac{\partial \Psi^I_t}{\partial t} = H'_t \Psi^I_t\,, \qquad \frac{\mathrm{d}A^I_t}{\mathrm{d}t} = \frac{\partial A^I_t}{\partial t} - \frac{i}{\hbar}[A^I_t, H_0]$$

where the $\partial A/\partial t$ is a partial derivative with respect to the explicit time-dependence in A_t,

$$\frac{\partial A^I_t}{\partial t} = W^*_t \left(\frac{\mathrm{d}A^S_t}{\mathrm{d}t}\right) W_t,\,,$$

and H'_t is the interaction representation of H'

$$H'_t = W^*_t (H^S - H^S_0) W_t\,.$$

Proof: First note that $U_s U_t = U_{s+t}$ and so $U_t U^*_t = I$ implies $U^*_t = U_{-t}$ with similar identities for W_t. By definition

$$i\hbar\frac{\partial U_t}{\partial t} = HU_t\,, \quad i\hbar\frac{\partial W_t}{\partial t} = H_0 W_t\,, \quad i\hbar\frac{\partial W^*_t}{\partial t} = i\hbar\frac{\partial W_{-t}}{\partial t} = -H_0 W_{-t} = -W^*_t H_0\,.$$

Thus

$$\begin{aligned} i\hbar\frac{\mathrm{d}}{\mathrm{d}t}\Psi^I_t &= i\hbar\frac{\mathrm{d}}{\mathrm{d}t}\left(W^*_t U_t \Psi_0\right) \\ &= \left(W^*_t(-H_0)U_t + W^*_t H U_t\right)\Psi_0 \\ &= W^*_t (H - H_0) U_t \Psi_0 \\ &= H'_t \Psi_t \end{aligned}$$

as required. Similarly

$$\begin{aligned}
i\hbar\frac{\mathrm{d}}{\mathrm{d}t}A^I_t &= i\hbar\frac{\mathrm{d}}{\mathrm{d}t}\left(W^*_tA^SW_t\right)\\
&= i\hbar\left(W^*_t\frac{\mathrm{d}A^S}{\mathrm{d}t}W_t\right)+\left(i\hbar\frac{\mathrm{d}}{\mathrm{d}t}W^*_t\right)A^SW_t+W^*_tA^S\left(i\hbar\frac{\mathrm{d}}{\mathrm{d}t}W_t\right)\\
&= i\hbar\frac{\partial}{\partial t}A^I_t+W^*_t(-H_0)A^SW_t+W^*_tA^SW_tH_0\\
&= i\hbar\frac{\partial}{\partial t}A^I_t-[H_0,A^I_t]
\end{aligned}$$

as required, using $H_0W_t=W_tH_0$. □

This reduces to a proof for the Heisenberg equations when $H=H_0$.

8 Time dependent perturbation theory

We use the interaction picture to compare the evolution of an interacting system with Hamiltonian $H=H_0+H'$ to that of the free hamiltonian H_0. However, we cannot now simply exponentiate H' to obtain the evolution in the interaction picture because in the interaction picture, H'_t is now time dependent even if it wasnt before. We can nevertheless solve the equation as follows.

We integrate the evolution equation in the interaction picture to find

$$\Psi_t=\Psi_0+\frac{1}{i\hbar}\int_0^t H'_s\Psi_s\,\mathrm{d}s\,.$$

We can use this to set up an iteration scheme with

$$\begin{aligned}
\Psi^0_t &= \Psi_0\\
\Psi^{(N)}_t &= \Psi_0+\frac{1}{i\hbar}\int_0^t H'_s\Psi^{(N-1)}_s\,\mathrm{d}s
\end{aligned}\tag{43}$$

Iterating (as in Picard) we obtain

$$\Psi^{(N)}_t=\Psi_0+\sum_{n=1}^N(i\hbar)^{-n}\int_0^t\dots\int_0^{t_2}H'_{t_n}H'_{t_{n-1}}\dots H'_{t_1}\Psi_0\,\mathrm{d}t_1\,\mathrm{d}t_2\dots\,\mathrm{d}t_n\,.$$

Since H'_s is small, we can hope that this will converge and, although this is not always the case, we will here assume that it does. Since $\Psi_t = W_t^* U_t \Psi_0$, this gives the solution for the evolution operator U_t as

$$W_t^* U_t = 1 + \sum_{n=1}^{\infty} (i\hbar)^{-n} \int_0^t \dots \int_0^{t_2} H'_{t_n} H'_{t_{n-1}} \dots H'_{t_1} \, \mathrm{d}^n t \, .$$

This is known as a path-ordered exponential integral because, for H'_t independent of time we find for the nth term

$$\left(\frac{H'}{i\hbar}\right)^n \int_0^t \dots \int_0^{t_2} \mathrm{d}^n t = \frac{1}{n!} \left(\frac{H't}{i\hbar}\right)^n ,$$

the nth term in the expansion of an exponential. This particular idea has wide application.

If we now multiply by W_t and substitute in $H'_t = W_t^* H' W_t$ we obtain in the Schrodinger representation

Proposition 8.1 (The Feynman-Dyson expansion)

$$U_t = W_t + \sum_{n=1}^{\infty} \frac{1}{(i\hbar)^n} \int_0^t \dots \int_0^{t_2} W_{t-t_n} H' W_{t_n - t_{n-1}} H' \dots W_{t_2 - t_1} H' W_{t_1} \, \mathrm{d}^n \mathbf{t} \, . \tag{44}$$

where $\mathbf{t} = (t_1, \dots, t_n)$.

Proof: This is mainly driven by the 'group law' $W_t W_s = W_{t+s}$ and $W_t^* = W_{-t}$ since together these give

$$H'_{t_i} H'_{t_{i-1}} = W_{t_i}^* H' W_{t_i} W_{t_{i-1}}^* H' W_{t_{i-1}} = W_{t_n}^* H' W_{t_i - t_{i-1}} H' W_{t_{i-1}} ,$$

and so on. □

The leading order contribution is often known as the Born approximation, at least in the scattering context.

The nth term contribution can be thought of as an integral over the n times of a contribution in which the particle propagates according to the free Hamiltonian H_0 for the open time intervals between the t_i and is then hit by H' at each of the times $t_1 \leq t_2 \leq \dots \leq t_n$. This works even when the Schrodinger Hamiltonian has some explicit dependence on time. Indeed such time dependence is usually incorporated in the following section's treatment of Fermi's golen rule.

8.1 Fermi's golden rule

We wish to understand the probability as a function of time for a particle to change energy level associated with the time dependent disturbance. We have an undisturbed time-independent hamiltonian H_0 with energy Eigenstates ϕ_i, $H_0\phi_i = E_i\phi_i$. We addresss the question:

Problem: What is the leading order probability of transition from ϕ_1 to ϕ_2 at time t as a consequence of a time dependent disurbance H' when $\langle\phi_2|H'|\phi_1\rangle = V_{12}\mathrm{e}^{i\omega t}$, with V_{21} time independent so that the matrix element has frequency ω.

The probability we wish to compute is

$$P(t) := |\langle\phi_2|U_t\phi_1\rangle|^2 \,.$$

To zeroth and first order we have from the Feynman-Dyson expansion

$$\langle\phi_2|U_t|\phi_1\rangle = \langle\phi_2|W_t\phi_1\rangle + \frac{1}{i\hbar}\int_0^t \langle\phi_2|W_{t-s}H'W_s\phi_1\rangle \,\mathrm{d}s + \dots \,.$$

Firstly, since $W_t = \mathrm{e}^{H_0 t/i\hbar}$, we have

$$W_t\phi_i = \mathrm{e}^{E_i t/i\hbar}\phi_i. \tag{45}$$

Thus the zeroth order term $\langle\phi_2|W_t|\phi_1\rangle = \mathrm{e}^{E_1 t/i\hbar}\langle\phi_2|\phi_1\rangle = 0$ as the ϕ_i are orthogonal to each other. Thus the leading order contribution comes from the next term in the expansion:

$$\begin{aligned}
\frac{1}{i\hbar}\int_0^t \langle\phi_2|W_{t-s}H'W_s\phi_1\rangle \,\mathrm{d}s &= \frac{1}{i\hbar}\int_0^t \mathrm{e}^{E_2(t-s)/i\hbar}\mathrm{e}^{E_1 s/i\hbar}\langle\phi_2|H'\phi_1\rangle \,\mathrm{d}s \\
&= \frac{1}{i\hbar}\int_0^t \mathrm{e}^{E_2(t-s)/i\hbar}\mathrm{e}^{i\omega s}\mathrm{e}^{E_1 s/i\hbar}V_{21}\,\mathrm{d}s \\
&= \frac{1}{i\hbar}\mathrm{e}^{E_2 t/i\hbar}V_{21}\int_0^t \mathrm{e}^{(E_1-E_2-\hbar\omega)s/i\hbar}\,\mathrm{d}s \\
&= \frac{\mathrm{e}^{(E_1-E_2-\hbar\omega)t/i\hbar}-1}{E_1-E_2+\hbar\omega}\mathrm{e}^{E_2 t/i\hbar}V_{21}
\end{aligned}$$

Taking the modulus squared yields the probability

$$P(t) = \frac{4\sin^2(\mathcal{E}t/2\hbar)}{\mathcal{E}^2}|V_{21}|^2 \,, \qquad \text{where} \quad \mathcal{E} = E_1 - E_2 - \hbar\omega \,.$$

Thus the leading contribution is second order. It is peaked as expected where $\hbar\omega = E_2 - E_1$ as for a beam of light of frequency ω on say an atom leading to a transition of energy eigenstate with energy of order $\hbar\omega$ according to Einstein's energy quantum hypothesis. Note the periodic time dependence and the complementarity between energy and time. This can be sharpened by observing that

$$\lim_{t\to\infty} \frac{2\hbar \sin^2(\mathcal{E}t/2\hbar)}{\pi t \mathcal{E}^2} = \delta(\mathcal{E})\,.$$

so that for large times the energy uncertainty tends to zero. Here this result follows by observing that away from $\mathcal{E} = 0$ the left hand side goes to zero as $t \to \infty$ whereas at $\mathcal{E} = 0$ de l'Hopital's rule gives for the left hand side $t/2\hbar\pi$. Furthermore, the change of variables $u = \mathcal{E}t/2\hbar$ gives the integral over $\mathcal{E}$ as

$$\int_{-\infty}^{\infty} \frac{2\hbar \sin^2(\mathcal{E}t/2\hbar)}{\pi t \mathcal{E}^2}\, \mathrm{d}\mathcal{E} = \frac{1}{\pi}\int_{-\infty}^{\infty} \frac{\sin^2 u}{u^2}\, \mathrm{d}u = 1\,.$$

With this approximation for large time we have that the rate of production for large t is

$$\lim_{t\to\infty} \frac{P(t)}{t} = \frac{2\pi}{\hbar} |V_{21}|^2 \delta(\mathcal{E})\,.$$