

# Time-Independent Perturbation Theory

Prof. Michael G. Moore, Michigan State University

## Atomic Physics Applications

### 1 Introduction

For many reasons it is important to understand the basic level-structure of atomic hydrogen. As the simplest atom, it is a good starting point to understand the various mechanisms at work inside atoms. Early atomic physics was focussed on measuring and explaining the various atomic spectra. In recent years, atomic physics has progressed into new areas such as precision measurement, quantum optics, and even quantum computation. Atomic level structure still plays an important role in modern atomic physics, however, particularly in the rapidly evolving field of laser-cooled and trapped atoms. In solid-state physics, the atomic properties of impurity and dopant atoms play a major role, and the properties of quantum dots (artificial atoms) closely mirror those of real atoms. In nuclear physics, precision atomic spectroscopy allows precise measurements of isotopic masses and other isotopic properties. In addition, as the most abundant element, much astronomical data is based on measuring the spectral lines of hydrogen. Lastly, hydrogen is a system with many degeneracies, in which physically important results can be obtained from low-order degenerate perturbation theory, so it is an excellent area to practice applying what we have learned.

We only have time to touch on atomic physics in this course, and we will focus only on the hydrogen atom for the time-being. This ignores the important and rich problem of electron-electron interaction, which dominates most of the periodic table of elements. The results from studying hydrogen, however, are readily generalized to other alkali metal (group I) atoms such as lithium (Li), sodium (Na), potassium (K), rubidium (Rb), cesium (Cs), and Francium (Fr). Because their optical properties are governed by the behavior of a single valence electron, they are currently the predominant elements used in laser-cooling and trapping experiments, and all but Francium have been evaporatively cooled into Bose-Einstein condensates, which are many-body states with a single macroscopic wavefunction analogous to laser-light, but with bosonic atomic isotopes taking the place of photons. Singly ionized group II elements (Be, Mg, Ca, Sr, Ba, Ra) are also hydrogen-like, and due to their similar optical properties, are commonly used in trapped-ion experiments, such as trapped-ion quantum computers. Lastly, exotic states such as positronium (electron-positron bound state) and muonic hydrogen (proton-muon bound state) also share the hydrogen level structure.

In this section we will consider four basic effects, two based on the response of hydrogen to external fields, and two based on internal effects related to the intrinsic spin of the electron and proton. These four effects are

1. **DC Stark Effect:** response of an atom to an applied static electric field.

2. **Atomic Fine-structure:** Interaction of the electron spin with the atom's internal magnetic field.
3. **Atomic Hyperfine-structure:** Interaction of the nuclear spin with the atom's internal magnetic field.
4. **Zeeman Effect:** response of an atom to an applied static magnetic field.

These lectures are based primarily on the treatment in the well-known quantum text by Griffiths, and 'Quantum Mechanics volume two' by Nobel Laureate Claude Cohen-Tannoudji.

We first briefly review the bare hydrogen problem without spin, which was treated last semester. The simplest isotope of hydrogen consists of an electron orbiting a proton, which interact via the coulomb interaction. The bare Hamiltonian for the relative motion of the electron and the nucleus in a hydrogen atom is

$$H_0 = \frac{P^2}{2m_e} + V_c(R), \quad (1)$$

where the coulomb potential is given by

$$V_c(R) = -\frac{e^2}{4\pi\epsilon_0 R}, \quad (2)$$

with  $R$  being the operator for the distance between the electron and the proton. Strictly speaking, the mass is not the electron mass,  $m_e$ , but rather the reduced mass  $\mu = \frac{m_e m_n}{m_e + m_n}$ , where  $m_n$  is the nuclear mass. To good approximation, we have  $\mu \approx m_e$ , but precision atomic spectroscopy can resolve the correction to this  $\mu \approx m_e(1 - m_e/m_n)$ . This leads to a detectable nuclear-mass dependent shift in the atomic spectrum known as the isotope effect. Thus all of the following results can be applied to Deuterium and Tritium, simply by replacing  $m_e$  with the appropriate reduced mass.

It is customary to use as a basis simultaneous eigenstates of  $H_0$ ,  $L^2$  and  $L_z$ . The corresponding eigenvalue equations are then

$$H_0|n\ell m_\ell^{(0)}\rangle = E_n^{(0)}|n\ell m_\ell^{(0)}\rangle, \quad (3)$$

$$L^2|n\ell m_\ell^{(0)}\rangle = \hbar^2 \ell(\ell + 1)|n\ell m_\ell^{(0)}\rangle, \quad (4)$$

and

$$L_z|n\ell m_\ell^{(0)}\rangle = \hbar m_\ell |n\ell m_\ell^{(0)}\rangle, \quad (5)$$

where the unperturbed energy eigenvalue is

$$E_n^{(0)} = -\frac{\hbar^2}{2m_e a_0^2} \frac{1}{n^2} = -\frac{13.6 \text{ eV}}{n^2} = -\frac{2.18 \times 10^{-18} \text{ J}}{n^2}, \quad (6)$$

with

$$a_0 = \frac{4\pi\epsilon_0 \hbar^2}{m_e e^2} \quad (7)$$

being the 'Bohr radius', which corresponds to the size of the ground-state electron wave-packet. The wavefunctions of the unperturbed orbitals are given by

$$\langle r\theta\phi | n\ell m_\ell^{(0)} \rangle = R_{n\ell}(r) Y_\ell^m(\theta, \phi). \quad (8)$$

Here  $Y_\ell^m(\theta, \phi) = (-1)^m \sqrt{\frac{2\ell+1}{4\pi} \frac{(\ell-m)!}{(\ell+m)!}} P_\ell^m(\cos \theta) e^{im\phi}$  is the standard spherical harmonic function, with  $P_n^a(x)$  being the generalized Legendre function. The radial wavefunction  $R_{n\ell}(r)$ <sup>1</sup> is given by

$$R_{n\ell}(r) = \left[ \left( \frac{2}{na_0} \right)^3 \frac{(n-\ell-1)!}{2n(n+\ell)!} \right]^{1/2} \exp\left(-\frac{r}{na_0}\right) \left( \frac{2r}{na_0} \right)^\ell L_{n-\ell-1}^{2\ell+1}\left(\frac{2r}{na_0}\right), \quad (9)$$

with  $L_q^p(x)$  being the associated Laguerre polynomial. These levels can be generalized to include electronic spin via  $|n\ell m_\ell^{(0)}\rangle \rightarrow |n\ell m_\ell m_s^{(0)}\rangle$ , where  $S_z |n\ell m_\ell m_s^{(0)}\rangle = \hbar m_s |n\ell m_\ell m_s^{(0)}\rangle$ , since  $[H_0, S_z] = 0$ , the energy levels and orbitals given above are not changed by the inclusion of electronic spin.

---

<sup>1</sup>Please take note that there are multiple typos in the formula for  $R_{n\ell}(r)$  in the appendix of Sakuari. The formula here is correct. Furthermore, the formula in Griffiths should also not be used, as it assumes a non-standard normalization for the Laguerre polynomials. Only the formula given here is compatible with the Mathematica function `LaguerreL[n, a, x]`.

## 2 DC Stark Effect

Here we consider what happens to the bare atomic levels when a uniform static electric field is applied. The perturbation operator takes the form

$$V_E = -eE_0\vec{e}_z \cdot \vec{R}, \quad (10)$$

where  $E_0$  is the magnitude of the electric field, and the direction of the applied field has defined a z-axis. Since  $[V_E, S_z] = 0$ , we can safely ignore spin. The first-order Stark-shift will be linear in the field amplitude  $E_0$ , and will be given by

$$E_n^{(1)} = V_{nn} = -eE_0\langle n\ell m_\ell^{(0)} | Z | n\ell m_\ell^{(0)} \rangle. \quad (11)$$

### 2.1 Non-degenerate case

First we note that the selection rule governing matrix-elements of  $V_E$  is

$$\langle n\ell m_\ell^{(0)} | Z | n'\ell' m_\ell'^{(0)} \rangle \propto \delta_{m',m} \delta_{\ell',\ell \pm 1}. \quad (12)$$

The constraint  $m' = m$  is clearly a consequence of the azimuthal symmetry of the problem, while the condition  $\ell' = \ell \pm 1$  arises from the Associated Legendre function recurrence relations [see Arfken, 3rd Edition, Eqs. 12.85 and in particular 12.210].

The ground state of hydrogen is not degenerate (ignoring spin) so that the first-order correction

$$E_1^{(1)} = -eE_0\langle 100^{(0)} | Z | 100^{(0)} \rangle = 0 \quad (13)$$

vanishes due to parity considerations. This is because the ground state is spherically symmetric and thus has no electric dipole moment. Note: the electric dipole moment is defined as  $\vec{d} = |e|\langle \vec{R} \rangle$ . The standard dipole moment for charges  $q$  and  $-q$  is  $q\vec{d}$ , where  $\vec{d}$  goes from the negative to positive charge. Thus we have  $q = -|e|$ , but we also need  $\vec{d} = -\vec{R}$ , so the minus signs cancel.

The first non-vanishing shift to the ground state is found at second-order, where taking the selection rules (12) into account leads to

$$E_1^{(2)} = \sum_{m \neq 1} \frac{|V_{m1}|^2}{E_{m1}} = -e^2 E_0^2 \sum_{n=2}^{\infty} \frac{|\langle n10^{(0)} | Z | 100^{(0)} \rangle|^2}{E_n^{(0)}(n^2 - 1)}. \quad (14)$$

The summation can be evaluated numerically, which we will not do at this time. We note, however, that it is non-zero. To estimate the order-of-magnitude, we can take  $\langle n\ell m_\ell^{(0)} | Z | n\ell m_\ell^{(0)} \rangle \sim a_0$ , giving

$$E_1^{(2)} \sim \frac{e^2 a_0^2}{E_1^{(0)}} E_0^2. \quad (15)$$

This effect is known as the quadratic Stark effect, and is due to a process where first, the applied field induces a dipole moment in the ground state atom (by mixing in the higher states), then the interaction between the induced dipole and the field takes the usual  $V = -\vec{d} \cdot \vec{E}$ . Since the induced dipole is parallel to the field, we can identify the induced dipole moment order-of-magnitude as

$$d \sim \frac{e^2 a_0^2 E_0}{E_1^{(0)}}. \quad (16)$$

## 2.2 Degenerate case

The Stark effect physics is very different when an energy level is degenerate. To illustrate this we will now apply perturbation theory to the  $n = 2$  manifold. Here the first task is to find the ‘good’ bare eigenstates. The  $n = 2$  manifold includes the states (neglecting spin, and writing the states as  $|n\ell m_\ell^{(b)}\rangle$ )  $\{|200^{(b)}\rangle, |210^{(b)}\rangle, |21(\pm 1)^{(b)}\rangle$ , where we have switched from  $(0)$  to  $(b)$  in anticipation of these being the ‘bad’ eigenstates. Taking into account the selection rules  $m'_\ell = m_\ell$  and  $\ell' = \ell \pm 1$ , we find that the reduced perturbation operator  $V_D$  is given in the ‘bad’ basis,  $\{|200^{(b)}\rangle, |21(-1)^{(b)}\rangle, |210^{(b)}\rangle, |211^{(b)}\rangle\}$ , by

$$V_D = \begin{pmatrix} 0 & 0 & v & 0 \\ 0 & 0 & 0 & 0 \\ v & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad (17)$$

where

$$v = \langle 200^{(b)} | V | 210^{(b)} \rangle = 3ea_0E_0. \quad (18)$$

Since  $V_D$  has only two non-zero elements, it is straightforward to determine the eigenvalues and eigenvectors by inspection.

The ‘good’ basis for perturbation theory are thus states with well-defined  $m_\ell$  but not necessarily  $\ell$ . Using the notation  $|nm_\ell m^{(0)}\rangle$ , where  $m = 1, 2, \dots$  distinguishes states with the same  $m_\ell$ , the ‘good’ basis becomes  $\{|2(-1)1^{(0)}\rangle, |201^{(0)}\rangle, |202^{(0)}\rangle, |211^{(0)}\rangle\}$ , where  $|2(-1)1^{(0)}\rangle = |21(-1)^{(b)}\rangle$  and  $|211^{(0)}\rangle = |211^{(b)}\rangle$ . The non-zero eigenvalues are  $v_{201} = -3ea_0E_0$  and  $v_{202} = 3ea_0E_0$ , with corresponding eigenvectors

$$|201^{(0)}\rangle = \frac{1}{\sqrt{2}} \left( |200^{(b)}\rangle - |210^{(b)}\rangle \right) \quad (19)$$

$$|202^{(0)}\rangle = \frac{1}{\sqrt{2}} \left( |200^{(b)}\rangle + |210^{(b)}\rangle \right). \quad (20)$$

From first-order perturbation theory we find

$$E_{201}^{(1)} = \langle 201^{(0)} | V | 201^{(0)} \rangle = -3ea_0E_0 \quad (21)$$

$$E_{202}^{(1)} = \langle 202^{(0)} | V | 202^{(0)} \rangle = 3ea_0E_0 \quad (22)$$

and

$$E_{2-11}^{(1)} = E_{211}^{(1)} = 0. \quad (23)$$

Thus we see that the applied electric field has partially lifted the quadruple degeneracy of the  $n = 2$  manifold, splitting it into three levels. The upper and lower levels have well-defined  $n$  and  $m_\ell$  quantum numbers, but are mixtures of different  $\ell$  states. The center level is not shifted at first-order, and is now doubly degenerate.

### 3 Fine structure

Even in the absence of external fields, the atoms own internal magnetic field perturbs the level structure via its interaction with the electronic and nuclear spins. Because of the mass difference, the intrinsic magnetic dipole moment of the electron is three orders of magnitude larger than that of the nucleus. For this reason, effects related to nuclear spin are much smaller than effects involving only the electron spin and orbital angular momentum. All of these internal effects turn out to be governed by the ‘fine-structure constant’  $\alpha$ , given by

$$\alpha = \frac{e^2}{(4\pi\epsilon_0)\hbar c} \approx \frac{1}{137}. \quad (24)$$

The connection between the fine-structure constant and special relativity is clear when you consider that  $\alpha^2$  is twice the ratio between the hydrogen ground-state energy and the electrons rest-mass energy,

$$\alpha^2 = \frac{2|E_1(0)|}{M_e c^2} \sim 10^{-5}. \quad (25)$$

Effects not involving the nuclear spin turn out to be proportional to  $\alpha^2$ , which together determine the hydrogen ‘fine structure’, while nuclear spin effects are proportional to  $\alpha^2(m_e/m_p)$  and are termed the ‘hyperfine structure’. Observing these line-splittings and shifts requires the ability to measure the hydrogen spectral lines with a high degree of wavelength resolution. Thus as the resolution is increased, first the fine structure will appear, and then at a higher resolution one would finally resolve the hyperfine structure.

We will first examine the hydrogen fine structure, which contains three corrections all of the same order. All of these effects should be considered as first-order relativistic corrections, as they are the first-order terms in the expansion of the relativistic Lagrangian in powers of  $(v/c)^2$ . The corresponding relativistic Hamiltonian for hydrogen, including terms up to order  $(v/c)^2$  is (note: this does not yet include nuclear spin)

$$H = M_e c^2 + \frac{P^2}{2M_e} + V_c(R) - \frac{P^4}{8M_e^3 c^2} + \frac{1}{2M_e^2 c^2} \frac{1}{R} \frac{dV_c(R)}{dR} \vec{L} \cdot \vec{S} + \frac{\hbar^2}{8M_e^2 c^2} \nabla^2 V_c(R) + \dots \quad (26)$$

This is of the form  $H = H_0 + V_{FS}$ , where

$$V_{FS} = V_M + V_{SO} + V_D. \quad (27)$$

Here  $V_M$  is the relativistic correction to the kinetic energy due to the increase in mass with velocity. It can be found by expanding the relativistic kinetic energy  $T = \sqrt{M^2 c^4 + c^2 P^2} - M c^2$  to second-order in  $P^2$  (i.e. fourth order in  $P$ ), which results in

$$V_M = -\frac{P^4}{8M^3 c^2}. \quad (28)$$

The next term  $V_{SO}$  is the spin-orbit coupling. From electrons viewpoint, the orbiting proton constitutes an electrical current which therefore sets up a magnetic field at the location of the electron. The interaction between the electrons intrinsic magnetic dipole moment and this magnetic field leads to the spin-orbit term

$$V_{SO} = \frac{1}{2M_e^2 c^2} \frac{1}{R} \frac{dV_c(R)}{dR} \vec{L} \cdot \vec{S}. \quad (29)$$

Lastly, the so-called ‘Darwin’ term  $V_D$  accounts for the fact that the Coulomb Potential is not strictly diagonal in coordinate representation, but is off-diagonal on the scale of the Compton wavelength  $\hbar/M_e c$ . Hence the electron really sees a slightly smeared out version of the coulumb potential, described by

$$V_D = \frac{\hbar^2}{8M_e^2 c^2} \nabla^2 V_c(R). \quad (30)$$

### 3.1 Relativistic Mass term

First we will briefly treat the relativistic mass term. This effect is governed by the operator  $V_M = -P^4/8M_e^3c^2$ , and can thus depend only on the quantum numbers  $n$  and  $\ell$ ,  $m_s$  is ruled out because spin is not involved, and  $m_s$  is ruled out due to spherical symmetry. The bare eigenstates are eigenstates of  $H_0$ ,  $L^2$ , and  $L_z$ . Since all of these operators commute with  $P^2$ , the bare eigenstates are already the ‘good’ eigenstates for perturbation theory. The first order relativistic mass effect is therefore

$$E_{n\ell}^{(1)} = -\frac{1}{8M_e^3c^2} \langle n\ell | P^4 | n\ell \rangle = -\frac{1}{8M_e^3c^2} |P^2 | n\ell \rangle|^2. \quad (31)$$

Now from the energy eigenvalue equation we have

$$E_n^{(0)} | n\ell \rangle = -\frac{1}{2M_e} P^2 | n\ell \rangle + V_C(R) | n\ell \rangle, \quad (32)$$

where  $V_C(R)$  is the Coulomb potential. Solving for  $P^2 | n\ell \rangle$  gives

$$P^2 | n\ell \rangle = 2M_e \left( V_C(R) - E_n^{(0)} \right) | n\ell \rangle. \quad (33)$$

Thus the first-order shift is

$$\begin{aligned} E_{n\ell}^{(1)} &= -\frac{1}{2M_e c^2} \langle n\ell | \left( V_C(R) - E_n^{(0)} \right)^2 | n\ell \rangle \\ &= -\frac{1}{2M_e c^2} \left( \left[ E_n^{(0)} \right]^2 - 2E_n^{(0)} \langle V_C(R) \rangle + \langle V_C^2(R) \rangle \right) \\ &= -\frac{1}{2M_e c^2} \left( \left[ E_n^{(0)} \right]^2 - 2E_n^{(0)} \frac{e^2}{4\pi\epsilon_0} \langle R^{-1} \rangle + \left( \frac{e^2}{4\pi\epsilon_0} \right)^2 \langle R^{-2} \rangle \right) \end{aligned} \quad (34)$$

The expectation values  $\langle R^{-1} \rangle$  and  $\langle R^{-2} \rangle$  can be computed via the Feynman-Hellman theorem,

$$\langle n(x) | \frac{\partial H(x)}{\partial x} | n(x) \rangle = \frac{\partial E_n(x)}{\partial x}, \quad (35)$$

where  $x$  can be any parameter that appears both in  $H$  and the energy eigenvalue  $E_n$ . Using  $H_0$  in place of  $H$ , and setting  $x = e$ , we get

$$\langle R^{-1} \rangle = -\frac{4\pi\epsilon_0}{2e} \frac{\partial E_n}{\partial e} = \frac{1}{n^2 a_0}. \quad (36)$$

Similarly taking  $x = \ell$ , and using  $n = n_r + \ell + 1$ , so that  $\partial n / \partial \ell = 1$  leads to

$$\langle R^{-2} \rangle = \frac{2m_e}{\hbar^2(2\ell + 1)} \frac{\partial E_n^{(0)}}{\partial n} = \frac{1}{(\ell + 1/2)n^3 a_0^3}. \quad (37)$$

With these results we find in the end that relativistic mass correction is

$$E_{n\ell}^{(1)} = -\frac{\left[ E_n^{(0)} \right]^2}{2M_e c^2} \left( \frac{4n}{\ell + 1/2} - 3 \right). \quad (38)$$

This lifts the degeneracy of the hydrogen levels with respect to  $\ell$ . If we compute the relative order-of-magnitude of this effect we find

$$\frac{E_{n\ell}^{(1)}}{E_n^{(0)}} \approx \frac{E_n^{(0)}}{M_e c^2} \sim \alpha^2. \quad (39)$$

### 3.2 Spin-Orbit interaction

The spin-orbit shift is the energy shift due to the magnetic-dipole interaction between the electron and the magnetic field generated by the proton. To estimate the effect, we can consider the interaction of the electron spin with the magnetic field as seen by the electron, which has a contribution from the orbital motion as well as the dipole-magnetic field generated by the proton spin,

$$\vec{B} = \vec{B}_{orbit} + \vec{B}_{proton}. \quad (40)$$

The field of the proton is smaller than the orbital field by a factor  $M_e/M_p \sim 10^{-3}$ , so we can ignore it here, and include it later when we calculate the hyperfine splitting. The perturbation operator is therefore

$$V_{SO} = -\vec{\mu}_e \cdot \vec{B}_{orbit}, \quad (41)$$

where  $\vec{\mu}_e = -\frac{|e|\hbar}{2m_e} \vec{S}$  is the electron intrinsic magnetic dipole moment. The field seen by the electron due to the orbital motion of the proton cannot be correctly computed from a naive circular orbit model, but one can estimate the order-of-magnitude of the field in this manner. The field of a circular loop of wire of radius  $R$  and current  $I$  is given by  $\vec{B} = I/(2\epsilon_0 c^2 R)$ . The current in a loop is given by the charge density times the flow velocity, which gives  $I = (e/2\pi R)v$ , so that with  $v = \vec{L}/m_e R$  we find  $\vec{B} = e\vec{L}/(4\pi\epsilon_0 M_e c^2 R^3)$ . The correct value, including a relativistic effect called ‘Thomas precession’ is reduced by a factor 2, giving

$$V_{SO} = \frac{e^2}{8\pi\epsilon_0 M_e^2 c^2} \frac{\vec{L} \cdot \vec{S}}{R^3}. \quad (42)$$

The bare eigenstates of hydrogen are simultaneous eigenstates of  $H$ ,  $L^2$ ,  $L_z$ , and  $S_z$ , but  $L_z$  and  $S_z$  do not commute with  $\vec{L} \cdot \vec{S}$ . However,  $J^2$  and  $J_z$  do, where  $\vec{J} = \vec{L} + \vec{S}$ . This means that eigenstates of  $L_z$  and  $S_z$  are not the ‘good’ eigenstates for perturbation theory. Instead, the states,  $\{|nljm_j^{(0)}\}$ , i.e. simultaneous eigenstates of  $H_0$ ,  $L^2$ ,  $J^2$  and  $J_z$ , are ‘good’ eigenstates. Fortunately we have already studied these states, and know that the good states can be expressed in terms of the bare states as  $|nljm_j^{(0)}\rangle = \sum_{m_\ell m_s} |nlm_\ell m_s\rangle \langle nlm_\ell m_s | nljm_j\rangle$ , where the components  $\langle nlm_\ell m_s | nljm_j\rangle$  are known as Clebsch-Gordon coefficients, which can be looked up in several reference books and are known by Mathematica as `SphericalHarmonicY`[ $\ell, m, \theta, \phi$ ].

The first-order spin-orbit energy shifts of the  $j$ -levels are

$$E_{nljm_j}^{(1)} = \frac{e^2}{8\pi\epsilon_0 M_e^2 c^2} \langle nljm_j^{(0)} | \frac{\vec{L} \cdot \vec{S}}{R^3} | nljm_j^{(0)} \rangle. \quad (43)$$

Now we have  $J^2 = L^2 + 2\vec{L} \cdot \vec{S} + S^2$ , so that

$$\vec{L} \cdot \vec{S} = \frac{1}{2}(J^2 - L^2 - S^2), \quad (44)$$

which leads to

$$\langle \ell j | \vec{L} \cdot \vec{S} | \ell j \rangle = \frac{\hbar^2}{2} (j(j+1) - \ell(\ell+1) - 3/4). \quad (45)$$

which vanishes for  $L = 0$ . It can also be shown via *Kramer’s relation* (homework exercise) that

$$\langle R^{-3} \rangle = \frac{1}{\ell(\ell+1/2)(\ell+1)n^3 a_0^3}, \quad (46)$$



so that the spin-orbit shift is given by

$$\begin{aligned}
E_{n\ell j}^{(1)} &= \frac{\hbar^2 e^2}{16\pi\epsilon_0 M_e^2 c^2} \frac{(j(j+1) - \ell(\ell+1) - 3/4)}{\ell(\ell+1/2)(\ell+1)n^3 a_0^3} \\
&= \frac{[E_n^{(0)}]^2}{M_e c^2} \frac{n(j(j+1) - \ell(\ell+1) - 3/4)}{\ell(\ell+1/2)(\ell+1)}.
\end{aligned} \tag{47}$$

Thus the relative magnitude of the spin-orbit shift is

$$\frac{E_{\ell j}^{(1)}}{E_n^{(0)}} \sim \frac{E_n^{(0)}}{M_e c^2} \sim \alpha^2. \tag{48}$$

We note that for  $\ell = 0$  we find the indeterminate expression of zero divided by zero, so it is not clear what the  $\ell = 0$  shift is. The numerator vanishes due to an exact symmetry, whereas the denominator vanishes due to the fact that  $\langle R^{-3} \rangle = \infty$  for  $\ell = 0$ . This divergence is due to the contribution from very small radius. Before we get to  $r = 0$  there must be new physics for at least two reasons. At  $r = 10^{-15}\text{m}$  the electron would enter the nucleus, within which, the coulomb potential drops back to zero at  $r = 0$ . This would constrain  $\langle R^{-3} \rangle$  to a large but finite value. In addition, at very small  $r$  the electron velocity becomes highly relativistic, since  $KE(r) = E_n - V_c(r) \xrightarrow{r \rightarrow 0} \infty$ . Thus at small  $r$ , the relativistic correction to the orbital must be non-perturbative. An exact relativistic calculation of the  $\ell = 0$  orbital may go to zero on a longer scale than  $10^{-15}\text{m}$ , and even further reduce the value of  $\langle R^{-3} \rangle$ . Thus we should conclude that the spin-orbit term vanishes exactly for  $\ell = 0$ , due to spherical symmetry.

### 3.3 The Darwin term

The Coulomb potential is an effective potential that describes the exchange of virtual photons between charged particles. Describing such processes with ‘potentials’, i.e. operators that are diagonal in  $\vec{R}$ , is a very good approximation, but is not rigorous. The true interaction has an off-diagonal form, so that the term in the Schrödinger equation takes the form

$$\int d^3 r' V\left(\frac{\vec{r} + \vec{r}'}{2}\right) f(\vec{r} - \vec{r}') \psi(\vec{r}'), \tag{49}$$

where the kernel  $f(\rho)$  is sharply peaked at  $\rho = 0$ . The width of  $f(\rho)$  is a measure of the off-diagonality of the interaction. In physical terms, it represents the difference between the position the particle when it emits a virtual photon and the position at which it re-absorbs the photon. This is due to the fact that the particle is moving at finite velocity, and the virtual photon takes time to scatter from another charged particle due to the finite speed of light. For a hydrogen atom, we can therefore estimate this time delay as  $T = a_0/c$ , while the velocity of the electron is of the order  $v = \frac{\hbar}{M_e a_0}$ . Thus the distance the electron travels during time  $T$  is  $r_C = vT = \frac{\hbar}{M_e c}$ , which is known as the Compton radius.

We can therefore estimate the order-of-magnitude of the Darwin term if we modify the Coulomb potential as

$$V_C(R) \rightarrow \int d^3 \rho f(\rho) V_C(\vec{R} + \vec{\rho}), \tag{50}$$

where  $f(\rho)$  is a normalized distribution centered at  $\rho = 0$ , which is only non-zero over a region of volume  $\left(\frac{\hbar}{M_e c}\right)^3$ . Expanding  $V_C(R + \rho)$  around  $\rho = 0$  gives

$$V_C(\vec{R} + \vec{\rho}) = V_C(R) + \frac{1}{2}\nabla^2 V_C(R) \int d^3\rho f(\rho)\rho^2, \quad (51)$$

where the first-order term vanishes due to the spherical symmetry of  $V_C$ . From a dimensional analysis we can estimate

$$\int d^3\rho f(\rho)\rho^2 \approx \frac{\hbar^2}{M_e^2 c^2}. \quad (52)$$

And we know that the Laplacian of  $1/R$  is  $4\pi\delta^3(\vec{R})$ , which gives

$$V_C(\vec{R} + \vec{\rho}) \approx V_C(R) + \frac{\hbar^2 e^2}{2\epsilon_0 M^2 c^2} \delta^3(\vec{R}). \quad (53)$$

The exact result,

$$V_D = \frac{\hbar^2 e^2}{8\epsilon_0 M_e^2 c^2} \delta^3(\vec{R}), \quad (54)$$

shows we are only off by a factor of 1/4. To first-order, the darwin shift depends only on  $n$  and  $\ell$  and is given by

$$E_{n\ell}^{(1)} = \frac{\hbar^2 e^2}{8\epsilon_0 M_e^2 c^2} |\psi_{n\ell m_\ell}(0)|^2 \quad (55)$$

Since  $|\psi_{n\ell m_\ell}(0)|^2 \propto R_{n\ell}^2(0)$ , and  $\lim_{r \rightarrow 0} R_{n\ell}(r) \propto r^\ell$ , we find that the Darwin term vanishes for  $\ell \neq 0$ . For  $\ell = 0$  we have  $|\psi_{n00}(0)|^2 = 1/\pi(na_0)^3$ , so that the Darwin shift is

$$E_{n0}^{(1)} = \frac{\hbar^2 e^2}{8\pi\epsilon_0 M_e^2 c^2 a_0^3 n^3} = \frac{[E_n^{(0)}]^2}{m_e c^2} 2n. \quad (56)$$

Again we find that the relative shift is of the order of  $\alpha^2$

$$\frac{E_n^{(1)}}{E_n^{(0)}} = \frac{E_n^{(0)}}{M_e c^2} 2n \sim \alpha^2. \quad (57)$$

### 3.4 Combined fine structure

For  $\ell = 0$ , which implies  $j = 1/2$ , we keep only the mass correction and Darwin terms. Adding them together gives

$$E_{n0\frac{1}{2}}^{(1)} = -\frac{[E_n^{(0)}]^2}{M_e c^2} \left(2n - \frac{3}{2}\right). \quad (58)$$

For the case  $\ell \neq 0$  we keep only the mass correction and the spin-orbit terms, giving

$$E_{n\ell j}^{(1)} = -\frac{[E_n^{(0)}]^2}{M_e c^2} \left[ \frac{n(3\ell(\ell+1) - j(j+1) + 3/4)}{\ell(\ell+1/2)(\ell+1)} - \frac{3}{2} \right]. \quad (59)$$

For all values of  $\ell$  and  $j$ , the following expression will give the correct fine-structure shift (relying on the fact that for  $\ell \neq 0$ , either  $j = \ell + 1/2$  or  $j = \ell - 1/2$ ),

$$E_{n\ell j}^{(1)} = -\frac{[E_1^{(0)}]^2}{M_e c^2} \frac{1}{n^4} \left( \frac{2n}{j + 1/2} - \frac{3}{2} \right), \quad (60)$$

which we in actuality only depends on  $n$  and  $j$ , and not  $\ell$ . In terms of the ground-state binding energy,  $|E_1^{(0)}|$ , and the fine structure constant, this can be written as

$$E_{n\ell j}^{(1)} = -|E_1^{(0)}| \frac{\alpha^2}{2n^4} \left( \frac{2n}{j + 1/2} - \frac{3}{2} \right), \quad (61)$$

from which we see that the fine structure shift scales as  $\alpha^2 \approx 10^{-5}$  times the bare level binding energy. Using  $2|E_1^{(0)}| = m_e c^2 \alpha^2$ , we can write the fine-structure shift in terms of fundamental constants as

$$E_{n\ell j}^{(1)} = -m_e c^2 \frac{\alpha^4}{4n^4} \left( \frac{2n}{j + 1/2} - \frac{3}{2} \right). \quad (62)$$

For  $\ell = 0$  states, only  $j = 1/2$  is allowed, so these levels are shifted down, but not split, by the fine-structure effects. The 1s ground state is shifted down by

$$E_{10\frac{1}{2}}^{(1)} = -\frac{1}{8} m_e c^2 \alpha^4. \quad (63)$$

## 4 Hyperfine structure

Prior to computing the fine structure, the energy depended only on the quantum number  $n$ . Fine-structure effects split each  $n$  level into sublevels, which are eigenstates of  $L^2$  and  $J^2$ . The energy shift of each  $|nlj\rangle$  state depends on the quantum numbers  $n$ , and  $j$ , but is smaller than the magnitude of the bare energy of the  $n$ th level by roughly a factor  $\alpha^2 10^{-4}$ . On an energy scale smaller still by a factor  $m_e/m_p \sim 10^{-3}$  we find that the  $|nlj\rangle$  states are again split into sublevels. This is due to the effects of the nuclear spin  $\vec{I}$ . For some atomic isotopes, the nuclear ground-state will have zero nuclear spin,  $i = 0$ , and there will be no hyperfine splitting. For Hydrogen, we have a single proton, so that  $i = 1/2$ . Other atomic species/isotopes can have nuclear spin values such as  $i = 1/2, 1/3/2, \dots$ . Each nuclear isotope also has its own magnetic dipole moment, which is not equal to the proton  $g$ -factor of  $g_p = 5.585694713$ . Thus the results for simple hydrogen can be extended to other hydrogen-like systems by putting in the correct nuclear mass and magnetic dipole moment.

For hydrogen, we have a nuclear magnetic dipole moment of  $\vec{\mu}_p = g_p |e|/(2m_p)$ . The hyperfine interaction includes then two effects, 1.) the interaction of the electron's spin with the magnetic field of the nucleus, and 2.) the interaction of the nuclear spin with the field generated by the electrons orbital motion. Thus the hyperfine perturbation operator is approximately given by

$$V_{hf} = -\vec{\mu}_p \cdot \vec{B}_{orbit} - \vec{\mu}_e \cdot \vec{B}_{proton} \quad (64)$$

The orbital field seen by the nucleus is the same as the nuclear field seen by the electron, but with opposite sign. Both see each other moving in the same sense (e.g. clockwise), but correspond to opposite currents due to having opposite charge. This gives

$$\vec{B}_{orbit} = -\frac{e\vec{L}}{4\pi\epsilon_0 M_e c^2 R^3}. \quad (65)$$

The magnetic field created by the proton's spin-magnetic moment is,

$$\vec{B}_{proton} = \frac{1}{4\pi\epsilon_0 c^2} \frac{[3(\vec{\mu}_p \cdot \vec{e}_r)\vec{e}_r - \vec{\mu}_p]}{R^3} + \frac{2}{3\epsilon_0 c^2} \delta^3(\vec{R}) \quad (66)$$

where  $\vec{e}_r$  is the radial unit vector, given in cartesian coordinates by

$$\vec{e}_r = \sin \Theta \cos \Phi \vec{e}_x + \sin \Theta \sin \Phi \vec{e}_y + \cos \Theta \vec{e}_z, \quad (67)$$

where  $\Theta$  and  $\Phi$  are the angle operators. When the electron is outside the proton, it sees a standard dipolar magnetic field, given by the first term in (66). The delta-function term in (66) is called the 'contact term' and should not be understood as a literal delta-function, but rather as a sharply peaked function whose width is small compared to  $a_0$ . This contact term can be interpreted as the interaction of the electron spin with the nuclear spin while the electron is *inside* the proton. This is possible because protons are not elementary particles and have a finite radius. Since the quarks inside are point-like elementary particles, there is plenty of space for the electron. If we assume a uniform current density inside the proton, then the interior magnetic field would be uniform, and parallel to the nuclear spin  $\vec{I}$ . Note that the contact term will only contribute to the hyperfine splitting of s-orbitals ( $\ell = 0$ ), as orbitals with non-zero orbital angular momentum have zero probability density at  $r = 0$ .

The hyperfine interaction operator is thus

$$V_{hf} = \frac{g_p e^2}{8\pi\epsilon_0 M_p M_e c^2} \left[ \frac{\vec{I} \cdot \vec{L}}{R^3} + \frac{[3(\vec{S} \cdot \vec{e}_r)(\vec{I} \cdot \vec{e}_r) - \vec{S} \cdot \vec{I}]}{R^3} + \frac{8\pi}{3} \vec{S} \cdot \vec{I} \delta^3(\mathbf{R}) \right] \quad (68)$$

The pre-factor can alternatively be expressed in terms of the Bohr magneton  $\mu_B = |e|\hbar/2M_e$  and the ‘nuclear magneton’,  $\mu_n = |e|\hbar/2M_p$ , as

$$\frac{g_p e^2}{8\pi\epsilon_0 M_p M_e c^2} = \frac{1}{4\pi\epsilon_0} \frac{g_e \mu_B g_p \mu_n}{\hbar^2}, \quad (69)$$

which might be slightly easier to remember.

The hyperfine interaction commutes with  $L^2$ ,  $S^2$ , and  $I^2$ , so  $\ell$ ,  $s$ , and  $i$  are good quantum numbers. As with the fine-structure interaction, the hyperfine interaction does not commute with  $L_z$ ,  $S_z$  or  $I_z$ , indicating that  $m_\ell$ ,  $m_s$ , and  $m_i$  are not good quantum numbers. Of course the fine-structure already eliminated  $m_\ell$  and  $m_s$ , leaving  $j$  and  $m_j$  as good quantum numbers. The hyperfine interaction, however, does not commute with  $J^2$  or  $J_z$ , so that  $j$  and  $m_j$  are no longer a good quantum numbers for perturbation theory. The good quantum numbers are therefore  $\ell$ ,  $s$ , and  $i$ . Note that because states with different  $j$  are non-degenerate, they do not affect each other at first-order in perturbation theory. Thus at first-order,  $j$  is still a good quantum number, but the second-order correction will mix slightly the different  $j$  states.

For a system with no external fields applied, we know that the total angular momentum is always a good quantum number. We can therefore define the total angular momentum operator,  $\vec{F} = \vec{J} + \vec{I} = \vec{L} + \vec{S} + \vec{I}$ , after which it is straightforward to show that  $F^2$  commutes with the hyperfine interaction, so that  $f$  is a good quantum number. According to the theory of addition of angular momentum, the value of  $f$  ranges from  $|j - i|$  to  $j + i$ . Thus the good states for perturbation theory are  $|n, \ell, s, i, j, f, m_f^{(0)}\rangle$ . Using Clebsch-Gordan coefficients,  $C(jm; j_1 m_1 j_2 m_2) = \langle j_1 j_2 j m | j_1 j_2 m_1 m_2 \rangle$  for the angular momentum addition  $\vec{J} = \vec{J}_1 + \vec{J}_2$ , we can relate the various basis sets we have encountered via

$$\begin{aligned} |nlsifm_f\rangle &= \sum_{m_j m_i} \delta_{m_j+m_i, m_f} |nlsjm_j\rangle \otimes |im_i\rangle C(fm_f; jm_j im_i) \\ &= \sum_{m_j m_i m_\ell m_s} \delta_{m_j+m_i, m_f} \delta_{m_\ell+m_s, m_j} |n\ell m_\ell\rangle \otimes |sm_s\rangle \otimes |im_i\rangle C(fm_f; jm_j im_i) C(jm_j; \ell m_\ell sm_s), \end{aligned} \quad (70)$$

where the delta functions are added to incorporate the standard selection rule and therefore avoid summing over terms which are zero. We note that the final answer for the hyperfine energy shift cannot depend on  $m_f$  due to the spherical symmetry of the atom. Thus the choice of quantization axis is arbitrary and should be chosen for convenience. In the end, the hyperfine interaction will lift the degeneracy of the  $j$  levels, splitting each level into a set of levels which have well-defined  $j$  and  $f$  values. The number of new hyperfine sub-levels can vary widely, as it depends on both  $j$  and  $i$ .

#### 4.1 Hyperfine splitting of an s-orbital

Computing the hyperfine splitting of an arbitrary energy level can be very difficult for  $\ell > 0$ . Here we will compute the hyperfine splitting of an arbitrary  $\ell = 0$  state. For the case  $\ell = 0$ , we have  $\vec{J} = \vec{S}$  and  $m_\ell = 0$ . From the radial wavefunctions,  $R_{n,00}(r)$ , we find

$$\langle R^{-3} \rangle = \infty, \quad (71)$$

which, as discussed in Section 3.2, this should be interpreted as a very large, but finite, number. We will see that due to spherical symmetry, each of the  $R^{-3}$  terms is exactly zero, so that the  $\ell = 0$  hyperfine splitting

is entirely due to the contact term in (68). The  $\vec{L} \cdot \vec{I}$  term vanishes since  $\ell = 0$ , while the dipole-dipole term in (68) requires the evaluation of

$$\begin{aligned} \langle (\vec{S} \cdot \vec{e}_r)(\vec{I} \cdot \vec{e}_r) \rangle &= \langle \sin^2 \Theta \cos^2 \Phi S_x I_x + \sin^2 \Theta \sin^2 \Phi S_y I_y + \cos^2 \Theta S_z I_z + \sin^2 \Theta \sin \Phi \cos \Phi (S_x I_y + I_y S_x) \\ &\quad + \sin \Theta \cos \Theta \cos \Phi (S_x I_z + S_z I_x) + \sin \Theta \cos \Theta \sin \Phi (S_z I_z + S_z I_y) \rangle \end{aligned} \quad (72)$$

Terms which go as  $\sin \Phi$ ,  $\cos \Phi$ , or  $\sin \Phi \cos \phi$  will vanish under phi-integration. The remaining term must be symmetric under cyclic permutation of  $x, y, z$ , as there is no preferred z-axis. Thus we will just evaluate the term  $\langle \sin^2 \Theta \cos^2 \Phi S_z I_z \rangle$ , and generate the other two terms by taking  $z \rightarrow y$  and then  $y \rightarrow x$ . Since  $\ell = 0$  and  $m_\ell = 0$ , the  $S_z I_z$  term is

$$\begin{aligned} \langle \ell = 0; m_\ell = 0 | \cos^2 \Theta | \ell = 0; m_\ell = 0 \rangle &= \int_0^\pi d\theta \int_0^{2\pi} d\phi \sin \theta \cos^2 \theta |Y_0^0(\theta, \phi)|^2 \\ &= \frac{1}{2} \int_{-1}^1 du (u^2) = \frac{1}{3}, \end{aligned} \quad (73)$$

so that the full expression becomes

$$\langle (\vec{S} \cdot \vec{e}_r)(\vec{I} \cdot \vec{e}_r) \rangle = \frac{1}{3} \langle S_x I_x \rangle + \frac{1}{3} \langle S_y I_y \rangle + \frac{1}{3} \langle S_z I_z \rangle = \frac{1}{3} \langle \vec{S} \cdot \vec{I} \rangle. \quad (74)$$

At this point we see that the dipole-dipole term is exactly zero, leaving only the contact term.

From the radial wavefunction, we have

$$\langle \delta^3(\vec{R}) \rangle = |\psi_{n00}(0)|^2 = \frac{1}{\pi n^3 a_0^3}, \quad (75)$$

and for the spin components we use the fact that, for  $\ell = 0$ ,  $\vec{J} = \vec{S}$  to find

$$\langle \vec{S} \cdot \vec{I} \rangle = \langle \vec{J} \cdot \vec{I} \rangle = \frac{\hbar^2}{2} (f(f+1) - i(i+1) - s(s+1)) = \frac{\hbar^2}{2} (f(f+1) - 3/2). \quad (76)$$

This gives for the hyperfine shift

$$E_{hf}^{(1)} = \frac{g_p e^2}{3\epsilon_0 M_p M_e c^2} \frac{1}{\pi n^3 a_0^3} \frac{\hbar^2}{2} (f(f+1) - 3/2). \quad (77)$$

This is often expressed in terms of fundamental constants of nature, giving

$$E_{nlf}^{(1)} = \frac{4}{3} g_p \frac{M_e}{M_p} (M_e c^2) \alpha^4 \frac{1}{n^3} (f(f+1) - 3/2). \quad (78)$$

In terms of the bare ground state binding energy, we can use  $m_e c^2 \alpha_2 = 2|E_1^{(0)}|$  to get

$$E_{hf}^{(1)} = |E_1^{(0)}| \frac{8}{3} g_p \alpha^2 \frac{M_e}{M_p} \frac{1}{n^3} (f(f+1) - 3/2). \quad (79)$$

For  $\ell = 0$ , with  $s = i = 1/2$ , there is only one fine-structure level, with  $j = 1/2$ . The hyperfine interaction then splits this level into two  $f$  levels, with  $f = 0, 1$ . The  $f = 0$  state is shifted down by

$$E_{100}^{(1)} = -\frac{\hbar^2 g_p e^2}{4\pi \epsilon_0 M_p M_e c^2 a_0^3 n^3} \quad (80)$$

whereas the  $f = 1$  state is shifted up by

$$E_{101}^{(1)} = \frac{\hbar^2 g_p e^2}{12\pi\epsilon_0 M_p M_e c^2 a_0^3 n^3}. \quad (81)$$

The up shift of the  $f = 1$  state is 1/3 smaller than the down shift of the  $f = 0$  state. The splitting between these two levels is therefore

$$\Delta E_{n0} = E_{101}^{(1)} - E_{100}^{(1)} = \frac{\hbar^2 g_p e^2}{3\pi\epsilon_0 M_p M_e c^2 a_0^3 n^3}. \quad (82)$$

For the 1s hydrogen ground state, this gives

$$\Delta E_{10} = 9.411706178314 \times 10^{-25} \text{J}. \quad (83)$$

Note that the ratio of the 1s hyperfine splitting to the fine-structure level shift is  $32g_p M_e / 3M_p \approx 1/16$ , so it is just over one order-of-magnitude smaller.

The hyperfine splitting of the hydrogen ground state is the physical quantity which is known experimentally to the highest number of significant figures. This is due to a device called the maser. The maser is essentially a laser which operates on this transition. Converting this to energy to a wavelength via  $\lambda = 2\pi c\hbar/\Delta E$  gives  $\lambda = 21\text{cm}$ . This means that a hydrogen atom in the electronic ground-state dropping from the  $f = 1$  to  $f = 0$  hyperfine levels will emit a photon with a wavelength of 21cm. Note that gas of hydrogen will contain both  $f = 0$  and  $f = 1$  atoms if  $KT$  is greater than the hyperfine splitting. This occurs at  $T = 34\text{mK}$ . By contrast, the  $n = 2$  level becomes thermally activated at  $T = 10^5\text{K}$ . Because of this, interstellar hydrogen ( $T = 2.75\text{K}$ ) consists only of equal numbers  $n = 1, f = 1$  and  $n = 1, f = 0$  atoms. Such a system will spontaneously emit light at 21cm wavelength as the  $f = 1$  atoms fall to the  $f = 0$  state. For this reason most information we have about interstellar hydrogen are detected by radioastronomy via this radiation.

## 4.2 Zeeman Effect

Here we consider the effect of a uniform static magnetic field on the hydrogen level structure. Here we should think of the atom as a tiny magnetic dipole, with dipole moment  $\vec{\mu}$ , so that the interaction operator takes the form

$$V_Z = -\vec{\mu} \cdot \vec{B}. \quad (84)$$

With the applied magnetic field defining the z-axis, we have  $\vec{B} = B_0 \vec{e}_z$ , so that

$$V_Z = -\mu_z B_0 \quad (85)$$

There are three sources of the atomic magnetic dipole moment: (1) the moment generated by the electrons orbital motion, (2) the moment due to the electron spin, and (3) the moment due to the nuclear spin, so that

$$\vec{\mu} = \vec{\mu}_{orbit} + \vec{\mu}_{electron} + \vec{\mu}_{nucleus}. \quad (86)$$

The orbital magnetic dipole moment can be correctly determined from a simplistic Bohr-type model where the electron revolves in a circular orbit with quantized angular momentum. An ideal magnetic dipole is a tiny ring with area  $A = \pi R^2$  carrying current  $I$ . In the limit as  $I \rightarrow \infty$  and  $A \rightarrow 0$ , the field approaches that of a perfect dipole with dipole moment  $\mu = IA$ , with a direction given by the right-hand-rule with respect to the current flow direction. Electrical current is defined as charge density times velocity, which gives  $I = \frac{e}{2\pi R} v$ , where  $v$  is the electron velocity. Multiplying by  $A$  and using  $\vec{L} = MvR$  to eliminate  $v$ , leads to

$$\vec{\mu}_{orbit} = -\frac{|e|\hbar}{2M_e} \vec{L}, \quad (87)$$

where the minus sign is due to the fact that the current flows in the opposite direction to the electrons orbit. The intrinsic dipole moment of the electron cannot be derived from such simple arguments, but the value is

$$\vec{\mu}_{electron} = -\frac{g_e |e|\hbar}{2M_e} \vec{S} \approx -\frac{e\hbar}{M_e} \vec{S}, \quad (88)$$

where  $g_e = 2.0023193043617$  is the experimental value of the electron  $g$ -factor. For the simplest hydrogen atom, the nucleus consists of a single proton. The proton magnetic dipole moment is

$$\vec{\mu}_{proton} = \frac{g_p |e|\hbar}{2M_p} \vec{S}_p, \quad (89)$$

where  $g_p = 5.585694713$  as the proton  $g$ -factor, and  $\vec{S}_p$  being the proton spin operator. Thus the orbital and electronic dipole moments are of equal order, while the nuclear dipole moment is a factor  $g_p M_e / M_p \sim 10^{-3}$  smaller. Hence we will ignore the nuclear spin contribution relative to the  $\vec{L}$  and  $\vec{S}$  contributions. The Zeeman interaction operator then becomes

$$V_B = \omega_0 (L_z + 2S_z), \quad (90)$$

where  $\omega_0 = \frac{|e|\hbar B_0}{2M_e}$  is known as the ‘Larmor frequency’.

In the section we will consider the Zeeman effect in three distinct regimes.



1. **Strong-field regime:** When the Zeeman energy  $\hbar\omega_0 \propto B_0$  is large compared to the fine-structure shift,  $E_{nlj}^{(1)} \approx |E_1^{(0)}|\alpha^2$ , but still small compared to the bare-energy spacing  $\sim |E_1^{(0)}|$ , then the fine structure should be considered as a perturbation of the Zeeman structure. While the Zeeman structure is treated as a perturbation on the bare Hamiltonian. The Strong field domain is therefore

$$|E_1^{(0)}|\alpha^2 \ll \hbar\omega_0 \ll |E_1^{(0)}|. \quad (91)$$

For stronger fields, the hydrogen atom will be ionized.

2. **Weak-field regime:** In the opposite extreme, when the Zeeman energy is small compared to the hyperfine splitting,

$$\hbar\omega_0 \ll |E_1^{(0)}|\alpha^2 \frac{M_e}{M_p}, \quad (92)$$

then the Zeeman shift should be treated as a perturbation on the hyperfine  $|nljm_f\rangle$  levels.

3. **Intermediate-field regime:** When the Zeeman energy is small compared to the fine-structure splitting, but large compared to the hyperfine splitting, then the Zeeman shift should be treated as a perturbation on the fine-structure  $|nljm_j\rangle$  levels

### 4.3 Wigner-Eckert theorem

Before we compute the Zeeman shifts for the various regimes, we need learn the Wigner-Eckert theorem. This theorem deals with the concept of vector operators, and the fact that they must be very closely related to the total-angular-momentum operator for a given system. Before defining a vector operator, we should define a scalar operator. Both operators are defined with respect to rotational transformations. In quantum mechanics, rotations are described by unitary transformations, with the unitary rotation operator given by

$$U_R(\vec{\theta}) = e^{-i\vec{J}\cdot\vec{\theta}/\hbar}, \quad (93)$$

which describes a rotation of angle  $\theta = |\vec{\theta}|$  about the axis  $\vec{e}_\theta = \vec{\theta}/\theta$ . The operator  $\vec{J}$  is a placeholder for the *total angular momentum operator* for the system under consideration. In this context it should not be confused with the single electron operator  $\vec{L} + \vec{S}$ . For example, in a two-spin system you would substitute  $\vec{S}_1 + \vec{S}_2$  for  $\vec{J}$  in the rotation operator, while for a hydrogen atom you would substitute  $\vec{F} = \vec{L} + \vec{S} + \vec{I}$  in place of  $\vec{J}$ .

**Definition: Scalar operator:**  $A$  is a scalar operator when, for all possible  $\vec{\theta}$ , one has

$$A' = U_R^\dagger(\vec{\theta})AU_R(\vec{\theta}) = A. \quad (94)$$

This has the same meaning as a usual scalar quantity, i.e. it must be invariant under rotation. It should be clear that  $[A, \vec{J}] = 0$  is a necessary and sufficient for  $A$  to be a scalar operator.

**Definition: Vector operator:**  $\vec{V}$  is a vector operator when, for all possible  $\vec{\theta}$ , one has

$$\vec{V}' = U_R^\dagger(\vec{\theta})\vec{V}U_R(\vec{\theta}) = \mathbf{M}_R(\vec{\theta})\vec{V}, \quad (95)$$

where  $\mathbf{M}_R(\vec{\theta})$  is the standard  $3 \times 3$  rotation matrix for ordinary (non-operator valued) vectors. For example, with  $\vec{\theta} = \phi \vec{e}_z$ , we have

$$\mathbf{M}_R(\phi \vec{e}_z) = \begin{pmatrix} \cos \theta & -\sin \theta & 0 \\ \sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{pmatrix}. \quad (96)$$

For the defining equation  $U_R(\vec{\theta})\vec{V}U_R^\dagger(\vec{\theta}) = \mathbf{R}(\vec{\theta})\vec{V}$  to be valid for any  $\vec{\theta}$  obviously requires a very strong relationship between  $\vec{V}$  and  $\vec{J}$ , which generates rotations. To see what exactly this relationship is, we need only consider infinitesimal rotations about the cartesian axes, as all possible rotations can be build up as sequences of these rotations. For an infinitesimal rotation  $\delta\phi$  about  $\vec{e}_z$ , and looking only at the z-component, we must have

$$(1 - i\delta\phi J_z/\hbar)V_z(1 + i\delta\phi J_z/\hbar) = V_z. \quad (97)$$

To be true to first-order in  $\delta\phi$  this leads directly to

$$[J_z, V_z] = 0. \quad (98)$$

In the end, similar calculations with infinitesimal rotations will yield the following nine relations that must be satisfied by vector operators:

$$\begin{aligned} [J_x, V_x] &= 0 & [J_y, V_y] &= 0 & [J_z, V_z] &= 0 \\ [J_x, V_y] &= i\hbar V_z & [J_y, V_z] &= i\hbar V_x & [J_z, V_x] &= i\hbar V_y \\ [J_x, V_z] &= -i\hbar V_y & [J_y, V_x] &= -i\hbar V_z & [J_z, V_y] &= -i\hbar V_x \end{aligned} \quad (99)$$

From these vector-operator relations it is straightforward to show that.

$$[J_z, V_\pm] = \pm\hbar V_\pm \quad (100)$$

$$[J_\pm, V_\pm] = 0 \quad (101)$$

$$[J_\pm, V_\mp] = \pm 2\hbar V_z \quad (102)$$

where  $V_\pm = V_x \pm iV_y$ . From these equations it is possible to derive the Wigner-Eckert theorem for vector operators. There are also Wigner Eckert theorems for scalar and tensor operators, which we do not require at present.

Let  $\{|k, j, m\rangle\}$  be a universal basis for a system with total angular momentum quantum number  $j$  and corresponding z-component  $m$ . The index  $k$  is a place-holder for whatever other quantum numbers describe the system. For example, for the hyperfine interaction we would make the replacements  $j \rightarrow f$ ,  $m_j \rightarrow m_f$ , and  $k \rightarrow n\ell s j$ .

**Vector Theorem 1**, selection rules:

$$\langle k j m | V_z | k' j' m' \rangle = 0 \quad \text{unless } m = m', \quad (103)$$

$$\langle k j m | V_\pm | k' j' m' \rangle = 0 \quad \text{unless } m = m' \pm 1, \quad (104)$$

where  $V_\pm = V_x \pm iV_y$ .

**Vector Theorem 2**, the ‘projection postulate’:

$$\vec{V}_{kj} = a(k, j) \vec{J}_{kj}, \quad (105)$$

where

$$I_{kj} = \sum_m |kjm\rangle\langle kjm|, \quad (106)$$

and

$$\begin{aligned} a(j, k) &= \frac{\text{Tr}\{I_{kj} \vec{J} \cdot \vec{V}\}}{J^2} \\ &= \frac{1}{\hbar^2 j(j+1)} \sum_m \langle kjm | \vec{J} \cdot \vec{V} | kjm \rangle. \end{aligned} \quad (107)$$

This theorem is very powerful, because it says that within a subspace with well-defined  $k$  and  $j$  quantum numbers, you can replace *any* vector operator with a constant times the total angular momentum operator.

### 4.3.1 Proof of vector theorem 1

From the commutator  $[J_z, V_z] = 0$ , it follows that  $[J_z, V_z]|kjm\rangle = 0$ . Expanding this out gives

$$J_z (V_z |kjm\rangle) = \hbar m (V_z |kjm\rangle), \quad (108)$$

which shows that  $V_z |kjm\rangle$  is an eigenstate of  $J_z$  with eigenvalue  $\hbar m$ .<sup>2</sup> Thus it follows from the orthogonality of the eigenstates of a Hermitian operator that  $\langle kjm | V_z |k'j'm'\rangle = 0$  unless  $m = m'$ .

From the commutator  $[J_z, V_\pm] = \pm \hbar V_\pm |kjm\rangle$ , we similarly find

$$J_z (V_\pm |kjm\rangle) = \hbar(m \pm 1) V_\pm |kjm\rangle \quad (109)$$

which shows that  $V_\pm |kjm\rangle$  is an eigenstate of  $J_z$  with eigenvalue  $\hbar(m \pm 1)$ . Again from the properties of Hermitian operators, it then follows that  $\langle kjm | V_\pm |k'j'm'\rangle = 0$  unless  $m = m' \pm 1$ .

### 4.3.2 Proof of vector theorem 2

STEP 1: From  $[J_\pm, V_\pm] = 0$ , it follows that  $\langle kj(m \pm 2) | [J_\pm, V_\pm] |kjm\rangle = 0$ . Expanding this out gives

$$\langle kj(m \pm 2) | J_\pm V_\pm |kjm\rangle = \langle kj(m \pm 2) | V_\pm J_\pm |kjm\rangle, \quad (110)$$

which, after making use of the relation  $J_\pm |kjm\rangle = \hbar \sqrt{j(j+1) - m(m \pm 1)} |kj(m \pm 1)\rangle$ , becomes

$$\hbar \sqrt{j(j+1) - (m \pm 2)(m \pm 1)} \langle kj(m \pm 1) | V_\pm |kjm\rangle = \hbar \sqrt{j(j+1) - m(m \pm 1)} \langle kj(m \pm 2) | V_\pm |kj(m \pm 1)\rangle. \quad (111)$$

Recognizing that  $\langle kj(m \pm 2) | J_\pm |kj(m \pm 1)\rangle = \hbar \sqrt{j(j+1) - (m \pm 1)(m \pm 2)}$  and  $\langle kj(m \pm 1) | J_\pm |kjm\rangle = \hbar \sqrt{j(j+1) - m(m \pm 1)}$ , we can rewrite Eq. (111) as

$$\frac{\langle kj(m \pm 1) | V_\pm |kjm\rangle}{\langle kj(m \pm 1) | J_\pm |kjm\rangle} = \frac{\langle kj(m \pm 2) | V_\pm |kj(m \pm 1)\rangle}{\langle kj(m \pm 2) | J_\pm |kj(m \pm 1)\rangle}. \quad (112)$$

---

<sup>2</sup>Note that this does not imply that  $V_z |kjm\rangle \propto |kjm\rangle$ , as it may be a superposition of states with different  $k$  and  $j$  eigenvalues.

With the definition

$$a_{\pm}(k, j, m) := \frac{\langle kj(m \pm 1) | V_{\pm} | kjm \rangle}{\langle kj(m \pm 1) | J_{\pm} | kjm \rangle}, \quad (113)$$

we see that

$$a_{\pm}(k, j, m) = a_{\pm}(k, j, m + 1). \quad (114)$$

For this to be true for any  $m$  requires  $a_{\pm}$  to be independent of  $m$ , so that we can take  $a_{\pm}(k, j, m) \rightarrow a_{\pm}(k, j)$ , which gives

$$\langle kjm | V_{\pm} | kj(m \mp 1) \rangle = a_{\pm}(k, j) \langle kjm | J_{\pm} | kj(m \mp 1) \rangle. \quad (115)$$

based on the second selection rule from theorem 1. We can re-write this as

$$\langle kjm | V_{\pm} | kjm' \rangle = a_{\pm}(k, j) \langle kjm | J_{\pm} | kjm' \rangle, \quad (116)$$

as both sides vanish unless  $m' = m \mp 1$ , which is equivalent to the statement

$$I_{kj} V_{\pm} I_{kj} = a_{\pm}(k, j) I_{kj} J_{\pm} I_{kj}. \quad (117)$$

STEP 2: From the commutation relation  $[J_{\mp}, V_{\pm}] = \mp 2\hbar V_z$ , it follows that

$$\langle kjm | J_{\mp} V_{\pm} | kjm' \rangle - \langle jkm | V_{\pm} J_{\mp} | kjm' \rangle = \mp 2\hbar \langle kjm | V_z | kjm' \rangle, \quad (118)$$

which leads to

$$\sqrt{j(j+1)-m(m\pm 1)} \langle kj(m \pm 1) | V_{\pm} | kjm' \rangle - \sqrt{j(j+1)-m'(m' \mp 1)} \langle jkm | V_{\pm} | kj(m' \mp 1) \rangle = \mp 2 \langle kjm | V_z | kjm' \rangle. \quad (119)$$

Making use of (117) then leads to

$$\langle kjm | V_z | kjm' \rangle = a_{\pm}(k, j) \hbar m \delta_{m', m}. \quad (120)$$

As the r.h.s. is the same for either choice of '+' or '-', it follows that  $a_+(k, j) = a_-(k, j) \rightarrow a(k, j)$ . From this, together with (117), it follows that

$$I_{kj} V_x I_{kj} = \frac{1}{2} (I_{kj} V_+ I_{kj} + I_{kj} V_- I_{kj}) = a(k, j) I_{kj} J_x I_{kj}, \quad (121)$$

and

$$I_{kj} V_y I_{kj} = \frac{1}{2i} (I_{kj} V_+ I_{kj} - I_{kj} V_- I_{kj}) = a(k, j) I_{kj} J_y I_{kj}, \quad (122)$$

so that

$$\vec{V}_{kj} = a(k, j) \vec{J}_{kj}, \quad (123)$$

which is the main result of the Wigner-Ekert theorem.

To find the coefficient,  $a(k, j)$ , we start from the identity

$$\begin{aligned} \vec{J} \cdot \vec{V} &= J_x V_x + J_y V_y + J_z V_z \\ &= \frac{1}{2} (J_+ V_- + J_- V_+) + J_z V_z. \end{aligned} \quad (124)$$

Applying  $I_{kj}$  and taking the trace gives

$$\text{Tr}\{I_{kj} \vec{J} \cdot \vec{V}\} = \frac{1}{2} \sum_m \langle kjm | (J_+ V_- + J_- V_+ + 2J_z V_z) | kjm \rangle. \quad (125)$$

Applying the  $J$ -operators to the left, and making use of Wigner-Ekert theorem 2 (123) then gives

$$\frac{\text{Tr}\{I_{kj}\vec{J}\cdot\vec{V}\}}{2j+1} = a(k, j)\hbar^2 j(j+1). \quad (126)$$

Because the inner product  $\vec{J}\cdot\vec{V}$  is a scalar, its expectation value with respect to a  $|kjm\rangle$  state cannot depend on  $m$ . Thus we can use a more compact notation, writing

$$\langle\vec{J}\cdot\vec{V}\rangle_{kj} = \alpha(k, j)\hbar^2 j(j+1), \quad (127)$$

where the expectation value is taken with respect to any state in the subspace spanned by  $I_{kj}$ , Solving for  $a(k, j)$  then gives

$$a(k, j) = \frac{\langle\vec{J}\cdot\vec{V}\rangle_{kj}}{\hbar^2 j(j+1)}, \quad (128)$$

which completes the proof of the Vector Wigner-Ekert theorem.

#### 4.4 Strong-field Zeeman effect

In the strong-field regime,  $\hbar\omega_0 \gg |E_1^{(0)}|\alpha^2$ , the bare Hamiltonian is just  $H_0 = \frac{1}{2M_e}P^2 + V_c(R)$ , and the perturbation is  $V_Z = \omega_0(L_z + 2S_z)$ . The original bare eigenstates,  $\{|n\ell m_\ell m_s\rangle\}$  are then already a ‘good’ basis, so we can compute the Zeeman shift of each magnetic sublevel as simply

$$E_{m_\ell m_s}^{(1)} = \hbar\omega_0(m_\ell + 2m_s), \quad (129)$$

or in terms of the magnetic field strength as

$$E_{m_\ell m_s} = \mu_B B_0(m_\ell + 2m_s), \quad (130)$$

where  $\mu_B$  is the ‘Bohr magneton’

$$\mu_B = -\frac{\hbar|e|}{2M_e}. \quad (131)$$

For  $\ell = 0$  levels we have  $m_\ell = 0$  and  $m_s = \pm 1/2$ , so the strong field splits the 1s level into two magnetic sublevels. The upper level, corresponding to the state  $|n\ell m_\ell m_s\rangle = |n00\frac{1}{2}\rangle$  is shifted up by  $\hbar\omega_0$ , and the lower level,  $|n00(-\frac{1}{2})\rangle$ , is shifted down by  $-\hbar\omega_0$ . The Zeeman level spacing is therefore

$$\Delta E_{n0} = E_{n00\frac{1}{2}}^{(1)} - E_{n00-\frac{1}{2}}^{(1)} = 2\hbar\omega_0. \quad (132)$$

For  $\ell = 1$ , we have  $\ell = 0, 1$ , so that  $m_\ell = -1, 0, 1$ , the possible values of  $m_\ell + 2m_s$  are therefore  $-2, -1, 0, 1, 2$ . The p-levels will therefore split into five sublevels, with level spacing:

$$\Delta E_{n1} = \hbar\omega_0. \quad (133)$$

The  $\ell = 2$  levels and corresponding eigenstates are

Table 1: default

level	degeneracy	state(s)
$2\hbar\omega_0$	1	$ n11\frac{1}{2}\rangle$
$\hbar\omega_0$	1	$ n10\frac{1}{2}\rangle$
0	2	$ n11 - (\frac{1}{2})\rangle,  n10\frac{1}{2}\rangle$
$-\hbar\omega_0$	1	$ n10(-\frac{1}{2})\rangle$
$-2\hbar\omega_2$	1	$ n1(-1)(-\frac{1}{2})\rangle$

In general, the  $n^{\text{th}}$  energy level is split into  $(2n + 1 - \delta_{n,1})$  Zeeman sub-levels, that can be labeled as  $|n, m\rangle$ , where  $m \in \{-n, -n+1, \dots, n\}$ . The first-order energy shift of the  $m^{\text{th}}$  sub-level is then  $E_{n,m}^{(1)} = \hbar\omega_0 m$ , with corresponding degeneracy  $d_{n,m} = 2(n - |m|) + \delta_{|m|,n} - 2\delta_{m,0}$ . The  $n = 1$  level, containing only an  $\ell = 0$  state, is split into just two-sublevels, labeled as  $|n, m\rangle$ , where  $m \in \{-1, 1\}$ , i.e. the  $m = 0$  state does not exist. This is consistent with the formula for  $d_{n,m}$ , which gives  $d_{1,0} = 0$ .

## 4.5 Intermediate-field Zeeman effect

For the case  $|E_1^{(0)}|\alpha^2 \gg \hbar\omega_0 \gg |E_1^{(0)}|^2\alpha^2\frac{M_e}{M_p}$ , we can use as a good basis  $\{|nljm_j\rangle\}$ , which means that the Zeeman level shifts will just be added to the fine-structure shifts that we calculated in Section 3. Evaluating the matrix elements  $\langle nljm_j|L_z + 2S_z|nljm_j\rangle$  by brute force, would require expanding the  $|nljm_j\rangle$  states in terms of  $|n\ell m_\ell m_s\rangle$  using Clebsch Gordan coefficients. To avoid this cumbersome approach, we will instead use the Wigner Eckert theorem.

First we note that

$$\langle nlsjm_j|L_z + 2S_z|nlsjm_j\rangle = \langle nlsjm_j|I_{nlsj}(L_z + 2S_z)I_{nlsj}|nlsjm_j\rangle, \quad (134)$$

and that  $L_z + 2S_z$  is the z-component of the vector operator  $\vec{L} + 2\vec{S}$ . The Wigner-Eckert theorem then tells us

$$I_{nlsj}(L_z + 2S_z)I_{nlsj} = a(n, \ell, s, j)I_{nlsj}\vec{J}I_{nlsj}. \quad (135)$$

From this we can see that the Zeeman shift is

$$E_{nlsjm_j}^{(1)} = \omega_0 a(n, \ell, s, j) \langle nlsjm_j|J_z|nlsjm_j\rangle = \hbar\omega_0 a(n, \ell, s, j) m_j, \quad (136)$$

Now

$$\begin{aligned} \vec{J} \cdot \vec{L} &= L^2 + \vec{S} \cdot \vec{L} \\ &= L^2 + \frac{1}{2}(J^2 - L^2 - S^2) \\ &= \frac{1}{2}(J^2 + L^2 - S^2) \end{aligned} \quad (137)$$

and likewise we find

$$\vec{J} \cdot \vec{S} = \frac{1}{2}(J^2 - L^2 + S^2). \quad (138)$$

This then gives

$$\vec{J} \cdot (\vec{L} + 2\vec{S}) = \frac{3}{2}J^2 - \frac{1}{2}L^2 + \frac{1}{2}S^2, \quad (139)$$

so that

$$a(n, \ell, s, j) = \frac{3j(j+1) - \ell(\ell+1) + s(s+1)}{2j(j+1)} \quad (140)$$

Note that this formula is valid for any number of electrons, where  $\vec{S}$  is the total electron spin, so we need not take  $s = 1/2$ .

Thus we conclude that the magnetic field lifts the degeneracy of the  $m_j$  levels, splitting them into  $2j + 1$  Zeeman sub-levels. The energy of each sublevel depends only on  $\ell$ ,  $j$ ,  $s$ , and  $m_j$ , and is given by

$$E_{\ell j s m_j}^{(1)} = \hbar\omega_0 g_j m_j, \quad (141)$$

or in terms of the magnetic field strength

$$E_{\ell j s m_j} = \mu_B B_0 g_j m_j, \quad (142)$$

where we have introduced the fine-structure g-factor,  $g_j = a(n, \ell, s, j)$ , i.e.

$$g_j = \frac{3}{2} + \frac{s(s+1) - \ell(\ell+1)}{2j(j+1)}. \quad (143)$$

## 4.6 Weak-field Zeeman effect

In the weak-field regime,  $\hbar\omega_0 \ll |E_1^{(0)}| \alpha^2 \frac{M_e}{M_p}$ , we start from the unperturbed hyperfine basis  $\{|n\ell s j f m_f\rangle\}$ . At first-order, the Zeeman shift is

$$E_{n\ell s j f m_f}^{(1)} = \hbar\omega_0 \langle n\ell s j f m_j | (L_z + 2S_z) | n\ell s j f m_f \rangle. \quad (144)$$

We are in a subspace of fixed  $n, \ell, j$  and  $s$ , so we can use the Wigner-Eckert theorem to make the replacement

$$L_z + 2S_z \rightarrow a(n, \ell, s, j) J_z, \quad (145)$$

provided we never compute matrix elements between states with different  $j$ . Since we are also in a subspace with fixed  $f$ , we can use Wigner-Eckert again to take

$$J_z \rightarrow a(n, \ell, s, j, f) F_z, \quad (146)$$

where

$$a(n, \ell, s, j, f) = I_{n\ell s j f} \frac{\vec{F} \cdot \vec{J}}{F^2} I_{n\ell s j f}. \quad (147)$$

This means we can use the replacement

$$L_z + 2S_z \rightarrow g_j a(n, \ell, s, j, f) F_z, \quad (148)$$

where  $g_j$  is the fine-structure g-factor.

Using

$$\vec{F} \cdot \vec{J} = J^2 + \vec{J} \cdot \vec{I} = J^2 + \frac{1}{2}(F^2 - J^2 - I^2) = \frac{1}{2}(F^2 + J^2 - I^2), \quad (149)$$

we arrive at

$$a(n, \ell, s, j, f) = \frac{1}{2} + \frac{j(j+1) - i(i+1)}{2f(f+1)}, \quad (150)$$

which leads us to define the hyperfine g-factor as

$$g_f = g_j \frac{1}{2} \left[ 1 + \frac{j(j+1) - i(i+1)}{f(f+1)} \right]. \quad (151)$$

where  $g_j$  is given by Eq. (143).

Thus we conclude that a weak magnetic field will lift the degeneracy of the hyperfine  $f$  levels, splitting each into  $2f + 1$  evenly-spaced sublevels. The Zeeman energy of each sub-level depends only on  $\ell, j, f$ , and  $m_f$ , and is given by

$$E_{\ell j f m_f}^{(1)} = \hbar\omega_0 g_f m_f. \quad (152)$$

In terms of the magnetic field strength this becomes

$$E_{\ell j f m_f}^{(1)} = \mu_B B_0 g_f m_f. \quad (153)$$