7. Atoms

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1 **H** 1.008																	2 **He** 4.0026
3 **Li** 6.94	4 **Be** 9.0122											5 **B** 10.81	6 **C** 12.011	7 **N** 14.007	8 **O** 15.999	9 **F** 18.998	10 **Ne** 20.180
11 **Na** 22.990	12 **Mg** 24.305											13 **Al** 26.982	14 **Si** 28.085	15 **P** 30.974	16 **S** 32.06	17 **Cl** 35.45	18 **Ar** 39.948
19 **K** 39.098	20 **Ca** 40.078	21 **Sc** 44.956	22 **Ti** 47.867	23 **V** 50.942	24 **Cr** 51.996	25 **Mn** 54.938	26 **Fe** 55.845	27 **Co** 58.933	28 **Ni** 58.693	29 **Cu** 63.546	30 **Zn** 65.38	31 **Ga** 69.723	32 **Ge** 72.630	33 **As** 74.922	34 **Se** 78.97	35 **Br** 79.904	36 **Kr** 83.798
37 **Rb** 85.468	38 **Sr** 87.62	39 **Y** 88.906	40 **Zr** 91.224	41 **Nb** 92.906	42 **Mo** 95.95	43 **Tc** (98)	44 **Ru** 101.07	45 **Rh** 102.91	46 **Pd** 106.42	47 **Ag** 107.87	48 **Cd** 112.41	49 **In** 114.82	50 **Sn** 118.71	51 **Sb** 121.76	52 **Te** 127.60	53 **I** 126.90	54 **Xe** 131.29
55 **Cs** 132.91	56 **Ba** 137.33	57-71 *	72 **Hf** 178.49	73 **Ta** 180.95	74 **W** 183.84	75 **Re** 186.21	76 **Os** 190.23	77 **Ir** 192.22	78 **Pt** 195.08	79 **Au** 196.97	80 **Hg** 200.59	81 **Tl** 204.38	82 **Pb** 207.2	83 **Bi** 208.98	84 **Po** (209)	85 **At** (210)	86 **Rn** (222)
87 **Fr** (223)	88 **Ra** (226)	89-103 #	104 **Rf** (265)	105 **Db** (268)	106 **Sg** (271)	107 **Bh** (270)	108 **Hs** (277)	109 **Mt** (276)	110 **Ds** (281)	111 **Rg** (280)	112 **Cn** (285)	113 **Uut** (286)	114 **Fl** (289)	115 **Uup** (289)	116 **Lv** (293)	117 **Uus** (294)	118 **Uuo** (294)

* Lanthanide series	57 **La** 138.91	58 **Ce** 140.12	59 **Pr** 140.91	60 **Nd** 144.24	61 **Pm** (145)	62 **Sm** 150.36	63 **Eu** 151.96	64 **Gd** 157.25	65 **Tb** 158.93	66 **Dy** 162.50	67 **Ho** 164.93	68 **Er** 167.26	69 **Tm** 168.93	70 **Yb** 173.05	71 **Lu** 174.97
# Actinide series	89 **Ac** (227)	90 **Th** 232.04	91 **Pa** 231.04	92 **U** 238.03	93 **Np** (237)	94 **Pu** (244)	95 **Am** (243)	96 **Cm** (247)	97 **Bk** (247)	98 **Cf** (251)	99 **Es** (252)	100 **Fm** (257)	101 **Md** (258)	102 **No** (259)	103 **Lr** (262)

The periodic table is one of the most iconic images in science. All elements are classified in groups, ranging from metals on the left that go bang when you drop them in water through to gases on the right that don't do very much at all.

However, the periodic table contains plenty of hints that it is not the last word in science. There are patterns and order that run through it, all hinting at some deeper underlying structure. That structure, we now know, is quantum mechanics.

The most important pattern is also the most obvious: the elements are ordered, labelled by an integer, Z. This is the *atomic number* which counts the number of protons in the nucleus. The atomic number is the first time that the integers genuinely play a role in physics. They arise, like most other integers in physics, as the spectrum of a particular Schrödinger equation. This equation is rather complicated and we won't describe it in this course but, for what it's worth, it involves a Hamiltonian which describes the interactions of quarks and is known as the theory of *quantum chromodynamics*.

While the atomic number is related to the quantum mechanics of quarks, all the other features of the periodic table arise from the quantum mechanics of the electrons.

The purpose of this section is to explain some of the crudest features of the table from first principles. We will answer questions like: what determines the number of elements in each row? Why are there gaps at the top, and two rows at the bottom that we can't fit in elsewhere? What's special about the sequence of atomic numbers $2, 10, 18, 26, 54, 86, \ldots$ that label the inert gases?

We will also look at more quantitative properties of atoms, in particular their energy levels, and the *ionization energy* needed to remove a single electron. In principle, all of chemistry follows from solving the Schrödinger equation for some number of electrons. However, solving the Schrödinger equation for many particles is hard and there is a long path between "in principle" and "in practice". In this section, we take the first steps down this path.

7.1 Hydrogen

We're going to start by looking at a very simple system that consists of a nucleus with just a single electron. This, of course, is hydrogen.

Now I know what you're thinking: you already solved the hydrogen atom in your first course on quantum mechanics. But you didn't quite do it properly. There are a number of subtleties that were missed in that first attempt. Here we're going to explore these subtleties.

7.1.1 A Review of the Hydrogen Atom

We usually treat the hydrogen atom by considering an electron of charge $-e$ orbiting a proton of charge $+e$. With a view to subsequent applications, we will generalise this slightly: we consider a nucleus of charge Ze, still orbited by a single electron of charge $-e$. This means that we are also describing ions such as He^+ (for $Z = 2$) or Li^{2+} (for $Z = 3$). The Hamiltonian is

$$H = -\frac{\hbar^2}{2m}\nabla^2 - \frac{1}{4\pi\epsilon_0}\frac{Ze^2}{r} \tag{7.1}$$

The mass m is usually taken to be the electron mass m_e but since this is a two-body problem it's more correct to think of it as the *reduced mass*. (See, for example, Section 5.1.5 of the lectures on Dynamics and Relatvity.) This means that $m = m_e M/(m_e + M) \approx m_e - m_e^2/M$ where M is the mass of the nucleus. The resulting m is very close to the electron mass. For example, for hydrogen where the nucleus is a single proton, $M = m_p \approx 1836 m_e$.

The Schrödinger equation is the eigenvalue problem

$$H\psi = E_n\psi$$

This is the problem that you solved in your first course. The solutions are

$$\psi_{n,l,m}(r,\theta,\phi) = R_{n,l}(r)Y_{l,m}(\theta,\phi) \tag{7.2}$$

where $R_{n,l}(r)$ are the (generalised) Laguerre polynomials and $Y_{l,m}(\theta,\phi)$ are spherical harmonics. with energy eigenvalues. The states are labelled by three quantum numbers, n, l and m, which take integer values in the range

$$n = 1, 2, 3, \dots \quad , \quad l = 0, 1, \dots, n-1 \quad , \quad m = -l, \dots, +l$$

(Don't confuse the quantum number m with the mass m! Both will appear in formulae below, but it should be obvious which is which.) Importantly, the energy eigenvalue only depends on the first of these quantum numbers n,

$$E_n = -\left(\frac{Ze^2}{4\pi\epsilon_0}\right)^2 \frac{m}{2\hbar^2}\frac{1}{n^2} \qquad n \in \mathbf{Z}$$

where, just in case you weren't sure, it's the mass m that appears in this formula. This is sometimes written as

$$E_n = -\frac{Z^2\, Ry}{n^2}$$

where $Ry \approx 13.6\ eV$ is known as the *Rydberg energy*; it is the binding energy the ground state of hydrogen. Alternatively, it is useful to write the energy levels as

$$E_n = -\frac{(Z\alpha)^2 mc^2}{2n^2} \quad \text{where} \quad \alpha = \frac{e^2}{4\pi\epsilon_0\hbar c} \tag{7.3}$$

This may appear slightly odd as we've introduced factors of the speed of light c which subsequently cancel those in α. The reason for writing it in this way is that the quantity α is dimensionless. It is called the *fine structure constant* and takes the approximate value $\alpha \approx 1/137$. It should be thought of as the way to characterise the strength of the electromagnetic force.

Some Definitions

This energy spectrum can be seen experimentally as spectral lines. These are due to excited electrons dropping from one state n to a lower state $n' < n$, emitting a photon of fixed frequency $\hbar\omega = E_n - E_{n'}$. When the electron drops down to the ground state with $n' = 1$, the resulting lines are called the *Lyman series*. When the electron drops to higher states $n' > 1$, the sequences are referred to as *Balmer*, *Paschen* and so on.

Instead of using the angular momentum quantum number l to label the state, they are sometimes referred to as letters. $l = 0, 1, 2, 3$ are called s, p, d and f respectively. The names are old fashioned and come from the observed quality of spectral lines; they stand for *sharp*, *principal*, *diffuse* and *fundamental*, but they remain standard when describing atomic structure.

Degeneracy

The fact that the energy depends only on n and not on the angular momentum quantum numbers l and m means that each energy eigenvalue is degenerate. For fixed l, there are $2l + 1$ states labelled by m. Which means that for a fixed n, the total number of states is

$$\text{Degeneracy} = \sum_{l=0}^{n-1} 2l + 1 = n^2$$

Moreover, each electron also carries a spin degree of freedom. Measured along a given axis, this spin can either be up (which means $m_s = \frac{1}{2}$) or down ($m_s = -\frac{1}{2}$). Including this spin, the total degeneracy of states with energy E_n is

$$\text{Degeneracy} = 2n^2$$

The main reason for revisiting the quantum mechanics of hydrogen is to understand what becomes of this degeneracy. Before we proceed, it's worth first thinking about where this degeneracy comes from. Usually in quantum mechanics, any degeneracy is related to a conservation law which, in turn, are related to symmetries. The hydrogen atom is no exception.

The most subtle degeneracy to explain is the fact that the energy does not depend on l. This follows from the fact that the Hamiltonian (7.1) has a rather special conserved symmetry known as the Runge-Lenz vector. (We've met this in earlier courses in classical and quantum mechanics.) This follows, ultimately, from a hidden $SO(4)$ symmetry in the formulation of the hydrogen atom. We therefore expect that any deviation from (7.1) will lift the degeneracy in l.

Meanwhile, the degeneracy in m follows simply from rotational invariance and the corresponding conservation of angular momentum $\mathbf{L}$. We don't, therefore, expect this to be lifted unless something breaks the underlying rotational symmetry of the problem.

Finally, the overall factor of 2 comes, of course, from the spin **S**. The degeneracy must, therefore, follow from the conservation of spin. Yet there is no such conservation law; spin is just another form of angular momentum. The only thing that is really conserved is the total angular momentum $\mathbf{J} = \mathbf{L} + \mathbf{S}$. We would therefore expect any addition to the Hamiltonian (7.1) which recognises that only **J** is conserved to lift this spin degeneracy.

We'll now see in detail how this plays out. As we'll show, there are a number of different effects which split these energy levels. These effects collectively go by the name of *fine structure* and *hyperfine structure*.

7.1.2 Relativistic Motion

The "fine structure" corrections to the hydrogen spectrum all arise from relativistic corrections. There are three different relativistic effects that we need to take into account: we will treat the first here, and the others in Sections 7.1.3 and 7.1.4

You can run into difficulties if you naively try to incorporate special relativity into quantum mechanics. To things properly, you need to work in the framework of Quantum Field Theory and the *Dirac equation*, both of which are beyond the scope of this course. However, we're only going to be interested in situations where the relativistic effects can be thought of as small corrections on our original result. In this situation, it's usually safe to stick with single-particle quantum mechanics and use perturbation theory. That's the approach that we'll take here. Nonetheless, a number of the results that we'll derive below can only be rigorously justified by working with the Dirac equation.

The first, and most straightforward, relativistic shift of the energy levels comes simply from the fact that the effective velocity of electrons in an atom is a substantial fraction of the speed of light. Recall that the energy of a relativistic particle is

$$E = \sqrt{\mathbf{p}^2 c^2 + m^2 c^4} \;\approx\; mc^2 + \frac{\mathbf{p}^2}{2m} - \frac{\mathbf{p}^4}{8m^3c^2} + \ldots$$

The constant term mc^2 can be neglected and the next term is the usual non-relativistic kinetic energy which feeds into the Hamiltonian (7.1). Here we'll treat the third term as a perturbation of our hydrogen Hamiltonian

$$\Delta H = -\frac{\mathbf{p}^4}{8m^3c^2}$$

At first glance, it looks as if we're going to be dealing with degenerate perturbation theory. However, this particular perturbation is blind to both angular momentum quantum numbers l and m, as well as the spin m_s. This follows straightforwardly from the fact that $[\Delta H, \mathbf{L}^2] = [\Delta H, L_z] = 0$. If we denote the states (7.2) as $|nlm\rangle$, then it's simple to show that

$$\langle nlm|\Delta H|nl'm'\rangle = 0 \qquad \text{unless } l = l' \text{ and } m = m'$$

This means that the energy shifts are

$$(\Delta E_1)_{n,l} = \langle \Delta H \rangle_{n,l}$$

where we're introduced the notation $\langle \Delta H \rangle_{n,l} = \langle nlm|\Delta H|nlm\rangle$ and we've used the fact that perturbation preserves $SO(3)$ rotational invariance to anticipate that the change of energy won't depend on the quantum number m. We want to compute this overlap. In fact, it's simplest to massage it a little bit by writing

$$\Delta H = -\frac{1}{2mc^2}\left[H - V(r)\right]^2$$

where $V(r) = Ze^2/4\pi\epsilon_0 r$. This gives us the expression

$$(\Delta E_1)_{n,l} = -\frac{1}{2mc^2}\left[E_n^2 + 2E_n \langle V(r)\rangle_{n,l} + \langle V(r)^2\rangle_{n,l}\right] \tag{7.4}$$

and our new goal is to compute the expectation values $\langle 1/r\rangle_{n,l}$ and $\langle 1/r^2\rangle_{n,l}$ for the hydrogen atom wavefunctions.

The first of these follows from the virial theorem (see Section 6.1.3) which tells us that the relative contribution from the kinetic energy and potential energy is $2\langle T\rangle = -\langle V\rangle$, so that $\langle E\rangle = \langle T\rangle + \langle V\rangle = \frac{1}{2}\langle V\rangle$. Then,

$$\left\langle \frac{1}{r}\right\rangle_{n,l} = -\frac{1}{Z\alpha\hbar c}\langle V\rangle_{n,l} = -\frac{1}{Z\alpha\hbar c}\,2E_n = \frac{Z}{a_0}\frac{1}{n^2} \tag{7.5}$$

where $a_0 = \hbar/\alpha mc$ is the Bohr radius, the length scale characteristic of the hydrogen atom.

Next up is $\langle 1/r^2\rangle$. Here there's a cunning trick. For any quantum system, if we took the Hamiltonian H and perturbed it to $H + \lambda/r^2$, then the leading order correction to the energy levels would be $\langle \lambda/r^2\rangle$. But, for the hydrogen atom, such a perturbation can be absorbed into the angular momentum terms,

$$\frac{\hbar^2}{2m}\frac{l(l+1)}{r^2} + \frac{\lambda}{r^2} = \frac{\hbar^2}{2m}\frac{l'(l'+1)}{r^2}$$

But this is again of the form of the hydrogen atom Hamiltonian and we can solve it exactly. The only difference is that l' is no longer an integer but some function $l'(\lambda)$. The exact energy levels of the Hamiltonian with l' follow from our first course on quantum mechanics: they are

$$E(l') = -mc^2(Z\alpha)^2 \frac{1}{2(k+l'+1)^2}$$

Usually we would define the integer $n = k + l + 1$ to get the usual spectrum E_n given in (7.3). Here, instead, we Taylor expand $E(\lambda)$ around $\lambda = 0$ to get

$$\begin{aligned} E(l') &= E_n + (Z\alpha)^2 mc^2 \left[\frac{1}{(k+l'+1)^3} \frac{dl'}{d\lambda} \right] \Bigg|_{\lambda=0} \lambda + \dots \\ &= E_n + \frac{Z^2}{a_0^2} \frac{2\lambda}{n^3(2l+1)} + \dots \end{aligned}$$

From this we can read off the expectation value that we wanted: it is the leading correction to our exact result,

$$\left\langle \frac{1}{r^2} \right\rangle_{n,l} = \frac{Z^2}{a_0^2} \frac{2}{n^3(2l+1)} \tag{7.6}$$

The two expectation values (7.5) and (7.6) are what we need to compute the shift of the energy levels (7.4). We have

$$(\Delta E_1)_{n,l} = -\frac{(Z\alpha)^4 mc^2}{2} \left(\frac{n}{l+1/2} - \frac{3}{4} \right) \frac{1}{n^4} \tag{7.7}$$

As anticipated above, the relativistic effect removes the degeneracy in the quantum number l.

Notice that the size of the correction is of order $(Z\alpha)^4$. This is smaller than the original energy (7.3) by a factor of $(Z\alpha)^2$. Although we may not have realised it, $(Z\alpha)^2$ is the dimensionless ratio which we're relying on to be small so that perturbation theory is valid. (Or, for higher states, $(Z\alpha/n)^2$).

It's worth asking why we ended up with a perturbation to the energy which is smaller by a factor of $(Z\alpha)^2$. Since this was a relativistic correction, we expect it to be of order v^2/c^2 where v is the characteristic velocity of the electron. We can understand this by invoking the virial theorem which, in general, states that the expectation value of the kinetic energy $\langle T \rangle$ is related to the expectation value of the energy $V \sim r^n$ by $2\langle T \rangle = n\langle V \rangle$. For the hydrogen atom, this means that $\langle T \rangle = \frac{1}{2}m\langle v^2 \rangle = -\frac{1}{2}\langle V \rangle$. Since, from the ground state energy (7.3), we know that $E_1 = \langle T \rangle + \langle V \rangle = mc^2(Z\alpha)^2/2$ we have $\langle v^2 \rangle = (Z\alpha)^2 c^2$ which confirms that $(Z\alpha)^2$ is indeed the small parameter in the problem.

7.1.3 Spin-Orbit Coupling and Thomas Precession

The second shift of the energy levels comes from an interaction between the electron spin $\mathbf{S}$ and its angular momentum $\mathbf{L}$. This is known as the *spin-orbit coupling*.

The origin of the spin-orbit coupling is an interaction between the electron spin and a magnetic field and was described in Section 1.5. Here we review the basics. The fact first we will need is that spin endows the electron with a magnetic dipole moment given by

$$\mathbf{m} = -g\frac{e}{2m}\mathbf{S} \tag{7.8}$$

The coefficient of proportionality is called the *gyromagnetic ratio* or, sometimes, just the *g-factor*. As we explained in Section 1.5, to leading order $g = 2$ for the electron, a fact which follows from the Dirac equation.

The second fact that we need is that the energy of a magnetic moment $\mathbf{m}$ in a magnetic field $\mathbf{B}$ is given by

$$E = -\mathbf{B}\cdot\mathbf{m}$$

This is something we derived in Section 3 of the lectures on Electromagnetism.

The final fact is the Lorentz transformation of the electric field: as electron moving with velocity $\mathbf{v}$ in an electric field $\mathbf{E}$ will experience a magnetic field

$$\mathbf{B} = \frac{\gamma}{c^2}\mathbf{v}\times\mathbf{E}$$

This was derived in Section 5 of the lectures on Electromagnetism.

We now apply this to the electron in orbit around the nucleus. The electron experiences a radial electric field given by $\mathbf{E} = -\nabla\phi(r)$ with $\phi(r) = Ze/4\pi\epsilon_0 r$. Putting everything together, the resulting magnetic field interacts with the spin, giving rise to a correction to the energy of the electron

$$\Delta E = -\frac{e\gamma}{mc^2}(\mathbf{v}\times\nabla\phi)\cdot\mathbf{S} = -\frac{e}{(mc)^2}\frac{\partial\phi}{\partial r}\left(\mathbf{p}\times\hat{\mathbf{r}}\right)\cdot\mathbf{S} = \frac{e}{(mc)^2}\frac{1}{r}\frac{\partial\phi}{\partial r}\,\mathbf{L}\cdot\mathbf{S}$$

where $\mathbf{p} = m\gamma\mathbf{v}$ is the momentum and $\mathbf{L} = \mathbf{r}\times\mathbf{p}$ is the angular momentum. This is the promised spin-orbit coupling, in a form which we can promote to an operator. Thus the spin-orbit correction to the Hamiltonian is

$$\Delta H_{SO} = \frac{e}{(mc)^2}\frac{1}{r}\frac{\partial\phi}{\partial r}\,\mathbf{L}\cdot\mathbf{S} \tag{7.9}$$

Except. . . .

Thomas Precession

It turns that the interaction (7.9) is actually incorrect by a factor of 1/2. This is because of a subtle, relativistic effect known as *Thomas precession*.

Thomas precession arises because the electron orbiting the nucleus is in a non-inertial frame. As we will now explain, this means that even if the electron experienced no magnetic field, its spin would still precess around the orbit.

The basic physics follows from the structure of the Lorentz group. (See Section 7 of the lectures on Dynamics and Relativity.) Consider a Lorentz boost $\Lambda(v)$ in the x-direction, followed by a Lorentz boost $\Lambda'(v')$ in the y-direction. Some simple matrix multiplication will convince you that the resulting Lorentz transformation cannot be written solely as a boost. Instead, it is a boost together with a rotation,

$$\Lambda'(v')\Lambda(v) = R(\theta)\Lambda''(v'')$$

where $\Lambda''(v'')$ is an appropriate boost while $r(\theta)$ is a rotation in the $x - y$ plane. This rotation is known as the *Wigner rotation* (or sometimes the *Thomas rotation*). Although we will not need this fact below, you can check that $\cos\theta = (\gamma + \gamma')/(\gamma\gamma' + 1)$ with γ and γ' the usual relativistic factors.

Now we're going to apply this to a classical electron in orbit around the nucleus. At a fixed moment in time, it is moving with some velocity $\mathbf{v}$ relative to the nucleus. At some moment of time later, $\mathbf{v} + \delta\mathbf{v}$. The net effect of these two boosts is, as above, a boost together with a rotation.

If the electron were a point particle, the Wigner rotation would have no effect. However, the electron is not a point particle: it carries a spin degree of freedom $\mathbf{S}$ and this is rotated by the Wigner/Thomas effect. The cumulative effect of these rotations is that the spin precesses as the electron orbits the nucleus. We would like to calculate how much.

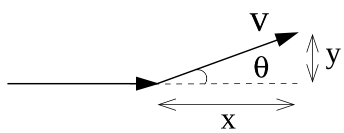

Figure 75:

The correct way to compute the precession is to integrate up the consecutive, infinitesimal Lorentz transformations as the electron orbits the nucleus. Here, instead, we present a quick and dirty derivation. We approximate the circular orbit of the electron by an N-sided polygon. Clearly in the lab frame, at the end of each segment the electron shifts it velocity by an angle $\theta = 2\pi/N$. However, in the electron's frame there is a Lorentz contraction along the direction parallel to the electron's motion. This means that the electron thinks it rotates by the larger angle $\tan\theta' = x/(\gamma/y)$ which, for N large, is $\theta' \approx 2\pi\gamma/N$. The upshot it that, by the time the electron has completed a full orbit, it thinks that it has rotated by an excess angle of

$$\Delta\theta = 2\pi(\gamma - 1) \approx \frac{2\pi v^2}{2c^2}$$

where we have expanded the relativistic factor $\gamma = (1 - v^2/c^2)^{-1/2} \approx 1 + v^2/2c^2$.

This is all we need to determine the precession rate, ω_T. If the particle traverses the orbit with speed v and period T, then

$$\omega_T = \frac{\Delta\theta}{T} \approx \frac{2\pi v^2}{2c^2 T} = \frac{av}{2c^2}$$

where, in the final step, we've replaced the period T with the acceleration $a = v^2/R = 2\pi v/T$.

Our derivation above tells us the angular precession. But what does this mean for a vector like $\mathbf{S}$? A little thought shows that the component of $\mathbf{S}$ that lies perpendicular to the plane of the orbit remains unchanged, while the component that lies within the plane precesses with frequency ω_T. In other words,

$$\frac{\partial \mathbf{S}}{\partial t} = \boldsymbol{\omega}_T \times \mathbf{S} \quad \text{with} \quad \boldsymbol{\omega}_{\mathrm{T}} = \frac{\mathbf{v} \times \mathbf{a}}{2\mathrm{c}^2} \tag{7.10}$$

This is *Thomas precession*. The effect is purely kinematic, due to the fact that the electron is not in an inertial frame. It can be thought of as a relativistic analog of the Coriolis force.

Finally, note that in several places above, we needed the assumption that v/c is small. Correspondingly, our final result (7.10) is only the leading order answer. The correct answer turns out to be

$$\omega_T = \frac{\gamma^2}{\gamma + 1} \frac{\mathbf{v} \times \mathbf{a}}{2c^2}$$

However, (7.10) will suffice for our purposes.

Thomas Precession and the Spin-Orbit Coupling

Let's now see how the existence of Thomas precession affects the spin orbit coupling. Again, we'll start with some basics. Classically, the energy $E = -(e/m)\mathbf{B} \cdot \mathbf{S}$ means that a spin will experience a torque when placed in a magnetic field. This, in turn, will cause it to precesss

$$\frac{\partial \mathbf{S}}{\partial t} = \mathbf{B} \times \mathbf{S}$$

However, we've seen that Thomas precession (7.10) gives a further contribution to this. So the correct equation should be

$$\frac{\partial \mathbf{S}}{\partial t} = \mathbf{B} \times \mathbf{S} + \boldsymbol{\omega}_T \times \mathbf{S}$$

The energy functional which gives rise to this is

$$E = \frac{e}{m}\mathbf{B} \cdot \mathbf{S} + \boldsymbol{\omega}_T \cdot \mathbf{S}$$

Working to leading order in v/c, we massage the second term as

$$\boldsymbol{\omega}_T \cdot \mathbf{S} = \frac{e}{2mc^2}(\mathbf{v} \times \nabla\phi) \cdot \mathbf{S} = -\frac{e}{2(mc)^2}\frac{1}{r}\frac{\partial \phi}{\partial r}\,\mathbf{L} \cdot \mathbf{S}$$

where we've used Newton's second law to write $m\mathbf{a} = e\nabla\phi$. We see that comes with the opposite sign and half the magnitude of the original contribution (7.9) to the energy. Adding the two together gives the final result for the correction to the Hamiltonian due to the spin-orbit coupling

$$\Delta H_{SO} = \frac{e}{2(mc)^2}\frac{1}{r}\frac{\partial \phi}{\partial r}\,\mathbf{L} \cdot \mathbf{S} \tag{7.11}$$

with $\phi(r)$ the electrostatic potential which, for us, is $\phi = Ze/4\pi\epsilon_0 r$.

Computing the Spin-Orbit Energy Shift

Before our perturbation, the electron states were labelled by $|nlm\rangle$, together with the spin $\pm 1/2$. The spin-orbit coupling will split the spin and angular momentum l degeneracy of the spectrum. To anticipate this, we should label these states by the total angular momentum

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

which takes quantum numbers $j = l \pm 1/2$ with $l = 0, 1, \ldots$. (When $l = 0$, we only have $j = 1/2$.) Each state can therefore be labelled by $|n, j, m_j; l\rangle$ where $|m_j| \leq j$ and the additional label l is there to remind us where these states came from.

We want to compute the eigenvalue of $\mathbf{L}\cdot\mathbf{S}$ acting on these states. The simplest way to do this is to consider $\mathbf{J}^2 = \mathbf{L}^2 + \mathbf{S}^2 + 2\mathbf{L}\cdot\mathbf{S}$, which tells us that

$$\begin{aligned} \mathbf{L}\cdot\mathbf{S}\,|n,j,m_j;l\rangle &= \frac{\hbar^2}{2}\left(j(j+1) - l(l+1) - \frac{3}{4}\right)|n,j,m_j;l\rangle \\ &= \frac{\hbar^2}{2}\begin{cases} -(l+1)|n,j,m_j;l\rangle & j = l-\frac{1}{2} \qquad (l\neq 0) \\ l|n,j,m_j;l\rangle & j = l+\frac{1}{2} \end{cases} \end{aligned} \tag{7.12}$$

As in Section 7.1.2, when computing degenerate perturbation theory with $|n,j,m_j;l\rangle$, the off-diagonal matrix elements vanish. We are left with the shift of the energy eigenvalues given by

$$(\Delta E_2)_{n,j;l} = \langle \Delta H_{SO}\rangle_{n,j;l}$$

where $\langle \Delta H_{SO}\rangle_{n,j;l} = \langle n,j,m_j;l|\Delta H_{SO}|n,j,m_j;l\rangle$.

With ΔH_{SO} given in (7.9), and $\phi(r) = Ze/4\pi\epsilon_0 r$, the shift of energy levels are

$$(\Delta E_2)_{n,j;l} = -\frac{Ze^2\hbar^2}{4.4\pi\epsilon_0(mc)^2}\begin{Bmatrix} -(l+1) \\ l \end{Bmatrix}\left\langle \frac{1}{r^3}\right\rangle_{n,j;l}$$

where, as in (7.12), the upper entry in $\{\cdot\}$ corresponds to $j = l - \frac{1}{2}$ (with $l \neq 0$) and the lower entry corresponds to $j = l + \frac{1}{2}$. Note that when $l = 0$, we have $\Delta E_2 = 0$ because there is no angular momentum for the spin to couple to.

In the previous section, we needed to compute $\langle 1/r\rangle$ and $\langle 1/r^2\rangle$. We see that now we need to compute $\langle 1/r^3\rangle$. Once again, there is a cute trick. This time, we introduce a new "radial momentum" observable

$$\tilde{p} = -i\hbar\left(\frac{\partial}{\partial r} + \frac{1}{r}\right)$$

It's simple to check that the radial part of the Hamiltonian can be written as

$$\begin{aligned} H &= -\frac{\hbar^2}{2m}\left(\frac{1}{r^2}\frac{\partial}{\partial r}r^2\frac{\partial}{\partial r}\right) + \frac{\hbar^2 l(l+1)}{2mr^2} - \frac{Ze^2}{4\pi\epsilon_0 r} \\ &= \frac{\tilde{p}^2}{2m} + \frac{\hbar^2 l(l+1)}{2mr^2} - \frac{Ze^2}{4\pi\epsilon_0 r} \end{aligned}$$

A quick computation shows that

$$[\tilde{p}, H] = -i\hbar\left(-\frac{\hbar^2 l(l+1)}{mr^3} + \frac{Ze^2}{4\pi\epsilon_0 r^2}\right)$$

Clearly this commutator doesn't vanish. However, when evaluated on an energy eigenstate, we must have $\langle [\tilde{p}, H] \rangle_{n,j,l} = 0$. From our expression above, this tells us that

$$\left\langle \frac{1}{r^3} \right\rangle_{n,j;l} = \frac{Z}{a_0} \frac{1}{l(l+1)} \left\langle \frac{1}{r^2} \right\rangle_{n,j;l} = \left(\frac{Z}{a_0}\right)^3 \frac{1}{l(l+\frac{1}{2})(l+1)} \frac{1}{n^3} \qquad (l \neq 0)$$

where we've used our earlier result (7.6) and, as before, $a_0 = \hbar/\alpha mc$ is the Bohr radius. Putting this together, and re-writing the resulting expression in terms of j rather than l, we find that the shift of energy levels due to spin-orbit coupling is

$$(\Delta E_2)_{n,j;l} = \frac{(Z\alpha)^4 mc^2}{4} \left\{ \begin{array}{c} -\frac{1}{j+1} \\ \frac{1}{j} \end{array} \right\} \frac{1}{j+\frac{1}{2}} \frac{1}{n^3}$$

This is the same order of magnitude as the first fine-structure shift (7.7) which, re-written in terms of $j = l \pm \frac{1}{2}$, becomes

$$(\Delta E_1)_{n,l} = -\frac{(Z\alpha)^4 mc^2}{2} \left(\left\{ \begin{array}{c} \frac{1}{j+1} \\ \frac{1}{j} \end{array} \right\} - \frac{3}{4n} \right) \frac{1}{n^3}$$

Combining these results, we get an expression which happily looks the same regardless of the minus sign in $j = l \pm \frac{1}{2}$. It is

$$(\Delta E_1)_{n,l} + (\Delta E_2)_{n,j;l} = \frac{(Z\alpha)^4 mc^2}{2} \left(\frac{3}{4n} - \frac{2}{2j+1} \right) \frac{1}{n^3} \tag{7.13}$$

where we should remember that for $l = 0$, $(\Delta E_2)_{n,j;l} = 0$ and we only get the $(\Delta E_1)_{n,l}$ term.

7.1.4 Zitterbewegung and the Darwin Term

There is one final contribution to the fine structure of the hydrogen atom. This one is somewhat more subtle than the others and a correct derivation really requires us to use the Dirac equation. Here we give a rather hand-waving explanation.

One of the main lessons from combining quantum mechanics with special relativity is that particles are not point-like. A particle of mass m has a size given by the *Compton wavelength*,

$$\lambda = \frac{\hbar}{mc}$$

For the electron, $\lambda \approx 3 \times 10^{-11}$ *cm*. Roughly speaking, if you look at a distance smaller than this you will see a swarm of particle and anti-particles and the single particle that you started with becomes blurred by this surrounding crowd.

Quantum field theory provides the framework to deal with this. However, within the framework of quantum mechanics it is something that we have to put in by hand. In this context, it is sometimes called *Zitterbewegung*, or "trembling motion". Suppose that a particle moves in a potential $V(\mathbf{r})$. Then, if the particle sits at position $\mathbf{r}_0$, it will experience the average of the potential in some region which is smeared a distance $\sim \lambda$ around $\mathbf{r}_0$. To include this, we Taylor expand the potential

$$V(\mathbf{r}) = V(\mathbf{r}_0) + \langle \Delta \mathbf{r} \rangle \cdot \frac{\partial V}{\partial \mathbf{r}} + \frac{1}{2} \langle \Delta r_i \Delta r_j \rangle \frac{\partial^2 V}{\partial r_i \partial r_j} + \ldots$$

By rotational symmetry, $\langle \Delta \mathbf{r} \rangle = 0$. Meanwhile, we take

$$\langle \Delta r_i \Delta r_j \rangle = \left(\frac{\lambda}{2} \right)^2 \delta_{ij}$$

I don't have an argument for the factor of $1/2$ on the right-hand-side of this expectation value. You will have to resort to the Dirac equation to see this. This gives a further contribution to the Hamiltonian, known as the *Darwin term*

$$\Delta H_{\text{Darwin}} = \frac{\hbar^2}{8m^2c^2} \nabla^2 V$$

For the Coulomb potential, this becomes

$$\Delta H_{\text{Darwin}} = \frac{Z\alpha\hbar^3}{8m^2c}\, 4\pi \delta^3(\mathbf{r})$$

However, all wavefunctions with $l > 0$ are vanishing at the origin and so unaffected by the Darwin term. Only those with $l = 0$, have a correction to their energy given by

$$(\Delta E_3)_{n,l} = \langle \Delta H_{\text{Darwin}} \rangle_{n,l} = \frac{Z\alpha\hbar^3\pi}{2m^2c}\, |\psi_{nlm}(\mathrm{r} = 0)|^2$$

The normalised wavefunction takes the form $\psi_{nlm}(\mathbf{r}) = R_{nl}(r) Y_{lm}(\theta, \phi)$. For $l = 0$, we have $Y_{00} = 1/\sqrt{4\pi}$ and the radial wavefunction take the form

$$R_{n,l=0}(r) = -\sqrt{\left(\frac{2Z}{na_0} \right)^3 \frac{(n-1)!}{2n(n!)^3}}\; e^{-r/na_0}\, L_n^1(2r/na_0)$$

Now we need to dig out some properties of Laguerre polynomials. We will need the facts that $L_n^1(x) = dL_n(x)/dx$ and $L_n(x) \approx n! - n!nx + \mathcal{O}(x^2)$ so that $L_n^1(0) = n!n$. The wavefunction at the origin then becomes

$$|\psi_{n,l=0}(0)|^2 = \frac{Z^3}{a_0^3 \pi n^3} \tag{7.14}$$

From this we get

$$(\Delta E_3)_{n,l} = \frac{(Z\alpha)^4 mc^2}{2} \frac{1}{n^3} \delta_{l0} \tag{7.15}$$

7.1.5 Finally, Fine-Structure

It's been quite a long journey. Our fine structure calculations have revealed three contributions, the first two given by (7.13) and the third by (7.15). Recall that the spin-orbit coupling in (7.13) gave vanishing contribution when $l = 0$. Rather curiously, the Darwin term gives a contribution only when $l = 0$ which coincides with the formal answer for the spin-orbit coupling when $l = 0$, $j = 1/2$. The upshot of this is that the answer (7.13) we found before actually holds for all l. In other words, adding all the contributions together, $(\Delta E)_{n,j} = (\Delta E_1)_{n,l} + (\Delta E_2)_{n,j;l} + (\Delta E_3)_{n,l}$, we have our final result for the fine structure of the hydrogen atom

$$(\Delta E)_{n,j} = \frac{(Z\alpha)^4 mc^2}{2}\left(\frac{3}{4n} - \frac{2}{2j+1}\right)\frac{1}{n^3}$$

We learn that the energy splitting depends only on j. This didn't have to be the case. There is no symmetry that requires states with $j = |l \pm \frac{1}{2}|$ to have the same energy. We refer to this as an *accidental degeneracy*. Meanwhile, the energy of each state is independent of the remaining angular momentum quantum number $m \leq l$. This is not accidental: it is guaranteed by rotational invariance.

To describe the states of hydrogen, we use the notation $n\#_j$ where we replace # with the letter that denotes the orbital angular momentum l. The ground state is then $1s_{1/2}$. This is doubly degenerate as there is no angular momentum, so the spin states are not split by spin-orbit coupling. The first excited states are $2s_{1/2}$ (two spin states) which is degenerate with $2p_{1/2}$ (three angular momentum states). Similarly, as we go up the spectrum we find that the $3p_{3/2}$ and $3d_{3/2}$ states are degenerate and so on.

The Result from the Dirac Equation

Our fine structure calculations have all treated relativistic effects perturbatively in v^2/c^2. As we explained in Section 7.1.2, for the hydrogen atom this is equivalent to an expansion in $1/(Z\alpha)^2$. In fact, for this problem there is an exact answer. The derivation of this requires the Dirac equation and is beyond the scope of this course; instead we simply state the answer. The energy levels of the relativistic hydrogen atom are given by

$$E_{n,j} = mc^2\left(1 + \left(\frac{Z\alpha}{n - j - \frac{1}{2} + \sqrt{(j+\frac{1}{2})^2 - (Z\alpha)^2}}\right)^2\right)^{-1/2} \tag{7.16}$$

Expanding in $1/(Z\alpha)$ gives

$$E_{n,j} = mc^2\left(1 - (Z\alpha)^2\frac{1}{2n^2} + (Z\alpha)^4\left(\frac{3}{4n} - \frac{2}{2j+1}\right)\frac{1}{2n^3} + \ldots\right)$$

The first term is, of course, the rest mass of the electron. The second term is the usual hydrogen binding energy, while the final term is the fine structure corrections that we've laboriously computed above.

The Lamb Shift

It turns out that the "exact" result (7.16) is not exact at all! In 1947, Willis Lamb reported the experimental discovery of a splitting between the $2s_{1/2}$ and $2p_{1/2}$ states. For this, he won the 1955 Nobel prize. The effect is now referred to as the *Lamb shift*.

The Lamb shift cannot be understood using the kind of single-particle quantum mechanics that we're discussing in this course. It is caused by quantum fluctuations of the electromagnetic field and needs the full machinery of quantum field theory, specifically *quantum electrodymamics*, or *QED* for short. Historically the experimental discovery of the Lamb shift was one of the prime motivations that led people to develop the framework of quantum field theory.

7.1.6 Hyperfine Structure

Both the fine structure corrections and the QED corrections treat the nucleus of the atom as a point-like object. This means that, although the corrections are complicated, the problem always has rotational symmetry.

In reality, however, the nucleus has structure. This structure effects the atomic energy levels, giving rise to what is called *hyperfine structure*. There are a number of different effects that fall under this heading.

The most important effects come from the magnetic dipole moment of the nucleus. Each constituent neutron and proton is a fermion, which means that they have an internal intrinsic spin 1/2. This is described by the quantum operator $\mathbf{I}$. This, in turn, gives the nucleus a magnetic dipole moment

$$\mathbf{m}_N = g_N \frac{Ze}{2M}\mathbf{I}$$

This takes the same form as (7.8) for the electron magnetic moment. Here M is the mass of the nucleus while g_N is the nucleus gyromagnetic factor.

The Dirac equation predicts that every fundamental fermion has $g = 2$ (plus some small corrections). However, neither the proton nor the neutron are fundamental particles. At a cartoon level, we say that each is made of three smaller particles called quarks. The reality is much more complicated! Each proton and neutron is made of many hundreds of quarks and anti-quarks, constantly popping in an out of existence, bound together by a swarm of further particles called gluons. It is, in short, a mess. The cartoon picture of each proton and neutron containing three quarks arises because, at any given time, each contains three more quarks than anti-quarks.

The fact that the protons and neutrons are not fundamental first reveals itself in their anomalously large gyromagnetic factors. These are

$$g_{\text{proton}} \approx 5.56 \quad \text{and} \quad g_{\text{neutron}} \approx -3.83$$

The minus sign for the neutron means that a neutron spin precesses in the opposite direction to a proton spin. Moreover, the spins point in opposite directions in their ground state.

Now we can describe the ways in which the nuclear structure affects the energy levels of the atom

- Both the electron and the nucleus carry a magnetic moment. But we know from our first course on Electromagnetism that the is an interaction between nearby magnetic moments. This will lead to a coupling of the form $\mathbf{I} \cdot \mathbf{S}$ between the nucleus and electron spins.
- The orbital motion of the electron also creates a further magnetic field, parallel to $\mathbf{L}$. This subsequently interacts with the magnetic moment of the nucleus, resulting in a coupling of the form $\mathbf{I} \cdot \mathbf{L}$.
- The nucleus may have an electric quadrupole moment. This means that the electron no longer experiences a rotationally invariant potential.

For most purposes, the effects due to the nuclear magnetic moment are much larger than those due to its electric quadrupole moment. Here we restrict attention to s-wave states of the electron, so that we only have to worry about the first effect above.

To proceed, we first need a result from classical electromagnetism. A magnetic moment $\mathbf{m}_N$ placed at the origin will set up a magnetic field

$$\mathbf{B} = \frac{2\mu_0}{3}\mathbf{m}_N \delta^3(0) + \frac{\mu_0}{4\pi r^3}\left(3(\mathbf{m}_N \cdot \hat{\mathbf{r}})\hat{\mathbf{r}} - \mathbf{m}_N\right) \tag{7.17}$$

The second term is the long-distance magnetic field and was derived in Section 3 of the Electromagnetism lectures. The first term is the magnetic field inside a current loop, in the limit where the loop shrinks to zero size, keeping the dipole moment fixed. (It actually follows from one of the problem sheet questions in the Electromagnetism course.)

The electron spin interacts with this nuclear magnetic field through the hyperfine Hamiltonian

$$\Delta H = -\mathbf{m} \cdot \mathbf{B} = \frac{e}{m} \mathbf{S} \cdot \mathbf{B}$$

For the s-wave, the contribution from the second term in (7.17) vanishes and we only have to compute the first term. Writing the magnetic moments in terms of the spin, and using the expression (7.14) for the s-wave wavefunction at the origin, the hyperfine Hamiltonian becomes

$$\begin{aligned} \Delta H &= \frac{2\mu_0 g_N Z e^2}{6Mm} \, |\psi_{n,l=0}(0)|^2 \, \mathbf{S} \cdot \mathbf{I} \\ &= \frac{4}{3} \frac{m}{M} (Z\alpha)^4 mc^2 \frac{1}{n^3} \frac{1}{\hbar^2} \, \mathbf{S} \cdot \mathbf{I} \end{aligned} \tag{7.18}$$

where, in the second line, we've used our previous expression (7.14) for the value of the wavefunction at the origin, $|\psi_{n,l=0}(0)|^2 = Z^3/a_0^3 \pi n^3$, together with the usual definitions $a_0 = \hbar/\alpha mc$ and $\alpha = e^2/4\pi\epsilon_0 \hbar c$

We see that the hyperfine splitting (7.18) has the same parametric form as the fine structure, with the exception that it is further suppressed by the ratio of masses m/M. For hydrogen with $Z = 1$, we should take $M = m_p$, the proton mass, and $m/m_p \approx 1/1836$. So we expect this splitting to be three orders of magnitude smaller than the fine structure splitting.

We can evaluate the eigenvalues of the operator $\mathbf{S} \cdot \mathbf{I}$ in the same way as we dealt with the spin orbit coupling in Section 7.1.3. We define the total spin as $\mathbf{F} = \mathbf{S} + \mathbf{I}$. For hydrogen, where both the electron and proton have spin 1/2, we have

$$\frac{1}{\hbar^2} \mathbf{S} \cdot \mathbf{I} = \frac{1}{2\hbar^2} \left(\mathbf{F}^2 - \mathbf{S}^2 - \mathbf{I}^2 \right) = \frac{1}{2} \left(F(F+1) - \frac{3}{2} \right) = \frac{1}{2} \begin{cases} -\frac{3}{2} & F = 0 \\ \frac{1}{2} & F = 1 \end{cases} \tag{7.19}$$

This gives rise to the splitting between the spin up and spin down states of the electron. Or, equivalently, between the total spin $F = 0$ and $F = 1$ of the atom.

The 21cm Line

The most important application of hyperfine structure is the splitting of the $1s_{1/2}$ ground state of hydrogen. As we see from (7.19), the $F = 0$ spin singlet state has lower energy than the $F = 1$ spin state. The energy difference is

$$\Delta E_{1s_{1/2}} = \frac{4\alpha^4 m^2 c^2}{3M} \approx 9.39 \times 10^{-25} \ J$$

This is small. But its not that small. The temperature of the cosmic microwave background is $T \approx 2.7\ K$ which corresponds to an energy of $E = k_B T \approx 3.7 \times 10^{-23}\ J > \Delta E_{1s_{1/2}}$. This means that the hydrogen that is spread throughout space, even far from stars and galaxies, will have its $F = 1$ states excited by the background thermal bath of the universe.

When an electron drops from the $F = 1$ state to the $F = 0$ state, it emits a photon with energy $\Delta E_{1s_{1/2}}$. This has frequency $\sim 1400\ MHz$ and wavelength ~ 21 cm. This is important. The wavelength is much longer than the size of dust particles which float around in space, blocking our view. This means that, in contrast to visible light, the 21cm emission line from hydrogen can pass unimpeded through dust. This makes it invaluable in astronomy and cosmology.

For example, the hydrogen line allowed us to discover that our home, the Milky way, is a spiral galaxy. In this case, the velocity of the hydrogen gas in the spiral arms could be detected by the red-shift of the 21cm line. Similarly, the 21cm line has allowed us to map the distribution of hydrogen around other galaxies. It shows that hydrogen sitting in the outskirts of the galaxies is rotating much to fast to be held in place by the gravity from the visible matter alone. This is one of the key pieces of evidence for dark matter. An example from the KAT7 telescope, a precursor to the square kilometer array, is shown on the right. The green contours depict the hydrogen, as measured by the 21cm line, stretching far beyond the visible galaxy.

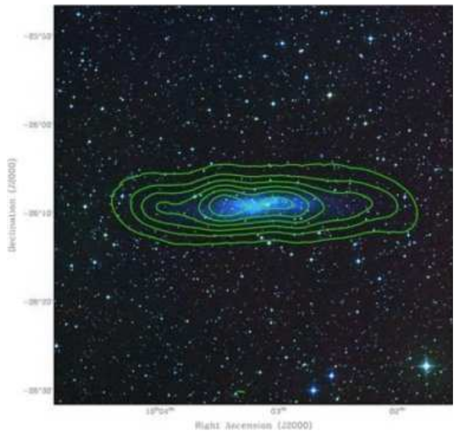

Figure 76:

Looking forwards, there is optimism that the 21cm line will allow us to see the "dark ages" of cosmology, the period several hundreds of millions of years between when the fireball of the Big Bang cooled and the first stars appeared.

Caesium

Caesium has atomic number 55 and. Its nucleus has spin $I = 7/2$. The mixing with the outer electron spin results in a hyperfine splitting of the ground state into two states, one with $F = 3$ and the other with $F = 4$. The frequency of the transition between these is now used as the definition of a time. A second is defined as 9192631770 cycles of the hyperfine transition frequency of caesium 133.

7.2 Atomic Structure

In this section, we finally move away from hydrogen and discuss atoms further up the periodic table. The Hamiltonian for N electrons orbiting a nucleus with atomic number Z is

$$H = \sum_{i=1}^{N}\left(-\frac{\hbar^2}{2m}\nabla_i^2 - \frac{Ze^2}{4\pi\epsilon_0}\frac{1}{r_i}\right) + \sum_{i<j}\frac{e^2}{4\pi\epsilon_0}\frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} \tag{7.20}$$

For a neutral atom, we take $N = Z$. However, in what follows it will be useful to keep N and Z independent. For example, this will allows us to describe ions.

We can, of course, add to this Hamiltonian relativistic fine structure and hyperfine structure interactions of the kind we described in the previous section. We won't do this. As we will see, the Hamiltonian (7.20) will contain more than enough to keep us busy. Our goal is the find its energy eigenstates. Further, because electrons are fermions, we should restrict ourselves to wavefunctions which are anti-symmetric under the exchange of any two electrons.

It is a simple matter to write down the Schrödinger equation describing a general atom. It is another thing to solve it! No exact solutions of (7.20) are known for $N \geq 2$. Instead, we will look at a number of different approximation schemes to try to understand some aspects of atomic structure. We start in this section by making the drastic assumption that the electrons don't exert a force on each other. This is not particularly realistic, but it means that we can neglect the final interaction term in (7.20). In this case, the Hamiltonian reduces to N copies of

$$H_0 = -\frac{\hbar^2}{2m}\nabla^2 - \frac{Ze^2}{4\pi\epsilon_0}\frac{1}{r}$$

This, of course, is the Hamiltonian for the hydrogen atom, albeit with the proton charge $+e$ replaced by Ze. And, as reviewed in Section 7.1.1, we know everything about the solutions with this Hamiltonian.

7.2.1 A Closer Look at the Periodic Table

Ignoring the interaction between electrons gives us an eminently solvable problem. The only novelty comes from the Pauli exclusion principle which insists that no two electrons can sit in the same state. The ground state of a multi-electron atom consists of filing the first Z available single-particle states of the hydrogen atom.

However, as we've seen above, there is a large degeneracy of energy levels in the hydrogen atom. This means that, for general Z, the rule above does not specify a unique ground state for the atom. Nonetheless, when Z hits certain magic numbers, there will be a unique ground state. This occurs when there are exactly the right number of electrons to fill energy levels. Those magic numbers are:

n	l	Degeneracy	N
1	0	2	2
2	0,1	$2 \times (1+3) = 8$	$2+8=10$
2	0,1,2	$2 \times (1+3+5) = 18$	$2+10+18=28$
3	0,1,2,3	$2 \times (1+3+5+7) = 32$	$2+8+18+32=60$

This simple minded approach suggests that at the magic numbers $Z = 2, 10, 28, 60, \ldots$ the atoms will have a full shell of electrons. If we were to add one more electron it would have to sit in a higher energy level, so would be less tightly bound. We might, then, want to predict from our simple minded non-interacting model that atoms with these special values of Z will be the most chemically stable.

A look at the periodic table shows that our prediction is not very impressive! We learn in school that the most chemically stable elements are the inert Noble gases on the far right. We can quantify this by looking at the ionization energies of atoms as a function of Z, as shown on the right which shows that the most stable elements have $Z = 2, 10, 18, 36, 54, 86$ and 118.

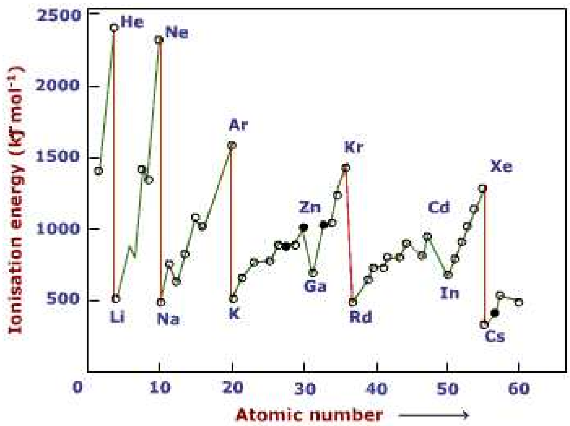

Figure 77:

We see that our non-interacting model gets the first two numbers right, but after that it all goes pear shaped. In particular, we predicted that $Z = 28$ would be special but this corresponds to nickel which sits slap in the middle of the transition metals! Meanwhile, we missed argon, a stable Noble gas with $Z = 18$. Of course, there's no secret about what we did wrong. Our task is to find a way to include the interactions between electrons to explain why the Noble gases are stable.

Before we return to the Schrödinger equation, we will build some intuition by looking more closely at the arrangement of electrons that arise in the periodic table. First some notation. We describe the *configuration* of electrons by listing the hydrogen orbitals that are filled, using the notation $n\#^p$ where # is the letter (s, p, d, f, etc.) denoting the l quantum number and p is the number of electrons in these states.

The electrons which have the same value of n are said to sit in the same *shell*. Electrons that have the same value of n and l are said to sit in the same *sub-shell*. Each sub-shell contains $2(l+1)$ different states. Electrons which sit in fully filled shells (or sometimes sub-shells) are said to be part of the *core* electrons. Those which sit in partially filled shells are said to form the *valence* electrons. The valence electrons lie farthest from the nucleus of the atom and are primarily responsible for its chemical properties.

There are only two elements with electrons lying in the $n = 1$ shell. These are hydrogen and helium

Z	1	2
Element	H	He
Electrons	$1s^1$	$1s^2$

The elements with electrons in the first two shells are

Z	3	4	5	6	7	8	9	10
	Li	Be	B	C	N	O	F	Ne
[He]+	$2s^1$	$2s^2$	$2s^22p^1$	$2s^22p^2$	$2s^22p^3$	$2s^22p^4$	$2s^22p^5$	$2s^22p^6$

where the notation in the bottom line means that each element has the filled $n = 1$ shell of helium, together with the extra electrons listed. We see that the atoms seem to be following a reasonable pattern but, already here, there is a question to answer that does not follow from our non-interacting picture: why do the electrons prefer to first fill up the $2s$ states, followed by the $2p$ states?

The next set of atoms in the periodic table have electrons in the third shell. They are

Z	11	12	13	14	15	16	17	18
	Na	Mg	Al	Si	P	S	Cl	Ar
[Ne]+	$3s^1$	$3s^2$	$3s^23p^1$	$3s^23p^2$	$3s^23p^3$	$3s^23p^4$	$3s^23p^6$	$3s^23p^6$

where now the electrons fill the $2s^22p^6$ states of neon, together with those listed on the bottom line. Again, we see that the $3s$ level fills up before the $3p$, something which we will later need to explain. But now we see that it's sufficient to fill the $3p$ states to give a chemically inert element. This suggests that there is a big energy gap between between $3p$ and $3d$, something that we would like to understand.

In the next row of elements, we see another surprise. We have

Z	19	20	21	22	...	30	31	...	36
	K	Ca	Sc	Ti	...	Zn	Ga	...	Kr
[Ar]+	$4s^1$	$4s^2$	$3d^14s^2$	$3d^24s^2$	...	$3d^{10}4s^2$	$3d^{10}4s^24p^1$	...	$3d^{10}4s^24p^6$

We see that we fill the $4s$ states *before* the $3d$ states. This is now in direct contradiction to the non-interacting model, which says that $4s$ states should have greater energy that $3d$ states.

There is a simple rule which chemists employ to explain the observed structure. It is called the *aufbau principle* and was first suggested by Bohr. It says that you should write all possible $n\#$ energy levels in a table as shown to the right. The order in which the energy levels are filled is set by the arrows: first $1s$, followed by $2s$, $2p$, $3s$, and then $3p$, $4s$, $3d$, $4p$ and so on. This explains the observed filling above. Our task in these lectures is to explain where the aufbau principle comes from, together with a number of further rules that chemists will wheel out.

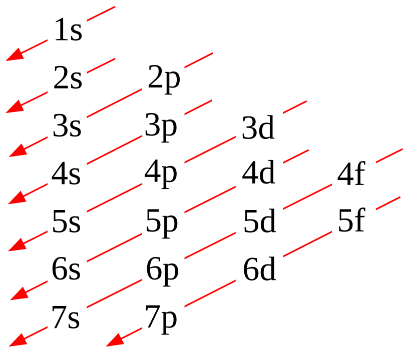

Figure 78: Aufbau

The aufbau principle also explains why the periodic table needs those two extra lines at the bottom: after we fill $6s$ (Cs and Ba) we move $4f$ which has 14 states. These are elements $Z = 58$ to $Z = 71$. Element $Z = 57$ is La and is an exception to the aufbau principle, with electron configuration $[\text{Xe}]5d^16s^2$.

In fact, the "*aufbau principle*" is more an "*aufbau rule of thumb*". As we go to higher values of Z there are an increasing number of anomalies. Some of these are hidden in the ... in the last table above. Vanadium with $Z = 23$ has electron configuration $[\text{Ar}]3d^34s^2$, but it is followed by chromium with $Z = 24$ which has $[\text{Ar}]3d^54s^1$. We see that the $4s$ state became depopulated, with an extra electron sitting in $3d$. By the time we get to manganese at $Z = 26$, we're back to $[\text{Ar}]3d^54s^2$, but the anomaly occurs again for copper with $Z = 29$ which has $[\text{Ar}]3d^{10}4s^1$.

Even scandium, with $Z = 21$, hides a failure of the aufbau principle. At first glance, it would seem to be a poster child for aufbau, with its configuration $[\mathrm{Ar}]3d^1 4s^2$. But if we strip off an electron to give the ion Sc^+, we have $[\mathrm{Ar}]3d^1 4s^1$. Stripping off a further electron, Sc^{++} has $[\mathrm{Ar}]3d^1$. Neither of these follow aufbau. These anomalies only get worse as we get to higher Z. There are about 20 neutral atoms which have anomalous fillings and many more ions.

We will not be able to explain all these anomalies here. Indeed, even to derive the aufbau principle we will have to resort to numerical results at some stage. We will, however, see that multi-electron atoms are complicated! In fact, it is rather surprising that they can be accurately described using 1-particle states at all. At the very least you should be convinced that there need not be a simple rule that governs all of chemistry.

7.2.2 Helium and the Exchange Energy

We're going to start by looking at the simplest example of a multi-electron atom: helium. This will start to give some physical intuition for the aufbau principle. It will also help reveal the role that the spin of the electron plays in the energy of states.

The Ground State of Helium

We've already discussed the ground state of Helium in Section 6.1.2 as an example of the variational method. Let's first recap the main results of that analysis.

In the ground state, both electrons sit in the $1s$ state, so that their spatial wavefunction takes the form

$$\Psi(\mathbf{r}_1, \mathbf{r}_2) = \psi_{1,0,0}(\mathbf{r}_1)\psi_{1,0,0}(\mathbf{r}_2) \quad \text{with} \quad \psi_{1,0,0}(\mathbf{r}) = \sqrt{\frac{Z^3}{\pi a_0^3}}\, e^{-Zr/a_0} \tag{7.21}$$

Here $a_0 = 4\pi\epsilon_0\hbar^2/me^2$ is the Bohr radius. For helium, we should pick $Z = 2$.

Since the spatial wavefunction is symmetric under exchange of the particles, we rely on the spin degrees of freedom to provide the necessary anti-symmetry of the full wavefunction. The spins must therefore sit in the singlet state

$$|0,0\rangle = \frac{|\uparrow\rangle|\downarrow\rangle - |\downarrow\rangle|\uparrow\rangle}{\sqrt{2}} \tag{7.22}$$

Computing the shift of energy is a simple application of first order perturbation theory. The interaction Hamiltonian is

$$H_{\text{int}} = \frac{e^2}{4\pi\epsilon_0}\frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|} \tag{7.23}$$

and, correspondingly, the shift of the ground state energy is given by

$$\Delta E = \frac{e^2}{4\pi\epsilon_0}\int d^3r_1 d^3r_2\; \frac{|\psi_{1,0,0}(\mathbf{r}_1)|^2|\psi_{1,0,0}(\mathbf{r}_2)|^2}{|\mathbf{r}_1 - \mathbf{r}_2|}$$

We showed how to compute this integral in Section 6.1.2 and found $\Delta E = \frac{5}{4}Z\,Ry$. This then gives a total ground state energy of $E_0 \approx -74.8\ eV$ which, given the lack of control of perturbation theory, is surprisingly close to the true value $E_0 \approx -79\ eV$.

We also learned in Section 6.1.2 that we can do better using a variational ansatz. Although we will not employ this technique below, there is a physics lesson that it's useful to highlight. In the variational method, we again work with the form of the wavefunction (7.21), but this time allow the atomic number Z of the nucleus to be our variational parameter. We found that we can achieve a lower ground state energy, $E_0 \approx -77.5\ eV$ — one which is closer to the true value — if instead of setting $Z = 2$ in the wavefunction, we take

$$Z = 2 - \frac{5}{16}$$

There is some physical intuition behind this result. Each electron sees the charge $Z = 2$ of the nucleus reduced somewhat by the presence of the other electron. This is called *screening* and it is the basic phenomenon which, ultimately, underlies much of the physics of atomic structure.

Excited States of Helium

Let's now extend our discussion to the first excited state of helium. From our non-interacting model, there are two possibilities which, as far as non-interacting electrons are concerned, are degenerate. These are $1s^1 2s^1$ and $1s^1 2p^1$. We would like to understand which of these has lowest energy.

In fact, there is a further splitting of each of these states due to the spin-degrees of freedom. To understand this splitting, we need to recall the following:

- The Hamiltonian is blind to the spin degrees of freedom. This means that the wavefunction takes the form of a tensor product of a spatial state with a spin state.
- Electrons are fermions. This means that the overall wavefunction must be anti-symmetric under exchange of the two particles.

There are two ways to achieve the anti-symmetry: we either make the spatial wavefunction symmetric and the spin wavefunction anti-symmetric, or vice versa. The two possibilities for the spatial wavefunction are

$$\Psi_{ab\pm}(\mathbf{r}_1, \mathbf{r}_2) = \frac{1}{\sqrt{2}} \left(\psi_a(\mathbf{r}_1)\psi_b(\mathbf{r}_2) \pm \psi_a(\mathbf{r}_2)\psi_b(\mathbf{r}_1)\right)$$

where we're using the notation a, b to denote the triplet of quantum numbers of (n, l, m). For the first excited states, we should take $a = (1, 0, 0)$. Then $b = (2, 0, 0)$ for the $1s^1 2s^1$ state and $b = (2, 1, m)$ for the triplet of $1s^1 2p^1$ states, with $m = -1, 0, 1$

The symmetric wavefunctions $\Psi_{ab,+}$ must be combined with the anti-symmetric spin-singlet (7.22) which we write as

$$|ab; s = 0\rangle = \Psi_{ab,+}(\mathbf{r}_1, \mathbf{r}_2) \otimes |0, 0\rangle \tag{7.24}$$

where $|0, 0\rangle$ is the spin singlet state defined in (7.22). Note that we shouldn't confuse the $s = 0$ spin with the label "s" used to denote the $l = 0$ atomic orbital. They are different! Also, I've been a bit lax about my notation for wavefunctions: the expression above should really read $|ab; s = 0\rangle = |\Psi_{ab,+}\rangle \otimes |0, 0\rangle$ where the fermionic two-particle state $|\Psi_+\rangle$ has overlap $\Psi_{ab,+}(\mathbf{r}_1, \mathbf{r}_2) = \langle \mathbf{r}_1, \mathbf{r}_2 | \Psi_{ab,+}\rangle$ with the position basis of two-particle states $|\mathbf{r}_1, \mathbf{r}_2\rangle$. This, more precise, notation turns out to be somewhat more cumbersome for our needs.

Similarly, the anti-symmetric wavefunction must be paired with the symmetric spin states. There is a triplet of such states, $|s = 1; m_s\rangle$,

$$|1, 1\rangle = |\uparrow\rangle|\uparrow\rangle \quad , \quad |1, 0\rangle = \frac{|\uparrow\rangle|\downarrow\rangle + |\downarrow\rangle|\uparrow\rangle}{\sqrt{2}} \quad , \quad |1, -1\rangle = |\downarrow\rangle|\downarrow\rangle \tag{7.25}$$

The total wavefunctions are again anti-symmetric,

$$|ab; s = 1\rangle = \Psi_{ab\,-}(\mathbf{r}_1, \mathbf{r}_2) \otimes |1, m_s\rangle \qquad m_s = -1, 0, 1 \tag{7.26}$$

For both $\Psi_{ab,+}$ and $\Psi_{ab,-}$ we take a to be the $1s$ state and b to be either the $2s$ or $2p$ state. The upshot of this analysis is that there are 4 possible $1s^1 2s^1$ states: a spin-singlet and a spin-triplet. There 12 possible $1s^1 2p^1$ states: 3 spin-singlets and 9 spin-triplets, the extra factor of 3 coming from the orbital angular momentum $m = -1, 0, 1$. Notice how fast the number of states grows, even for the simplest multi-electron atom! For the first excited state, we already have 16 options. This fast growth in the dimension of the Hilbert space is one of the characteristics of quantum mechanics.

Fortunately, we don't have to do degenerate perturbation theory with 16×16 dimensional matrices! The matrix elements of the interaction Hamiltonian (7.23) are already diagonal in the basis $|ab; s\rangle$ that we've described above already. This follows on symmetry grounds. The interaction Hamiltonian preserves rotational invariance, so the total orbital angular momentum must remain a good quantum number. Further, it doesn't mix spin states and $\langle 0, 0|1, m\rangle = 0$. This means that the states (7.24) and (7.26) are guaranteed to be energy eigenstates, at least to first order in perturbation theory.

In summary, we are looking for four energy levels, corresponding to the states $|1s^1 2s^1; s\rangle$ and $|1s^1 2p^1; s\rangle$ where $s = 0$ or 1. The question we would like to ask is: what is the ordering of these states?

We can make some progress with this question without doing any calculations. The interaction Hamiltonian (7.23) is a repulsive potential between the electrons. Clearly the states with lowest energy will be those where the electrons try to stay apart from each other. But the anti-symmetric wavefunction Ψ_{ab-} has the property that it vanishes when $\mathbf{r}_1 = \mathbf{r}_2$ and the electrons sit on top of each other. This strongly suggests that Ψ_{ab-} will have lower energy than Ψ_{ab+} and, correspondingly, the spin-triplet versions of a state will have lower energy than the spin-singlets.

We can see this mathematically. The energy splitting is

$$\Delta E_{ab\pm} = \frac{1}{4\pi\epsilon_0}\int d^3r_1 d^3r_2\, \frac{|\Psi_{ab\pm}(\mathbf{r}_1, \mathbf{r}_2)|^2}{|\mathbf{r}_1 - \mathbf{r}_2|} = J_{ab} \pm K_{ab}$$

where J_{ab} and K_{ab} are given by

$$\begin{aligned} J_{ab} &= \frac{1}{4\pi\epsilon_0}\int d^3r_1 d^3r_2\, \frac{1}{2}\frac{|\psi_a(\mathbf{r}_1)\psi_b(\mathbf{r}_2)|^2 + |\psi_a(\mathbf{r}_2)\psi_b(\mathbf{r}_1)|^2}{|\mathbf{r}_1 - \mathbf{r}_2|} \\ &= \frac{1}{4\pi\epsilon_0}\int d^3r_1 d^3r_2\, \frac{|\psi_a(\mathbf{r}_1)\psi_b(\mathbf{r}_2)|^2}{|\mathbf{r}_1 - \mathbf{r}_2|} \end{aligned} \tag{7.27}$$

where the second line follows because the integrand is symmetric under exchange $\mathbf{r}_1 \leftrightarrow \mathbf{r}_2$. Meanwhile, we have

$$\begin{aligned} K_{ab} &= \frac{1}{4\pi\epsilon_0}\int d^3r_1 d^3r_2\, \frac{1}{2}\frac{\psi_a^\star(\mathbf{r}_1)\psi_b^\star(\mathbf{r}_2)\psi_a(\mathbf{r}_2)\psi_b(\mathbf{r}_1) + \psi_a^\star(\mathbf{r}_2)\psi_b^\star(\mathbf{r}_1)\psi_a(\mathbf{r}_1)\psi_b(\mathbf{r}_2)}{|\mathbf{r}_1 - \mathbf{r}_2|} \\ &= \frac{1}{4\pi\epsilon_0}\int d^3r_1 d^3r_2\, \frac{\psi_a^\star(\mathbf{r}_1)\psi_b^\star(\mathbf{r}_2)\psi_a(\mathbf{r}_2)\psi_b(\mathbf{r}_1)}{|\mathbf{r}_1 - \mathbf{r}_2|} \end{aligned} \tag{7.28}$$

The contribution J_{ab} is called the *direct integral*; K_{ab} is called the *exchange integral* or, sometimes, the *exchange energy*. Note that it involves an integral over the position of the particle $\mathbf{r}_1$, weighted with both possible states $\psi_a(\mathbf{r}_1)$ and $\psi_b(\mathbf{r}_1)$ that the electron can sit in.

Both J_{ab} and K_{ab} are positive definite. This is not obvious for K_{ab}, but is intuitively true because the integral is dominated by the region $\mathbf{r}_1 \approx \mathbf{r}_2$ where the numerator is approximately $|\psi_a(\mathbf{r})|^2|\psi_b(\mathbf{r})|^2$. Since the shift in energy is $\Delta E_{ab\pm} = J_{ab} \pm K_{ab}$ we see that, as expected, the spin-triplet states with spatial anti-symmetry have lower energy.

We've learned that each of the spin-triplet states is lower than its spin-singlet counterpart. But what of the ordering of $1s^1 2s^1$ vs $1s^1 2p^1$? For this, we have to do the integrals J and K. One finds that the pair of $2s$ energy levels have lower energy than the pair of $2p$ energy levels. This, of course, is the beginning of the aufbau principle: the $2s$ levels fill up before the $2p$ levels. The resulting energy levels are shown in the diagram.

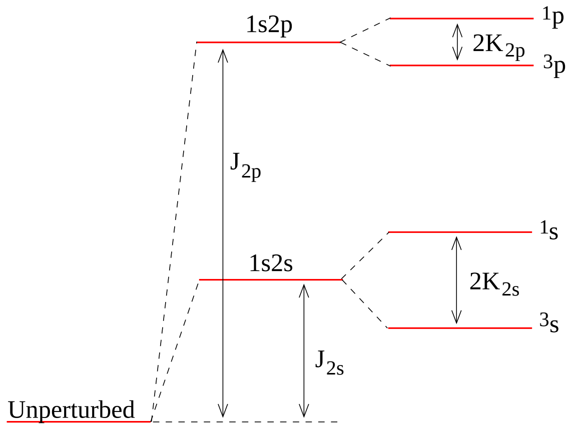

Figure 79:

Taken literally, our calculation suggests that the $2s$ state has lower energy because it does a better job at avoiding the original $1s$ electron. This is misleading: it's more an artefact of our (not particularly good) perturbative approach to the problem, rather than a way to good description of the underlying physics. One could do a better job by introducing variational wavefunctions, similar to those we looked at for the ground state. This approach would highlight the reason why states of higher l have higher energy. This reason is screening.

To understand this, you need to spend some time staring at the hydrogen wavefunctions. For fixed n, you can compare states with different l. Naively, it looks as if the low l states sit further from the origin. For example, the average radial position for an electron is $\langle r \rangle = a_0(3n^2 - l(l+1))/2$. But this average hides an important fact about the spread of the wavefunctions. The lower l wavefunctions not only extend further from the origin, but also extend closer to the origin. In contrast, the higher l wavefunctions are more localised around the average value $\langle r \rangle$. This means that the lower l wavefunctions spend more time near the nucleus. In contrast, the higher l wavefunctions remain far from the nucleus which is effectively screened by the inner electrons. The result is that, in a given shell, the higher l electrons experience a smaller nuclear charge and are correspondingly more weakly bound. This is why the states with smaller angular momentum full up first.

As we've seen, excited states of helium sit in both spin-singlets and spin-triplets. Parity means that transitions between these two states can only occur through the emission of two photons which makes these transitions much rarer. The lifetime of the $1s2s$ state turns out to be around 2.2 hours. This is very long on atomic timescales; indeed, it is the longest lived of all excited states of neutral atoms. It is said to be *meta-stable*. Before these transitions were observed, it was thought that there were two different kinds of helium atoms: those corresponding to spin-singlet states and those corresponding to spin-triplets. Historically the spin-singlet states were referred to as *parahelium*, the spin-triplet states as *orthohelium*.

The punchline from the story above is that spatially anti-symmetric wavefunctions are preferred since these come with a negative exchange energy. The fermionic nature of electrons means that these wavefunctions sit in a spin-triplet states. This fact plays an important role in many contexts beyond atomic physics. For example, the spins in solids often have a tendency to align, a phenomenon known as *ferromagnetism*. This too can be be traced the exchange integral for the Coulomb repulsion between atoms preferring the spins to sit in a triplet state. This results in the kind of $\mathbf{S}_1 \cdot \mathbf{S}_2$ spin-spin interaction that we met in the Statistical Physics course when discussing the Ising model.

7.3 Self-Consistent Field Method

As we've seen from our attempts to understand helium, a naive application of perturbation theory is not particularly effective. Not only does it become complicated as the number of possible states grows, but it also fails to capture the key physics of screening.

In this section, we will develop a variational approach to multi-electron atoms where, as we will see, the concept of screening sits centre stage. The idea is to attempt to reduce our multi-particle problem to a single-particle problem. But we don't do this merely by ignoring the effects of the other particles; instead we will alter our Hamiltonian in a way that takes these other particles into account. This method is rather similar to the *mean field theory* approach that we met in Statistical Physics; in both cases, one averages over many particles to find an effective theory for a single particle.

7.3.1 The Hartree Method

We start by considering a variational ansatz for the multi-particle wavefunction. For now, we will forget that the electrons are fermions. This means that we won't impose the requirement that the wavefunction is anti-symmetric under the exchange of particles, nor will we include the spin degrees of freedom. Obviously, this is missing something important but it will allow us to highlight the underlying physics. We will fix this oversight in Section 7.3.3 when we discuss the Hartree-Fock method.

We pretend that the electrons are independent and take as our ansatz the product wavefunction

$$\Psi(\mathbf{r}_1, \ldots, \mathbf{r}_N) = \psi_{a_1}(\mathbf{r}_1)\, \psi_{a_2}(\mathbf{r}_2) \ldots \psi_{a_N}(\mathbf{r}_N) \tag{7.29}$$

Here the labels a_i denote various quantum numbers of the one-particle states. We will ultimately see that the states $\psi_a(\mathbf{r})$ are eigenstates of a rotationally invariant Hamiltonian, albeit one which is different from the hydrogen Hamiltonian. This means that we can label each state by the usual quantum numbers

$$a = (n, l, m)$$

Although we haven't imposed anti-symmetry of the wavefunction, we do get to choose these quantum numbers for the states. This means that we can, for example, use this ansatz to look at the 3-particle states that lie in the shell $1s^2 2s^1$ as an approximation for the ground state of lithium.

We will view (7.29) as a very general variational ansatz, where we get to pick anything we like for each $\psi_a(\mathbf{r})$. We should compare this to the kind of variational ansatz (7.21) where we allowed only a single parameter Z to vary. For the Hartree ansatz, we have an infinite number of variational parameters.

The multi-electron Hamiltonian is

$$H = \sum_{i=1}^{N} \left(-\frac{\hbar^2}{2m} \nabla_i^2 - \frac{Ze^2}{4\pi\epsilon_0} \frac{1}{r_i} \right) + \sum_{i<j} \frac{e^2}{4\pi\epsilon_0} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|}$$

Evaluated on our ansatz (7.29), the average energy is

$$\begin{aligned} \langle E \rangle = & \sum_{i=1}^{N} \int d^3r\ \psi_{a_i}^\star(\mathbf{r}) \left(-\frac{\hbar^2}{2m} \nabla^2 - \frac{Ze^2}{4\pi\epsilon_0} \frac{1}{r} \right) \psi_{a_i}(\mathbf{r}) \\ & + \frac{e^2}{4\pi\epsilon_0} \sum_{i<j} \int d^3r\, d^3r'\ \frac{\psi_{a_i}^\star(\mathbf{r})\, \psi_{a_j}^\star(\mathbf{r}')\, \psi_{a_i}(\mathbf{r})\, \psi_{a_j}(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} \end{aligned} \tag{7.30}$$

The last term is an example of the kind of "direct integral" (7.27) that we met when discussing helium.

To find the best approximation to the ground state within the product ansatz (7.29), we minimize $\langle E \rangle$ over all possible states. However, there's a catch: the states $\psi_a(\mathbf{r})$ must remain normalised. This is easily achieved by introducing Lagrange multipliers. To this end, consider the functional

$$F[\Psi] = \langle E \rangle - \sum_i \epsilon_i \left(\int d^3r \; |\psi_{a_i}(\mathbf{r})|^2 - 1 \right)$$

with ϵ_i the N Lagrange multipliers imposing the normalisation condition.

Because the wavefunction is complex, we can vary its real and imaginary parts independently. Since we have N independent wavefunctions, this gives rise to $2N$ real conditions. It's not too hard to convince yourself that this is formally equivalent to the treating $\psi(\mathbf{r})$ and $\psi^\star(\mathbf{r})$ as independent and varying each of them, leaving the other fixed. Minimizing $F[\Psi]$ then requires us to solve

$$\frac{\delta F[\Psi]}{\delta \psi^\star_{a_i}(\mathbf{r})} = 0 \quad \text{and} \quad \frac{\delta F[\Psi]}{\delta \psi_{a_i}(\mathbf{r})} = 0$$

The first of these is N complex conditions, while the second is simply the conjugate of the first. These N complex conditions are called the *Hartree equations*,

$$\left[-\frac{\hbar^2}{2m}\nabla^2 - \frac{Ze^2}{4\pi\epsilon_0}\frac{1}{r} + \frac{e^2}{4\pi\epsilon_0}\sum_{j\neq i}\int d^3r' \; \frac{|\psi_{a_j}(\mathbf{r}')|^2}{|\mathbf{r}-\mathbf{r}'|} \right] \psi_{a_i}(\mathbf{r}) = \epsilon_i \psi_{a_i}(\mathbf{r}) \qquad (7.31)$$

These equations look tantalisingly similar to the Schrödinger equation. The only difference — and it is a big difference — is that the effective potential for $\psi_{a_i}(\mathbf{r})$ depends on the wavefunctions for all the other electrons, through the contribution

$$U_{a_i}(\mathbf{r}) = \frac{e^2}{4\pi\epsilon_0}\sum_{j\neq i}\int d^3r' \; \frac{|\psi_{a_j}(\mathbf{r}')|^2}{|\mathbf{r}-\mathbf{r}'|} \qquad (7.32)$$

Physically this is clear: the potential $U_{a_i}(\mathbf{r})$ is the electrostatic repulsion due to all the other electrons. Note that each electron experiences a different effective Hamiltonian, with a different $U_{a_i}(\mathbf{r})$. The catch is that each of the $\psi_{a_j}(\mathbf{r})$ that appears in the potential $U(\mathbf{r})$ is also determined by one of the Hartree equations.

The Hartree equations (7.31) are not easy to solve. They are N, coupled non-linear integro-differential equations. We see that there's a certain circularity needed to get to the solution: the potentials $U_{a_i}(\mathbf{r})$ determine the wavefunctions but are also determined by the wavefunctions. In this sense, the ultimate solution for $U_{a_i}(\mathbf{r})$ is said to be "self-consistent".

The usual techniques that we use for the Schrödinger equation do not work for the Hartree equations. Instead, we usually proceed iteratively. We start by guessing a form for the potentials $U_{a_i}(\mathbf{r})$ which we think is physically realistic. Often this involves making the further approximation that $U(\mathbf{r})$ is spherically symmetric, so we replace

$$U_{a_i}(\mathbf{r}) \rightarrow U_{a_i}(r) = \int \frac{d\Omega}{4\pi}\, U_{a_i}(\mathbf{r})$$

Then, with this potential in hand, we solve the Schrödinger equations

$$\left[-\frac{\hbar^2}{2m}\nabla^2 - \frac{Ze^2}{4\pi\epsilon_0}\frac{1}{r} + U_{a_i}(r)\right]\psi_{a_i}(\mathbf{r}) = \epsilon_i \psi_{a_i}(\mathbf{r}) \tag{7.33}$$

This can be done numerically. We then substitute the resulting wavefunctions back into the definition of the potential (7.32) and then play the whole game again. If we chose a good starting point, this whole process will being to converge.

Suppose that we've done all of this. What is the answer for the ground state energy of the atom? From (7.31), the Lagrange multipliers ϵ_i look like the energies of individual particles. We can write

$$\epsilon_i = \int d^3r\ \psi^\star_{a_i}(\mathbf{r})\left[-\frac{\hbar^2}{2m}\nabla^2 - \frac{Ze^2}{4\pi\epsilon_0}\frac{1}{r} + \frac{e^2}{4\pi\epsilon_0}\sum_{j\neq i}\int d^3r'\ \frac{|\psi_{a_j}(\mathbf{r}')|^2}{|\mathbf{r}-\mathbf{r}'|}\right]\psi_{a_i}(\mathbf{r})$$

Summing these gives an expression that is almost the same as the expected energy (7.30), except that the sum $\sum_i\sum_{j\neq i}$ is twice the sum $\sum_{i<j}$. The evaluated on the solutions to the Hartree equations is then given by

$$\langle E\rangle = \sum_i \epsilon_i - \frac{e^2}{4\pi\epsilon_0}\sum_{j\neq i}\int d^3r\, d^3r'\ \frac{|\psi_{a_j}(\mathbf{r}')|^2|\psi_{a_i}(\mathbf{r})|^2}{|\mathbf{r}-\mathbf{r}'|}$$

By the usual variational arguments, this gives an upper bound for the ground state energy.

An Example: Potassium

We won't describe in detail the numerical solutions to the Hartree equations (nor to the more sophisticated Hartree-Fock equations that we will meet shortly). We can, however, use this approach to offer some hand-waving intuition for one of the more surprising features of the aufbau principle: why does the $4s$ state fill up before the $3d$ state?

This question first arises in potassium, an alkali metal with electron configuration $1s^2 2s^2 2p^6 3s^2 3p^6 4s^1$. Why is the last electron in $4s$ rather than $3d$ as the non-interacting picture of electrons would suggest?

In the Hartree approach, we see that the electron experiences an effective potential with Schrödinger equation (7.33). The key piece of physics which determines $U(r)$ is, once again, screening. When the electron is far away, the nuclear charge Ze is expected to be almost entirely screened by the other $Z - 1$ electrons. In contrast, when the electron is close to the nucleus, we expect that it feels the full force of the Ze charge. On these grounds, the total effective potential should be

$$-\frac{Ze^2}{4\pi\epsilon_0 r} + U(r) = -\frac{Z(r)e^2}{4\pi\epsilon_0 r}$$

where $Z(r)$ is some function which interpolates between $Z(r) \to Z$ as $r \to 0$ and $Z(r) \to 1$ as $r \to \infty$.

We should now solve the Schrödinger equation with this potential. All quantum states are labelled by the usual triplet (n, l, m), but as the potential is no longer simply $1/r$ the energy levels will depend on both n and l. The basic physics is the same as we described for the excited states of helium. The $l = 0$ s-wave states extend to the origin which causes their energy to be lower. In contrast, the higher l states experience an angular momentum barrier which keeps them away from the origin and raises their energy. This explains why $3s$ fills up before $3p$. But this same screening effect also lowers the $4s$ states below that of $3d$.

7.3.2 The Slater Determinant

The Hartree ansatz (7.29) is not anti-symmetric under the exchange of particles. As such, it is not a physical wavefunction in the Hilbert space of fermions. We would like to remedy this.

Our task is a simple one: given a collection of 1-particle states, how do we construct a multi-particle wavefunction for fermions that are anti-symmetric under the exchange of any pair of particles? This general question arises in many contexts beyond the spectrum of atoms.

We will use the notation $|\psi_i(j)\rangle$ to mean the particle j occupies the one-particle state $|\psi_i\rangle$. Then we can build a suitably anti-symmetrised N-particle wavefunction by using the *Slater determinant*,

$$|\Psi\rangle = \frac{1}{\sqrt{N!}} \begin{vmatrix} |\psi_1(1)\rangle & |\psi_1(2)\rangle & \dots & |\psi_1(N)\rangle \\ |\psi_2(1)\rangle & |\psi_2(2)\rangle & \dots & |\psi_2(N)\rangle \\ \vdots & & \ddots & \\ |\psi_N(1)\rangle & |\psi_N(2)\rangle & \dots & |\psi_N(N)\rangle \end{vmatrix}$$

Expanding out the determinant gives $N!$ terms that come with plus and minus signs. The overall factor of $1/\sqrt{N!}$ ensures that the resulting state is normalised. The plus and minus signs provide the anti-symmetry that we need for fermions. In fact, we can see this quickly without expanding out: swapping the first and second particle is tantamount to swapping the first and second rows of the matrix. But we know that this changes the determinant by a minus sign. In particular, if two particles sit in the same state then the rows of the matrix become linearly dependent and the determinant vanishes. In this way, the Slater determinant enforces the Pauli exclusion principle.

One can build the Slater determinant for any states $|\psi_i\rangle$ which span an N-dimensional Hilbert space. It will be convenient to choose the states $|\psi_i\rangle$ to form an orthogonal basis.

An Example: Helium

For helium, we take the set of one-particle states to be the hydrogen wavefunctions for $Z = 2$, so $|\psi_a\rangle = \psi_{(n,l,m)}(\mathbf{r}) \otimes |m_s\rangle$ where the spin quantum number $m_s = \pm\frac{1}{2}$ is usually replaced by the notation $|\frac{1}{2}\rangle = |\uparrow\rangle$ and $|-\frac{1}{2}\rangle = |\downarrow\rangle$.

For the ground state we place both particles in the $1s$ state with different spins. The corresponding Slater determinant is

$$\frac{1}{\sqrt{2}} \begin{vmatrix} \psi_{1s}(\mathbf{r}_1) \otimes |\uparrow\rangle & \psi_{1s}(\mathbf{r}_2) \otimes |\uparrow\rangle \\ \psi_{1s}(\mathbf{r}_1) \otimes |\downarrow\rangle & \psi_{1s}(\mathbf{r}_2) \otimes |\downarrow\rangle \end{vmatrix} = \psi_{1s}(\mathbf{r}_1)\psi_{1s}(\mathbf{r}_2) \otimes |0,0\rangle$$

where $|0,0\rangle$ is the spin-singlet state (7.22). This is the ground state of helium that we used previously.

When particles sit in different hydrogenic states, there are more possibilities for the Slater determinant. For example, for the $1s^1 2s^1$ excited state, there are four Slater determinants. Two of these sit in spin eigenstates

$$\frac{1}{\sqrt{2}}\begin{vmatrix} \psi_{1s}(\mathbf{r}_1)\otimes|\uparrow\rangle & \psi_{1s}(\mathbf{r}_2)\otimes|\uparrow\rangle \\ \psi_{2s}(\mathbf{r}_1)\otimes|\uparrow\rangle & \psi_{2s}(\mathbf{r}_2)\otimes|\uparrow\rangle \end{vmatrix} = \Psi_-(\mathbf{r}_1,\mathbf{r}_2)\otimes|1,1\rangle$$

$$\frac{1}{\sqrt{2}}\begin{vmatrix} \psi_{1s}(\mathbf{r}_1)\otimes|\downarrow\rangle & \psi_{1s}(\mathbf{r}_2)\otimes|\downarrow\rangle \\ \psi_{2s}(\mathbf{r}_1)\otimes|\downarrow\rangle & \psi_{2s}(\mathbf{r}_2)\otimes|\downarrow\rangle \end{vmatrix} = \Psi_-(\mathbf{r}_1,\mathbf{r}_2)\otimes|1,-1\rangle$$

where $\Psi_\pm(\mathbf{r}_1,\mathbf{r}_2) = \frac{1}{\sqrt{2}}(\psi_{1s}(\mathbf{r}_1)\psi_{2s}(\mathbf{r}_2) \pm \psi_{1s}(\mathbf{r}_2)\psi_{2s}(\mathbf{r}_1))$ and $|1,m\rangle$ are the spin-triplet states (7.25). Meanwhile, the other Slater determinants are

$$\frac{1}{\sqrt{2}}\begin{vmatrix} \psi_{1s}(\mathbf{r}_1)\otimes|\uparrow\rangle & \psi_{1s}(\mathbf{r}_2)\otimes|\uparrow\rangle \\ \psi_{2s}(\mathbf{r}_1)\otimes|\downarrow\rangle & \psi_{2s}(\mathbf{r}_2)\otimes|\downarrow\rangle \end{vmatrix} = \frac{1}{\sqrt{2}}(\Psi_+(\mathbf{r}_1,\mathbf{r}_2)\otimes|0,0\rangle + \Psi_-(\mathbf{r}_1,\mathbf{r}_2)\otimes|1,0\rangle)$$

$$\frac{1}{\sqrt{2}}\begin{vmatrix} \psi_{1s}(\mathbf{r}_1)\otimes|\downarrow\rangle & \psi_{1s}(\mathbf{r}_2)\otimes|\downarrow\rangle \\ \psi_{2s}(\mathbf{r}_1)\otimes|\uparrow\rangle & \psi_{2s}(\mathbf{r}_2)\otimes|\uparrow\rangle \end{vmatrix} = \frac{1}{\sqrt{2}}(\Psi_+(\mathbf{r}_1,\mathbf{r}_2)\otimes|0,0\rangle - \Psi_-(\mathbf{r}_1,\mathbf{r}_2)\otimes|1,0\rangle)$$

We see that the Slater determinants do not necessarily give spin eigenstates.

This is one of the short-comings of the Slater determinant. In general, one can show that the state $|\Psi\rangle$ can always be guaranteed to be an eigenstate of angular momentum L_z and spin S_z. But it is not always an eigenstate of $\mathbf{L}^2$ and $\mathbf{S}^2$.

7.3.3 The Hartree-Fock Method

The *Hartree-Fock* method is a repeat of the Hartree method, but now with the fully anti-symmetrised wavefunction

$$|\Psi\rangle = \frac{1}{\sqrt{N!}}\begin{vmatrix} |\psi_{a_1}(1)\rangle & |\psi_{a_1}(2)\rangle & \ldots & |\psi_{a_1}(N)\rangle \\ |\psi_{a_2}(1)\rangle & |\psi_{a_2}(2)\rangle & \ldots & |\psi_{a_2}(N)\rangle \\ \vdots & & \ddots & \\ |\psi_{a_N}(1)\rangle & |\psi_{a_N}(2)\rangle & \ldots & |\psi_{a_N}(N)\rangle \end{vmatrix} \tag{7.34}$$

Further, we will take the quantum numbers a_i to include both the (n,l,m) information about the orbital angular momentum state, as well as the spin degrees of freedom of the electron. (Had we included spin in the original Hartree ansatz, it simply would have dropped out of the final answer; but now that we have anti-symmetry the spin wavefunctions are correlated with the spatial wavefunctions.)

Repeating the Hartree story, we find that the average energy in the state $|\Psi\rangle$ contains one extra term

$$
\begin{aligned}
\langle E\rangle = \sum_{i=1}^{N} \int d^3r\ \psi^\star_{a_i}(\mathbf{r}) \left(-\frac{\hbar^2}{2m}\nabla^2 - \frac{Ze^2}{4\pi\epsilon_0}\frac{1}{r}\right)\psi_{a_i}(\mathbf{r}) \\
+\frac{e^2}{4\pi\epsilon_0}\sum_{i<j}\int d^3r\, d^3r'\ \frac{\psi^\star_{a_i}(\mathbf{r})\,\psi^\star_{a_j}(\mathbf{r}')\,\psi_{a_i}(\mathbf{r})\,\psi_{a_j}(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} \\
-\frac{e^2}{4\pi\epsilon_0}\sum_{i<j}\int d^3r\, d^3r'\ \frac{\psi^\star_{a_i}(\mathbf{r})\,\psi^\star_{a_j}(\mathbf{r}')\,\psi_{a_i}(\mathbf{r}')\,\psi_{a_j}(\mathbf{r})\,\delta_{m_{s\,i},m_{s\,j}}}{|\mathbf{r}-\mathbf{r}'|}
\end{aligned}
$$

The last term is an exchange integral of the kind we met when discussing the helium atom (7.28). The delta function ensures that it only contributes if the a_i and a_j spin states coincide.

While the direct integral clearly captures the electrostatic repulsion between electrons, it is somewhat harder to drape comforting classical words around the exchange term. It is a purely quantum effect arising from the Pauli exclusion principle. Nonetheless, we can extract some physics from it, in particular from the fact that the delta function means that the exchange term lowers the energy only when spins are aligned. This means that, all else being equal, the spins will wish to align. This is the first of three *Hund's rules.* (The other two describe the preferential order to fill degenerate states with quantum numbers $\mathbf{L}$ and $\mathbf{J} = \mathbf{L} + \mathbf{S}$; we won't discuss these second two rules in these lectures.)

In practice, this does nothing for a filled shell. In this case, half the electrons have spin up and the other half spin down. However, when we start to fill a shell, the exchange term means that it's preferable for all the spins to point in the same direction. This suggests that half-filled shells should be particularly stable. Moreover, the next electron to go in after half-filling should have a noticeably larger energy and the atom will, correspondingly, have a smaller ionization energy.

We can see evidence for this by looking again at the ionization data. The ionization energy does not increase monotonically between Li and Ne: there are two glitches. The first of these is the jump from beryllium ($2s^2$) to boron ($2s^2 2p^1$) where we jump to another shell. The other is the jump from nitrogen ($1s^2 2s^2 2p^3$) to oxygen ($1s^2 2s^2 2p^4$). Nitrogen has a half-filled $2p$ sub-shell, where all three electrons have spin up to benefit from the exchange energy. But for oxygen one electron is spin down, and the benefit from from the exchange energy is less. This means that the next electron costs higher energy and, correspondingly, the ionization energy is smaller. The same behaviour is

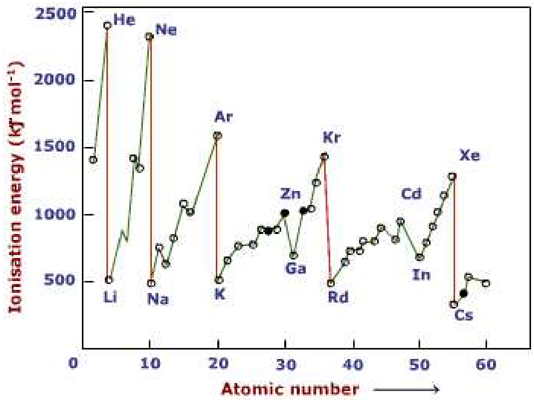

Figure 80: Ionization data again.

seen to disrupt the linear growth between Na and Ar. The two glitches occur between manganese ([Ne]$3s^2$) and aluminium ([Ne]$3s^23p^1$) where we jump to the next shell, and between phosphorus ([Ne]$3s^23p^3$) to sulphur ([Ne]$3s^23p^3$) where we cross the half-filled sub-shell.

The exchange energy also lies behind one of the exceptions to the aufbau principle. Recall that chromium has electron configuration [Ar]$3d^54s^1$ as opposed to the aufbau-predicted [Ar]$3d^44s^2$. The former configuration has lower energy because it allows all spins to point up and so benefits more from the exchange term.

Minimising the energy $\langle E\rangle$ gives us N coupled equations

$$\left[-\frac{\hbar^2}{2m}\nabla^2 - \frac{Ze^2}{4\pi\epsilon_0}\frac{1}{r} + U(\mathbf{r})\right]\psi_{a_i}(\mathbf{r}) - \int d^3r'\; U^{\text{ex}}_{a_i}(\mathbf{r},\mathbf{r}')\psi_{a_i}(\mathbf{r}') = \epsilon_i\psi_{a_i}(\mathbf{r}) \tag{7.35}$$

where $U(\mathbf{r})$ is given by

$$U(\mathbf{r}) = \frac{e^2}{4\pi\epsilon_0}\sum_{j=1}^{N}\int d^3r'\;\frac{|\psi_{a_j}(\mathbf{r}')|^2}{|\mathbf{r}-\mathbf{r}'|}$$

This differs from the Hartree expression (7.32) because we sum over all states $\sum_j$ rather than $\sum_{j\neq i}$. This is a simplification because it means that all electrons feel the same potential. However, it is also puzzling because it would appear to suggest that we need to include a "self-interaction" between the electrons. But this $i=j$ term is an artefact of the way we've written things: it cancels the corresponding term in the exchange integral, which is given by

$$U_{a_i}^{\rm ex}(\mathbf{r},\mathbf{r}') = \frac{e^2}{4\pi\epsilon_0}\sum_{j=1}^N \int d^3r' \ \frac{\psi_{a_j}^\star(\mathbf{r}')\,\psi_{a_j}(\mathbf{r})}{|\mathbf{r}-\mathbf{r}'|}\delta_{m_{si},m_{sj}}$$

This is sometimes referred to as a *non-local potential.* This term does depend on the state a_i, but only through the spin dependence. This means that each electron experiences one of two different exchange potentials, $U_\uparrow^{\rm ex}$ or $U_\downarrow^{\rm ex}$.

The set of equations (7.35) are known as the *Hartree-Fock equations.* It should come as no surprise to learn that they are no easier to solve than the Hartree equations. Indeed, the presence of the exchange term makes even numerical solutions considerably harder to come by. Nonetheless, this scheme has some success in reproducing the properties of atoms observed in the periodic table, in particular the aufbau principle.

Limitations of Hartree-Fock

We finish with a warning. Throughout this section, we've used the language of one-particle states to describe atoms. Indeed, the basic idea that we've focussed on is that atoms are made by filling successive shells of states. This is something that is often taught in high school and, over time, becomes so familiar that we don't question it. The Hartree-Fock method panders to this idea because it looks for states within the anti-symmetrised product ansatz (7.34). However, the vast majority of states in the Hilbert space are not of the product form and, for complicated atoms, it's quite possible, indeed likely, that the true ground state is a superposition of such states. In this case the very language of filing shells become inappropriate since there's no way to say that any electron sits in a given state.