1
ATOMIC STRUCTURE

1.1 Ground state of phosphorus

One of the most important topics in atomic physics is the description of atomic energy levels. The study of atomic structure continues to be an exciting field, with increasingly precise measurements and improved calculational tools allowing ever more detailed comparisons between experiment and theory.

The first few problems in this chapter deal with some of the basic features of atomic energy levels in multi-electron atoms. In the simplest atom, hydrogen, most of the splitting between energy levels comes from the difference between the principal quantum numbers n for different states – the energy E_n of an electron in the hydrogen atom is given, approximately, by the famous Bohr formula,

$$E_n \approx -\frac{me^4}{2\hbar^2}\frac{1}{n^2}\,, \tag{1.1}$$

where m is the mass of the electron and e is the absolute value of its charge. As a first approximation, in more complex atoms we can consider the states of individual electrons as if they were moving in an effective centrally symmetric field created by the nucleus and the other electrons (the *central field approximation*). In this case we can assign a principal quantum number n and an orbital angular momentum l to each electron (the distribution of the electrons among the states with various n's and l's is known as the *electron configuration*), and, as in hydrogen, the differences between the principal quantum numbers for different configurations are an important source of energy level splittings. In contrast to hydrogen, however, the energy of a particular configuration for a multi-electron atom is also dependent on l. This is because electrons with larger l values are, on average, further from the nucleus due to the centrifugal barrier, and the other electrons screen the nuclear charge. Based on these two general ideas, one expects that configurations for which electrons have the lowest possible values of n and l have the lowest energies. For s and p orbitals, it is most important to have the smallest value of n ("regular" configurations). However, one finds that in some cases, for

d and f orbitals it can be energetically favorable for electrons to have higher n so that they can occupy states with lower l ("irregular" configurations[1]).

Next one must consider in more detail the mutual electrostatic repulsion between the electrons. Already, the central field approximation has accounted for the spherically symmetric part of the potential due to the electron-electron interaction which is responsible for the screening of the nucleus. There is also a nonspherically symmetric part of the electron-electron potential related to the fact that it is energetically favorable for electrons to be as far apart as possible.

For a given configuration, the atomic states can be specified by the total orbital angular momentum $\vec{L} = \sum \vec{l}_i$ and the total spin $\vec{S} = \sum \vec{s}_i$ (this characterization of atomic states, valid for low to intermediate Z atoms, is known as the *Russell-Saunders* or *L-S* coupling scheme[2]). Since the average separation between electrons are different for states with different L and S, these states are split in energy (see Problem 1.2). The states of a given configuration identified with particular L and S are known as the *term*.

There are empirical rules, known as *Hund's rules* [see, for example, Bransden and Joachain (2003), Landau and Lifshitz (1977), or Herzberg (1944)], for determining which term has the lowest energy for a given configuration (in the Russell-Saunders coupling scheme). Hund's rules state that

- the energy of a term decreases with increasing S (so the term with the largest S has the lowest energy), and
- for a given S, the larger the total orbital angular momentum L the lower the energy.

These rules ensure that the ground state electrons are, on average, as far apart as possible, which minimizes the electrostatic repulsion between them.

For both single and multi-electron atoms, there is also the spin-orbit interaction (see Problem 1.3) which causes splitting of states with different values of the total electronic angular momentum J (spin S coupled to orbital L). Since this splitting is typically considerably smaller than the energy differences from the previously discussed mechanisms, it is known as the *fine-structure* splitting.

[1] An example of such an irregular configuration is the ground state of potassium, which is $1s^2 2s^2 2p^6 3s^2 3p^6 4s$ instead of $1s^2 2s^2 2p^6 3s^2 3p^6 3d$.

[2] In atoms with large Z, the spin-orbit energy (arising from relativistic effects) can become more important than the residual, nonspherical electrostatic interaction between the electrons. In such a case, it is useful to specify individual electron states by their total angular momentum $\vec{j} = \vec{l} + \vec{s}$. This is known as the j-j coupling scheme. In each coupling scheme, we are trying to begin our considerations with what are nearly the energy eigenfunctions for the atom, but in the general case neither L-S nor j-j coupling give the correct energy eigenfunctions.

Typically, the L, S, and J for a particular state are designated in *spectroscopic notation* (see, for example, Appendix C):

$$^{2S+1}L_J \ . \tag{1.2}$$

In this problem, we consider all of the aforementioned interactions in order to determine the energy level structure for the terms corresponding to the ground state configuration of phosphorus (P).

(a) What is the ground state configuration of P, $Z = 15$?

(b) What is the ground term and J value for P according to Hund's rules?

(c) What other terms are possible for the ground state configuration? Which of the terms has the highest energy?

(d) For the term with the highest energy corresponding to the ground state configuration, identified in part (c), can one say what value of J has the highest energy using first-order perturbation theory?

Hints

In part (b), we can make use of the fact that electrons in the filled subshells $(1s, 2s, 2p, 3s)$ have total spin and total orbital angular momentum equal to zero, and so we need only consider the three outer electrons in the $3p$ orbital to determine the ground term.

For the ground term, first consider the maximum spin and maximum projection of the spin along the quantization axis for three electrons (according to Hund's first rule, this should be S for the ground term). What must be the orbital wavefunction in this case?

Part (d) is tricky! One can think of a subshell (which can hold a total of $2(2l + 1)$ electrons) containing N electrons as consisting of N electrons *or* $2(2l + 1) - N$ holes. The spin-orbit splitting (to first order) has opposite sign for electrons and holes.

Solution

(a) Phosphorus has sufficiently low Z that each shell fills regularly (there are no d or f states involved). A subshell can contain $2(2l + 1)$ electrons, where l is the orbital angular momentum quantum number for the individual electrons.

Therefore, the ground state configuration is

$$\boxed{1s^2 2s^2 2p^6 3s^2 3p^3.} \tag{1.3}$$

(b) The maximum total spin that three electrons can possess is $S = 3/2$. The $S = 3/2$ spin wavefunctions are symmetric under particle interchange. This can be seen from the fact that the stretched state $|S, M_s = S\rangle$ is obviously symmetric

$$|3/2, 3/2\rangle = |+\rangle_1 |+\rangle_2 |+\rangle_3 \,, \tag{1.4}$$

and all the other $S = 3/2$ states can be obtained by application of the lowering operator $S_- = S_{1-} + S_{2-} + S_{3-}$, which is symmetric and so cannot change the exchange symmetry of the spin states.

To satisfy the *spin-statistics theorem*,[3] we must choose a spatial wavefunction which is totally antisymmetric with respect to particle interchange. With three p electrons, we have $l_1 = 1$, $l_2 = 1$, and $l_3 = 1$, in principle allowing $L = 3, 2, 1$ and 0, where L is the total angular momentum of all three valence electrons. However, in order to construct a totally antisymmetric wavefunction, one needs at least as many states as there are particles. This can be seen by considering the *Slater determinant* [see, for example, Bransden and Joachain (2003)], which provides a simple method for finding a totally antisymmetric wavefunction for a system of particles. In this case we have three particles in, let us say, states α, β, and γ. Then the totally antisymmetric wavefunction is given by:

$$\Psi_{AS} = \frac{1}{\sqrt{3!}} \begin{vmatrix} \alpha(1) & \beta(1) & \gamma(1) \\ \alpha(2) & \beta(2) & \gamma(2) \\ \alpha(3) & \beta(3) & \gamma(3) \end{vmatrix} . \tag{1.5}$$

This wavefunction is clearly exchange antisymmetric, since interchanging two particles is equivalent to interchanging two rows of the matrix, which changes the sign of the determinant. Furthermore, the determinant vanishes if two columns are the same, which occurs if any two of the states α, β, or γ are the same. Thus the *Pauli exclusion principle* is seen to be a consequence of the fact that electrons must be in exchange antisymmetric states.

Therefore, for $S = 3/2$, the allowed spatial wavefunction $|\psi_{\text{space}}\rangle$ must involve a superposition of electrons in the $m_l = 1$, $m_l = 0$, and $m_l = -1$ states (denoted

[3] There is an important distinction between the *symmetrization postulate* and the *spin-statistics theorem*. The symmetrization postulate states that any wavefunction for a system of identical particles must be either exchange symmetric or exchange antisymmetric (i.e., the wavefunction must be an eigenstate of the permutation operator). The spin-statistics theorem states that integer spin particles (bosons) must be in an exchange symmetric state and half-integer spin particles (fermions) are in exchange antisymmetric states.

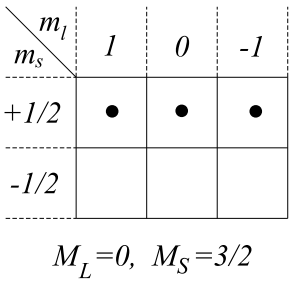

FIG. 1.1 A simple way to figure out the ground state term according to Hund's rules for the case of phosphorus.

$|1\rangle, |0\rangle, |-1\rangle$), namely [using (1.5)],

$$|\psi_{\text{space}}\rangle = \frac{1}{\sqrt{6}}\Big[|1\rangle_a|0\rangle_b|-1\rangle_c + |0\rangle_a|-1\rangle_b|1\rangle_c + |-1\rangle_a|1\rangle_b|0\rangle_c - |-1\rangle_a|0\rangle_b|1\rangle_c - |1\rangle_a|-1\rangle_b|0\rangle_c - |0\rangle_a|1\rangle_b|-1\rangle_c\Big] , \tag{1.6}$$

so that the total projection of the orbital angular momentum along the z-axis is $M_L = 0$. Thus the only possible value for the orbital angular momentum is $L = 0$.

Here is another simple, graphical way to arrive at this conclusion. Consider the chart shown in Fig. 1.1. Each box corresponds to an individual electron state with unique quantum numbers. We want to think about the stretched state with maximum S. The Pauli exclusion principle demands that we can only put one electron in each box (since otherwise we could not form an exchange antisymmetric state). We put all electrons in $m_s = 1/2$ boxes to maximize M_S (and hence S). Then the maximum projection of M_L consistent with $M_S = 3/2$ is zero. This tells us that $L = 0$ for the ground state of phosphorus.

The final step is to figure out what J should be. Luckily, there is only one choice in this case, $J = 3/2$. Consequently, the ground state for phosphorus is

$$\boxed{{}^4S_{3/2}\ .} \tag{1.7}$$

(c) As we have seen in part (b), ${}^4S_{3/2}$ is the only possible term when the electrons have total spin $S = 3/2$. But three electrons can also have total spin $S = 1/2$, so there are additional, higher-energy terms possible for the given configuration. Unlike the $S = 3/2$ spin wavefunctions which are exchange symmetric, the $S = 1/2$ spin wavefunctions possess no definite exchange symmetry, so the approach we used in part (b) in which we first considered the exchange symmetry of the spin wavefunctions and then the exchange symmetry of the orbital wavefunctions fails. In order to construct totally antisymmetric wavefunctions for the three electrons

TABLE 1.1 All possible single particle states for the ground state configuration of phosphorus grouped according to M_L and M_S. For completeness, we have written out all the states, but because of the symmetry between the M_L and $-M_L$ states and the M_S and $-M_S$ states, only one corner of the chart is actually necessary for specifying all the states.

M_S	$M_L = 2$	$M_L = 1$	$M_L = 0$	$M_L = -1$	$M_L = -2$
$+\frac{3}{2}$			$(1^+)(0^+)(-1^+)$		
$+\frac{1}{2}$	$(1^+)(0^+)(1^-)$	$(1^+)(-1^+)(1^-)$ $(1^+)(0^+)(0^-)$	$(-1^+)(0^+)(1^-)$ $(-1^+)(1^+)(0^-)$ $(1^+)(0^+)(-1^-)$	$(1^+)(-1^+)(-1^-)$ $(1^+)(-1^+)(-1^-)$	$(-1^+)(0^+)(-1^-)$
$-\frac{1}{2}$	$(1^+)(0^-)(1^-)$	$(1^+)(-1^-)(1^-)$ $(0^+)(1^-)(0^-)$	$(-1^+)(0^-)(1^-)$ $(1^+)(-1^-)(0^-)$ $(0^+)(1^-)(-1^-)$	$(-1^+)(1^-)(-1^-)$ $(-1^+)(1^-)(-1^-)$	$(-1^+)(0^-)(-1^-)$
$-\frac{3}{2}$			$(1^-)(0^-)(-1^-)$		

for these higher-energy terms, we must consider products of spatial and spin states for each individual electron.

As pointed out in the discussion surrounding Eq. (1.5), in order to construct a totally antisymmetric wavefunction, one needs at least as many individual electron states as there are electrons. To determine the additional terms, we can write out all the possible states using the shorthand notation

$$(m_l^{m_s})_a(m_l^{m_s})_b(m_l^{m_s})_c \ ,$$

which actually refers to the completely antisymmetric wavefunction built from the three states according to Eq. (1.5). To satisfy the Pauli exclusion principle, we require that no two of the states for the three electrons are the same. This notation is equivalent to the chart representation described in part (b) – the chart shown in Fig. 1.1 corresponds to $(1^+)(0^+)(-1^+)$. Now we can make a table of all possible states grouped together by the projection of their total orbital angular momentum M_L and the projection of their total spin M_S (Table 1.1).

The states described in Table 1.1 are eigenstates of the operators L_z and S_z, and by forming appropriate linear combinations one can construct eigenstates of L^2, S^2, L_z, and S_z. The states corresponding to a given term are eigenstates of the operators $\left\{L^2, S^2, J^2, J_z\right\}$. Since both sets of eigenstates form a complete basis for our system, we know that there must be the same number of eigenstates in each set. To figure out the additional terms possible for the ground state configuration, we begin with a stretched state ($M_L = L$ and $M_S = S$) for a particular term, and count off how many states are in that term. We continue this process until we account for all 20 states in Table 1.1.

Let us begin with the term we already know about, ${}^4S$. There are four states in this term (since $J = 3/2$), all with $M_L = 0$, so that accounts for four states in the

$M_L = 0$ column of our table. Next, we see that there is one state $[(1^+)(0^+)(1^-)]$ with $M_L = 2$ and $M_S = 1/2$. This is the stretched state of a ^2D term. The possible values of the total electronic angular momentum for the ^2D term are $J = 5/2$ and $J = 3/2$, so the ^2D term accounts for 10 states (six with $J = 5/2$ and four with $J = 3/2$), two in each M_L column, one corresponding to $M_S = 1/2$ and the other with $M_S = -1/2$ (since $S = 1/2$ for the ^2D term).

There is another state with $M_L = 1$ and $M_S = 1/2$ not accounted for by the ^4S or ^2D terms. This is the stretched state (the state with the maximum possible projection of all angular momenta along the quantization axis) for a ^2P term, which accounts for the six remaining states (for the ^2P term we have $J = 3/2$ and $J = 1/2$). This covers all the terms:

$$\boxed{^4S,\ ^2D,\ \text{and}\ ^2P\ .} \tag{1.8}$$

As for which term in this configuration has the highest energy, we can use arguments based on Hund's rules. The highest energy term should be the term with the smallest value of S and the smallest value of L, which is the ^2P term.

(d) As mentioned in the hint, to first order the spin-orbit interaction has opposite sign for electrons and holes. Since the p^3 configuration can be thought of as consisting of three holes or three electrons, we have that the energy splitting due to the spin-orbit interaction ΔE_{LS} satisfies

$$\Delta E_{LS} \approx -\Delta E_{LS}\ , \tag{1.9}$$

so

$$\boxed{\Delta E_{LS} \approx 0\ .} \tag{1.10}$$

To first order, the $^2P_{1/2}$ and the $^2P_{3/2}$ states have the same energy.

1.2 Exchange interaction

In the nonrelativistic limit, the Hamiltonian describing the interaction of electrons with the nucleus is independent of the electron and nuclear spins. Nonetheless, as we have seen in Problem 1.1, energy levels in multi-electron atoms do depend on the spin state of the electrons. Consider a two-electron atom. The nonrelativistic

Hamiltonian H is given by:

$$H = H_0 + H_1 \ , \tag{1.11}$$

where

$$H_0 = -\sum_{i=1}^{2} \left(\frac{\hbar^2}{2m_e} \nabla_{r_i}^2 + \frac{Ze^2}{r_i} \right) \tag{1.12}$$

and

$$H_1 = \frac{e^2}{|\vec{r}_1 - \vec{r}_2|} \ . \tag{1.13}$$

Here H_0 contains the Coulomb attraction of the electrons to the nucleus and H_1 describes the Coulomb repulsion between the electrons, where $\vec{r}_i$ is the position of the i-th electron, and $|\vec{r}_1 - \vec{r}_2|$ is the distance between the electrons.

The overall electron wavefunction Ψ is the product of a spatial wavefunction ψ and a spin function χ. Fermi statistics demands that Ψ must be antisymmetric with respect to particle interchange for identical particles with half-integer spin. Thus if χ is a triplet state (symmetric), then ψ must be antisymmetric; if χ is a singlet state (antisymmetric), then ψ must be symmetric. Thus the electron spin state dictates the symmetry of the spatial part of the wavefunction. It turns out that due to the Coulomb repulsion between the electrons (H_1), symmetric and antisymmetric spatial wavefunctions have different energies – this is called the *exchange interaction* [see, for example, Griffiths (1995) or Landau and Lifshitz (1977)].

(a) For the case where one electron is in the ground state and another electron is in an excited state with quantum numbers (n, l, m_l), which spin state, the triplet or singlet, has higher energy? How does this relate to Hund's rules (discussed in Problem 1.1)?

(b) As a simple illustration of the exchange interaction, consider two electrons in a one-dimensional simple harmonic oscillator (SHO) potential. For the case where one of the electrons is in the ground state and the other is in the first excited state, calculate $\langle (x_2 - x_1)^2 \rangle$ for the triplet and singlet spin states, where x_1 and x_2 are the positions of the two electrons.

Solution

(a) Consider the case where one electron is in the ground state $\psi_{100}(\vec{r}_1)$ and another electron is in an excited state $\psi_{nlm_l}(\vec{r}_2)$. The symmetric ψ_s and antisymmetric ψ_a spatial wavefunctions are given by:

$$\psi_s = \frac{1}{\sqrt{2}}[\psi_{100}(\vec{r}_1)\cdot\psi_{nlm_l}(\vec{r}_2) + \psi_{100}(\vec{r}_2)\cdot\psi_{nlm_l}(\vec{r}_1)]$$
$$\psi_a = \frac{1}{\sqrt{2}}[\psi_{100}(\vec{r}_1)\cdot\psi_{nlm_l}(\vec{r}_2) - \psi_{100}(\vec{r}_2)\cdot\psi_{nlm_l}(\vec{r}_1)]\ . \tag{1.14}$$

If the electrons are in ψ_a, they can never occupy the same position since if $\vec{r}_1 = \vec{r}_2$ then $\psi_a = 0$. However, $\psi_s \neq 0$ if $\vec{r}_1 = \vec{r}_2$, so it is possible for the two electrons to be located at the same point. Thus it turns out that, on average, electrons in ψ_s are closer to one another than electrons in ψ_a. Therefore H_1 causes a state with a symmetric spatial wavefunction to have a higher energy than one with an antisymmetric wavefunction, so the spin singlet states (χ antisymmetric $\rightarrow$ ψ symmetric) have higher energies than the triplet ones (χ symmetric $\rightarrow$ ψ antisymmetric). Based on this reasoning, we can plausibly argue that, in general, terms with higher total spin S have lower energy since they have more "antisymmetrical" spatial wavefunctions, leading to less spatial overlap of the electron wavefunctions. This is the basis for Hund's first rule.

(b) We can illustrate this principle more explicitly by considering the expectation value $\langle(x_2 - x_1)^2\rangle$ for two electrons in a 1D simple harmonic oscillator potential. For the 1D SHO we have energy eigenstates $|n\rangle$ with energy eigenvalues $\hbar\omega(n + 1/2)$. Suppose that one electron is in the ground state $|0\rangle$ and the other electron is in the first excited state $|1\rangle$. For electrons in the triplet spin states, the spatial wavefunction must be exchange antisymmetric

$$|\psi_a\rangle = \sqrt{\frac{1}{2}}(|0\rangle_1|1\rangle_2 - |1\rangle_1|0\rangle_2)\ , \tag{1.15}$$

while for the singlet spin state, the spatial wavefunction must be exchange symmetric

$$|\psi_s\rangle = \sqrt{\frac{1}{2}}(|0\rangle_1|1\rangle_2 + |1\rangle_1|0\rangle_2)\ . \tag{1.16}$$

We can evaluate the expectation value of the distance between the electrons by expressing the position operators in terms of the raising and lowering operators for

the two electrons [see, for example, Griffiths (1995)]:

$$\hat{x}_i = \sqrt{\frac{\hbar}{2m\omega}}\left(\hat{a}_i + \hat{a}_i^\dagger\right) , \tag{1.17}$$

where

$$\begin{aligned} \hat{a}_i|n\rangle_i &= \sqrt{n}\,|n-1\rangle_i \,, \\ \hat{a}_i^\dagger|n\rangle_i &= \sqrt{n+1}\,|n+1\rangle_i \,. \end{aligned} \tag{1.18}$$

For the operator describing the distance between the electrons, we find

$$(\hat{x}_1 - \hat{x}_2)^2 = \hat{x}_1^2 - 2\hat{x}_1\hat{x}_2 + \hat{x}_2^2 \,, \tag{1.19}$$

where, by using Eq. (1.17), we obtain

$$\hat{x}_i^2 = \frac{\hbar}{2m\omega}\left(\hat{a}_i\hat{a}_i + \hat{a}_i^\dagger\hat{a}_i + \hat{a}_i\hat{a}_i^\dagger + \hat{a}_i^\dagger\hat{a}_i^\dagger\right) , \tag{1.20}$$

$$\hat{x}_1\hat{x}_2 = \frac{\hbar}{2m\omega}\left(\hat{a}_1\hat{a}_2 + \hat{a}_1^\dagger\hat{a}_2 + \hat{a}_1\hat{a}_2^\dagger + \hat{a}_1^\dagger\hat{a}_2^\dagger\right) . \tag{1.21}$$

Using the above relations, it is straightforward to show that

$$\boxed{\langle (x_2 - x_1)^2 \rangle_{\text{triplet}} = \frac{3\hbar}{m\omega} \,,} \tag{1.22}$$

while

$$\boxed{\langle (x_2 - x_1)^2 \rangle_{\text{singlet}} = \frac{\hbar}{m\omega} \,.} \tag{1.23}$$

Thus we see that, indeed, electrons in the triplet spin states are, on average, significantly further apart than those in the singlet states.

1.3 Spin-orbit interaction

Beyond the electrostatic attraction between the electrons and the nucleus and the electrostatic repulsion between the electrons (Problems 1.1 and 1.2), the next most important cause of energy level splittings in low Z atoms is the relativistic effects, which are responsible for what is known as the *fine structure* of atomic spectra.

For low-lying states in hydrogen, the fine structure splitting of the energy levels is a factor of $\sim \alpha^2$ smaller than the Bohr energies [Eq. (1.1)]. There are two

causes of this splitting: (1) the interaction of the magnetic moment of the electrons with the effective magnetic field the electrons see due to their motion around the nucleus, and (2) relativistic corrections to the kinetic and potential energies of the electrons. In the following problem we consider only the splitting due to the spin-orbit interaction (1), which is the principal cause of fine-structure splitting for heavier atoms.

Consider the term ^5D for a multi-electron atom.

(a) What are the possible values of J?

(b) Show the splitting of the ^5D term due to the spin-orbit interaction on an energy level diagram. Indicate the value of J and the energy for each level in terms of A, where the spin-orbit interaction is described by the Hamiltonian[4]

$$H' = A\vec{L} \cdot \vec{S} \, . \tag{1.24}$$

Assume A is positive, as is the case for subshells which are less than half-filled.

(c) What effect does the spin-orbit interaction have on the "center of gravity" (the mean perturbation of all the states of the term)?

Hints

For part (c), you may find useful the following formulae for summations (which can be proved by induction):

$$\sum_{J=0}^{N} J = \frac{1}{2} N(N+1) \, , \tag{1.25}$$

$$\sum_{J=0}^{N} J^2 = \frac{1}{6} N(N+1)(2N+1) \, , \tag{1.26}$$

[4] In a multi-electron atom, the "single electron" spin-orbit coupling (meaning that the interaction between one electron's spin and another's orbital motion is ignored, which is a good approximation for elements located in the middle and end of the periodic system) is given by

$$H' = \sum_i a\vec{l}_i \cdot \vec{s}_i \, .$$

Assuming that the spin-orbit coupling is small compared to the electrostatic repulsion between the electrons, the Russell-Saunders coupling scheme (see Problem 1.1) is still valid and, to first order, we need only calculate the diagonal matrix elements $\langle L, S, J, M_J | H' | L, S, J, M_J \rangle$. Seeking the overall shift of an atomic state with total orbital angular momentum $\vec{L}$ and total spin $\vec{S}$, we note that the *average* orbital angular momentum $\vec{l}$ of an electron in such a state is $\propto \vec{L}$, and the *average* spin $\vec{s}$ is $\propto \vec{S}$ (a consequence of the Wigner-Eckart theorem, see Appendix F), allowing us to write H' as is done in Eq. (1.24) [see, for example, Condon and Shortley (1970) or Landau and Lifshitz (1977)].

and

$$\sum_{J=0}^{N} J^3 = \frac{1}{4}N^2(N+1)^2 \,. \tag{1.27}$$

Solution

(a) Here we use the general rule for addition of angular momenta $\vec{L} + \vec{S} = \vec{J}$:

$$|L - S| \leq J \leq L + S \,, \tag{1.28}$$

which is known as the triangle inequality. Using (1.28), we find that the possible values of J are:

$$\boxed{{}^5D: \quad S = 2,\ L = 2 \ \rightarrow\ J = 4, 3, 2, 1, 0.} \tag{1.29}$$

(b) For the total angular momentum J we have the relation:

$$\vec{J} = \vec{S} + \vec{L} \,, \tag{1.30}$$

and therefore, by squaring both sides, we have

$$J^2 = S^2 + L^2 + 2\vec{L} \cdot \vec{S} \,. \tag{1.31}$$

Therefore the spin-orbit Hamiltonian (1.24) is

$$H' = \frac{A}{2}\left(J^2 - S^2 - L^2\right) \,. \tag{1.32}$$

The states described by the spectral terms ${}^{2S+1}L_J$ are eigenstates of J^2, S^2, and L^2. Therefore, they are also eigenstates of H' with eigenvalues

$$\Delta E = \frac{A\hbar^2}{2}[J(J+1) - S(S+1) - L(L+1)] \,. \tag{1.33}$$

Using Eq. (1.33) and setting $\hbar = 1$ we can evaluate the energy shifts for states with various J's with $S = 2,\ L = 2$:

$$\Delta E(J = 4) = 4A \tag{1.34}$$

$$\Delta E(J = 3) = 0 \tag{1.35}$$

$$\Delta E(J = 2) = -3A \tag{1.36}$$

$$\Delta E(J = 1) = -5A \tag{1.37}$$

$$\Delta E(J = 0) = -6A \,. \tag{1.38}$$

The energy level structure of this system is shown in Fig. 1.2.

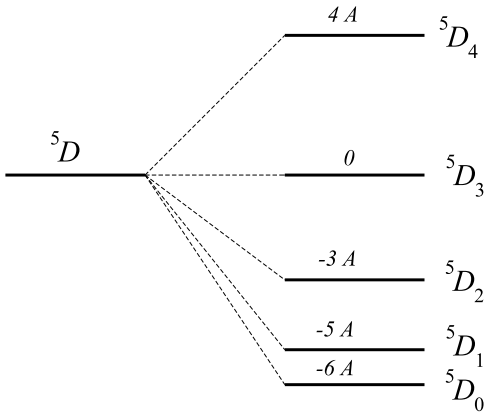

FIG. 1.2 Splitting of the ^5D term due to the spin-orbit interaction $H' = A\vec{L}\cdot\vec{S}$, with $A > 0$. We set $\hbar = 1$ for convenience.

Note that the energy difference between adjacent components is given by

$$\Delta E(J) - \Delta E(J-1) = AJ\ . \tag{1.39}$$

This formula is known as the *Landé interval rule*.

(c) The "center of gravity" of a term does not change due to the spin-orbit interaction. This answer can be guessed, since we expect the average of $\vec{L}\cdot\vec{S}$ over all possible orientations of $\vec{L}$ and $\vec{S}$ to be zero.

Alternatively, one can use the summation formulae [Eqs. (1.25)-(1.27)] to evaluate the shift of the center of gravity. For each J there are $(2J+1)$ Zeeman sublevels, so that the average energy shift $\langle\Delta E\rangle$ is given by the sum:

$$\langle\Delta E\rangle = \frac{A}{2}\sum_{J=|L-S|}^{L+S}(2J+1)[J(J+1) - S(S+1) - L(L+1)] = 0\ . \tag{1.40}$$

1.4 Hyperfine structure and Zeeman effect in hydrogen

In this classic problem, we are interested in what is known as the *hyperfine structure*, which in general arises due to the interaction of atomic electrons with the electric and magnetic multipole fields of the nucleus (the most important being the magnetic dipole and electric quadrupole). The transition between the hyperfine levels in the ground state of hydrogen is responsible for the famous 21-cm line in radio astronomy (the wavelength of the radiation is 21 cm), and the splitting between these levels has been measured extremely precisely with the hydrogen

maser. The transition between the ground state hyperfine levels of cesium is used for atomic clocks and this transition frequency defines the second.

(a) For the ground state of hydrogen ($^2S_{1/2}$), calculate the splitting of the $F = 1$ and $F = 0$ hyperfine levels (in MHz). What is the form of the Hamiltonian describing the hyperfine interaction?

(b) Consider the effect of a uniform magnetic field $\vec{B} = B\hat{z}$ on the ground state energy levels of hydrogen (the effects of external fields on atoms are considered in more detail in Chapters 2 and 4). For now, neglect the interaction of the proton magnetic moment with the external magnetic field. Calculate the energies of the ground-electron-state levels of the hydrogen atom as a function of the applied magnetic field B.

(c) If one includes the interaction of the proton magnetic moment with the magnetic field, two of the energy levels cross at a certain magnetic field value. Which levels cross and at what magnetic field does the crossing occur?

Hint

For part (a), since the electron has no orbital angular momentum, one can think of the magnetic field from the electron $\vec{B}_e$ being generated by a magnetization $\vec{M}_e(r)$:

$$\vec{M}_e(r) = -g_e\mu_0\vec{S}\,|\psi_{100}(r)|^2 \ , \tag{1.41}$$

where $g_e = 2$ is the Landé g-factor for the electron,[5] and $\psi_{100}(r)$ is the $n = 1$, $l = 0$, $m_l = 0$ ground state wavefunction of hydrogen.

Solution

(a) The $\psi_{100}(r)$ wavefunction is spherically symmetric, so we can envision the average magnetization produced by the electron (1.41) to consist of the contributions of a series of concentric spherical balls each with constant magnetization $\vec{M}_i$, so that

$$\sum_i \vec{M}_i = \vec{M}_e(r) \ . \tag{1.42}$$

[5] Note that the standard sign convention for the Bohr magneton is positive, so the magnetic dipole moment of the electron is $\mu_e = -g_e\mu_0$.

Recalling from classical electromagnetism that the magnetic field inside a spherical ball with constant magnetization $\vec{M}$ is given by (Griffiths 1999)

$$\vec{B} = \frac{8\pi}{3}\vec{M} \,, \tag{1.43}$$

we have for the field at $r = 0$

$$\vec{B}(0) = \frac{8\pi}{3}\sum_i \vec{M}_i = \frac{8\pi}{3}\vec{M}_e(0) \,, \tag{1.44}$$

from which we can calculate the magnetic field seen by the proton using Eq. (1.41).

We assume that $|\psi_{100}(r)|^2 = |\psi_{100}(0)|^2$ over the volume of the proton,[6] so

$$\vec{B}_e = -\frac{16\pi}{3}\mu_0 \left|\psi_{100}(0)\right|^2 \vec{S} = -\frac{16}{3a_0^3}\mu_0 \vec{S} \,, \tag{1.45}$$

where we have made use of the fact that

$$|\psi_{100}(0)|^2 = \frac{1}{\pi a_0^3} \,. \tag{1.46}$$

The Hamiltonian $H_{\rm hf}$ describing the interaction of the magnetic moment of the proton $\vec{\mu}_p$ with this magnetic field is thus

$$H_{\rm hf} = -\vec{\mu}_p \cdot \vec{B}_e = \frac{16}{3a_0^3} g_p \mu_N \mu_0 \vec{I} \cdot \vec{S} \,, \tag{1.47}$$

where $g_p = 5.58$ is the proton g-factor and μ_N is the nuclear magneton.

Using the same trick employed in the derivation of the fine structure splitting in Problem 1.3, we find that the Hamiltonian has the form

$$H_{\rm hf} = a\vec{I} \cdot \vec{S} = \frac{a}{2}\left(F^2 - I^2 - S^2\right) \,. \tag{1.48}$$

In units where $\hbar = 1$,

$$a \approx 5.58 \frac{16}{3a_0^3}\mu_N \mu_0 \approx 1420 \text{ MHz} \,, \tag{1.49}$$

and in terms of the eigenvalues of the angular momentum operators,

$$\boxed{H_{\rm hf} = \frac{a}{2}[F(F+1) - I(I+1) - S(S+1)] \,.} \tag{1.50}$$

Therefore the hyperfine splitting in the ground state of hydrogen is

$$\boxed{\Delta E_{\rm hf} \approx 1420 \text{ MHz} \,,} \tag{1.51}$$

which corresponds to electromagnetic radiation of wavelength $\lambda = 21 \text{ cm}$.

[6] An important point is that the hyperfine interaction in this case arises due to the wavefunction overlap between the proton and electron. This is somewhat subtle, as can be seen by comparing this analysis to that carried out in Problem 2.5 for a small ball carved out of a uniformly magnetized ball.

(b) From Eq. (1.50), we see that the energy eigenstates for the Hamiltonian describing the hyperfine interaction are also eigenstates of the operators $\left\{F^2, F_z, I^2, S^2\right\}$. Therefore if we write out a matrix for this Hamiltonian in the coupled basis, it is diagonal. However, the Hamiltonian H_B for the interaction of the magnetic moment of the electron with the external magnetic field

$$H_B = -\vec{\mu}_e \cdot \vec{B} = 2\mu_0 B S_z \tag{1.52}$$

is diagonal in the uncoupled basis (which is made up of eigenstates of the operators $\left\{I^2, I_z, S^2, S_z\right\}$).

The relationship between the coupled and uncoupled bases is as follows

$$|F=1, M_F=1\rangle = |+\rangle_S|+\rangle_I \,, \tag{1.53}$$

$$|F=1, M_F=0\rangle = \frac{1}{\sqrt{2}}(|+\rangle_S|-\rangle_I + |-\rangle_S|+\rangle_I) \,, \tag{1.54}$$

$$|F=1, M_F=-1\rangle = |-\rangle_S|-\rangle_I \,, \tag{1.55}$$

$$|F=0, M_F=0\rangle = \frac{1}{\sqrt{2}}(|+\rangle_S|-\rangle_I - |-\rangle_S|+\rangle_I) \,. \tag{1.56}$$

Employing Eqs. (1.50) and (1.52), one finds for the matrix $\mathbf{H}$ for the overall Hamiltonian ($H_{\rm hf} + H_B$) in the coupled basis:

	$\lvert1,1\rangle$	$\lvert1,-1\rangle$	$\lvert1,0\rangle$	$\lvert0,0\rangle$
$\langle1,1\rvert$	$\frac{a}{4}+\mu_0 B$	0	0	0
$\langle1,-1\rvert$	0	$\frac{a}{4}-\mu_0 B$	0	0
$\langle1,0\rvert$	0	0	$\frac{a}{4}$	$-\mu_0 B$
$\langle0,0\rvert$	0	0	$-\mu_0 B$	$-\frac{3a}{4}$

We can use this matrix to solve for the energies of the states as a function of B by employing the Schrödinger equation

$$\mathbf{H}|\psi\rangle = E|\psi\rangle \tag{1.57}$$

which implies that

$$(\mathbf{H} - E\mathbf{1})|\psi\rangle = 0 \,, \tag{1.58}$$

where $\mathbf{1}$ is the identity matrix. If $(\mathbf{H} - E\mathbf{1})$ had an inverse, then we could multiply both sides of Eq. (1.58) by $(\mathbf{H} - E\mathbf{1})^{-1}$ to show that $|\psi\rangle = 0$. Assuming $|\psi\rangle \neq 0$, in order to satisfy Eq. (1.58), the matrix $(\mathbf{H} - E\mathbf{1})$ must be singular. This implies

that its determinant is zero:

$$\begin{vmatrix} \frac{a}{4}+\mu_0 B - E & 0 & 0 & 0 \\ 0 & \frac{a}{4}-\mu_0 B - E & 0 & 0 \\ 0 & 0 & \frac{a}{4} - E & -\mu_0 B \\ 0 & 0 & -\mu_0 B & -\frac{3a}{4} - E \end{vmatrix} = 0 \, . \tag{1.59}$$

The above expression is known as the *secular equation*. The matrix is block diagonal, so the energies are obtained by solving

$$\frac{a}{4} + \mu_0 B - E = 0 \, , \tag{1.60}$$

$$\frac{a}{4} - \mu_0 B - E = 0 \, , \tag{1.61}$$

$$\left(\frac{a}{4} - E\right)\left(-\frac{3a}{4} - E\right) - \mu_0^2 B^2 = 0 \, . \tag{1.62}$$

This gives the following energies

$$E_1 = \frac{a}{4} + \mu_0 B \, , \tag{1.63}$$

$$E_2 = \frac{a}{4} - \mu_0 B \, , \tag{1.64}$$

$$E_3 = -\frac{a}{4} + \frac{a}{2}\sqrt{1 + 4\frac{\mu_0^2 B^2}{a^2}} \, , \tag{1.65}$$

$$E_4 = -\frac{a}{4} - \frac{a}{2}\sqrt{1 + 4\frac{\mu_0^2 B^2}{a^2}} \, , \tag{1.66}$$

which are plotted as a function of B in Fig. 1.3.

(c) If we include the effect of the proton's magnetic moment, we have

$$\vec{\mu} = \vec{\mu}_e + \vec{\mu}_p \, , \tag{1.67}$$

so

$$H_B = -\vec{\mu} \cdot \vec{B} = g_e \mu_0 B S_z - g_p \mu_N B I_z \, . \tag{1.68}$$

In the high field limit we expect that the highest energy state should be the $|+\rangle_S|-\rangle_I$ state. In the low field limit, the $|1,1\rangle = |+\rangle_S|+\rangle_I$ state is the highest energy state, so these two levels must cross at some magnetic field.

In part (b), where we neglected the proton magnetic moment, for sufficiently high fields ($2\mu_0 B/a \gg 1$), the difference in energy between the two highest lying energy levels is [see Eqs. (1.63) and (1.65)]:

$$E_1 - E_3 \approx \frac{a}{2} \, . \tag{1.69}$$

When the difference in energy between $|+\rangle_I$ and $|-\rangle_I$ due to the interaction of the proton's magnetic moment with the magnetic field is equal to this energy

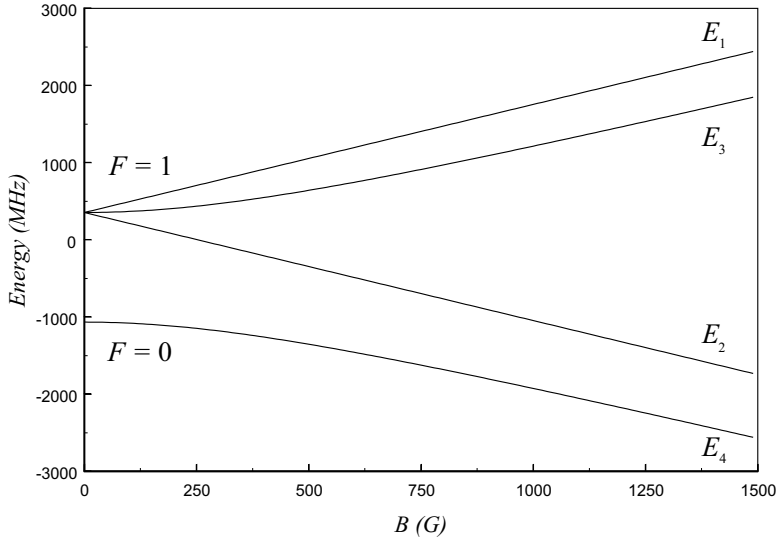

FIG. 1.3 Energies of the ground-state hyperfine manifold of hydrogen as a function of applied magnetic field. Such a plot is known as the *Breit-Rabi diagram*. At low fields, the system is well described in the coupled basis ($F = 1, 0$), while at high fields the energy eigenstates are best approximated by the uncoupled basis. The energies of the $|F = 1, M_F = 1\rangle$ and $|F = 1, M_F = -1\rangle$ states are linear in the magnetic field because they are not mixed with other states by the magnetic field [see Eqs. (1.53) and (1.55)].

difference, then the levels will cross. This occurs for the magnetic field:

$$\boxed{B \approx \frac{a}{2 \times 5.58 \times \mu_N} \approx 167 \text{ kG} .} \tag{1.70}$$

1.5 Hydrogenic ions

Hydrogen is an attractive object for the study of atomic structure because its simplicity allows accurate theoretical calculations which can be compared to experiment. A number of features in the energy-level structure of hydrogen are more pronounced in hydrogenic ions (atoms consisting of one electron bound to a nucleus with $Z > 1$) due to the larger nuclear charge. Hydrogenic ions are of interest for precision experiments testing quantum electrodynamics (Silver 2001), measuring the mass of the electron (Quint 2001), determining the fine structure constant (Quint 2001), and testing the Standard Model of electroweak interactions (Zolotorev and Budker 1997), to name a few.

For hydrogenic ions with nuclear charge Z, find the scaling with Z of:

(a) the expectation values of r, $1/r$, and $1/r^3$, where r is the distance of the electron from the nucleus,

(b) the expectation value of the potential energy V,

(c) the total energy E,

(d) the probability to find the electron at the origin, $|\psi(r=0)|^2$,

(e) $\left|\frac{\partial}{\partial r}\psi(r=0)\right|^2$,

(f) the fine structure energy splitting (see Problem 1.3), and

(g) the hyperfine structure energy splitting due to the magnetic dipole moment of the nucleus (see Problem 1.4). In this part, neglect the nonsystematic dependence of the nuclear dipole moments on Z.

Hints

You should not have to use any explicit wavefunctions – just consider the dimensions of the quantities of interest.

Solution

First, let us consider how the natural length scale of hydrogen, the Bohr radius $a_0 = \hbar^2/me^2$, compares to the natural length scale of a hydrogenic ion. The Hamiltonian for a hydrogenic ion is

$$H = -\frac{\hbar^2}{2m}\nabla^2 - \frac{Ze^2}{r} \,. \tag{1.71}$$

If we replace r with $\rho = r/Z$ in the Hamiltonian (1.71), taking into account that

$$\nabla^2 \to Z^2\nabla_\rho^2 \,, \tag{1.72}$$

we have

$$H = Z^2\left(-\frac{\hbar^2}{2m}\nabla_\rho^2 - \frac{e^2}{\rho}\right) . \tag{1.73}$$

Thus the Hamiltonian for a hydrogenic ion can be put into one-to-one correspondence with the Hamiltonian for hydrogen by rescaling r by a factor of Z^{-1}, so the

natural length scale for a hydrogenic ion is

$$a = \frac{a_0}{Z} \, . \tag{1.74}$$

One can also see from Eq. (1.73) that the total energy scales as Z^2 compared to hydrogen, which is the answer to part (c).

(a) As mentioned above, $a = a_0/Z$ is the only length scale for the system. Therefore any quantity with units $[\text{length}]^n$ must scale as Z^{-n}. Thus, for any n,

$$\boxed{\langle r^n \rangle \propto Z^{-n} \, .} \tag{1.75}$$

(b) The potential energy of an electron in a hydrogenic ion is

$$V(r) = -\frac{Ze^2}{r} \, , \tag{1.76}$$

and, based on part (a), since $r^{-1} \propto Z$, it is clear that

$$\boxed{\langle V \rangle \propto Z^2 \, .} \tag{1.77}$$

(c) Although we have already seen from Eq. (1.73) that $\langle E \rangle \propto Z^2$, we can also obtain this result by relating the total energy to the potential energy. According to the virial theorem, for two particles interacting via a central, conservative potential $V(r) = Cr^n$, the expectation value of the kinetic energy $\langle T \rangle$ is given by

$$\langle T \rangle = \frac{n}{2} \langle V \rangle \, . \tag{1.78}$$

For the electrostatic attraction between the nucleus and the electron, $n = -1$. Therefore

$$\langle E \rangle = \langle T \rangle + \langle V \rangle = \frac{1}{2} \langle V \rangle \, , \tag{1.79}$$

so we find

$$\boxed{\langle E \rangle \propto Z^2 \, .} \tag{1.80}$$

(d) Since the hydrogenic wavefunctions are normalized,

$$\int_0^\infty |\psi(r)|^2 \, d^3r = 1 \, , \tag{1.81}$$

and it is apparent that $|\psi(r=0)|^2$ has the dimensions $[\text{length}]^{-3}$. Thus

$$\boxed{|\psi(r=0)|^2 \propto Z^3 \ .} \tag{1.82}$$

(e) $\left|\frac{\partial}{\partial r}\psi(r=0)\right|^2$ has units of $[\text{length}]^{-5}$, so evidently

$$\boxed{\left|\frac{\partial}{\partial r}\psi(r=0)\right|^2 \propto Z^5 \ .} \tag{1.83}$$

(f) From the point-of-view of the electron, the nucleus of charge Z is orbiting around it with velocity $v \approx Z\alpha c$ (see Appendix A). Since the electric field $\mathcal{E}$ due to the nucleus is

$$\mathcal{E} = \frac{Ze}{r^2} \ , \tag{1.84}$$

the magnetic field B due to the relative motion of the electron and nucleus is

$$B = \left|\frac{\vec{v}\times\vec{\mathcal{E}}}{c}\right| \approx \frac{Z^2\alpha e}{r^2} \ . \tag{1.85}$$

This magnetic field interacts with the spin magnetic dipole moment of the electron to induce an energy shift

$$\Delta E_f \approx \mu_0 B \propto \frac{Z^2}{r^2} \ . \tag{1.86}$$

Since the expectation value of $1/r^2$ scales as Z^2, the fine-structure splitting scales as

$$\boxed{\Delta E_f \propto Z^4 \ .} \tag{1.87}$$

(g) In part (a) of Problem 1.4, we saw that the hyperfine energy splitting ΔE_{hf} for an s state is $\propto |\psi(0)|^2$, because the hyperfine shift is due to the interaction of the nuclear dipole moment with the magnetic field generated by the magnetic dipole moment associated with the electron spin. Using the result of part (d), we see that

$$\boxed{\Delta E_{\text{hf}} \propto Z^3 \ .} \tag{1.88}$$

We get the same result for higher angular momentum states. Indeed, here one can view the hyperfine shift as arising from the interaction of the electron magnetic dipole moment with the magnetic field due to the nuclear dipole moment

(the problem can be solved from either perspective). Since the magnetic field of a dipole falls off as $1/r^3$, according to part (a) of this problem, $\Delta E_{\rm hf} \propto \langle r^{-3} \rangle \propto Z^3$.

Note that the contribution to the hyperfine energy splitting from the nuclear quadrupole moment Q_{ij}, which arises from the interaction of Q with the electric field from the electron $E = e/r^2$, scales in the same way:

$$\Delta E_{\rm hf}{}^{(Q)} = Q_{ij} \cdot \frac{\partial E_i}{\partial x_j} \sim Q \frac{\partial E}{\partial r} \sim \frac{Qe}{r^3} \propto Z^3 \ . \tag{1.89}$$

1.6 Geonium

A beautiful and very useful "atomic system" was invented and perfected by Hans Dehmelt and co-workers (Dehmelt 1989). It consists of a single electron (or positron) confined inside a Penning trap (Fig. 1.4). Dehmelt called this system "geonium" (Van Dyck, Jr. *et al.* 1976), since it is essentially an electron bound to the Earth. The geonium atom has enabled precise measurements of the electron and positron g-factors (Van Dyck, Jr. *et al.* 1987) which, in combination with the calculations of Kinoshita (1996), constitute one of the best tests of the fundamental theory of quantum electrodynamics and one of the most accurate determinations of the fine structure constant.

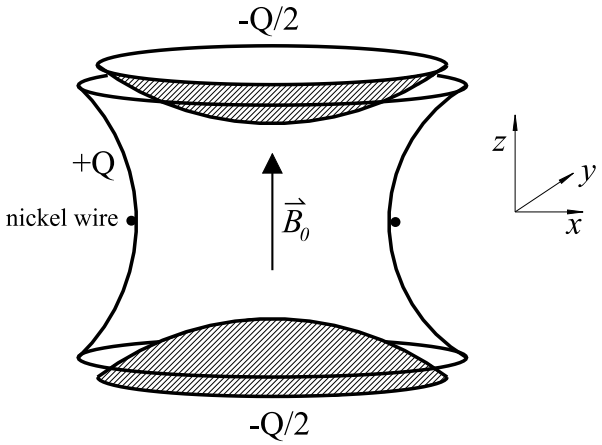

FIG. 1.4 Schematic picture of a Penning trap. The top and bottom conducting caps are charged to $-Q/2$ and the ring is charged to $+Q$, which generates a quadrupolar electrostatic field. A strong, homogeneous magnetic field $\vec{B}_0$ is applied in the $\hat{z}$-direction. A nickel wire (magnetized to saturation by $\vec{B}_0$) is wrapped around the center of the ring electrode to provide the "bottle field" used to measure properties of the trapped electron.

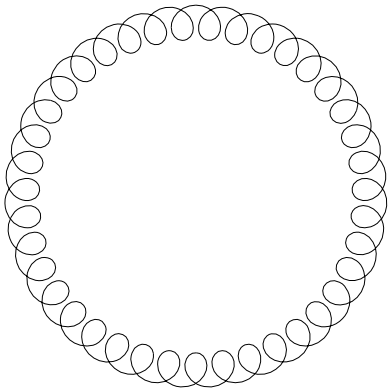

FIG. 1.5 Path of an electron in the midplane of a Penning trap executing magnetron and cyclotron motion. The path is a superposition of tight orbits associated with the cyclotron motion and an orbit with larger radius which is the magnetron motion. For illustrative purposes, the ratio of the period of the cyclotron motion to that of the magnetron motion has been increased by a factor of $\sim 10^5$ compared to that for typical experimental conditions.

The quadrupolar electrostatic field produced by the Penning trap's electrodes is described by the scalar potential

$$\Phi = \aleph\big(x^2 + y^2 - 2z^2\big) \; . \tag{1.90}$$

In addition to the electrostatic field, there is a homogeneous leading magnetic field $\vec{B}_0 = B_0\hat{z}$. These fields can confine an electron to a region near the axis of the trap for many months! The electron's motion can be subdivided into axial oscillations (the electron moves along the symmetry axis of the trap (z), reversing direction when it comes too close to one of the negatively charged caps) and motion in the xy-plane (the cyclotron motion and a slow drift referred to as the magnetron motion).

(a) What is the axial oscillation frequency ω_z for motion of the electron along z (assume the electron is on-axis, $x = y = 0$)?

(b) Ignoring the effect of the electric quadrupole field, calculate the cyclotron frequency ω_c with which the electron orbits about the trap axis.

(c) The Hamiltonian for the cyclotron motion can be put in one-to-one correspondence with the Hamiltonian for a one-dimensional simple harmonic oscillator (1D SHO). Show how this can be done (this was first realized by Landau in 1930, and the corresponding energy levels are known as *Landau levels*). Continue to ignore the electric quadrupole field.

(d) Consider an electron in the midplane of the trap ($z = 0$). Including both the homogeneous, z-directed magnetic field B_0 and the electric force, calculate the frequencies of circular motion.

The faster frequency is the shifted cyclotron frequency ω_c' and the slow frequency ω_m describes the *magnetron motion* (Fig. 1.5). The magnetron motion is a specific case of what is generally known as the $\vec{E} \times \vec{B}$ drift [see, for example, Jackson (1975)], where a charged particle in a combination of nonparallel magnetic and electric fields tends to "drift" in a direction orthogonal to both fields. Since in the present case, the magnetic field is in the $\hat{z}$-direction and the electric field is in the radial direction, the electron drifts in a circular path[7] centered about the trap axis.

It turns out that the magnetron motion can be described using the formalism of the 1D SHO just like the cyclotron motion, except that as the magnetron quantum number q increases, the magnetron energy decreases (Brown and Gabrielse 1986). This is because the electrostatic potential for an electron in the midplane of the trap is $\Phi = \aleph(x^2 + y^2)$, which means that the larger the radius of the magnetron orbit, the lower the energy. Therefore the magnetron motion is unbound, in the sense that as the electron dissipates energy, q gets larger and larger until the particle collides with the ring electrode.

(e) For the Penning trap used in Dehmelt's experiments at the University of Washington, the axial oscillation frequency was $\omega_z = 2\pi \times 60$ MHz. A 5 T (50,000 G) magnetic field was applied along the trap axis.

Sketch an energy level diagram for geonium with these experimental parameters. What is the energy difference between the geonium states $|m = +1/2, n, k, q\rangle$ and $|m = -1/2, n + 1, k, q\rangle$, where m is the quantum number describing the projection of the electron spin along z, n and q are the cyclotron and magnetron quantum numbers, and k is the quantum number describing axial oscillations?

(f) In addition to the fields discussed in the introduction to this problem, a weak "bottle" magnetic field $\vec{B}_b$ is applied to the electron, where

$$\vec{B}_b = -\beta\left[zx\hat{x} + zy\hat{y} - \left(z^2 - \frac{x^2 + y^2}{2}\right)\hat{z}\right]. \tag{1.91}$$

The bottle field allows measurement of the cyclotron, spin, and magnetron quantum numbers using what is known as the "continuous Stern-Gerlach" technique (Dehmelt 1989). The inhomogeneous bottle field interacts with the spin and motion of the electron (the magnetron and cyclotron motions, because they involve a charged particle moving in a circular orbit, produce magnetic dipole moments

[7] Here we have decomposed the motion in the xy-plane into cyclotron and magnetron motion, both of which are circular.

which interact with the bottle field). These interactions shift the axial oscillation frequency, which is measured with a radio-frequency resonant circuit.

In early geonium experiments, the magnetic bottle field was produced by a nickel wire (magnetized to saturation by $\vec{B}_0$) wound about the ring electrode (Fig. 1.4).

Show that the correction to the axial oscillation frequency ω_z due to the spin orientation, cyclotron motion, and magnetron motion is given by:

$$\delta\omega_z \approx \left(m + n + \frac{1}{2} + \frac{\omega_m}{\omega_c} q\right) \frac{2\mu_0 \beta}{m_e \omega_z} \,. \tag{1.92}$$

Hints

In part (c), use the fact that the kinetic momentum $m\vec{v}$ of an electron in the presence of a magnetic vector potential $\vec{A}$ is related to the canonical momentum $\vec{p}$ in the following way [see, for example, Griffiths (1999) or Landau and Lifshitz (1987)]:

$$m_e \vec{v} = \vec{p} - \frac{e}{c}\vec{A} \,, \tag{1.93}$$

and this is the momentum that enters the kinetic energy term in the Hamiltonian. For example, choose the vector potential to be

$$\vec{A} = -B_0 y \hat{x} \,, \tag{1.94}$$

and show that the effective Hamiltonian H_c' describing motion in the xy-plane can be written

$$H_c' = \frac{1}{2m_e} p_y^2 + \frac{1}{2} m_e \omega_c^2 (y - y_0)^2 \,, \tag{1.95}$$

where y_0 is a constant.

In part (f), an effective magnetic dipole moment μ can be assigned to the magnetron and cyclotron motion by setting the energy of the Landau levels equal to μB_0.

Solution

(a) On the axis of the Penning trap ($x = y = 0$), there is a restoring force

$$F_z = e\frac{\partial \Phi}{\partial z} = -4e\aleph z \,. \tag{1.96}$$

The effective spring constant for this simple harmonic motion is $4e\aleph$, so the axial oscillation frequency is

$$\boxed{\omega_z = \sqrt{\frac{4e\aleph}{m_e}}\ .} \tag{1.97}$$

(b) The force on the electron due to B_0 must balance the centrifugal force in order to hold the electron in orbit:

$$m_e\omega_c^2 r = e\frac{v}{c}B_0 = e\frac{\omega_c r}{c}B_0\ , \tag{1.98}$$

where $r = \sqrt{x^2+y^2}$ is the radius of the electron's orbit and $v = \omega_c r$ is the electron's velocity. The cyclotron frequency is

$$\boxed{\omega_c = \frac{eB_0}{m_e c}\ .} \tag{1.99}$$

(c) The vector potential $\vec{A}$ for the uniform field in the $\hat{z}$-direction can be written, for example, as:

$$\vec{A} = -B_0 y\hat{x}\ . \tag{1.100}$$

Using Eq. (1.93), the Hamiltonian H_c governing the cyclotron motion is

$$H_c = \frac{1}{2m_e}\left[\left(p_x - \frac{e}{c}B_0 y\right)^2 + p_y^2 + p_z^2\right]\ . \tag{1.101}$$

Note that the Hamiltonian does not contain the coordinates x and z, so it commutes with the momenta in the x and z directions:

$$[H_c, p_x] = [H_c, p_z] = 0\ . \tag{1.102}$$

Therefore motion in the axial direction can be decoupled from the cyclotron motion in this case, and p_x can be treated as a constant. We can rewrite the effective Hamiltonian H_c' describing motion in the xy-plane as

$$\boxed{H_c' = \frac{1}{2m_e}p_y^2 + \frac{1}{2}m_e\omega_c^2(y-y_0)^2\ ,} \tag{1.103}$$

where $y_0 = cp_x/(eB_0)$. Therefore the cyclotron motion can be described with the formalism developed for the 1D simple harmonic oscillator, even though it is a two-dimensional problem.

(d) In the midplane of the trap, the force $\vec{F}_e$ due to the electric quadrupole field is radial,

$$\vec{F}_e = e\frac{\partial \Phi}{\partial r}\hat{r} = 2e\aleph r\hat{r} = \frac{m_e}{2}\omega_z^2 r\hat{r}\ , \tag{1.104}$$

where we employed Eq. (1.97). This force subtracts from the radial force due to the magnetic field [Eq. (1.98)],

$$m_e\omega^2 r = e\frac{\omega r}{c}B_0 - \frac{m_e}{2}\omega_z^2 r\ . \tag{1.105}$$

The above relation yields [using Eq. (1.99)] a quadratic equation for the new frequencies ω of circular motion in the midplane:

$$\omega^2 - \omega_c\omega + \frac{\omega_z^2}{2} = 0\ . \tag{1.106}$$

We obtain two roots of this equation, giving us the shifted cyclotron frequency ω_c' and the magnetron frequency ω_m:

$$\boxed{\omega_c' \approx \omega_c - \frac{\omega_z^2}{2\omega_c}} \tag{1.107}$$

and

$$\boxed{\omega_m \approx \frac{\omega_z^2}{2\omega_c}\ .} \tag{1.108}$$

These are approximate roots under the assumption that $\omega_c \gg \omega_z$, as is the case in the experiment.

(e) From parts (a) and (c) and the statement of part (d), we know that the axial, cyclotron, and magnetron motions can be described using the formalism of the 1D SHO. We must also recall that there is an energy splitting between the electron spin up and spin down states due to the magnetic field B_0. Therefore, the energy levels of geonium are specified by four quantum numbers: m, denoting the spin projection along the z-axis, n, describing the cyclotron motion, k, the axial quantum number, and q, the magnetron quantum number. The spin quantum number can take on the values $\pm 1/2$, while n, k, and q can equal $0, 1, 2, 3, ...$

The energy of a particular state $|n, m, k, q\rangle$ of geonium is given by

$$E_{nmkq} = g_e \mu_0 B_0 m + \hbar\omega_c' \left(n + \frac{1}{2}\right) + \hbar\omega_z \left(k + \frac{1}{2}\right) - \hbar\omega_m \left(q + \frac{1}{2}\right) , \tag{1.109}$$

where g_e is the electron g-factor. The values of the energy splittings, for the experimental conditions listed in the statement of the problem, are

$$g_e \mu_0 B_0 \approx 2 \times 1.4 \text{ MHz/G} \times 50,000 \text{ G} = 140 \text{ GHz} , \tag{1.110}$$

$$\frac{\omega_c'}{2\pi} \approx \frac{eB_0}{2\pi m_e c} \approx \frac{4.8 \times 10^{-10} \text{ esu} \times 50,000 \text{ G}}{2\pi \times 9.1 \times 10^{-28} \text{ g} \times 3 \times 10^{10} \text{ cm/s}} = 140 \text{ GHz} , \tag{1.111}$$

$$\frac{\omega_z}{2\pi} = 60 \text{ MHz} , \tag{1.112}$$

$$\frac{\omega_m}{2\pi} = \frac{\omega_z^2}{4\pi\omega_c} \approx \frac{(60 \text{ MHz})^2}{2 \times 140 \text{ GHz}} \approx 13 \text{ kHz} , \tag{1.113}$$

where we have made use of the fact that $\omega_c' \approx \omega_c$.

This gives us a basic idea of the energy levels for geonium. It is, of course, no accident that the frequencies given by Eqs. (1.110) and (1.111) are nearly the same. Let us take a closer look at the nearly identical splittings between the different spin states and cyclotron states. Since

$$\mu_0 = \frac{e\hbar}{2m_e c} , \tag{1.114}$$

the difference in energy ΔE between the geonium states $|m = +1/2, n, k, q\rangle$ and $|m = -1/2, n + 1, k, q\rangle$ is

$$\boxed{\Delta E \approx (g_e - 2)\mu_0 B_0 \approx 164 \text{ MHz} .} \tag{1.115}$$

Thus, measuring this frequency difference enables a measurement of $g_e - 2$. The g-factor of the electron differs from 2 by a small amount $\approx \alpha/(2\pi)$ which can be accurately calculated using the theory of quantum electrodynamics.

A schematic sketch of the geonium energy levels is shown in Fig. 1.6.

(f) Here we again consider the axial oscillation frequency on-axis ($x = y = 0$), now taking into account the bottle field in conjunction with all the various motions. The Hamiltonian H governing the axial motion is given by:

$$H = \frac{p_z^2}{2m_e} - \vec{\mu} \cdot \vec{B}_b - e\Phi = \frac{p_z^2}{2m_e} - \mu_{\text{eff}} \beta z^2 + 2e\aleph z^2 , \tag{1.116}$$

where μ_{eff} is the effective magnetic moment of the electron due to its spin, cyclotron and magnetron motion (discussed in detail below). Equation (1.116) is

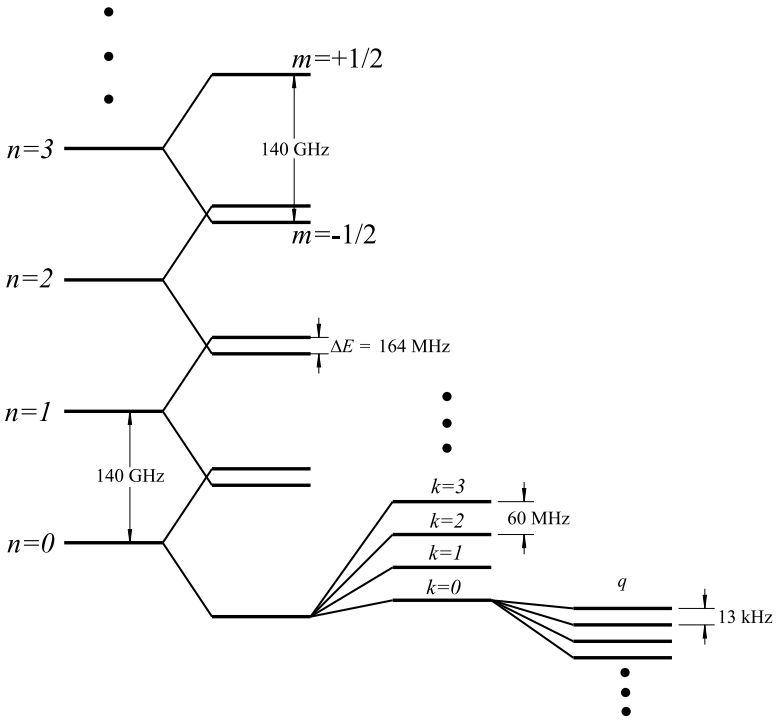

FIG. 1.6 Geonium energy levels (not to scale). Adapted from Van Dyck, Jr. *et al.* (1978).

the Hamiltonian for a 1D SHO:

$$H = \frac{p_z^2}{2m_e} + \frac{1}{2}m_e\omega_z'^2 z^2 \,, \tag{1.117}$$

where the new axial oscillation frequency ω_z' is given by:

$$\omega_z'^2 = -\frac{2\mu_{\rm eff}\beta}{m_e} + \frac{4e\aleph}{m_e} \,. \tag{1.118}$$

The second term in Eq. (1.118) is the square of the unperturbed axial oscillation frequency ω_z (1.97), so the axial oscillation frequency in the presence of the bottle field is given by:

$$\omega_z' = \left(-\frac{2\mu_{\rm eff}\beta}{m_e} + \omega_z^2\right)^{1/2} \approx \omega_z - \frac{\mu_{\rm eff}\beta}{m_e\omega_z} \,. \tag{1.119}$$

The correction to the axial oscillation frequency is thus given by

$$\delta\omega_z \approx -\frac{\mu_{\rm eff}\beta}{m_e\omega_z} \,. \tag{1.120}$$

Now we need to determine $\mu_{\rm eff}$, which is given by:

$$\mu_{\rm eff} = \mu_s + \mu_c + \mu_m, \tag{1.121}$$

where μ_s is the spin magnetic moment, μ_c is the cyclotron magnetic moment, and μ_m is the magnetic moment due to magnetron motion. The spin magnetic moment is:

$$\mu_s = -g_e \mu_0 m \approx -2\mu_0 m \; . \tag{1.122}$$

If we assign an energy $-\mu_c B_0$ to each of the Landau levels for the cyclotron motion, we can associate the cyclotron frequency with a magnetic moment according to:

$$\hbar\omega_c\left(n + \frac{1}{2}\right) = \frac{e\hbar}{m_e c}\left(n + \frac{1}{2}\right)B_0 = -\mu_c B_0 \; , \tag{1.123}$$

so that

$$\mu_c = -2\mu_0\left(n + \frac{1}{2}\right) \; . \tag{1.124}$$

The magnetron motion can be described in the same manner:

$$\mu_m = -2\mu_0\left(\frac{\omega_m}{\omega_c}\right)\left(q + \frac{1}{2}\right) \; . \tag{1.125}$$

Employing the above relations in Eq. (1.120), we find:

$$\boxed{\delta\omega_z \approx \frac{2\mu_0\beta}{\omega_z m_e}\left(m + n + \frac{1}{2} + (q + 1/2)\frac{\omega_m}{\omega_c}\right)} \tag{1.126}$$

agreeing with the equation in the statement of the problem, if we neglect the 1/2 associated with the magnetron motion (since under typical experimental conditions $q \gg 1$).

The axial oscillation frequency, therefore, is sensitive to the cyclotron, spin, and magnetron quantum numbers. Whenever the electron makes a quantum jump between energy levels, this is observed as a sudden shift in ω_z. In practice, to measure ω_z, an oscillating drive voltage is applied to the caps of the electrode (Fig. 1.4), and the amplitude of the axial oscillation greatly increases when the drive frequency hits resonance. This method of detection uses what is known as the "continuous Stern-Gerlach effect," since it is analogous to the classic experiment [see, for example, Griffiths (1995) or Bransden and Joachain (2003)] in which the trajectories of silver atoms in an atomic beam are perturbed by the interaction of their dipole moments with a magnetic field gradient.

A clever technique is employed to measure the $g_e - 2$ anomaly [discussed in part (e)]. An inhomogeneous, radio-frequency magnetic field at $\omega_{\rm rf} \approx 2\pi \times 164$ MHz is applied to the electron. If the inhomogeneous field was dc, then an electron moving through the field would see a magnetic field oscillating at the cyclotron frequency. Because the field is oscillating with frequency $\omega_{\rm rf}$, the electron sees (in its rest frame) sidebands at $\omega = \omega_c \pm \omega_{\rm rf}$. Thus when $\omega_{\rm rf}$ corresponds to the difference ΔE from Eq. (1.115) between the energy level splitting between different spin orientations and that between different cyclotron levels, spin flips are induced (since $\omega = \omega_c + \omega_{\rm rf} = g_e \mu_0 B_0$). The increased rate of spin flips is detected using the continuous Stern-Gerlach effect. This constitutes a direct measurement of the $g_e - 2$ anomaly.

1.7 The Thomas-Fermi model (T)

A starting point for precise numerical calculations of atomic energy levels in complex, multi-electron atoms is the theory developed independently by Thomas and Fermi [see, for example, Bransden and Joachain (2003) or Landau and Lifshitz (1977)]. The Thomas-Fermi model assumes that the electron cloud is a zero temperature Fermi gas. The central result of this model can be derived by balancing electrostatic forces with the gradient of pressure ($\vec{\nabla} P$) produced as a consequence of the Pauli exclusion principle:

$$\vec{\nabla} P(r) = \rho(r) \vec{\mathcal{E}}(r) = -en(r) \left[-\vec{\nabla} \phi(r) \right] , \tag{1.127}$$

where $\rho(r) = -en(r)$ is the charge density, $n(r)$ is the electron number density, $\vec{\mathcal{E}}(r) = -\vec{\nabla}\phi(r)$ is the electric field, and $\phi(r)$ is the electrostatic potential.

Equation (1.127) is the condition for hydrostatic equilibrium. This means that in the Thomas-Fermi model, we treat the electron cloud as a fluid, much like the atmosphere of the Earth, except here electrostatic forces due to the nucleus hold the fluid in rather than gravity. To describe the electron cloud in this manner, we assume that the electrons can be treated semiclassically and assume that we may apply statistical arguments. Such an approach is justified if the atom has a large number of electrons ($N \gg 1$).

(a) Discuss why the condition $N \gg 1$ allows one to use the semiclassical approximation.

Solution

If $N \gg 1$, because of the Pauli exclusion principle, many of the electrons must occupy states with large radial quantum numbers n. The key requirement for

employing the semiclassical approximation is that there must be many oscillations of the wavefunction over the regions of space where the potential changes appreciably. In other words, the deBroglie wavelength $\lambda_{\text{dB}} \sim \hbar/p$ must change slowly with respect to the distance from the nucleus:

$$\frac{\partial \lambda_{\text{dB}}}{\partial r} \ll 1 \, . \tag{1.128}$$

The typical angular momentum of an electron with radial quantum number n is given by

$$L = rp \sim n\hbar \, , \tag{1.129}$$

which implies that $\lambda_{\text{dB}} \sim r/n$. Therefore, the condition (1.128) requires that

$$n \gg 1 \, , \tag{1.130}$$

which is true for most electrons if $N \gg 1$.

(b) Calculate the Fermi momentum p_F (the momentum of the electron with the highest energy) for an electron gas in a small volume V. Use the fact that since the electron wavefunctions are semiclassical, they can be approximated by plane waves. Also assume that the number of electrons within V is large enough that one can apply statistical arguments.

Solution

The number of electron states dN with momenta between p and $p + dp$ occupying a small volume V is given by

$$dN = \frac{1}{\pi^2} \frac{p^2 dp}{\hbar^3} V \, , \tag{1.131}$$

where we have taken into account that the number of possible states per unit volume is doubled for electrons due to their spin (compared to a particle without spin). The number of available electron states per unit volume (at zero temperature, this is equal to the electron density since each state is occupied) can be found by integrating Eq. (1.131) from 0 to p_F (the Fermi momentum), yielding

$$\frac{N}{V} = \frac{1}{3\pi^2} \frac{p_F^3}{\hbar^3} \, . \tag{1.132}$$

From Eq. (1.132), we have

$$\boxed{p_F = \hbar \left(\frac{3\pi^2 N}{V} \right)^{1/3} \, .} \tag{1.133}$$

(c) Calculate the total kinetic energy K of the electron gas.

Solution

The total kinetic energy K can be obtained by multiplying dN from Eq. (1.131) by the kinetic energy per electron $p^2/2m$ (where m is the electron mass) and integrating from 0 to p_F:

$$K = \frac{p_F^5 V}{10\pi^2 m\hbar^3} \, . \tag{1.134}$$

Substituting the value of p_F from Eq. (1.133) into our expression for kinetic energy (1.134) we obtain:

$$\boxed{K = \frac{\hbar^2}{10\pi^2 m}\frac{\left(3\pi^2 N\right)^{5/3}}{V^{2/3}} \, .} \tag{1.135}$$

(d) Using the thermodynamic relation for the Fermi pressure $P = -dK/dV$ [valid at zero temperature, see Reif (1965)], obtain the pressure of the electron gas as a function of density. The Thomas-Fermi model assumes V is small compared to the volume of the atom, so that P is the local pressure at a particular distance from the nucleus.

Solution

Using the thermodynamic relation $P = -dK/dV$:

$$P = \frac{\hbar^2}{15\pi^2 m}\left(3\pi^2\frac{N}{V}\right)^{5/3} \, , \tag{1.136}$$

which yields the pressure as a function of electron density $n(r)$ in the atom

$$\boxed{P(r) = \frac{3^{2/3}\pi^{4/3}}{5}\frac{\hbar^2}{m}[n(r)]^{5/3} \, .} \tag{1.137}$$

(e) Now use Eq. (1.137) in conjunction with Eq. (1.127) to obtain a relationship between the electrostatic potential $\phi(r)$ and the electron density $n(r)$.

Solution

Employing Eq. (1.137) in Eq. (1.127) and integrating gives

$$\boxed{\frac{\hbar^2}{2m}\left[3\pi^2 n(r)\right]^{2/3} = e[\phi(r) - \phi_0] \, ,} \tag{1.138}$$

where ϕ_0 is a constant of integration. Equation (1.138) is the central result of the Thomas-Fermi model.

Equation (1.138) can be combined with the Poisson equation,

$$\nabla^2\phi(r) = 4\pi e n(r) \ , \tag{1.139}$$

to obtain two independent equations for the two unknown functions $n(r)$ and $\phi(r)$, and values of these functions for various r can be obtained numerically, taking into account the appropriate boundary conditions [see, for example, Bransden and Joachain (2003), Landau and Lifshitz (1977), or Messiah (1966)].

1.8 Electrons in a shell

In this problem (Budker 1998a), intended to illustrate the basic principles of the Thomas-Fermi method, we consider the case of a large number of electrons at zero temperature placed inside a spherical cavity of radius a with impenetrable walls. This is like an atom without the nucleus (of course, the walls are necessary to keep the electrons from flying apart due to the electrostatic repulsion). In the following problem we ignore all numerical factors, such as 4π, in order to simplify expressions and concentrate on the scaling of various effects. Assume $a \gg a_0$, where $a_0 = \hbar^2/(me^2)$ is the Bohr radius. Note that under the stated conditions, the Thomas-Fermi model (Problem 1.7) is applicable.

(a) Argue that the electrons collect in a thin shell of thickness δ at the edge of the spherical cavity. Determine the scaling of δ with respect to the number of electrons N and radius of the cavity a.

(b) For what N does the assumption that the electrons are nonrelativistic break down?

(c) What is the lower bound on N for which the assumptions of the Thomas-Fermi model are satisfied? Estimate δ for the case of low N.

Hint

For part (a), suppose that the shell contains half the electrons in the cavity.

Solution

(a) The boundary conditions in this problem are, of course, quite different from those for an atom: here there is no positively charged nucleus at the center and there is an infinitely high potential barrier at $r = a$. Choosing the electrostatic

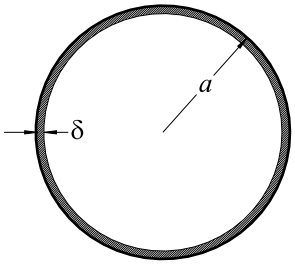

FIG. 1.7 Electrons collect in a shell of thickness δ near the wall.

potential as, for example, $\phi = 0$ at the boundary, the functions $\phi(r)$ and $n(r)$ can be obtained by numerical integration of the Thomas-Fermi equation.

However, without resorting to detailed calculations, we can obtain a general idea of the spatial distribution of the electrons in the cavity. Consider a shell of thickness δ bounded by the wall of the cavity containing half of the electrons (we will see that $\delta \ll a$). The Coulomb repulsion from the other electrons tends to compress the shell as a whole, pushing the shell toward the wall. We define the Coulomb pressure P_C as the repulsion force ($\sim N^2e^2/a^2$) per unit surface area of the shell:

$$P_C \sim \frac{N^2e^2}{a^4} \, . \tag{1.140}$$

In equilibrium, this Coulomb pressure is balanced by the Fermi pressure P from Eq. (1.136) in Problem 1.7,

$$P \sim \frac{\hbar^2}{m}\left(\frac{N}{a^2\delta}\right)^{5/3} , \tag{1.141}$$

where we have substituted $V \sim a^2\delta$. Setting $P = P_C$, we find that

$$\boxed{\delta \sim \frac{a^{2/5}a_0^{3/5}}{N^{1/5}} \, .} \tag{1.142}$$

As can be seen from Eq. (1.142), $\delta \ll a$, so indeed the electrons collect in a thin shell at the edge of the cavity. It is particularly interesting to note that the thickness of the shell decreases as more electrons are placed in the cavity.

(b) The assumption that the electrons are nonrelativistic will break down when the Fermi momentum p_F becomes of order mc. By substituting Eq. (1.142) into the expression for p_F [Eq. (1.133)], and supposing that the volume of the shell is

$\sim a^2\delta$ we obtain an expression for the Fermi momentum in terms of N:

$$p_F \sim \hbar \frac{N^{2/5}}{a^{4/5} a_0^{1/5}} \, . \tag{1.143}$$

Setting $p_F = mc$, we find that the critical number of electrons N^* for which the nonrelativistic approximation breaks down is

$$\boxed{N^* \sim \left(\frac{a^2}{a_0^2}\right) \alpha^{-5/2} \, ,} \tag{1.144}$$

where $\alpha = e^2/\hbar c$ is the fine structure constant.

(c) In order to see at which values of N the semiclassical approximation is valid, suppose that we gradually increase the number of electrons inside the cavity. At low densities, all electrons occupy the lowest radial state, which means that the semiclassical approximation is invalid (Problem 1.7). Therefore, for low densities, most of the kinetic energy of the electrons is due to radial motion.

In this regime, we can estimate the thickness of the shell containing the electrons δ by finding the minimum of the total energy $E = K + U$, where K is the kinetic energy and U is the potential energy due to the Coulomb repulsion between the electrons. For the δ where the energy is minimized,

$$\frac{\partial K}{\partial \delta} + \frac{\partial U}{\partial \delta} = 0 \, . \tag{1.145}$$

Based on the Heisenberg uncertainty relation:

$$\Delta p_r \delta \sim \hbar \, , \tag{1.146}$$

where Δp_r is the uncertainty in the radial component of the electron momentum. This tells us that the kinetic energy of radial motion, which as noted above is the dominant contribution to K, is

$$N(\Delta p_r)^2/m \sim \frac{N\hbar^2}{m\delta^2} \sim K \, , \tag{1.147}$$

so

$$\frac{\partial K}{\partial \delta} \sim -\frac{N\hbar^2}{m\delta^3} \, . \tag{1.148}$$

The Coulomb energy of the shell can be estimated from the work required to move the electrons in from the edge of the cavity by $\sim \delta$:

$$U \sim \frac{N^2 e^2}{a^2} \delta \, , \tag{1.149}$$

from which we find

$$\frac{\partial U}{\partial \delta} \sim \frac{N^2 e^2}{a^2} \ . \tag{1.150}$$

From Eqs. (1.146), (1.148), and (1.150) it follows that

$$\boxed{\delta \sim \frac{a^{2/3} a_0^{1/3}}{N^{1/3}} \ ,} \tag{1.151}$$

which can be compared to Eq. (1.142).

In order to satisfy the Pauli principle, states of different angular momenta l are excited as the number of electrons in the cavity is increased. Setting the total number of electron states equal to the number of particles N, we have [see Eq. (1.25)]:

$$N = \sum_{l=0}^{L} (2l+1) \sim L^2 \ , \tag{1.152}$$

where L is the highest value of l which is excited.

As we keep increasing N, at some point the kinetic energy due to the orbital motion of electrons with angular momentum L becomes equal to the kinetic energy of radial motion:

$$\frac{\hbar^2 L^2}{m a^2} \sim \frac{\hbar^2}{m \delta^2} \ . \tag{1.153}$$

At higher values of N, it is energetically favorable to excite higher radial modes, corresponding to the transition to the Thomas-Fermi regime discussed in part (a). From Eqs. (1.152), (1.151), and (1.153), we obtain the lower bound N^{**} on the value of N for which the Thomas-Fermi considerations apply:

$$\boxed{N^{**} \sim \frac{a^2}{a_0^2} \ .} \tag{1.154}$$

It is curious to note that this corresponds to a surface density of about one electron per Bohr radius squared.

The situation analyzed in this problem is difficult to realize experimentally because all materials are made of atoms, and so the impenetrable wall is unrealistic. However, such a situation could arise with quasiparticles in condensed matter physics.

1.9 Isotope shifts and the King plot

In atomic spectra, there appear small shifts of the transition energies for different isotopes. Such *isotope shifts* arise due to differences between the masses and the volumes of the nuclei. A commonly used method to separate experimentally measured isotope shifts into mass and volume (or field) shift contributions is based on the so-called King plot (King 1963).

Consider two spectral lines A and B. Due to the isotope shifts, the resonance frequencies are slightly different for each of the isotopes. For pairs of isotopes, where the difference in neutron number ΔN is always the same[8] and ΔN is much smaller than the atomic mass, one can assume that

- the isotope shift due to the finite nuclear volume can be expressed as a product of an electronic factor E (different for lines A and B, but the same for each isotope pair in a given spectral line) times a nuclear volume factor V (different for each pair of isotopes, but independent of which line is used), and
- the mass effects M are the same, in any one line, for all pairs of isotopes.[9]

From these assumptions, it follows that if one makes a plot in which the isotope shifts for isotope pairs in line B are plotted against the isotope shifts for the same isotope pairs in line A, the points will fall on a straight line.

An example of original isotope shift data (Brand *et al.* 1978) is shown in Table 1.2. In this experiment, isotope shifts were measured in a number of spectral lines of samarium using laser spectroscopy and an atomic beam. The laser and atomic beams intersected at right angles, and laser-induced fluorescence was detected in a third orthogonal direction. This setup minimizes the Doppler broadening (see, for example, Problem 3.6) of the spectral lines.

(a) Derive expressions for the slope and intercept of this line in terms of E_A, E_B, M_A, and M_B.

(b) Using the data in Table 1.2, make a King plot with the 598.97 nm $^7F_0 \rightarrow {}^7D_1$ transition as line A and the 562.18 nm $^7F_1 \rightarrow {}^7H_2$ transition as line B.

(c) The mass shift term M consists of two contributions:

$$M = M^{(\mathrm{nms})} + M^{(\mathrm{sms})} \,, \tag{1.155}$$

where the abbreviations stand for *normal mass shift* (nms) and *specific* (or anomalous) *mass shift* (sms). The former is due to the fact that all atomic energy intervals

[8] If the difference in neutron number is not the same, we can always normalize the isotope shifts to, for example, $\Delta N = 1$ or 2.

[9] To be exact, the variation of the mass shift with nuclear mass should be considered. Since here we assume that ΔN is much smaller than the atomic mass, this can be done by introducing a small correction factor which can be ignored for the purpose of this problem.

TABLE 1.2 Isotope shifts (IS) for various pairs of Sm isotopes for the 562.18 nm $^7F_1 \rightarrow {}^7H_2$ transition and the 598.97 nm $^7F_0 \rightarrow {}^7D_1$ transition. Data from Brand *et al.* (1978). The table indicates the difference in the measured transition frequencies, for example, the resonance frequency for the 562.18 nm line was $3093.6(16)$ MHz higher for ^{144}Sm than for ^{148}Sm.

Isotope pair	IS (MHz) for 562.18 nm	IS (MHz) for 598.97 nm	ΔN
(144,148)	3093.6(16)	-2794.4(17)	4
(148,150)	1938.3(15)	-1641.2(13)	2
(150,152)	2961.0(15)	-2308.0(19)	2
(152,154)	1362.3(11)	-1242.4(17)	2
(147,148)	970.4(7)	-826.0(4)	1
(148,149)	473.3(4)	-493.8(4)	1

are proportional to the reduced mass rather than just the electron mass. The latter is due to the correlation between momenta of various electrons. One can think of the specific mass shift as arising from the formation of multi-electron "quasiparticles" which move about the nucleus (King 1984).

Using the empirical information that the sms contribution in the 598.97 nm $^7F_0 \rightarrow {}^7D_1$ transition is negligible, evaluate the mass shift for the 562.18 nm $^7F_1 \rightarrow {}^7H_2$ transition. Compare its magnitude and sign to the expected normal mass shift.

Solution

(a) We let the isotope shifts for line A be the independent variable x and the isotope shifts for line B be the dependent variable y. Based on the assumptions stated in the problem, we have for the isotope shifts:

$$x = E_A V_i + M_A \,, \tag{1.156}$$

$$y = E_B V_i + M_B \,, \tag{1.157}$$

where i is the label for the isotope pair. Using Eq. (1.156), we can write V_i in terms of x, E_A, and M_A,

$$V_i = \frac{1}{E_A} x - \frac{M_A}{E_A} \,. \tag{1.158}$$

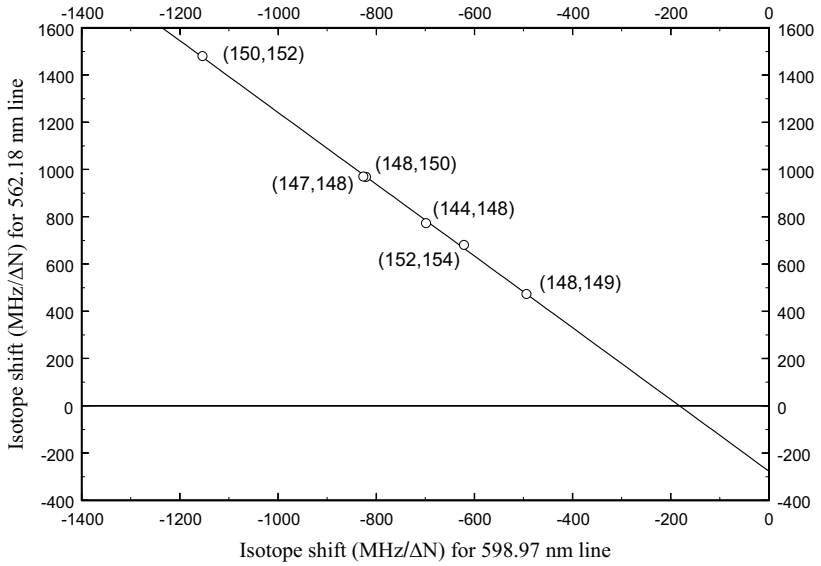

FIG. 1.8 King plot for samarium transitions listed in Table 1.2.

Substituting this expression into the equation for y (1.157) yields

$$\boxed{y = \frac{E_B}{E_A} x + M_B - \frac{E_B}{E_A} M_A \, .} \tag{1.159}$$

(b) Making sure to divide all isotope shifts by the difference in neutron number ΔN for the given pair (see Table 1.2), we obtain the King plot shown in Fig. 1.8. We see that the points indeed fall on a straight line.

(c) We begin by estimating the normal mass shift. For a single-electron atom, the finite mass of the nucleus can be taken into account by using the solution for an infinitely heavy nucleus and replacing the electron mass with the reduced mass $\mu_{\rm red}$:

$$\mu_{\rm red} = \frac{m M_N A}{m + M_N A} \, , \tag{1.160}$$

where M_N is the nucleon mass and A is the atomic mass number. Since energies of all atomic levels are proportional to this mass, a transition corresponding to an isotope with mass number A is shifted compared to that of a (fictitious) isotope

with infinite nuclear mass by:

$$\Delta\omega = \frac{\omega_0}{m}(\mu_{\rm red} - m) = \omega_0\left(\frac{M_N A}{m + M_N A} - 1\right)$$
$$\approx -\frac{m}{M_N A}\omega_0 \ . \tag{1.161}$$

Transitions corresponding to isotopes A and $A + \Delta N$ are shifted by

$$\Delta\omega' = \omega_0\left(-\frac{m}{M_N A + M_N \Delta N} + \frac{m}{M_N A}\right)$$
$$\approx \omega_0 \frac{m\Delta N}{M_N A^2} \ , \tag{1.162}$$

where we have made use of the fact that $\Delta N \ll A$. For the optical transitions near 600 nm considered here, with $\Delta N = 1$ as in our King plot (Fig. 1.8), this corresponds to

$$\boxed{\Delta\omega' \approx 5 \times 10^{14}\ {\rm Hz}\frac{1}{1836 \times 150^2} \approx 12\ {\rm MHz} \ .} \tag{1.163}$$

The mass shift for the 598.97 nm transition, M_A in our case, is given by

$$M_A = M_A^{(\rm nms)} + M_A^{(\rm sms)} \ . \tag{1.164}$$

We have shown that $M_A^{(\rm nms)} \approx 12$ MHz, while it is known empirically that $M_A^{(\rm sms)}$ is negligible. From the y-intercept on the King plot and Eq. (1.159), we find that the quantity $M_B - (E_B/E_A)M_A \approx -280$ MHz. The slope of the King plot shows that $(E_B/E_A) \approx -1.5$, so we obtain a mass shift of ≈ -300 MHz for the 562.18 nm transition, which is larger in magnitude and of opposite sign compared to the normal mass shift. Such large and originally unexpected specific mass shift contributions are often found in the rare earth elements when the number of f-electrons is different in the upper and lower states (the 562.18 nm transition is nominally between electron configurations $4f^6 6s^2$ and $4f^5 5d 6s^2$).

1.10 Crude model of a negative ion

Singly charged negative ions ($K = 1$) are very common (Massey 1976), and several doubly charged ($K = 2$) negative ions of atoms [see Massey (1976), Chapter 5.8] and clusters (Vandenbosch *et al.* 1997) have been observed as well. It is not entirely obvious that such systems should be bound – one might expect that the

Coulomb repulsion between an extra electron and the original Z electrons could overwhelm the attraction of the extra electron to the nucleus. Here we construct a crude model to see that such a conclusion need not be correct.

Using an electrostatic analogy, explain why it is possible to have bound states of a positive nucleus of charge $+Z$ with an electron cloud of charge $-(Z + K)$, $K > 0$.

Hint

As a very crude model of an atom, consider electrons as a conducting shell, and neglect the *exchange interaction* between electrons (see Problem 1.2). The exchange interaction is actually of great importance in many cases – nonetheless, this simple model illustrates the basic physical principle behind the existence of negative ions.

Solution

Let us formulate the following electrostatics problem. Consider a thin spherical conducting shell which is charged to total charge Q. Suppose the sphere is cut into two parts. Since each of the two resulting parts carries a charge of the same sign, the two parts will tend to fly apart. What is the charge that one needs to put in the center of the sphere in order to keep the two parts together?

In order to answer this question, we use the following facts from electrostatics:

- The pressure produced by an electric field E is $-E^2/8\pi$.
- The electric field inside a conductor is zero.

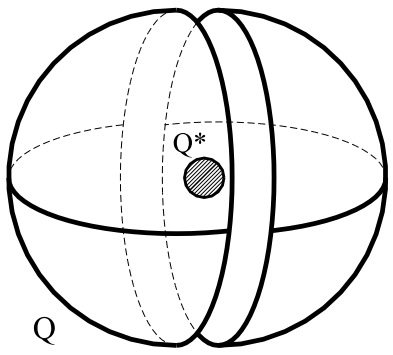

FIG. 1.9 Charge Q^* can hold together the parts of a conducting sphere carrying charge Q.

- The electric field from a uniformly charged spherical shell is zero inside the shell, and outside the shell, it is just as if all charge was in the center of the sphere.

If there is no charge in the center, there is no field inside the sphere, and the two parts are "sucked" apart by the negative pressure of the field outside. All we need to do to compensate for the negative outside pressure is put a charge (call it Q^*) in the center such that the magnitude of the field inside the shell (near its surface) is the same as outside. The field inside is Q^*/R^2 (R is the shell radius, which will, of course, cancel). Outside, the field is $(Q + Q^*)/R^2$. If we want to match the magnitudes of the fields inside and outside the shell, we need to have:

$$Q^* = -(Q + Q^*) \Rightarrow Q^* = -Q/2. \tag{1.165}$$

Any charge of sign opposite to that of Q and magnitude larger than $|Q|/2$, placed in the center, will hold the system together.

1.11 Hyperfine-interaction-induced mixing of states of different J

Two atomic fine structure levels in a multi-electron atom, $^2P_{3/2}$ and $^2P_{1/2}$, are separated by an energy gap ΔE. The nuclear spin of the system is $I = 1/2$.

Determine the admixture of the $^2P_{3/2}, F = 1$ state in the state that is nominally $^2P_{1/2}, F = 1$ if the energy separation between the $^2P_{3/2}, F = 2$ and $^2P_{3/2}, F = 1$ hyperfine states is $\Delta E_{\rm hf} \ll \Delta E$. Assume that the hyperfine interactions are dominated by the Hamiltonian term[10]

$$H_{\rm hf} = a\vec{I} \cdot \vec{S}\,. \tag{1.166}$$

Hyperfine-interaction-induced mixing of states with different J is important, for example, for understanding hyperfine structure splittings in situations where the fine structure intervals are relatively small. This occurs for certain excited states in He and He-like atoms [Bethe and Salpeter (1977), Section 44]. Another example

[10] In general, the expression for the hyperfine Hamiltonian describing the interaction of the magnetic moment of the nucleus with a single atomic electron is (Sobelman 1992, Section 6.2.2)

$$H_{\rm hf} = a_l \vec{l} \cdot \vec{I} - a_l[\vec{s} - 3(\vec{s} \cdot \hat{r})\hat{r}] \cdot \vec{I}\,,$$

where $\hat{r}$ is a unit vector along $\vec{r}$ and a_l is a constant proportional to $\langle r^{-3} \rangle$. For multi-electron atoms, the general expression for the hyperfine splitting is composed of two parts, one proportional to $\vec{I} \cdot \vec{S}$ and another proportional to $\vec{I} \cdot \vec{L}$. The relative importance of the two parts depends on the particular configuration. For example, in a two-electron atom where one electron is in the ground state and the other is in an excited state with $l \geq 1$, the term proportional to $\vec{I} \cdot \vec{S}$ dominates (Bethe and Salpeter 1977).

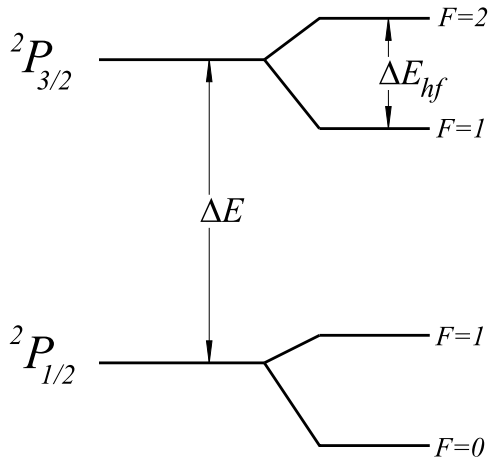

FIG. 1.10 Energy level diagram for two atomic fine structure levels, $^2P_{3/2}$ and $^2P_{1/2}$, with hyperfine splitting. Nuclear spin is $I = 1/2$.

is hyperfine-interaction-induced transitions between two levels of nominal $J = 0$ [see, for example, Fischer *et al.* (1997), Section 9.12, and Birkett *et al.* (1993)].

Hint

Note that the Hamiltonian describing the hyperfine interaction does not mix states of different F or M_F. This is because $H_{\rm hf}$ is a scalar operator (see Appendix F).

Solution

First we neglect the mixing of states with different J and relate $\Delta E_{\rm hf}$ to the coefficient a appearing in Eq. (1.166). Note that the $^2P_{3/2}, F = 2$ state corresponds to a maximum possible projection of the vectors $\vec{L}, \vec{S}$ and $\vec{I}$ onto each other. Therefore, the expectation value of $\vec{I} \cdot \vec{S}$ in this state is simply $\langle \vec{I} \cdot \vec{S} \rangle = 1/4$.

Even though the hyperfine splitting is dominated by the term $a\vec{I} \cdot \vec{S}$ in this problem, it is still the case that we can write, according to the usual formulae for hyperfine structure,

$$E_F = \frac{A}{2}[F(F+1) - J(J+1) - I(I+1)] \,, \tag{1.167}$$

$$\Delta E_{\rm hf} = E_F - E_{F-1} = AF \,, \tag{1.168}$$

where A is the hyperfine structure constant and E_F represent hyperfine energy shifts. This is because both $\vec{S}$ and $\vec{L}$ are vector operators (see Appendix F), so

$\langle\vec{S}\rangle \propto \langle\vec{J}\rangle$ and $\langle\vec{L}\rangle \propto \langle\vec{J}\rangle$, and we may write

$$E_F = \langle H_{\rm hf}\rangle = A\langle\vec{I}\cdot\vec{J}\rangle \,, \tag{1.169}$$

from which we obtain the formulae (1.167) and (1.168). In our case ($J = 3/2$, $I = 1/2$, $F = 2, 1$), so these formulae yield:

$$A = \Delta E_{\rm hf}/2 \,, \tag{1.170}$$

$$E_{F=2} = \frac{3}{8}\Delta E_{\rm hf} \,. \tag{1.171}$$

Now we use the fact that the dominant contribution to the hyperfine splitting comes from $a\vec{I}\cdot\vec{S}$ to write

$$E_{F=2} \approx a\langle\vec{I}\cdot\vec{S}\rangle \,, \tag{1.172}$$

which gives us

$$a \approx \frac{3}{2}\Delta E_{\rm hf} \,. \tag{1.173}$$

Now we evaluate the hyperfine interaction matrix element between the $F = 1$ states. As mentioned above, the interaction (1.166) can only mix states with the same value of F and M_F. We perform explicit calculations for $M_F = 1$; however, the result must not depend on M_F due to isotropy of space. We need to express the states in the $|L, M_L\rangle|M_S\rangle|M_I\rangle$ basis. We do this by first expanding into the $|J, M_J\rangle|M_I\rangle$ basis (using the Clebsch-Gordan coefficients):

$$|\,^2P_{3/2}, F = 1, M_F = 1\rangle = \frac{\sqrt{3}}{2}|3/2, 3/2\rangle|-\rangle_I - \frac{1}{2}|3/2, 1/2\rangle|+\rangle_I \,, \tag{1.174}$$

$$|\,^2P_{1/2}, F = 1, M_F = 1\rangle = |1/2, 1/2\rangle|+\rangle_I \,. \tag{1.175}$$

Next, we expand into the $|L, M_L\rangle|M_S\rangle|M_I\rangle$ basis:

$$\begin{aligned}|\,^2P_{3/2}, F = 1, M_F = 1\rangle = {}& \frac{\sqrt{3}}{2}|1, 1\rangle|+\rangle_S|-\rangle_I - \frac{1}{2\sqrt{3}}|1, 1\rangle|-\rangle_S|+\rangle_I \\ & - \frac{1}{\sqrt{6}}|1, 0\rangle|+\rangle_S|+\rangle_I \,,\end{aligned} \tag{1.176}$$

$$|\,^2P_{1/2}, F = 1, M_F = 1\rangle = \sqrt{\frac{2}{3}}|1, 1\rangle|-\rangle_S|+\rangle_I - \frac{1}{\sqrt{3}}|1, 0\rangle|+\rangle_S|+\rangle_I \,. \tag{1.177}$$

Next, we explicitly evaluate the matrix element of $\vec{I}\cdot\vec{S}$ between the states (1.176) and (1.177). One way to perform this calculation is to express $\vec{I}\cdot\vec{S}$ in

terms of the raising and lowering operators for the angular momenta $\vec{I}$ and $\vec{S}$ [see, for example, Griffiths (1995)]:

$$I_\pm = I_x \pm iI_y \,, \tag{1.178}$$

$$S_\pm = S_x \pm iS_y \,, \tag{1.179}$$

for which

$$I_\pm |I, M_I\rangle = \sqrt{I(I+1) - M_I(M_I \pm 1)}\, |I, M_I \pm 1\rangle \,, \tag{1.180}$$

$$S_\pm |S, M_S\rangle = \sqrt{S(S+1) - M_S(M_S \pm 1)}\, |S, M_S \pm 1\rangle \,. \tag{1.181}$$

Using Eqs. (1.178) and (1.179), we find that

$$\vec{I} \cdot \vec{S} = I_x S_x + I_y S_y + I_z S_z = \frac{1}{2}(I_+ S_- + I_- S_+) + I_z S_z \,. \tag{1.182}$$

Employing the expressions for the atomic states in terms of the $|L, M_L\rangle|M_S\rangle|M_I\rangle$ basis [Eqs. (1.176) and (1.177)] along with the relations (1.180) – (1.182), we obtain

$$\langle {}^2P_{3/2}, F = 1, M_F = 1 | a\vec{I} \cdot \vec{S} |\, {}^2P_{1/2}, F = 1, M_F = 1\rangle = \frac{2}{3}\frac{a}{\sqrt{2}} \,, \tag{1.183}$$

$$= \frac{\Delta E_{\rm hf}}{\sqrt{2}} \,. \tag{1.184}$$

Finally, we determine the mixed eigenstate $|\widetilde{{}^2P_{1/2}}, F = 1\rangle$ from first-order perturbation theory:

$$\boxed{|\widetilde{{}^2P_{1/2}}, F = 1\rangle \approx |\, {}^2P_{1/2}, F = 1\rangle + \frac{1}{\sqrt{2}}\frac{\Delta E_{\rm hf}}{\Delta E} |\, {}^2P_{3/2}, F = 1\rangle \,.} \tag{1.185}$$

We see that, up to a numerical coefficient, the amplitude for the mixing of states with different J is given by the ratio of the hyperfine-structure interval to the fine-structure interval.

1.12 Electron density inside the nucleus (T)

A number of phenomena in atomic physics depend on the electron density inside the nucleus – for example, hyperfine structure (see Problems 1.4 and 1.5), isotope shifts (Problem 1.9), and parity nonconservation (Problem 1.13). In this problem, we determine the scaling with Z of the electron density inside the nucleus ($r \approx$

0) for s- and p-wave valence electrons in heavy, neutral, multi-electron atoms. This result was first derived by Fermi and Segrè (1933), and the discussion below roughly follows that of Landau and Lifshitz (1977) and Khriplovich (1991).

(a) We begin by considering the properties of a valence electron.

Argue that a single valence electron in a heavy, multi-electron atom is found primarily at distances $\gtrsim a_0$ from the nucleus. Also note that the core electrons are found at distances $\lesssim a_0$.

Solution

Let us imagine that we build up the multi-electron atom by starting with a bare nucleus and adding one electron at a time. The first electron added will create a hydrogenic ion, and so the electron's average distance $\langle r \rangle$ from the nucleus will be $\approx a_0/Z$ as discussed in Problem 1.5. As we continue to add electrons, because the nucleus is shielded by the electrons already in place, each successive electron added is more weakly bound than the previous. According to this simple picture, we conclude that $\langle r \rangle$ has a larger value for the valence electron than for any of the other (core) electrons, i.e., the valence electron is indeed the "outer" electron.[11] Just before the final, valence electron is added, we have a nucleus of charge $+Ze$ surrounded by an approximately spherical distribution of $Z - 1$ core electrons – from a large distance, this looks just like the nucleus of a hydrogen atom. Therefore we expect the valence electron to orbit at a distance $\sim a_0$ from the nucleus.

Also, we can draw from our knowledge of chemical bonds and radii of different atoms to note that there is not a significant change in atomic radii as a function of Z and bond lengths are usually a few Bohr radii. Since it is the valence electrons that determine the chemical properties of an element and define the radius of the atom, we can say from these empirical facts that the valence electron spends most of its time at distances $\gtrsim a_0$.

This conclusion can also be arrived at using the Thomas-Fermi model (Problem 1.7).

(b) Based on part (a), what can one say about the wavefunction of the valence s-wave electron in the region $\gtrsim a_0$ from the nucleus?

[11] When there are multiple valence electrons, this argument does not necessarily hold for some of the electrons. For instance, in the complicated valence shells of the transition metals and rare earth atoms, valence d and f electrons are actually held more tightly to the nucleus than the valence s-wave electrons. This is the reason that the chemical properties of the rare earths are all rather similar (determined predominantly by the valence s-wave electrons).

Solution

We have learned that a valence electron spends most of its time beyond a_0 and most core electrons are at distances $< a_0$. The core electrons screen the nuclear charge in the region $r \gtrsim a_0$, so the wavefunction for a valence s electron beyond a_0 is similar to that for the hydrogen $1s$ state (since the nucleus plus core electrons in this range "appears" to be a nucleus with charge $+e$):

$$\boxed{\psi_s(r) \approx Ce^{-r/a_0} \ ,} \tag{1.186}$$

where C is a constant approximately independent of Z. The Z-independence of C follows from the fact that

$$4\pi \int_0^\infty |\psi_s(r)|^2 \, r^2 dr \sim \pi C^2 a_0^3 \sim 1 \ . \tag{1.187}$$

Although we will see that there is some nonzero probability for an electron to be found near the nucleus, because the valence electron is found predominantly at distances $\gtrsim a_0$ from the nucleus, equation (1.187) is generally a good approximation.

(c) Now we consider the region close to the nucleus. At what distance from the nucleus is the nuclear charge completely unscreened by core electrons? What is the form of the valence s electron wavefunction in this region?

Solution

We know that the radius of a hydrogenic ion is a_0/Z (see Problem 1.5), and additional electrons will have higher energy and therefore be $\gtrsim a_0/Z$ away from the nucleus, so for

$$\boxed{r \lesssim \frac{a_0}{Z}} \tag{1.188}$$

we can regard the nucleus as completely unscreened.[12] In this region,

$$\boxed{\psi_s(r) = Ae^{-Zr/a_0} \ ,} \tag{1.189}$$

where the constant A has Z dependence which will be determined later.

[12] In fact, the Thomas-Fermi model shows that nuclear screening is already negligible for $r \ll a_0 Z^{-1/3}$, a much larger radius, but Eq. (1.188) is sufficient for our purposes.

(d) In between these two regions, show that the electron is quasiclassical (satisfying the conditions for application of the WKB approximation[13]), and determine the form of the wavefunction in this region in terms of r and the momentum of the electron p.

Solution

In order to satisfy the assumptions of the WKB approximation, the wavefunction must oscillate many times within a distance where the potential energy changes significantly. This means that the deBroglie wavelength of the electron,

$$\lambda_{\text{dB}} = 2\pi\frac{\hbar}{p} \, , \tag{1.190}$$

must change slowly with respect to the distance from the nucleus:

$$\frac{\partial\lambda_{\text{dB}}}{\partial r} \ll 1 \, . \tag{1.191}$$

The WKB approximation does not apply for

$$r \gtrsim a_0 \, , \tag{1.192}$$

since in this region the electron wavefunction does not oscillate.

In the region $r \ll a_0$, the nuclear charge is not well screened by the core electrons, so we may crudely estimate that the effective nuclear charge $Z_{\text{eff}}(r) \sim Z$. Since the total energy of the valence electron is rather small (~ 1 eV), we estimate that the kinetic energy of the electron in this region is about equal to Ze^2/r. Employing the classical approximation for the momentum of the electron in this region

$$p(r) \sim \sqrt{\frac{mZe^2}{r}} \, . \tag{1.193}$$

Thus the deBroglie wavelength satisfies

$$\lambda_{\text{dB}}(r) \sim \sqrt{\frac{ra_0}{Z}} \, , \tag{1.194}$$

so that

$$\frac{\partial\lambda_{\text{dB}}}{\partial r} \sim \sqrt{\frac{a_0}{Zr}} \, . \tag{1.195}$$

[13] This approximation is named for Wentzel, Kramers, and Brillouin, who were the first to apply the method of "short wavelength asymptotics," commonly used in optics, to quantum mechanics [see, for example, Griffiths (1995), Bransden and Joachain (1989), or Landau and Lifshitz (1977)]. The method can, in fact, be applied to any wave system.

Comparing expressions (1.191) and (1.195), and noting again that the WKB approximation does not apply for $r \gtrsim a_0$ (1.192), we see that the condition for application of the WKB approximation (1.191) is satisfied as long as

$$\boxed{\frac{a_0}{Z} \ll r \lesssim a_0 \, .} \tag{1.196}$$

What does this quasiclassical wavefunction look like? If we introduce a radial function $u(r)$ such that

$$\psi_s(r) = \frac{u(r)}{r} \, , \tag{1.197}$$

the radial part of the three-dimensional Schrödinger equation reads

$$\frac{d^2u}{dr^2} + \frac{2m}{\hbar^2}[E - V_{\text{eff}}]u(r) = 0 \, , \tag{1.198}$$

which is just the one-dimensional Schrödinger equation. The quasiclassical solutions for the 1D case have the form [Griffiths (1995), for example]

$$u(r) \approx \frac{B}{\sqrt{p}} e^{\pm (i/\hbar) \int p(r) dr} \, , \tag{1.199}$$

where B is a constant. For this problem, we are not interested in the phase factor, so we say that in the quasiclassical region

$$\boxed{\psi_s(r) \sim \frac{B}{r\sqrt{p}} \, .} \tag{1.200}$$

(e) Using the results above, determine the scaling of $|\psi_s(0)|^2$ with Z.

Solution

To determine the Z-scaling of $|\psi_s(0)|^2$, we patch the three solutions for $\psi_s(r)$ obtained in parts (b), (c), and (d) together. At $r \sim a_0$, from Eqs. (1.186) and (1.200), we have

$$\frac{B}{a_0\sqrt{\alpha m c}} \sim Ce^{-1} \, . \tag{1.201}$$

Since C is independent of Z, B must also be independent of Z.

At $r \sim a_0/Z$, we have

$$\frac{ZB}{a_0\sqrt{Z\alpha mc}} \sim Ae^{-1} \,, \tag{1.202}$$

where we made use of the fact that $p \sim Z\alpha mc$ (the momentum for a hydrogenic ion) near $r \sim a_0/Z$. Equation (1.202) means that $A \propto \sqrt{Z}$, and consequently,

$$\boxed{|\psi_s(0)|^2 \approx \frac{Z}{a_0^3} \,.} \tag{1.203}$$

(f) Use similar reasoning to obtain the approximate form of $\psi_p(\vec{r})$ close to the nucleus.

Solution

For the p-state, we can apply similar reasoning to that employed in finding the wavefunction for the s-state. Ignoring angular factors and keeping only terms up to $\mathcal{O}(r)$ (since we are interested in the behavior of the wavefunction at $r \approx 0$), we see that the p wavefunction should behave as

$$\psi_p(r) \sim C_p \frac{r}{a_0} e^{-r/a_0} \qquad r \gtrsim a_0 \,, \tag{1.204}$$

$$\sim \frac{B_p}{r\sqrt{p}} \qquad \frac{a_0}{Z} \lesssim r \lesssim a_0 \,, \tag{1.205}$$

$$\sim A_p \frac{r}{a_0} e^{-Zr/a_0} \qquad r \lesssim \frac{a_0}{Z} \,. \tag{1.206}$$

As for the s-state [part (b) of this problem], because of the shielding of the nuclear charge by the core electrons, we expect C_p to be approximately independent of Z. Matching the solutions for $\psi_p(r)$ at $r \sim a_0$ shows that the constant B_p is also independent of Z. Matching the solutions for $\psi_p(r)$ at $r \sim a_0/Z$ requires that

$$A_p \sim \left(\frac{Z}{a_0}\right)^{3/2} \,. \tag{1.207}$$

Thus, for $r \ll a_0/Z$,

$$\boxed{\psi_p(r) \sim \left(\frac{Z}{a_0}\right)^{3/2} \frac{r}{a_0} e^{-Zr/a_0} \,.} \tag{1.208}$$

1.13 Parity nonconservation in atoms

Before the mid-1950s, physicists believed that the laws of Nature were invariant with respect to *spatial inversion* (also known as the parity (P) transformation, which reverses the directions of all three spatial axes). Spatial inversion changes the *handedness* of an object or a process (a term stemming from the fact that P turns a glove for the left hand into a glove for the right hand). Spatial inversion is an example of a *discrete* transformation, as opposed to a continuous transformation (for example, a finite rotation can be thought of as a succession of infinitely small rotations).

In 1956, the belief in P-invariance was shattered by a series of experiments, where violation of this symmetry at a $\sim 100\%$ level was discovered in *weak-interaction-induced* nuclear decays [see, for example, Trigg (1975), chapter 10].

When Glashow (1961), Weinberg (1967), and Salam (1968) developed what came to be known as the Standard Model of electroweak interactions, they predicted the existence of a neutral weak interaction mediated by a particle called the Z_0 boson. Even before the advent of the Standard Model, Zel'dovich (1959) had noted that if there were a parity-violating electron-nucleon weak neutral-current interaction, it would interfere with the regular electromagnetic interaction in an atom. Zel'dovich's estimate for the size of this effect, however, indicated that it would be too small to measure with experimental techniques available at that time.

Later, motivated by the new developments in weak interaction theory and the tremendous advances in laser spectroscopy, Bouchiat and Bouchiat (1974) reanalyzed the possibility of searching for parity nonconservation (PNC) in atomic systems. They found that in fact the effects were considerably enhanced in heavy atoms, and according to the predictions of the Standard Model, should be observable. Based on this new analysis, many groups throughout the world began extensive experimental efforts to search for PNC effects in atoms.

The first experiments to observe PNC effects in atoms were carried out by Barkov and Zolotorev (1978) in Novisibirsk using the technique of *optical rotation* in bismuth and by Commins and co-workers (Conti *et al.* 1979) at Berkeley using the *Stark-interference method* (see Problem 4.5) in thallium. These experiments provided crucial evidence that helped establish the existence of the neutral weak current and were the first indications of the parity-violating nature of the neutral weak interaction.

At present, atomic PNC experiments continue to serve as stringent tests of the fundamental theory of electroweak interactions and as sensitive probes for new physics (Khriplovich 1991; Bouchiat and Bouchiat 1997; Budker 1998b). The most precise measurement of PNC effects in an atomic system to date was performed in Boulder by Wieman and collaborators (Wood *et al.* 1997), using the Stark-interference technique with cesium. This experiment also was the first to definitively observe nuclear-spin-dependent PNC effects, primarily due to the

nuclear anapole moment (discussed in detail in Problem 1.15). In this problem we restrict our considerations to nuclear-spin-independent PNC effects.

(a) In the nonrelativistic approximation and the limit of infinite Z_0 mass, the Hamiltonian H_w describing the weak interaction between the nucleus and a single electron is given by (Bouchiat and Bouchiat 1974):

$$H_w = \frac{G_F}{\sqrt{2}} \frac{1}{2mc\hbar} Q_w \vec{s} \cdot \left[\vec{p}\delta^3(\vec{r}) + \delta^3(\vec{r})\vec{p} \right] , \tag{1.209}$$

where we have ignored nuclear spin-dependent effects and

$$G_F \approx 3 \times 10^{-12} mc^2 \left(\frac{\hbar}{mc} \right)^3 \tag{1.210}$$

is Fermi's constant, $\vec{s}$ is the electron spin, $\vec{p}$ is the electron momentum, and

$$Q_w = -N + (1 - 4\sin^2\theta_w)Z \tag{1.211}$$

is the dimensionless *weak nuclear charge* (N is the number of neutrons and $\sin^2\theta_w \approx 0.23$ where θ_w is the Weinberg mixing angle).

Show that H_w violates parity.

(b) From the mass of the Z_0 ($m_Z \approx 92.6\ \mathrm{GeV}/c^2$), estimate the range of the neutral weak interaction.

(c) Consider the mixing of the opposite parity states $|ns_{1/2}\rangle$ and $|n'p_{1/2}\rangle$ (these are single-particle states for the valence electron), where n, n' are the principal quantum numbers. Show that the mixing is enhanced in heavy atoms by a factor proportional to Z^3.

(d) What is the significance of the fact that the matrix element $\langle n'p_{1/2}|H_w|ns_{1/2}\rangle$ is imaginary?

(e) Calculate the PNC-induced mixing between the $2S_{1/2}$ and $2P_{1/2}$ states in hydrogen. Recall that the states are split in energy only by the Lamb shift ($\Delta E \approx 1058\ \mathrm{MHz}$).

(f) Based on the calculation in part (e) for hydrogen, estimate the order of magnitude of the PNC-induced mixing between the $6s_{1/2}$ and $6p_{1/2}$ states in Cs.

Hints

In part (d), consider the fact that although H_w violates parity it does respect time-reversal invariance (T). In particular, consider the application of an electric field to a state that is not an eigenstate of parity, and show that if the matrix element $\langle n'p_{1/2}|H_w|ns_{1/2}\rangle$ is not purely imaginary the system is T-violating.

Solution

(a) To prove that H_w violates parity, it is sufficient to show that it does not commute with the parity operator P. If $[H_w, P] \neq 0$, then the energy eigenstates of the atomic system are not in general also be eigenstates of the parity operator. In this case, a state with definite parity must be a superposition of different energy eigenstates. Therefore, if the system is left to evolve in time, the parity of the state will change – i.e., parity is not conserved.

How does H_w transform under the action of P? Because $H_w \propto \vec{s} \cdot \vec{p}$, and $\vec{s}$, like orbital angular momentum, is an axial vector (pseudovector) which does not change sign under P while $\vec{p}$ is a polar vector which does change sign:

$$P^\dagger H_w P = -H_w \ . \tag{1.212}$$

Operating P on both sides of Eq. (1.212), we find that

$$H_w P = -P H_w \ , \tag{1.213}$$

so H_w and P anticommute, which means that

$$[H_w, P] = H_w P - P H_w = 2 H_w P \neq 0 \ . \tag{1.214}$$

(b) In order for a Z_0 to be exchanged between an electron and the nucleus, it must occur on a time scale sufficiently short so as not to violate energy conservation. According to the Heisenberg uncertainty relation,

$$\Delta E \Delta t \sim \hbar \ . \tag{1.215}$$

Thus the range R of the neutral weak force is

$$R \sim c\Delta t \sim \frac{\hbar c}{\Delta E} \ . \tag{1.216}$$

We let ΔE be the minimum energy required for the creation of a Z_0, $m_Z c^2$. Then

$$\boxed{R \sim \frac{\hbar c}{m_Z c^2} \sim \frac{197.3 \text{ MeV} \cdot \text{fm}}{92.6 \times 10^3 \text{ MeV}} \sim 2 \times 10^{-3} \text{ fm} \ .} \tag{1.217}$$

Treating H_w as a point-like interaction is a very good approximation in atomic physics.

(c) To first order in time-independent perturbation theory, the weak Hamiltonian admixes some of the $|n'p\rangle$ state into an $|ns\rangle$ state according to

$$|\widetilde{ns}_{1/2}\rangle = |ns_{1/2}\rangle + \frac{\langle n'p_{1/2}|H_w|ns_{1/2}\rangle}{\Delta E}|n'p_{1/2}\rangle \ , \tag{1.218}$$

where $\Delta E = E_s - E_p$ is the energy separation between the states (where E_s is the energy of $|ns_{1/2}\rangle$ and E_p is the energy of $|n'p_{1/2}\rangle$). Because, as discussed in part

(b), the interaction is point-like, the mixing between the s and p states depends on the s and p wavefunctions near the nucleus. From Problem 1.12, we have for the s state in the region $r \lesssim a_0/Z$

$$\psi_s(r) \sim \frac{\sqrt{Z}}{a_0^{3/2}}\, e^{-Zr/a_0}\ , \tag{1.219}$$

where the factor of $a_0^{-3/2}$ gives the wavefunction the correct dimensions. For the p-state in the same region,

$$\psi_p(r) \sim \left(\frac{Z}{a_0}\right)^{3/2} \frac{r}{a_0} e^{-Zr/a_0}\ . \tag{1.220}$$

The weak Hamiltonian (1.209) has two terms. The second term does not contribute to the matrix element because

$$\langle n' p_{1/2} | \delta^3(\vec{r}) \vec{s} \cdot \vec{p} | n s_{1/2} \rangle = 0\ . \tag{1.221}$$

This follows from the fact that an electron in an s-wave state has zero momentum at the origin, as well as the fact that the p wavefunction has a value of zero at $r = 0$. Thus only the first term contributes, and so the matrix element between the $|n'p\rangle$ state and the $|ns\rangle$ state is given by

$$\langle n' p_{1/2} | H_w | n s_{1/2} \rangle \sim \frac{G_F Q_w}{mc} \int \psi_s(r) \left(\delta^3(r) \frac{\hbar}{i} \frac{\partial}{\partial r}\right) \psi_p(r)\ d^3 r \tag{1.222}$$

$$\sim -i \frac{G_F Q_w \hbar}{mc} \psi_s(0) \frac{\partial \psi_p}{\partial r}\Big|_0\ . \tag{1.223}$$

In the above estimate, we have set $\vec{s} \cdot \vec{p} = \hbar p_r/2$, since details of the angular distribution of $\vec{p}$ do not affect the Z-scaling.

From Eq. (1.220), we see that

$$\frac{\partial \psi_p}{\partial r}\Big|_0 \sim \frac{Z^{3/2}}{a_0^{5/2}} \tag{1.224}$$

and since

$$\psi_s(0) \sim \frac{\sqrt{Z}}{a_0^{3/2}}\ , \tag{1.225}$$

then using Eq. (1.223), we have

$$\langle n' p_{1/2} | H_w | n s_{1/2} \rangle \sim -i \frac{G_F Q_w \hbar}{mc a_0^4} Z^2\ . \tag{1.226}$$

The weak charge Q_w is roughly proportional to the atomic number, so

$$\boxed{\langle n'p_{1/2}|H_w|ns_{1/2}\rangle \propto Z^3 \ .} \tag{1.227}$$

Usually this result is understood as a consequence of the electron density inside the nucleus for s-states, $|\psi_s(0)|^2$, scaling as Z (Problem 1.12), the momentum of the electron (which enters H_w) near the nucleus (where there is no shielding) scaling as Z, and the scaling of Q_w with Z. This explanation is equivalent to our considerations above if one takes into account the fact that the momentum operator acting on the p-wave state yields

$$\frac{\hbar}{i}\frac{\partial\psi_p}{\partial r}\Big|_{r=0} \sim Z\frac{\hbar}{a_0}\psi_s(0) \ . \tag{1.228}$$

The observation that the PNC-induced mixing is inversely proportional to the energy difference between the states [Eq. (1.218)] has led to the suggestion to measure PNC in atoms with nearly degenerate opposite parity levels: the hydrogen $2s$-$2p$ system [Hinds (1988) and references therein] and heavy elements such as the rare earth atoms (Dzuba *et al.* 1986), e.g., samarium (Barkov *et al.* 1989; Wolfenden and Baird 1993), dysprosium (Budker *et al.* 1994; Nguyen *et al.* 1997), and ytterbium (DeMille 1995).

(d) The PNC-induced mixing between the $|n'p\rangle$ and $|ns\rangle$ states is imaginary [Eq. (1.226)] because the weak interaction violates parity but not time-reversal invariance (T).

Consider the application of a static electric field (along z) to an atom in the state $|\widetilde{ns}_{1/2}\rangle$ [see Eq. (1.218)]. The Hamiltonian governing the interaction of the electric field with the atom is (in the single-electron approximation)

$$H_1 = -\vec{d}\cdot\vec{E} = ezE_0 \ , \tag{1.229}$$

where $\vec{d}$ is the electric dipole operator and E_0 is the magnitude of the electric field $\vec{E}$. The first-order energy shift due to the electric field is

$$\Delta E^{(1)} = eE_0\langle\widetilde{ns}_{1/2}|z|\widetilde{ns}_{1/2}\rangle \tag{1.230}$$

$$= eE_0\big(\langle ns_{1/2}| - i\eta\langle n'p_{1/2}|\big)z\big(|ns_{1/2}\rangle + i\eta|n'p_{1/2}\rangle\big) \ , \tag{1.231}$$

where we have implicitly assumed that the states $|ns_{1/2}\rangle$ and $|n'p_{1/2}\rangle$ have the same projection of total angular momentum on the quantization axis [so that they can be coupled by the Hamiltonian H_1 (1.229)] and $i\eta$ is the PNC-induced mixing

amplitude [Eqs. (1.218) and (1.226)]:

$$i\eta = \frac{\langle n'p_{1/2}|H_w|ns_{1/2}\rangle}{\Delta E} \,. \tag{1.232}$$

Since z only connects states of opposite parity, we have

$$\Delta E^{(1)} = eE_0\langle ns_{1/2}|z|n'p_{1/2}\rangle(-i\eta + i\eta) = 0 \,, \tag{1.233}$$

where we have made use of the fact that

$$\langle n'p_{1/2}|z|ns_{1/2}\rangle = \langle ns_{1/2}|z|n'p_{1/2}\rangle \,. \tag{1.234}$$

Thus there is no linear Stark shift because $i\eta$ is purely imaginary.

If $i\eta$ had a real part, $\Delta E^{(1)}$ would be nonzero, which would violate T-invariance. To see this, we note that if $\Delta E^{(1)} = -\langle \vec{d}\cdot\vec{E}\rangle \neq 0$, then $\langle \vec{d}\rangle \neq 0$. The Wigner-Eckart theorem (Appendix F) states that any vector quantity must be proportional to the total angular momentum of the system, so:

$$\langle \vec{d}\rangle \propto \langle \vec{J}\rangle \,. \tag{1.235}$$

However, the time reversal operator T takes $\vec{d} \to \vec{d}$ while it takes $\vec{J} \to -\vec{J}$, so the existence of a "permanent" electric dipole moment (one that exists even in the absence of an applied electric field) violates T-invariance (see Problem 4.8 for further discussion).

(e) The weak Hamiltonian is a scalar operator, so it only couples states having the same total angular momentum J and the same projection of angular momentum M_J. Thus we will consider the PNC-induced mixing between the $M_J = 1/2$ states,

$$i\eta = \frac{\langle 2P_{1/2}\ M_J = 1/2|H_w|2S_{1/2}\ M_J = 1/2\rangle}{\Delta E} \,, \tag{1.236}$$

realizing that, because of rotational invariance, the $M_J = -1/2$ states have the same mixing amplitude.

To simplify the notation, taking into account Eq. (1.221), we may write

$$H_w = \beta\vec{s}\cdot\vec{p}\,\delta^3(\vec{r}) \,, \tag{1.237}$$

where

$$\beta = \frac{G_F}{\sqrt{2}}\frac{1}{2mc\hbar}Q_w \,. \tag{1.238}$$

The S and P states can be expressed as products of the spin wavefunctions, $|+\rangle$ and $|-\rangle$, and the spatial wavefunctions, $|n,l,m_l\rangle$, using Clebsch-Gordan coefficients:

$$|2S_{1/2}\ M_J = 1/2\rangle = |2,0,0\rangle|+\rangle\ , \tag{1.239}$$

$$|2P_{1/2}\ M_J = 1/2\rangle = \sqrt{\frac{2}{3}}\,|2,1,1\rangle|-\rangle - \sqrt{\frac{1}{3}}\,|2,1,0\rangle|+\rangle\ . \tag{1.240}$$

First we will find $\vec{s}\cdot\vec{p}\,|2P_{1/2}\ M_J = 1/2\rangle$. The operator $\vec{s}\cdot\vec{p}$, represented in the spin basis using the Pauli matrices, is given by

$$\vec{s}\cdot\vec{p} = s_x p_x + s_y p_y + s_z p_z \tag{1.241}$$

$$= \frac{\hbar}{2}\left[\begin{pmatrix} 0 & p_x \\ p_x & 0 \end{pmatrix} + \begin{pmatrix} 0 & -ip_y \\ ip_y & 0 \end{pmatrix} + \begin{pmatrix} p_z & 0 \\ 0 & -p_z \end{pmatrix}\right] \tag{1.242}$$

$$= \frac{\hbar}{2}\begin{pmatrix} p_z & p_x - ip_y \\ p_x + ip_y & -p_z \end{pmatrix}\ . \tag{1.243}$$

Thus, in the spin basis, we have

$$\vec{s}\cdot\vec{p}\,|2P_{1/2}\ M_J = 1/2\rangle = \frac{\hbar}{2}\begin{pmatrix} p_z & p_x - ip_y \\ p_x + ip_y & -p_z \end{pmatrix}\begin{pmatrix} -\sqrt{\frac{1}{3}}\,|2,1,0\rangle \\ \sqrt{\frac{2}{3}}\,|2,1,1\rangle \end{pmatrix}. \tag{1.244}$$

Substituting this into Eq. (1.236), we have

$$i\eta\Delta E = \frac{\hbar\beta}{2}\times \left(-\sqrt{\frac{1}{3}}\,\langle 2,1,0|p_z\delta^3(\vec{r})|2,0,0\rangle + \sqrt{\frac{2}{3}}\,\langle 2,1,1|(p_x+ip_y)\delta^3(\vec{r})|2,0,0\rangle\right). \tag{1.245}$$

In the above equation, note that $(\langle 2P_{1/2}\ M_J = 1/2|\ \vec{s}\cdot\vec{p})$, which appears in the formula for η (1.236) when expression (1.237) for H_w is used, is obtained from (1.244) by Hermitian conjugation. Next we take into account the fact that in the spherical basis (see Appendix F)

$$p_1 = -\frac{1}{\sqrt{2}}(p_x + ip_y) \tag{1.246}$$

$$p_0 = p_z \tag{1.247}$$

$$p_{-1} = \frac{1}{\sqrt{2}}(p_x - ip_y)\ , \tag{1.248}$$

so

$$i\eta\Delta E = -\frac{\hbar\beta}{2}\left(\frac{1}{\sqrt{3}}\,\langle 2,1,0|p_0\delta^3(\vec{r})|2,0,0\rangle + \frac{2}{\sqrt{3}}\,\langle 2,1,1|p_1\delta^3(\vec{r})|2,0,0\rangle\right). \tag{1.249}$$

Now we can apply the Wigner-Eckart theorem (Appendix F):

$$\langle 2,1,0|p_0\delta^3(\vec{r})|2,0,0\rangle = \frac{1}{\sqrt{3}}\langle 2,1||p\delta^3(\vec{r})||2,0\rangle\langle 0,0,1,0|1,0\rangle\ , \tag{1.250}$$

$$\langle 2,1,1|p_1\delta^3(\vec{r})|2,0,0\rangle = \frac{1}{\sqrt{3}}\langle 2,1||p\delta^3(\vec{r})||2,0\rangle\langle 0,0,1,1|1,1\rangle\ , \tag{1.251}$$

where $\langle l_1, m_1, \kappa, q|l_2, m_2\rangle$ is the appropriate Clebsch-Gordan coefficient and $\langle n_2, l_2||p\delta^3(\vec{r})||n_1, l_1\rangle$ is the reduced matrix element. Both Clebsch-Gordan coefficients in (1.251) equal one, so using the fact that

$$\langle 2,1,0|p_0\delta^3(\vec{r})|2,0,0\rangle = \langle 2,1,1|p_1\delta^3(\vec{r})|2,0,0\rangle\ , \tag{1.252}$$

in Eq. (1.249) we have

$$\eta\Delta E = -\frac{\hbar\beta\sqrt{3}}{2i}\langle 2,1,0|p_0\delta^3(\vec{r})|2,0,0\rangle\ . \tag{1.253}$$

Now we use the hydrogen wavefunctions in order to carry out the integral associated with the above matrix element. We begin by finding $p_0\psi_{210}(\vec{r})$. The momentum operator in the z-direction is given by

$$p_0 = p_z = \frac{\hbar}{i}\frac{\partial}{\partial z}\ . \tag{1.254}$$

In order to differentiate the hydrogen wavefunctions, which are most conveniently expressed in spherical coordinates, with respect to z, we need to express $\frac{\partial}{\partial z}$ in

spherical coordinates:[14]

$$\frac{\partial}{\partial z} = \cos\theta \frac{\partial}{\partial r} - \frac{\sin\theta}{r}\frac{\partial}{\partial \theta} \ . \tag{1.260}$$

The hydrogen wavefunction $\psi_{210}(r,\theta,\phi)$ is

$$\psi_{210}(r,\theta,\phi) = \frac{1}{4\sqrt{2\pi}} \frac{1}{a_0^{3/2}} \frac{r}{a_0} e^{-r/2a_0} \cos\theta \ . \tag{1.261}$$

Employing Eqs. (1.260) and (1.261), we find that

$$p_0 \psi_{210}(\vec{r}) = \frac{\hbar}{i}\frac{\partial}{\partial z}\psi_{210}(\vec{r}) = \frac{-i\hbar}{4\sqrt{2\pi}} \frac{1}{a_0^{5/2}} \left(1 - \frac{r}{2a_0}\cos^2\theta\right) e^{-r/2a_0} \ . \tag{1.262}$$

Now we are prepared to evaluate the integral in Eq. (1.253). Noting that we must take the complex conjugate of (1.262) and using the hydrogen wavefunction $\psi_{200}(r,\theta,\phi)$

$$\psi_{200}(r,\theta,\phi) = \frac{1}{2\sqrt{2\pi}} \frac{1}{a_0^{3/2}} \left(1 - \frac{r}{2a_0}\right) e^{-r/2a_0} \ , \tag{1.263}$$

we obtain

$$\eta \Delta E = -\frac{\hbar^2 \beta \sqrt{3}}{32\pi a_0^4} \int \left(1 - \frac{r}{2a_0}\cos^2\theta\right)\left(1 - \frac{r}{2a_0}\right) e^{-r/a_0} \delta^3(r) d^3r \tag{1.264}$$

$$= -\frac{\hbar^2 \beta \sqrt{3}}{32\pi a_0^4} \ . \tag{1.265}$$

[14] The gradient in spherical coordinates is given by

$$\vec{\nabla} = \frac{\partial}{\partial r}\hat{r} + \frac{1}{r}\frac{\partial}{\partial \theta}\hat{\theta} + \frac{1}{r\sin\theta}\frac{\partial}{\partial \phi}\hat{\phi} \, , \tag{1.255}$$

while in Cartesian coordinates it is

$$\vec{\nabla} = \frac{\partial}{\partial x}\hat{x} + \frac{\partial}{\partial y}\hat{y} + \frac{\partial}{\partial z}\hat{z} \, . \tag{1.256}$$

The spherical basis is related to the Cartesian basis by:

$$\hat{r} = \sin\theta\cos\phi\hat{x} + \sin\theta\sin\phi\hat{y} + \cos\theta\hat{z} \, , \tag{1.257}$$

$$\hat{\theta} = \cos\theta\cos\phi\hat{x} + \cos\theta\sin\phi\hat{y} - \sin\theta\hat{z} \, , \tag{1.258}$$

$$\hat{\phi} = -\sin\phi\hat{x} + \cos\phi\hat{y} \, . \tag{1.259}$$

Using the above relations in (1.255), we find for the z-component of the gradient the expression given in Eq. (1.260).

Using Eqs. (1.238) and (1.210) in (1.265), we find

$$\eta\Delta E = -\sqrt{\frac{3}{2}}\,\frac{G_F Q_w}{64\pi}\,\frac{\hbar}{mca_0^4} \tag{1.266}$$

$$= -\sqrt{\frac{3}{2}}\,\frac{3\times 10^{-12}}{64\pi} Q_w \frac{\hbar c}{a_0^4}\left(\frac{\hbar}{mc}\right)^3 \tag{1.267}$$

$$= -\sqrt{\frac{3}{2}}\,\frac{3\times 10^{-12}}{64\pi} Q_w \alpha^4 mc^2\ . \tag{1.268}$$

The weak charge for hydrogen, according to (1.211), is

$$Q_w = 1 - 4\sin^2\theta_w \approx 0.08\ . \tag{1.269}$$

Thus

$$\eta\Delta E \approx -2\times 10^{-18}\ \mathrm{eV}\ . \tag{1.270}$$

The energy splitting between the $2S_{1/2}$ and $2P_{1/2}$ states is given by the Lamb shift:

$$\Delta E \approx 1058\ \mathrm{MHz} \approx 4\times 10^{-6}\ \mathrm{eV}\ . \tag{1.271}$$

Therefore the PNC-induced mixing between the $2S_{1/2}$ and $2P_{1/2}$ states in hydrogen is

$$\boxed{\eta \approx -5\times 10^{-13}\ .} \tag{1.272}$$

Note that this effect is about an order of magnitude larger in deuterium, where the weak charge is $Q_w \approx -1$. Experimental efforts to measure PNC effects in hydrogen are reviewed by Hinds (1988).

(f) As a rough estimate, we will say that the only differences between the PNC-induced mixing of states in hydrogen and the mixing in cesium are Z, Q_w, and the energy level splitting. In this approximation, the amplitude $\eta(\mathrm{Cs})$ of the PNC-induced mixing between the $6s$ and $6p$ states in Cs is related to the amplitude of PNC-induced mixing of the $2s$ and $2p$ states in hydrogen, $\eta(H)$, by

$$\eta(\mathrm{Cs}) \sim \eta(\mathrm{H})\frac{Z^2 Q_w(\mathrm{Cs})}{Q_w(\mathrm{H})}\,\frac{\Delta E_{\mathrm{H}}(2s,2p)}{\Delta E_{\mathrm{Cs}}(6s,6p)}\ , \tag{1.273}$$

where $Z = 55$ is the atomic number of Cs, $\Delta E_{\mathrm{H}}(2s,2p) \approx 4\times 10^{-6}$ eV is the splitting of the $2S_{1/2}$ and $2P_{1/2}$ states in hydrogen due to the Lamb shift, and

$\Delta E_{\rm Cs}(6s, 6p)$ is the energy separation between the $6s$ and $6p$ levels in Cs,

$$\Delta E_{\rm Cs}(6s, 6p) \sim -10^4 \text{ cm}^{-1} \sim -1.25 \text{ eV} . \tag{1.274}$$

Based on Eq. (1.211), we estimate the weak charge for Cs, $Q_w(\rm Cs)$, to be

$$Q_w(\rm Cs) \approx -75 , \tag{1.275}$$

close to the experimental result of $-72.06(46)$ (Wood *et al.* 1997), which differs from -75 due to radiative corrections.[15] Using these values in (1.273), we get

$$\boxed{\eta(\rm Cs) \sim -5 \times 10^{-12} .} \tag{1.276}$$

A more detailed analysis for Cs leads to a result about twice as large[16] [see, for example, Khriplovich (1991)].

Given the exceedingly small mixing amplitude, it is indeed a remarkable achievement that PNC in cesium has been measured to a fraction of a percent (Wood *et al.* 1997).

1.14 Parity nonconservation in anti-atoms

As discussed in Problem 1.13, parity nonconservation due to the neutral weak interaction manifests itself in atomic transitions. For example, for the highly forbidden one-photon decay of unpolarized excited hydrogen

$$|2S\rangle \to |1S\rangle + \gamma, \tag{1.277}$$

the emitted photons have a preferred circular polarization. The effect is larger for deuterium, due to its larger weak charge Q_w [Eq. (1.211)]. For deuterium, the degree of circular polarization of the photons is $\sim 2 \times 10^{-4}$.

[15] *Radiative corrections* are modifications to interactions due to higher-order processes – in the language of Feynman diagrams (Appendix H), they arise from diagrams with the same initial and final states but with more intermediate vertices than the lowest-order process [see, for example, the book by Griffiths (1987)]. For example, in the case considered here, the expression for Q_w (1.211) is based on Z_0 exchange between the nucleus and an electron. One of the Feynman diagrams describing a radiative correction to Q_w involves a Z_0 propagating from the nucleus, then turning into a top quark and anti-top quark, which subsequently annihilate to create a Z_0 which then interacts with the electron.

[16] One may note that because all the p-states in Cs are relatively far from the $6s$ state in energy, a proper calculation of PNC effects should include a sum over all p-states. However, the amplitudes of the couplings to states other than the $6p$ and $7p$ are rather small, and one can obtain a reasonably accurate prediction for the PNC-induced mixing by only including the admixtures of the $6p$ and $7p$ states into $6s$. The PNC-induced mixing between $6s$ and $7p$ is a factor of ~ 4 smaller than the $6s-6p$ mixing, so our estimate should give close to the right result for the total mixing of p-states into the $6s$ state.

As it turns out, while P-invariance is violated, the symmetry is almost restored by performing the combined transformation of spatial inversion and *charge conjugation*,[17] C. So far, the only examples of CP-violation have been found in the decays of neutral K- and B-mesons, and for the purpose of this problem, we will assume that CP is a good symmetry.

The question is: if hydrogen preferentially emits right-circularly polarized (R) photons,[18] what is the sign of the preferred circular polarization for *antihydrogen*?

Antihydrogen has been produced at CERN (G. Baur *et al.* 1996) and FermiLab (G. Blanford *et al.* 1998) using particle accelerators. The antihydrogen produced in these experiments travel nearly at the speed of light with respect to the laboratory reference frame, making it exceedingly difficult to perform precision spectroscopy. Low-energy antiprotons and cold positrons can be stored simultaneously in nested Penning traps (Gabrielse *et al.* 1999), and quite recently this method has been successfully used to produce cold antihydrogen (Amoretti *et al.* 2002, Gabrielse *et al.* 2002). This appears to be a promising route toward producing antihydrogen useful for spectroscopic tests. Antihydrogen may serve as an interesting system with which to test CPT and Lorentz invariance [see, for example, Holzscheiter and Charlton (1999) and Gabrielse (2001)].

Solution

Consider the decay (1.277) as seen in the laboratory and in its mirror image (upper part of Fig. 1.11). Mirror reflection changes the R photons into L photons (L and R designate left- and right-circular polarization, respectively), so looking in the mirror, we see that excited hydrogen preferentially decays into L photons. This actually does *not* occur in reality, which is the essence of parity nonconservation: laws of nature are not the same in the real laboratory and its counterpart in the mirror. (In this case, the law of nature says that excited hydrogen preferentially decays into R photons.)

Now, instead of a usual (P) mirror that inverts coordinates,[19] consider its C-analog. A "C-mirror" reverses all charges converting particles into antiparticles, but does not affect coordinates. Looking at the original process in such a mirror (middle part of Fig. 1.11), we see antihydrogen preferentially decaying into an R photon, and again, this does not occur in reality (C-violation).

[17] This transformation changes matter into antimatter.

[18] In this book, we use the spectroscopists' convention for left and right circular polarization, where a σ_+ photon (one with positive *helicity*, with the photon spin along its direction of propagation) is said to be left-circularly polarized, and a σ_- photon is said to be right-circularly polarized.

[19] Actually, a mirror inverts only one out of three coordinates, but this is not essential since inversion of all three coordinate axes is equivalent to inversion of just one of them followed by a rotation by 180 degrees around the inverted axis.

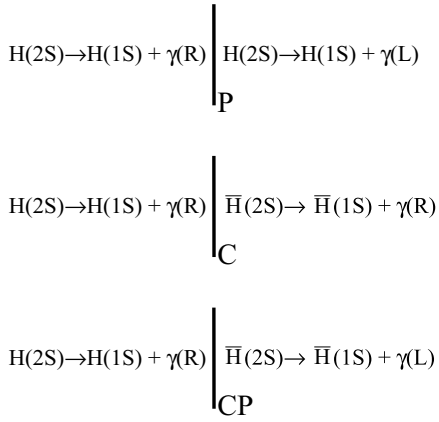

FIG. 1.11 Hydrogen $|2S\rangle \to |1S\rangle + \gamma$ decay as seen in the laboratory, and upon action of various transformations: P, C, and the combined CP-transformation. $\overline{\mathrm{H}}$ designates antihydrogen, γ(L) and γ(R) designate photons and their preferential circular polarization.

In order to determine how antihydrogen actually decays, we can use CP-invariance, and look at the original process in a "CP-mirror" (the lower part of Fig. 1.11). Because of CP-invariance, whatever we see in the "CP-mirror" actually corresponds to a real process that, at least in principle, can be reproduced in the laboratory. The conclusion is that antihydrogen preferentially decays into L photons, and the degree of circular polarization is the same as for hydrogen.

The same result can be obtained more formally. First consider hydrogen. Its Hamiltonian can be written as

$$H = H_0 + H_w, \tag{1.278}$$

where H_0 is the parity-conserving part of the Hamiltonian, and H_w is the parity-violating part. As we have seen in Problem 1.13, the parity-violating part of the Hamiltonian leads to mixing of eigenstates of H_0 which have definite parity. In particular, an eigenstate of the full Hamiltonian is now

$$|\widetilde{2S}\rangle \approx |2S\rangle + \frac{\langle 2P|H_w|2S\rangle}{E_{2S} - E_{2P}}|2P\rangle \equiv |2S\rangle \ + i\eta|2P\rangle. \tag{1.279}$$

Here we used first-order perturbation theory, E_{2S} and E_{2P} are the energies of the corresponding eigenstates of H_0, and we neglected mixing of $|2S\rangle$ with all states other than the nearby $|2P\rangle$. As we have seen in Problem 1.13, the constant $i\eta$ turns out to be imaginary as a consequence of T-invariance. The form of Eq. (1.279) corresponds to the so-called "spin-helix," whose sign dictates the preferred handedness of the decay photon [see (Khriplovich 1991), Chapter 2 for details].

Now, let us analyze antihydrogen. In quantum field theory, antihydrogen is automatically included in the Hamiltonian, so generalizing (1.278) we write:

$$H = H_0 + H_w + \overline{H}_0 + \overline{H}_w \,, \tag{1.280}$$

i.e., we explicitly include the parity-conserving and parity nonconserving terms for antihydrogen (the bar designates quantities related to antihydrogen). CP-invariance means that the full Hamiltonian (1.280) is invariant under CP, so

$$(CP)H_0(CP)^{-1} = \overline{H}_0; \quad (CP)H_w(CP)^{-1} = \overline{H}_w. \tag{1.281}$$

The analog of Eq. (1.279) for antihydrogen is:

$$|\widetilde{\overline{2S}}\rangle \approx |\overline{2S}\rangle \ + \overline{i\eta}\,|\overline{2P}\rangle. \tag{1.282}$$

In order to evaluate $\overline{i\eta}$, we can start with the definition of $i\eta$, and insert the identity $(CP)^{-1}(CP) = 1$ on both sides of the operator H_w:

$$i\eta = \frac{\langle 2P|H_w|2S\rangle}{E_{2S} - E_{2P}} = \frac{\langle 2P|(CP)^{-1}(CP)H_w(CP)^{-1}(CP)|2S\rangle}{E_{2S} - E_{2P}} \tag{1.283}$$

$$= \frac{\langle -\overline{2P}|(CP)H_w(CP)^{-1}|\overline{2S}\rangle}{E_{2S} - E_{2P}} \tag{1.284}$$

$$= -\frac{\langle \overline{2P}|\overline{H}_w|\overline{2S}\rangle}{E_{\overline{2S}} - E_{\overline{2P}}} = -\overline{\delta}\,. \tag{1.285}$$

Here we used the properties of the wavefunctions $CP|2S\rangle = |\overline{2S}\rangle$, $CP|2P\rangle = -|\overline{2P}\rangle$ (C just adds a bar, and P changes the sign of odd-parity wavefunctions), and the invariance of eigenenergies of $H_0 + \overline{H}_0$ with respect to CP. This shows that the sign of parity-violating mixing for antihydrogen is opposite to that for hydrogen, which implies preferential emission of L-photons, in agreement with our earlier conclusion.

1.15 The anapole moment (T)

The *anapole* (Zel'dovich 1958) is an electromagnetic moment[20] that appears, along with the more familiar moments such as the magnetic dipole, in the multipole expansion of the vector potential of an electromagnetic current distribution of finite spatial extent, e.g., an atomic nucleus.

[20] Different electromagnetic moments correspond to the different rank tensors necessary to describe charge and current distributions.

The anapole moment $\vec{a}$ of the system is defined as

$$\vec{a} = -\pi \int r^2 \vec{j}(\vec{r}) d^3 r, \tag{1.286}$$

where $\vec{j}(\vec{r})$ is the current density. From this definition, as we will show in this problem, it follows that the anapole's contribution to the vector potential is

$$\vec{A}_a\Big(\vec{R}\Big) = \vec{a}\delta(\vec{R}), \tag{1.287}$$

indicating that in order for a charged particle to "feel" the anapole moment it has to penetrate within the current distribution (*contact interaction*).

In order to visualize the simplest systems possessing an anapole moment, first consider a current loop offset from the origin (Flambaum and Murray 1997). Straight from the definition (1.286), there is a nonzero anapole moment pointing in the direction opposite to that of the current at a point furthest from the origin. A direct generalization of this is a toroidal winding (Fig. 1.12) where the current consecutively flows through a series of such loops. Note that in this case, in contrast to a single loop, the magnetic field produced by the current is confined within the

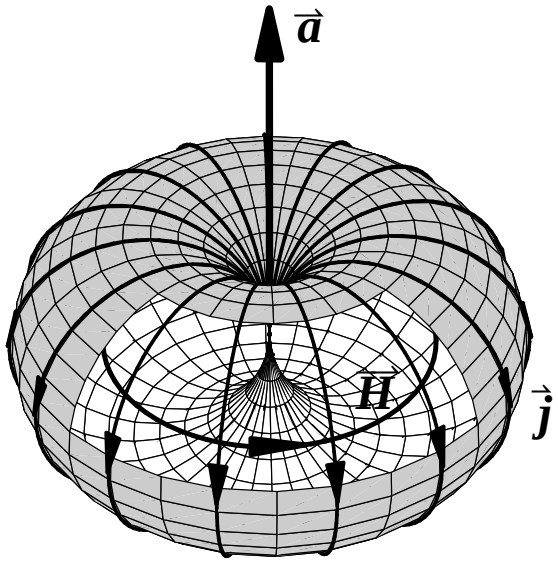

FIG. 1.12 The simplest system with nonvanishing "irreducible" anapole moment: a toroidal winding which can be thought of as a succession of current loops offset from the origin. Figure courtesy of S. M. Rochester.

torus. In addition, the value of the anapole moment does not depend on the choice of the origin. The situation here is analogous to a single charge offset from the origin and a dipole formed by two opposite charges. One may say that the toroidal current is the simplest example of an "irreducible" anapole.

Now to the problem.

(a) Explain why only nuclei with nonzero spin can have an anapole moment, and why an anapole can only arise due to parity-violating interactions.[21]

Solution

As is true for any rank-one tensor (vector) characteristic of a system (in this case an atomic nucleus) the anapole moment must be proportional to the total angular momentum of the system, $\vec{I}$ (see Appendix F). The parity-violating nature of the anapole moment becomes apparent when one considers the behavior of $\vec{a}$ and $\vec{I}$ under spatial inversion: while the latter is a pseudovector, the former is a normal vector just like $\vec{j}$ [see Eq. (1.286)]. Note, however, that the existence of an anapole moment does not violate time-reversal invariance because both $\vec{a}$ and $\vec{I}$ are T-odd.

(b) Now we will embark on the derivation of Eq. (1.287), which can be done similarly to the derivation of the "usual" moments (magnetic dipole, etc.), see, for example, Jackson (1975) or Landau and Lifshitz (1987). This part of the problem requires somewhat involved graduate-level mathematics (tensor algebra).

The derivation presented here follows that of Sushkov *et al.* (1984) and Khriplovich (1989,1991). We start from the general expression for the vector potential

$$\vec{A}\Big(\vec{R}\Big) = \frac{1}{c}\int \frac{\vec{j}(\vec{r})}{|\vec{R}-\vec{r}|}d^3r \ , \tag{1.288}$$

and expand the integrand in powers of r/R (assuming $R \gg r$).[22]

Consider the zeroth-order term in the expansion of the vector potential (1.288). Show that it vanishes for an atomic nucleus.

[21] The nuclear anapole moment, arising due to the weak interaction (Problem 1.13), was recently discovered in an atomic PNC experiment using cesium by Wood *et al.* (1997).

[22] A conscientious reader may question the validity of this approach, since we begin by expanding the vector potential in the limit where $R \gg r$ and end up with a delta function [Eq. (1.287)]. Nonetheless, more rigorous derivations (Flambaum and Khriplovich 1980; Flambaum and Hanhart 1993) confirm the conclusions presented here.

Solution

It is convenient to use the vector form of the Taylor expansion:

$$\frac{1}{|\vec{R}-\vec{r}|} = \frac{1}{R} - \left(\nabla_k \frac{1}{R}\right) r_k + \frac{1}{2}\left(\nabla_k \nabla_l \frac{1}{R}\right) r_k r_l + \ldots , \tag{1.289}$$

where we follow the usual convention in which summation is assumed over repeated indices.

The zeroth-order term in the expansion of the vector potential (1.288) is thus

$$\vec{A}^{(0)}\left(\vec{R}\right) = \frac{1}{cR}\int \vec{j}(\vec{r}) d^3 r. \tag{1.290}$$

This quantity vanishes for atomic nuclei because there cannot be any average current for a system of finite size in a steady state (for which the spatial distribution of charges does not change with time).

(c) Now consider the first-order term in the expansion of the vector potential. Show that it corresponds to the potential due to a magnetic dipole.

Solution

The integrand in the first-order term for each Cartesian component of the vector potential can be decomposed into a symmetric and an antisymmetric part:

$$\begin{aligned} A_i^{(1)}\left(\vec{R}\right) &= -\frac{1}{c}\left(\nabla_k \frac{1}{R}\right)\int j_i r_k d^3 r \\ &= -\frac{1}{c}\left(\nabla_k \frac{1}{R}\right)\int \left[\frac{1}{2}(j_i r_k + j_k r_i) + \frac{1}{2}(j_i r_k - j_k r_i)\right] d^3 r. \end{aligned} \tag{1.291}$$

Next, we use a common trick involving Gauss' theorem that says that the volume integral of a spatial derivative of an arbitrary well-behaved function f is equal to the surface integral of the function itself

$$\int \nabla_m f d^3 r = \int f dS_m \tag{1.292}$$

to show that the integral with the symmetric combination in Eq. (1.291) vanishes. Indeed, consider the integral

$$\int \nabla_m (r_i r_k j_m) d^3 r. \tag{1.293}$$

Converting the integral over volume to a surface integral according to Eq. (1.292) and choosing the integration surface outside the boundaries of the system (where

the current is zero), we see that the integral (1.293) vanishes. On the other hand, we also have

$$0 = \int \nabla_m (r_i r_k j_m) d^3r = \int \left(\delta_{mi} r_k j_m + r_i \delta_{mk} j_m + r_i r_k \vec{\nabla} \vec{j} \right) d^3r. \quad (1.294)$$

The last term in the integrand of (1.294) vanishes because the divergence of the current is zero in a steady state. The other two terms correspond to the symmetric combination in Eq. (1.291).

We now consider the antisymmetric part of the integral in Eq. (1.291):

$$A_i^{(1)}\left(\vec{R}\right) = \frac{1}{c}\frac{R_k}{R^3} \int \left\{ \frac{1}{2}(j_i r_k - j_k r_i) \right\} d^3r. \quad (1.295)$$

Because

$$R_k j_i r_k - R_k j_k r_i = j_i \vec{R} \cdot \vec{r} - r_i \vec{R} \cdot \vec{j} = \left\{ \vec{R} \times \left(\vec{j} \times \vec{r} \right) \right\}_i, \quad (1.296)$$

we recognize in Eq.(1.295) the familiar expression for the vector potential from a magnetic dipole $\vec{m}$

$$A_i^{(1)}\left(\vec{R}\right) = \frac{\vec{m} \times \vec{R}}{R^3}, \quad (1.297)$$

where the magnetic dipole moment is defined as

$$\vec{m} = \frac{1}{2c} \int \vec{r} \times \vec{j}(\vec{r}) d^3r. \quad (1.298)$$

(d) Finally, we turn to the second-order terms in the multipole expansion of the vector potential (1.288) that will yield us the anapole moment:

$$\vec{A}_i^{(2)}\left(\vec{R}\right) = \frac{1}{2c}\left(\nabla_k \nabla_l \frac{1}{R} \right) \int j_i r_k r_l d^3r. \quad (1.299)$$

Show that all symmetric terms in Eq. (1.299) vanish.

Solution

The integrand in Eq. (1.299) is a third-order tensor (i.e., a tensor resulting from combining three vectors) that can generally be decomposed into irreducible tensors of ranks 3, 2, 1, and 0, the first being totally symmetric with respect to the components of the constituent vectors and the last being completely antisymmetric.[23] We now use a trick similar to that used in Eq. (1.294) to show that the

[23] In order to see that the irreducible rank-three tensor is symmetric, it is possible to think of each of the constituent vectors (each described by three independent coordinates) as spin-one objects, and the

symmetric part of the tensor integrates to zero:

$$0 = \int \nabla_m (r_i r_k r_l j_m) d^3r = \int (\delta_{mi} r_k r_l j_m + r_i \delta_{mk} r_l j_m + r_i r_k \delta_{ml} j_m) d^3r$$
$$= \int (r_k r_l j_i + r_i r_l j_k + r_i r_k j_l) d^3r. \qquad (1.300)$$

The remaining parts of the tensor in the integrand of Eq.(1.299) are antisymmetric rank-two tensors and rank-one tensors (i.e., vectors).[24]

(e) Show that the antisymmetric rank-two part of the second order term in the vector potential expansion violates T-invariance.

Solution

Note that all of the remaining parts of the tensor in the integrand of Eq.(1.299) change sign under both spatial inversion (transforming $\vec{r} \to -\vec{r}$ and $\vec{j} \to -\vec{j}$) and time reversal (transforming $\vec{r} \to \vec{r}$ and $\vec{j} \to -\vec{j}$). By the Wigner-Eckart theorem (Appendix F), any second-rank tensor characteristic of the system (in this case, nucleus) with total angular momentum I has to be proportional to the only irreducible second-rank tensor that can be constructed out of the components of $\vec{I}$:

$$\frac{1}{2}(I_i I_k + I_k I_i) - \frac{1}{3} I(I+1)\delta_{ik} \ , \qquad (1.301)$$

which is even under both spatial inversion and time reversal. Therefore, the existence of the second-rank moment (called the *magnetic quadrupole moment*) would violate both parity (P) and time-reversal (T) invariance. In this problem, we restrict ourselves to T-conserving moments and will not further consider the magnetic quadrupole.

(f) Show that the remaining rank-one tensor in the second-order term of the vector potential expansion is a vector proportional to $r^2 \vec{j}$.

irreducible rank-three tensor as a spin-three object. It is clear that to get a spin-three by combining three spin-ones, it is necessary to construct a totally symmetric wavefunction. For rank-zero, the only scalar that can be built out of the three vectors $\vec{v}_{1-3}$ is their mixed product $\vec{v}_1 \cdot (\vec{v}_2 \times \vec{v}_3) = (v_1)_i (v_2)_j (v_3)_k \epsilon_{ijk}$, where ϵ_{ijk} is the totally antisymmetric (Levi-Civita) tensor. With two of the constituent vectors being the same, the rank-zero term vanishes.

[24] A general third-order tensor (27 independent components) is decomposed into one irreducible rank-three tensor (7 components), two irreducible rank-two tensors (2×5 components), three vectors (3×3 components), and a scalar (one component; see, for example, Varshalovich *et al.* (1988), Chapter 3.2.2); however in the present case the number of possible tensors is reduced because two of the constituent vectors are the same and only structures symmetric in the corresponding components are allowed.

Solution

Let us now turn to the vectors that can be constructed out of $r_i r_k j_l$. The only possibilities are: $r^2\vec{j}$ and $\vec{r}(\vec{r}\cdot\vec{j})$. It turns out that integrals over these two vectors are not independent. To see this, note that there is no summation over indices in Eq. (1.300), so the identity is valid for any set of components. If so, it is also true for sums of such components. In particular, the sum (called the *contraction of the tensor over two indices*) gives

$$0 = \delta_{kl}\int (r_k r_l j_i + r_i r_l j_k + r_i r_k j_l) d^3r = \int \Big(r^2 j_i + 2r_i(\vec{r}\cdot\vec{j})\Big) d^3r. \quad (1.302)$$

Thus, it is sufficient to only consider the $r^2\vec{j}$ structure.

(g) Using the above results, prove that the second order term in the expansion of the vector potential is given by

$$\vec{A}_i^{(2)}\Big(\vec{R}\Big) = -\frac{1}{3c}\Big(\nabla_k\nabla_l\frac{1}{R}\Big)\int \epsilon_{kip} V_p r_l d^3r\ , \quad (1.303)$$

where V_p is a *dual vector* for the antisymmetric second-order tensor $j_k r_i - j_i r_k$ (see, for example, Arfken (1985), Chapter 3.4):

$$(j_k r_i - j_i r_k) = \epsilon_{kip} V_p\ . \quad (1.304)$$

This form explicitly accounts for the antisymmetry. The explicit expression of the components of $\vec{V}$ is found by contracting both sides of Eq. (1.304) with ϵ_{kiq} and taking into account that

$$\epsilon_{kip}\epsilon_{kiq} = 2\delta_{pq}\ , \quad (1.305)$$

$$V_q = \frac{1}{2}\epsilon_{kiq}(j_k r_i - j_i r_k)\ . \quad (1.306)$$

Solution

Let us return to Eq. (1.299) and transform it taking into account Eq. (1.300) and the fact that the tensor $\big(\nabla_k\nabla_l\frac{1}{R}\big)$ is symmetric in indices k and l. First, from Eq. (1.300) we have:

$$\int j_i r_k r_l d^3r = -\int (j_k r_i r_l + j_l r_i r_k) d^3r. \quad (1.307)$$

Substituting this into Eq. (1.299), we have:

$$\vec{A}_i^{(2)}\left(\vec{R}\right) = -\frac{1}{2c}\left(\nabla_k\nabla_l\frac{1}{R}\right)\int (j_k r_i r_l + j_l r_i r_k)d^3r$$
$$= -\frac{1}{c}\left(\nabla_k\nabla_l\frac{1}{R}\right)\int j_k r_i r_l d^3r, \tag{1.308}$$

where in the last expression the symmetry of $\left(\nabla_k\nabla_l\frac{1}{R}\right)$ is used. One can combine Eqs. (1.299) and (1.308) (by adding (1.299) and (1.308) divided by two) to obtain a combination in the integrand that is antisymmetric in indices i and k:

$$\frac{3}{2}\vec{A}_i^{(2)}\left(\vec{R}\right) = -\frac{1}{2c}\left(\nabla_k\nabla_l\frac{1}{R}\right)\int (j_k r_i - j_i r_k)r_l d^3r. \tag{1.309}$$

Using Eq. (1.304) in Eq. (1.309), we obtain the desired result:

$$\vec{A}_i^{(2)}\left(\vec{R}\right) = -\frac{1}{3c}\left(\nabla_k\nabla_l\frac{1}{R}\right)\int \epsilon_{kip}V_p r_l d^3r. \tag{1.310}$$

(h) The term $\vec{A}_i^{(2)}\left(\vec{R}\right)$ corresponds to the vector potential due to the anapole moment. Use Eq. (1.310) to prove Eq. (1.287).

Solution

In order to complete the expansion of the integrand into irreducible structures, we can split the tensor $V_p r_l$ into its symmetric and antisymmetric parts. The symmetric part is a tensor of rank two and corresponds to the magnetic quadrupole moment. As we have discussed above, this moment violates time-reversal invariance and we do not consider it in this problem. We now turn to the antisymmetric part of $V_p r_l$ which we write, in analogy with Eq. (1.304), as

$$\frac{1}{2}(V_p r_l - V_l r_p) = \frac{1}{2}\epsilon_{plq}W_q. \tag{1.311}$$

Substituting this into Eq. (1.310), and using the identity

$$\epsilon_{kip}\epsilon_{plq} = \delta_{kl}\delta_{iq} - \delta_{kq}\delta_{il} \tag{1.312}$$

we write

$$\vec{A}^{(2)}_{i,Anapole}\left(\vec{R}\right) = -\frac{1}{6c}\left(\nabla_k\nabla_l\frac{1}{R}\right)\int (\delta_{kl}W_i - \delta_{il}W_k)d^3r. \tag{1.313}$$

We see that we have reduced the expression for the vector potential to the form that depends only on a single vector characteristic of the system, the integral of

$\vec{W}$. It is clear that it must be proportional to the anapole moment (1.286). In order to see this explicitly, we use Eq. (1.306) and its analog for the components of $\vec{W}$ and write:

$$\begin{aligned}\int W_i d^3r &= \frac{1}{2}\epsilon_{pli}\int (V_p r_l - V_l r_p) d^3r \\ &= \frac{1}{4}\epsilon_{pli}\int [\epsilon_{kqp}(j_k r_q - j_q r_k) r_l - \epsilon_{kql}(j_k r_q - j_q r_k) r_p] d^3r \\ &= \int \left[-j_i r^2 + r_i\left(\vec{j}\cdot\vec{r}\right)\right] d^3r, \end{aligned} \tag{1.314}$$

where once again we have used the identity (1.312). The expression (1.314) can be simplified with the help of Eq. (1.302):

$$\int W_i d^3r = -\frac{3}{2}\int j_i r^2 d^3r = \frac{3}{2\pi} a_i, \tag{1.315}$$

where in the latter equality we have substituted the definition (1.286).

We are now ready to substitute the result (1.315) into the expression for the vector potential (1.313):

$$\vec{A}^{(2)}_{i,Anapole}\left(\vec{R}\right) = -\frac{1}{4\pi c}\left(\nabla_k \nabla_l \frac{1}{R}\right)(\delta_{kl} a_i - \delta_{il} a_k)\ . \tag{1.316}$$

The first term in Eq. (1.316) contains

$$\nabla_k \nabla_l \frac{1}{R}\delta_{kl} = \nabla^2 \frac{1}{R} = -4\pi\delta\left(\vec{R}\right)\ . \tag{1.317}$$

The last equality is just the statement of the Laplace equation for the scalar potential of a point charge. Let us now discuss the second term in Eq. (1.316). It is proportional to

$$\left(\nabla_k \nabla_l \frac{1}{R}\right)\delta_{il} a_k = \nabla_i\left(\vec{a}\cdot\vec{\nabla}\frac{1}{R}\right)\ , \tag{1.318}$$

i.e., a gradient of a scalar function. Recall, however, that the vector potential is not unique: it is defined up to a gradient of an arbitrary scalar function; addition of such a function is called *gauge transformation* [see, for example, Landau and Lifshitz (1987)].[25] In fact, our starting point, Eq. (1.288), is written in a particular *Coulomb gauge* for which $\vec{\nabla}\cdot\vec{A} = 0$.

Skipping the second term in Eq. (1.316) because it can be eliminated by a gauge transformation and employing Eq. (1.317), we obtain the sought-for Eq. (1.287).

[25] This is because the observables are fields rather than potentials. To find the magnetic field, we take the curl of $\vec{A}$. Because the curl of a gradient is zero, the field is invariant with respect to a gauge transformation.