1. Introduction

The nucleus of an atom contains localized charge and current distributions, which produce electric and magnetic fields that can be decomposed into multipole fields much as in classical electrostatics or magnetostatics. Table 1 shows some of the lowest order fields that can be produced by nuclei.

Type	Pole	Numbers
Electric	Monopole	$1 = 2^0$
Magnetic	Dipole	$2 = 2^1$
Electric	Quadrupole	$4 = 2^2$
Magnetic	Octupole	$8 = 2^3$

Table 1. The lowest multipole moments that can be produced by a nucleus.

The first of the multipole moments, the electric monopole, is of course the Coulomb electrostatic field that holds the electrons in their orbits and produces the gross structure of the atom. The higher order multipole moments produce small corrections known generally as hyperfine effects.

The terminology "hyperfine" refers to the energy scale of these effects, which is smaller than that of the fine structure. The physics, however, of hyperfine effects is completely different from that of fine structure. In practice, the most important hyperfine effects are those due to the magnetic dipole and electric quadrupole fields of the nucleus. Higher multipole moments of the nucleus are important in nuclear physics, but not usually in atomic physics.

In these notes we will study hyperfine effects in atoms, concentrating mainly on the magnetic dipole field in the case of ordinary hydrogen. As we shall see, hyperfine effects couple the dynamics of the atomic electron to that of the nucleus, thereby enlarging the Hilbert space, introducing new quantum numbers, and shifting and splitting the energy levels.

In spite of the small energy scales involved, hyperfine effects are quite important in applications ranging from atomic physics to astrophysics. For example, the radiative transition between the two hyperfine levels of the ground state of the hydrogen atom produces the 21 centimeter line that is so important in radio astronomy.

For another example, atomic clocks usually use as their basic oscillator a hyperfine transition in the ground state of the heavy alkali atoms, rubidium or cesium. The reproducibility of the frequency of these transitions is so great that the second is now defined in terms of the hyperfine transition in the ^{133}Cs atom, that is, the second is defined as 9,192,631,770 periods of the photon emitted in this transition, exactly. The precision of atomic clocks keeps improving, and for the best clocks currently exceeds one part in 10^{15}.

The GPS (global positioning system) relies on radio timing signals emitted by satellites in high earth orbit containing atomic clocks. In order to achieve the desired precision of position measurements on the earth it is necessary to take into account the effects of both special and general relativity in analyzing the timing signals. Special relativity enters because of the time dilation of the clocks moving in their orbits, and general relativity because of the difference in the gravitational potential between the clock in orbit and the receiver on the surface of the earth (this is the gravitational red shift).

2. About Nuclear Multipole Moments

We denote the spin of the nucleus by $\mathbf{I}$, reserving the symbol $\mathbf{S}$ electron spin. We let the quantum number of the operator I^2 be i, so that I^2 has eigenvalues $i(i+1)\hbar^2$. We denote the nuclear Hilbert space by $\mathcal{E}_{\text{nucl}}$, a $(2i+1)$-dimensional space in which the standard basis is $\{|im_i\rangle, m_i = -i, \ldots, +i\}$. In actual stable nuclei, the spin ranges from $i = 0$ to $i = 15/2$. For example, the proton, the nucleus of ordinary hydrogen, has $i = 1/2$, while the deuteron, the heavier isotope of hydrogen, has $i = 1$. The isotope ^{133}Cs used in atomic clocks has $i = 7/2$.

In classical electrostatics or magnetostatics, fields produced by localized charge or current distributions can be decomposed into multipole fields, with the same sequence of fields (monopole, dipole, quadrupole, etc) in both cases, apart from the fact that the magnetic monopole field is normally considered to be nonexistent. The nomenclature uses a Latin or Greek prefix indicating a power of 2, which is attached to "pole," as seen in Table 1. We shall generally refer to these multipole terms as 2^k-poles.

Not all the multipole fields that occur classically are allowed in the case of a nucleus, however. Table 1 lists the lowest multipole moments of a nucleus that are normally considered to occur. Notice that the electric dipole, magnetic quadrupole, electric octupole, etc, are missing, all of which are allowed classically.

There are two rules governing the allowed multipole moments of the nucleus. The first is that a 2^k-pole can occur only if $k \leq 2i$. For example, the proton with $i = \frac{1}{2}$ can (and does) possess an electric monopole moment and a magnetic dipole moment, but not an electric quadrupole moment. The deuteron with $i = 1$ can (and does) possess an electric quadrupole moment, but the alpha particle with $i = 0$ can possess only the electric monopole moment. There are no hyperfine effects in ^{4}He, or with any other isotope with spin 0 (such as ^{12}C or ^{16}O).

Lying behind this rule is the fact that the operator representing the 2^k-pole on the nuclear

Hilbert space is, in fact, an order k irreducible tensor operator. We have already studied the examples of the magnetic dipole moment and the electric quadrupole moment, which are respectively $k = 1$ and $k = 2$ irreducible tensor operators. In fact, the expansion of classical electric and magnetic fields into multipoles is an example of the decomposition of a space of fields on three-dimensional space into irreducible subspaces under rotations, each of which has an angular momentum value ($\ell = 0$ for monopole, $\ell = 1$ for dipole, etc). The fields in question are classical fields, not quantum wave functions, but the transformation properties under rotations is exactly as presented in Notes 13 and 15.

But the maximum order of an irreducible tensor operator on the nuclear Hilbert space with spin i is $k = 2i$. To see this notice that any operator A that acts on the nuclear Hilbert space can be expanded as a linear combination of the operators $|m_i\rangle\langle m_i'|$, where $|m_i\rangle = |im_i\rangle$ (we suppress the index i, which is constant):

$$A = \sum_{m_i,m_i'} |m_i\rangle\langle m_i|A|m_i'\rangle\langle m_i'| = \sum_{m_i,m_i'} A_{m_i m_i'} |m_i\rangle\langle m_i'|, \tag{1}$$

where

$$A_{m_i m_i'} = \langle m_i|A|m_i'\rangle. \tag{2}$$

Thus, the $(2i + 1)^2$ operators $|m_i\rangle\langle m_i'|$ form a basis in the space of operators acting on the nuclear Hilbert space. But since the kets $|m_i\rangle$ transform under rotations according to the irreducible representation $j = i$, and since the bras $\langle m_i'|$ also transform according to the same irreducible representation, the operator $|m_i\rangle\langle m_i'|$ transforms according to

$$i \otimes i = 0 \oplus \ldots \oplus 2i. \tag{3}$$

Each k in the range $0 \le k \le 2i$ occurs precisely once in this list, so on the nuclear Hilbert space there is precisely one scalar operator (to within a multiplicative constant), one vector, etc, all the way up to one order $k = 2i$ irreducible tensor operator. Equivalently, any operator that acts on the nuclear Hilbert space can be represented as a sum of irreducible tensor operators of order not exceeding $k = 2i$. Operators with $k > 2i$ do not occur. See Prob. 18.2(a).

The second rule states that odd electric 2^k-poles and even magnetic 2^k-poles do not occur. This is because such 2^k-poles violate parity and/or time reversal. In fact, small symmetry violating effects may allow nuclei and other particles such as the electron to posses some of the moments excluded from Table 1. For example, the proton or the electron may have an electric dipole moment, but if such moments exist, they are certainly very small. There is currently considerable experimental interest in lowering the known upper bounds on these "forbidden" moments, or possibly even discovering a nonzero value for some of them, because a numerical value for one of these moments would shed light on physics beyond the standard model.

To understand how the moments that are forbidden in the case of a nucleus or an elementary particle can occur classically, see the discussion in Sec. 14.7 (it has to do with the fact that in the case of the nucleus, we are dealing with a single irreducible subspace under rotations).

3. About Magnetic Dipoles

The only hyperfine effect we shall examine in these notes is the one due to the magnetic dipole moment of the nucleus. It turns out that if the electric quadrupole moment exists, then its effects are of the same order of magnitude as those of the magnetic dipole, so any realistic treatment requires that the two be treated together. In the following we will treat magnetic dipole effects in a single electron atom with nuclear spin i, but to be realistic i should not exceed $\frac{1}{2}$. This condition holds in ordinary hydrogen with the proton as its nucleus, and at a certain point in the calculation we will specialize to that case.

We begin by considering magnetic dipole fields in classical magnetostatics. A point magnetic dipole of moment $\boldsymbol{\mu}$ situated at the origin of the coordinates produces a magnetic field $\mathbf{B} = \nabla \times \mathbf{A}$, where

$$\mathbf{A}(\mathbf{r}) = \frac{\boldsymbol{\mu} \times \mathbf{r}}{r^3} = -\boldsymbol{\mu} \times \nabla\left(\frac{1}{r}\right). \tag{4}$$

The vector potential $\mathbf{A}$ falls off as $1/r^2$ and the magnetic field $\mathbf{B}$ as $1/r^3$ at large r, both with a nontrivial angular dependence. Both fields have a singularity at the origin that must be treated with some care in the quantum mechanical analysis of hyperfine effects.

We will do this by smearing out the point dipole moment over a sphere of small radius a, producing a model in which the dipole is replaced by a uniformly magnetized sphere at the origin. Recall that in magnetostatics the magnetization $\mathbf{M}$ is the dipole moment per unit volume, so we require

$$\boldsymbol{\mu} = V\mathbf{M}, \tag{5}$$

where $\mathbf{M}$ is the constant magnetization inside the sphere and where

$$V = \frac{4}{3}\pi a^3 \tag{6}$$

is the volume of the sphere. This model produces fields that are well behaved everywhere in space. When we are done with the quantum mechanical calculation, we will take $a \to 0$ in such a way that $\boldsymbol{\mu}$ is constant, and we will find physical results that are finite and well behaved. Notice that in this limit, $\mathbf{M} \to \infty$.

The problem of the uniformly magnetized sphere is a standard one in courses on electromagnetism. The vector potential $\mathbf{A}$ at a field point $\mathbf{r}$ can be computed as an integral over the source distribution inside the volume of the sphere,

$$\mathbf{A}(\mathbf{r}) = \int_{\text{vol}} d^3\mathbf{r}' \, \frac{\mathbf{M} \times (\mathbf{r} - \mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|^3} = \mathbf{M} \times \left[-\nabla \int_{\text{vol}} d^3\mathbf{r}' \left(\frac{1}{|\mathbf{r} - \mathbf{r}'|} \right) \right]. \tag{7}$$

The field point $\mathbf{r}$ can be either inside or outside the sphere. The quantity in the square brackets is easily evaluated with some reasoning from electrostatics. The integral is the electrostatic potential produced at field point $\mathbf{r}$ due to a uniformly charged sphere of radius a and charge density $\rho = 1$. Thus, the negative gradient of that integral is the electric field produced by the same charge density.

But by Gauss's law, that electric field is $\hat{\mathbf{r}}/r^2$ times the amount of charge inside radius r. Thus we have

$$-\nabla \int_{\text{vol}} d^3\mathbf{r}'\left(\frac{1}{|\mathbf{r}-\mathbf{r}'|}\right) = \frac{\mathbf{r}}{r^3}\times\begin{cases}\dfrac{4\pi}{3}r^3, & r < a, \\[2mm] \dfrac{4\pi}{3}a^3, & r > a.\end{cases} \tag{8}$$

Combining this with Eqs. (5)–(7), we obtain

$$\mathbf{A}(\mathbf{r}) = \boldsymbol{\mu}\times\mathbf{r}\begin{cases}\dfrac{1}{a^3}, & r < a, \\[2mm] \dfrac{1}{r^3}, & r > a.\end{cases} \tag{9}$$

The exterior solution is identical to that of a point dipole, Eq. (4); in the region $r > a$, one cannot tell from the field alone whether it is produced by a point dipole or a uniformly magnetized sphere of the same strength.

By taking the curl we compute the magnetic field,

$$\mathbf{B}(\mathbf{r}) = \begin{cases}\dfrac{2\boldsymbol{\mu}}{a^3}, & r < a, \\[2mm] \dfrac{1}{r^3}\,\boldsymbol{\mu}\cdot\mathsf{T}, & r > a.\end{cases} \tag{10}$$

where T is the tensor with components

$$T_{ij} = \frac{3r_i r_j - r^2\,\delta_{ij}}{r^2}. \tag{11}$$

The notation $\boldsymbol{\mu}\cdot\mathsf{T}$ means the vector whose j-th component is $\sum_i \mu_i T_{ij}$. Notice that T_{ij} is a symmetric and traceless tensor, something that becomes a $k = 2$ irreducible tensor operator when reinterpreted as a quantum operator. The vector potential (9) is continuous at the radius $r = a$; the magnetic field (10), however, has a discontinuity there, due to a surface current.

To express both $\mathbf{A}$ and $\mathbf{B}$ in single formulas, we introduce the following functions:

$$\Delta(r) = \begin{cases}\dfrac{1}{a^3}, & r < a, \\[2mm] 0, & r > a,\end{cases} \tag{12}$$

and

$$f(r) = \begin{cases}0, & r < a, \\ 1, & r > a.\end{cases} \tag{13}$$

Then we can write

$$\mathbf{A}(\mathbf{r}) = (\boldsymbol{\mu}\times\mathbf{r})\left[\Delta(r) + \frac{f(r)}{r^3}\right], \tag{14}$$

and

$$\mathbf{B}(\mathbf{r}) = \boldsymbol{\mu}\cdot\left[2\Delta(r)\mathsf{I} + \frac{f(r)}{r^3}\mathsf{T}\right], \tag{15}$$

where I is the identity tensor.

When we take the limit $a \to 0$, $\Delta(r)$ becomes a function that is concentrated inside a small volume $(4\pi/3)a^3$ with a value that goes to ∞ in such a way that the integral of $\Delta(r)$ over all space is the constant $4\pi/3$. This is the behavior of a function whose limit is a delta function, so we can write

$$\lim_{a \to 0} \Delta(r) = \frac{4\pi}{3}\delta(\mathbf{r}). \tag{16}$$

As for $f(r)$, it has the limit,

$$\lim_{a \to 0} f(r) = 1. \tag{17}$$

4. The Hamiltonian and the Hilbert Space

We turn now to the atom in which the nucleus resides. In these notes we will be mainly interested in hyperfine effects in hydrogen and alkali atoms, the latter treated as atoms with a single electron (the valence electron) moving in a screened Coulomb potential.

In the following we use atomic units and take $g = 2$ for the g-factor of the electron, so $\mu_B = 1/2c$. The Hamiltonian for the atomic electron is

$$H = \frac{1}{2}\left(\mathbf{p} + \frac{1}{c}\mathbf{A}\right)^2 + V(r) + H_{\mathrm{FS}} + H_{\mathrm{Lamb}} + \frac{1}{c}\mathbf{S} \cdot \mathbf{B}, \tag{18}$$

where $V(r)$ is the central force potential. This could be the Hamiltonian for the Zeeman effect, except that in this case the fields $\mathbf{A}$ and $\mathbf{B}$ are not external fields, rather they are the magnetic dipole fields of the nucleus. Thus, they are given by Eqs. (14) and (15). We include in the Hamiltonian the fine structure terms H_{FS}, the sum of the three terms shown in Eq. (23.4), and terms responsible for the Lamb shift, since we wish to make a realistic treatment of hyperfine effects which are smaller than those just listed. We will not need to know much about the Lamb shift in the following analysis, except that it splits the energy levels in hydrogen to give them a dependence on ℓ.

The expressions (14) and (15) for $\mathbf{A}$ and $\mathbf{B}$ are the fields of a classical magnetic dipole at the origin, but now for use in the Hamiltonian (18) we must reinterpret $\boldsymbol{\mu}$ as an operator acting on the nuclear Hilbert space, given in terms of the nuclear spin by

$$\boldsymbol{\mu} = g_N \mu_N \mathbf{I}, \tag{19}$$

where g_N is the g-factor of the nucleus and μ_N is the nuclear magnetic moment (see Sec. 14.8). Thus, the Hamiltonian (18) must be interpreted as an operator acting the total Hilbert space

$$\mathcal{E} = \mathcal{E}_{\mathrm{elec}} \otimes \mathcal{E}_{\mathrm{nucl}}, \tag{20}$$

where $\mathcal{E}_{\mathrm{elec}}$, the electronic Hilbert space, is the product of its orbital and spin parts,

$$\mathcal{E}_{\mathrm{elec}} = \mathcal{E}_{\mathrm{orb}} \otimes \mathcal{E}_{\mathrm{spin}}. \tag{21}$$

Basis states in $\mathcal{E}$ can be defined as products of basis states in $\mathcal{E}_{\mathrm{elec}}$ and in $\mathcal{E}_{\mathrm{nucl}}$. For $\mathcal{E}_{\mathrm{elec}}$ the obvious basis is the eigenbasis of H_0, $\{|n\ell j m_j\rangle\}$ in the notation of Notes 23, which have energies

$E_{n\ell j}$. In alkalis the energies depend strongly on ℓ, and in hydrogen they depend on ℓ because of the Lamb shift. The obvious basis in $\mathcal{E}_{\mathrm{nucl}}$ is $\{|im_i\rangle\}$. Thus we define the basis states in $\mathcal{E}$,

$$|n\ell jm_j\rangle \otimes |im_i\rangle = |n\ell jm_jm_i\rangle, \tag{22}$$

where we suppress the index i in the shorthand notation on the right hand side, since it is a constant. We will call this the "uncoupled basis," because the electronic angular momentum $\mathbf{J}$ and nuclear angular momentum $\mathbf{I}$ are not coupled. Of course, orbital and spin angular momentum have already been coupled to produce $\mathbf{J}$, so the basis (22) is really only partially uncoupled. These uncoupled basis states are eigenstates of H_0 on the whole Hilbert space $\mathcal{E}$. The energy does not depend on the quantum numbers m_j or m_i, so the energy eigenstates are $(2j+1)(2i+1)$-fold degenerate.

Now we expand the kinetic energy in the Hamiltonian (18) as in Notes 24 and neglect the term in A^2, writing the result as $H = H_0 + H_1$, where

$$H_0 = \frac{p^2}{2} + V(r) + H_{\mathrm{FS}} + H_{\mathrm{Lamb}}, \tag{23}$$

and

$$H_1 = \frac{1}{c}(\mathbf{p} \cdot \mathbf{A} + \mathbf{S} \cdot \mathbf{B}). \tag{24}$$

Here we have used the fact that our $\mathbf{A}$ is in Coulomb gauge, $\nabla \cdot \mathbf{A} = 0$, so $\mathbf{p} \cdot \mathbf{A} = \mathbf{A} \cdot \mathbf{p}$. We write the two terms in H_1 as $H_{1,\mathrm{orb}}$ and $H_{1,\mathrm{spin}}$. For $H_{1,\mathrm{orb}}$ we use $1/c = 2\mu_B$ and Eqs. (19) and (14), rearranging the triple product,

$$\mathbf{p} \cdot (\mathbf{I}{\times}\mathbf{r}) = \mathbf{I} \cdot (\mathbf{r}{\times}\mathbf{p}) = \mathbf{I} \cdot \mathbf{L}. \tag{25}$$

This gives

$$H_{1,\mathrm{orb}} = \frac{1}{c}\mathbf{p} \cdot \mathbf{A} = k(\mathbf{I} \cdot \mathbf{L})\left[\Delta(r) + \frac{f(r)}{r^3}\right], \tag{26}$$

where we have set

$$k = 2g_N\mu_B\mu_N = g_e g_N\mu_B\mu_N \tag{27}$$

for the product of the effective magnetic moments of the electron and the nucleus (including g-factors). As for $H_{1,\mathrm{spin}}$, we rearrange it similarly, obtaining

$$H_{1,\mathrm{spin}} = \frac{1}{c}\mathbf{S} \cdot \mathbf{B} = k\left[2\Delta(r)(\mathbf{I} \cdot \mathbf{S}) + \frac{f(r)}{r^3}(\mathbf{I} \cdot \mathsf{T} \cdot \mathbf{S})\right]. \tag{28}$$

Notice that $H_{1,\mathrm{orb}}$ is a kind of spin-orbit interaction (but it is the spin of the nucleus, not the spin of the electron), and $H_{1,\mathrm{spin}}$ is a kind of spin-spin interaction.

Properly we should only take the limit $a \to 0$ after we have formed matrix elements, but in many books this limit is taken before, resulting in the formulas,

$$H_{1,\mathrm{orb}} = k(\mathbf{I} \cdot \mathbf{L})\left[\frac{4\pi}{3}\delta(\mathbf{r}) + \frac{1}{r^3}\right], \tag{29}$$

and

$$H_{1,\mathrm{spin}} = k\left[\frac{8\pi}{3}\delta(\mathbf{r})(\mathbf{I} \cdot \mathbf{S}) + \frac{1}{r^3}(\mathbf{I} \cdot \mathsf{T} \cdot \mathbf{S})\right]. \tag{30}$$

The terms involving $\delta(\mathbf{r})$ are called "Fermi contact terms," because they vanish except when the electron and nucleus are in contact with one another ($r = 0$), and in honor of Fermi, who first carried out the calculation presented in this set of notes.

5. The Perturbation Calculation

Both terms (26) and (28) in the perturbing Hamiltonian have the form of $\mathbf{I}$ dotted into a vector that is constructed out of electronic operators, that is, operators that act only the electronic Hilbert space. Since the angular momentum $\mathbf{J}$ generates rotations that rotate electronic vectors, while $\mathbf{I}$ generates rotations that rotate $\mathbf{I}$ itself, the dot products in H_1 are not invariant under either electronic rotations alone or under nuclear rotations alone. For this reason, H_1 does not commute with either J_z or I_z, and the uncoupled basis (22) is not the best one for carrying out the perturbation calculation (see Sec. 23.5).

The dot products in question, however, are invariant under total rotations of the system, electronic plus nuclear, which are generated by the total angular momentum of the system defined by

$$\mathbf{F} = \mathbf{J} + \mathbf{I} = \mathbf{L} + \mathbf{S} + \mathbf{I}. \tag{31}$$

This suggests that we couple together $\mathbf{J}$ and $\mathbf{I}$ to create eigenstates of F^2 and F_z. We will call the result the "coupled basis"; it is given by

$$|n\ell j f m_f\rangle = \sum_{m_j, m_i} |n\ell j m_j m_i\rangle \langle j i m_j m_i | f m_f\rangle, \tag{32}$$

where the final scalar products are the Clebsch-Gordan coefficients.

In the coupled basis the matrix elements we need to consider for degenerate perturbation theory are

$$\langle n\ell j f m_f | H_1 | n\ell j f' m_f' \rangle, \tag{33}$$

where the unprimed indices on the two sides specify the degenerate unperturbed energy level $E_{n\ell j}$, while the remaining indices specify the basis inside the degenerate eigenspace of the unperturbed system. But since $[\mathbf{F}, H_1] = 0$, and since if H_1 commutes with $\mathbf{F}$ it commutes with any function of $\mathbf{F}$ such as F^2, the matrix (33) is diagonal in both f and m_f, and the primes can be dropped. We need only compute the diagonal matrix elements to find the energy shifts, and once again we have succeeded in doing a degenerate perturbation calculation without diagonalizing any matrices. Thus we have

$$\Delta E = \langle n\ell j f m_f | H_1 | n\ell j f m_f \rangle. \tag{34}$$

We now evaluate these matrix elements for the case $\ell \neq 0$. In this case the contact terms do not contribute in the limit $a \to 0$, because the wave functions go as r^ℓ near $r = 0$, and thus vanish at $r = 0$ when $\ell \neq 0$. As for the other terms, we take the limit $a \to 0$ so $f(r) = 1$, and write the result in the form

$$\Delta E = k \langle n\ell j f m_f | \frac{1}{r^3} \mathbf{I} \cdot \mathbf{G} | n\ell j f m_f \rangle, \tag{35}$$

where

$$\mathbf{G} = \mathbf{L} + \mathsf{T} \cdot \mathbf{S} = \mathbf{L} - \mathbf{S} + \frac{3\mathbf{r}(\mathbf{r} \cdot \mathbf{S})}{r^2}, \tag{36}$$

where we use Eq. (11). Note that $\mathbf{G}$ is a purely electronic vector operator.

We simplify Eq. (35) by using the projection theorem (see Sec. 24.7 and Prob. 24.1), which tells us that a vector operator such as $\mathbf{G}$, sandwiched between eigenstates of J^2 with the same j value on both sides, can be replaced by an operator proportional to $\mathbf{J}$:

$$\mathbf{G} \to \frac{(\mathbf{G} \cdot \mathbf{J})\mathbf{J}}{j(j+1)}. \tag{37}$$

Thus

$$\Delta E = \frac{k}{j(j+1)} \langle n\ell j f m_f | \frac{1}{r^3} (\mathbf{I} \cdot \mathbf{J})(\mathbf{J} \cdot \mathbf{G}) | n\ell j f m_f \rangle. \tag{38}$$

Now the two dot products can be simplified. First, by Eq. (31), we have

$$\mathbf{I} \cdot \mathbf{J} = \frac{1}{2}(F^2 - J^2 - I^2). \tag{39}$$

Next, we use $\mathbf{r} \cdot \mathbf{J} = \mathbf{r} \cdot \mathbf{S}$ (since $\mathbf{r} \cdot \mathbf{L} = 0$) to write

$$\mathbf{G} \cdot \mathbf{J} = L^2 - S^2 + \frac{3(\mathbf{S} \cdot \mathbf{r})^2}{r^2}. \tag{40}$$

Then, since $\mathbf{S} = (1/2)\boldsymbol{\sigma}$ and since $\sigma_i \sigma_j = \delta_{ij} + i\epsilon_{ijk}\sigma_k$, the final term in Eq. (40) becomes

$$\frac{3(\mathbf{S} \cdot \mathbf{r})^2}{r^2} = \frac{3}{4}, \tag{41}$$

which cancels $S^2 = 3/4$. The result is simply

$$\mathbf{G} \cdot \mathbf{J} = L^2. \tag{42}$$

Then, with Eqs. (39) and (42), the energy shift (38) becomes

$$\Delta E = k \frac{f(f+1) - j(j+1) - i(i+1)}{2j(j+1)} \ell(\ell+1) \left\langle \frac{1}{r^3} \right\rangle, \tag{43}$$

where the final expectation value of $1/r^3$ can be reduced to a purely spatial matrix element with respect to the state $|n\ell 0\rangle$ exactly as in Sec. 23.7. Specializing to hydrogen and using Eqs. (27) and (23.33), we find

$$\Delta E = \frac{g_e g_N \mu_B \mu_N}{a_0^3} \frac{1}{n^3} \frac{f(f+1) - j(j+1) - i(i+1)}{j(j+1)(2\ell+1)}, \tag{44}$$

where we have introduced a factor of $1/a_0^3$ (a_0 is the Bohr radius) to make the answer have dimensions of energy (thus it is now valid in ordinary units).

Equation (44) is the final expression for the hyperfine structure energy shift in hydrogen in the case $\ell \neq 0$. However, one can show that the same answer holds also in the case $\ell = 0$, so Eq. (44) applies in all cases.

6. The New Energy Levels

The energy levels were $E_{n\ell j}$ before the hyperfine interactions were turned on, but since ΔE in Eq. (44) depends on f, they now have the form $E_{n\ell jf}$. The energy eigenstates are $|n\ell jfm_f\rangle$, and are $(2f+1)$-fold degenerate, since the energy does not depend on m_f. This is precisely the situation expected of a generic system invariant under rotations; the degeneracy present is explained by rotational invariance alone.

The energy shift (44) causes the fine structure levels in the Dirac picture of hydrogen to split, giving rise to hyperfine multiplets. For example, the ground state $1s_{1/2}$ splits into two levels $f = 0$ and $f = 1$, of which $f = 0$ is lower since by Eq. (44) the energies are an increasing function of f. This $f = 0$ level is the true ground state of hydrogen. It is nondegenerate. The $f = 1$ level is 3-fold degenerate, and lies above the ground state by an energy of approximately 1.42 GHz in frequency units, or 21 cm in wave length units, or 0.068 K in temperature units. Since $\ell = 0$ for these states, the total angular momentum $\mathbf{F}$ is really just the total spin $\mathbf{S} + \mathbf{I}$, and the $f = 0$ and $f = 1$ states are the spin singlet and triplet states, respectively. The electron and proton spins are antiparallel in the ground (singlet) state; when one of the spins is flipped to make them parallel (the triplet state), the energy is raised.

Similarly, other $j = 1/2$ states including the $2s_{1/2}$ and $2p_{1/2}$ states split into an $f = 0$ and $f = 1$ pair. The $2p_{3/2}$ state in the Dirac picture splits into an $f = 1$ and $f = 2$ pair.

Electric dipole transitions between the various hyperfine levels are governed by the matrix element,

$$\langle n\ell jfm_f|r_q|n'\ell'j'f'm'_f\rangle. \tag{45}$$

The selection rules for this matrix element follow from the Wigner-Eckart theorem and parity. The Wigner-Eckart theorem can be used three times, once for each of the three angular momenta $\mathbf{L}$, $\mathbf{J}$ and $\mathbf{F}$, since r_q is a $k = 1$ irreducible tensor operator with respect to rotations generated by any of these angular momenta. Thus, the matrix element (45) vanishes unless the following selection rules are satisfied:

$$m_f = m'_f + q,$$
$$\Delta f = 0, \pm 1 \quad \text{but } f = 0 \to f = 0 \text{ not allowed,}$$
$$\Delta j = 0, \pm 1,$$
$$\Delta \ell = \pm 1. \tag{46}$$

The exclusion of $f = 0 \to f = 0$ comes from the fact that $0 \otimes 1 = 1$, so if $f = f' = 0$, then the matrix element vanishes. The only reason we do not also exclude $j = 0 \to j = 0$ is that j is half-integral, and cannot take on the value 0. And the only reason we do not make a special case of excluding $\ell = 0 \to \ell = 0$ is that that case is already excluded by parity.

The hyperfine splitting of the $1s_{1/2}$ level in hydrogen is particularly interesting. The separation between the $f = 0$ and $f = 1$ hyperfine levels is 1.42 GHz, or 21 cm in wavelength units. Electric

dipole transitions between these two levels are forbidden by parity, but magnetic dipole transitions are allowed. (One can see this from another standpoint: since spins do not interact with electric fields, an electric dipole transition cannot flip the spins to convert the triplet to the singlet state. But magnetic dipole transitions can.)

The 21 cm line is quite important in radio astronomy. Spiral galaxies typically possess large clouds of atomic hydrogen, which radiate at the 21 cm wavelength. A population of the excited state $f = 1$ is maintained by collisions; the temperatures prevalent in the clouds are high enough that the populations of the ground state $f = 0$ and the first excited state $f = 1$ are determined mostly by the degeneracies (1 for $f = 0$ and 3 for $f = 1$, although there is some effect due to the Boltzmann factor). By measuring Doppler shifts, the state of motion of the clouds can be measured. In this way, it was first proven that the Milky Way is a spiral galaxy. The 21 cm line is also important in absorption spectra, which can be used to determine the temperature of the clouds of atomic hydrogen.

Molecular hydrogen has a completely different hyperfine structure from atomic hydrogen, arising from the spin-spin interaction of the two protons in the molecule. The transitions between the hyperfine levels of molecular hydrogen are in the megahertz range of frequencies.

Problems

1. A problem on the hyperfine interaction in hydrogen.

(a) Equation (44) was derived in the case $\ell \neq 0$. Show that it also applies in the case $\ell = 0$. Hint: Use the fact that the components of the tensor T_{ij}, defined in Eq. (25.11), are r^2 times linear combinations of the $Y_{2m}(\theta, \phi)$, for $m = -2, \ldots, +2$. This is related to the fact that T_{ij} is the Cartesian version of an order 2 irreducible tensor.

(b) Our analysis of the hyperfine interaction in hydrogen has included the energy of interaction of the electron with the magnetic dipole field produced by the proton, but it seems that we have not included the energy of interaction of the proton spin with the magnetic field produced by the electron. As seen by the proton, the electron produces a magnetic field for two reasons: first, it is a charge in motion, therefore a current, which makes a magnetic field. This is the magnetic field due to the orbital motion of the electron. Next, the electron has a magnetic moment of its own, which makes a dipole magnetic field. This is the magnetic field produced by the spin of the electron.

Work out an expression for the energy of interaction of the proton spin with the magnetic field produced by the orbital motion of the electron. Follow the analysis of the spin-orbit interaction in Sec. 23.3, but run it backwards. That is, putting primes on the fields in the electron rest frame and no primes on fields in the proton rest frame, use Coulomb's law to write down the field $\mathbf{E}'$ of the electron in its own rest frame, then Lorentz transform to the lab frame to get $\mathbf{B}$ (call this $\mathbf{B}_{\text{orb}}$, the magnetic field due to the orbital motion of the electron). Then the energy of interaction of the proton with this magnetic field is $-\boldsymbol{\mu}_p \cdot \mathbf{B}_{\text{orb}}$, where $\boldsymbol{\mu}_p$ is the proton magnetic moment. Notice that

unlike the analysis of Sec. 23.3, there is no factor of $\frac{1}{2}$ from Thomas precession, because the proton frame is not accelerated.

Now use Eq. (15) in the limit $a \to 0$ to obtain the magnetic field produced by the dipole moment of the electron at the position of the proton. Call this $\mathbf{B}_{\text{spin}}$, and write down an expression for the energy of interaction $-\boldsymbol{\mu}_p \cdot \mathbf{B}_{\text{spin}}$.

If you add these terms to the Hamiltonian (Eqs. (23) plus (24)), does it change the energy shifts (44)? These energy shifts are confirmed experimentally (for example, by the 21 cm line). What is wrong?

(c) Compute the hyperfine splitting of the ground state of positronium in wavelength units. Notice that in positronium, the fine structure and hyperfine structure are of the same order of magnitude.

Now some remarks about part (c). The interesting thing about this calculation is that the answer based on what you now know is actually wrong, because it omits a virtual process (a Feynman diagram) in which the positron and electron annihilate into a photon, which then materialize back into a positron and electron.

The analysis of this process requires quantum field theory. The Hamiltonians we usually use in atomic, molecular and solid state physics, expressed in terms of a finite number of particles and a finite number of degrees of freedom, are only valid up to a certain degree of accuracy, beyond which interactions with the infinite degrees of freedom in various fields (electromagnetic, electron-positron, strong interactions, ...) cannot be ignored. The first place where this occurs in hydrogen is with the Lamb shift. In positronium, it happens at the level of the fine structure (in positronium, the hyperfine structure is considered part of the fine structure).