17

Time-Independent Perturbation Theory

17.1. The Formalism

Time-independent perturbation theory is an approximation scheme that applies in the following context: we know the solution to the eigenvalue problem of the Hamiltonian H^0, and we want the solution to $H=H^0+H^1$, where H^1 is small compared to H^0 in a sense to be made precise shortly. For instance, H^0 can be the Coulomb Hamiltonian for an electron bound to a proton, and H^1 the addition due to an external electric field that is weak compared to the proton's field at the (average) location of the electron. One refers to H^0 as the *unperturbed Hamiltonian* and H^1 as the *perturbing Hamiltonian* or *perturbation*.

We proceed as follows. We assume that to every eigenket $|E_n^0\rangle \equiv |n^0\rangle$ of H^0 with eigenvalue E_n^0, there is an eigenket $|n\rangle$ of H with eigenvalue E_n.‡ We then assume that the eigenkets and eigenvalues of H may be expanded in a *perturbation series*§:

$$|n\rangle = |n^0\rangle + |n^1\rangle + |n^2\rangle + \cdots \tag{17.1.1}$$

$$E_n = E_n^0 + E_n^1 + E_n^2 + \cdots \tag{17.1.2}$$

The superscript k on each term gives the power of (the matrix element of) H^1 that it is expected to come out proportional to. A term with superscript equal to k is called a *kth-order term*. (Clearly a product like $E_n^k|n^{\bar{k}}\rangle$ is a term of order $k+\bar{k}$.) We hope that as the order increases, the terms get systematically smaller; this is when we can say that H^1 is small. When we find explicit formulas for $|n^k\rangle$ and E_n^k, these ideas will be sharpened.

To find the terms in the expansions for $|n\rangle$ and E_n, we start with the eigenvalue equation

$$H|n\rangle = E_n|n\rangle \tag{17.1.3}$$

---

‡ We are assuming that $|n^0\rangle$ is *nondegenerate*. The degenerate case follows.

§ We assume $|n^0\rangle$ is normalized (to unity). The norm of $|n\rangle$ will be discussed shortly.

or

$$(H^0+H^1)[|n^0\rangle+|n^1\rangle+\cdots]$$
$$=(E_n^0+E_n^1+\cdots)[|n^0\rangle+|n^1\rangle+\cdots] \qquad (17.1.4)$$

We approach these equations as we did the differential equation in the WKB approximation. Recall that there we had an equation with terms of order $\hbar^{-2}$, $\hbar^{-1}$, . . . , etc. We first ignored all but the $\hbar^{-2}$ terms and solved for ϕ_0. We then fed this into the $\hbar^{-1}$ part to determine ϕ_1. (We could have gone on this way, though we chose to stop there.) Likewise, in the present case, we first consider the zeroth-order terms of Eq. (17.1.4). We get the equation

$$H^0|n^0\rangle=E_n^0|n^0\rangle \qquad (17.1.5)$$

Notice that the zeroth-order quantities $|n^0\rangle$ and E_n^0 are indeed independent of H^1 (or, equivalently, they depend on the zeroth power of H^1). By assumption, this equation may be solved and the eigenvectors $|n^0\rangle$ and eigenvalues E_n^0 determined. So we move on to the first-order terms. We get the equation

$$H^0|n^1\rangle+H^1|n^0\rangle=E_n^0|n^1\rangle+E_n^1|n^0\rangle \qquad (17.1.6)$$

Let us dot both sides with $\langle n^0|$. Using $\langle n^0|H^0=\langle n^0|E_n^0$ and $\langle n^0|n^0\rangle=1$, we get

$$E_n^1=\langle n^0|H^1|n^0\rangle \qquad (17.1.7)$$

i.e., *the first-order change in energy is the expectation value of H^1 in the unperturbed state.* Notice that E_n^1 is indeed proportional to the first power of H^1. Let us next dot both sides of Eq. (17.1.6) with $\langle m^0|$, $m\neq n$, to get

$$\langle m^0|H^0|n^1\rangle+\langle m^0|H^1|n^0\rangle=E_n^0\langle m^0|n^1\rangle$$

or

$$\langle m^0|n^1\rangle=\frac{\langle m^0|H^1|n^0\rangle}{E_n^0-E_m^0} \qquad (17.1.8)$$

Since $m\neq n$, this equation determines all the components of $|n^1\rangle$ in the eigenbasis of H^0, except for the component parallel to $|n^0\rangle$, let's call it $|n_\parallel^1\rangle$. We determine it by the requirement that $|n\rangle$ is normalized to this order.‡ In obvious notation, we have

$$1=\langle n|n\rangle=(\langle n^0|+\langle n_\perp^1|+\langle n_\parallel^1|)(|n^0\rangle+|n_\perp^1\rangle+|n_\parallel^1\rangle) \qquad (17.1.9)$$

---

‡ Recall that even in eigenvalue problems that can be solved exactly, there is the arbitrariness in the norm of the vector. To this order, only $|n_\parallel^1\rangle$ alters the length of $|n^0\rangle$. [See Eq. (17.1.10).]

which leads to

$$1 = \langle n^0|n^0\rangle + \langle n^1_{\|}|n^0\rangle + \langle n^0|n^1_{\|}\rangle + \text{higher order}$$

or

$$0 = \langle n^1_{\|}|n^0\rangle + \langle n^0|n^1_{\|}\rangle + \text{higher order} \tag{17.1.10}$$

This means that

$$\langle n^0|n^1_{\|}\rangle = i\alpha, \qquad \alpha \text{ real} \tag{17.1.11}$$

Using

$$1 + i\alpha = e^{i\alpha} \quad \text{(to this order)} \tag{17.1.12}$$

we get

$$|n\rangle = |n^0\rangle\, e^{i\alpha} + \sum_m{}' \frac{|m^0\rangle\langle m^0|H^1|n^0\rangle}{E^0_n - E^0_m} \tag{17.1.13}$$

where the prime on $\sum'$ means that $m \neq n$. Since $|n\rangle$ has an arbitrariness in its overall phase, even after it is normalized, let us change its phase by the factor $e^{-i\alpha}$ in Eq. (17.1.13). This gets rid of the phase factor multiplying $|n^0\rangle$ and does nothing to the first-order piece, *to this order*. Calling the perturbed eigenket with the new phase also $|n\rangle$, we get the result to first order:

$$|n\rangle = |n^0\rangle + \sum_m{}' \frac{|m^0\rangle\langle m^0|H^1|n^0\rangle}{E^0_n - E^0_m} = |n^0\rangle + |n^1\rangle \tag{17.1.14}$$

Notice that $|n^1\rangle$ is orthogonal to $|n^0\rangle$ and proportional to the first power of H^1 (as anticipated). We determine E^2_n from the second-order part of Eq. (17.1.4):

$$H^0|n^2\rangle + H^1|n^1\rangle = E^0_n|n^2\rangle + E^1_n|n^1\rangle + E^2_n|n^0\rangle \tag{17.1.15}$$

Dotting with $\langle n^0|$ and using the results from lower order ($|n^1\rangle = |n^1_{\perp}\rangle$) we obtain

$$E^2_n = \langle n^0|H^1|n^1\rangle \tag{17.1.16}$$

$$= \sum_m{}' \frac{\langle n^0|H^1|m^0\rangle\langle m^0|H^1|n^0\rangle}{E^0_n - E^0_m}$$

$$= \sum_m{}' \frac{|\langle n^0|H^1|m^0\rangle|^2}{E^0_n - E^0_m} \tag{17.1.17}$$

We can go on to higher orders, but we choose to stop here.

Before we turn to examples, let us consider some general features of our results. First we note that the energy to a given order is determined by the state vector to the next lower order, see Eqs. (17.1.7) and (17.1.16). This is in accord with the remarks made in the study of the variational method. The physics behind this phenomenon will become clear when we consider a few examples. Next we ask under what conditions the perturbation expansion is good, namely, when the correction terms are small compared to the zeroth-order (unperturbed) results. The answer follows from Eq. (17.1.14). A necessary condition for $|n^1\rangle$ to be small compared to $|n^0\rangle$ is that

$$\left|\frac{\langle m^0|H^1|n^0\rangle}{E_n^0-E_m^0}\right| \ll 1 \tag{17.1.18}$$

Thus we see that the condition depends on (1) the absolute size of H^1 (i.e., if it is due to some external field, the magnitude of the field); (2) the matrix elements of H^1 between unperturbed states; and (3) the energy difference between the levels. If the unperturbed eigenstate is $|n^0\rangle$, the perturbation mixes in orthogonal states $|m^0\rangle$; this mixing is directly proportional to the matrix element $\langle m^0|H^1|n^0\rangle$ and inversely proportional to the energy difference between the two levels, which measures the "rigidity" of the system. If for any reason the above inequality is not fulfilled (say due to degeneracy, $E_n^0=E_m^0$) we must turn to an alternate formalism called *degenerate perturbation theory* to be described later in this chapter.

17.2. Some Examples

Consider a particle of charge q and mass m in a harmonic oscillator potential $V=\frac{1}{2}m\omega^2x^2$. Suppose we apply an external electric field of magnitude f along the positive x direction. This corresponds to an electrostatic potential $\phi=-fx$ and a potential energy $V=-qfx$. Thus

$$H=H^0+H^1=\frac{P^2}{2m}+\frac{1}{2}m\omega^2X^2-qfX \tag{17.2.1}$$

We wish to handle H^1 by perturbation theory. Let us first calculate the first-order shift in energy, given by

$$E_n^1=\langle n^0|H^1|n^0\rangle=-qf\langle n^0|X|n^0\rangle \tag{17.2.2}$$

where $|n^0\rangle$ is just the nth state of the unperturbed oscillator. We can see that E^1 vanishes in many ways. At a formal level, since

$$X=\left(\frac{\hbar}{2m\omega}\right)^{1/2}(a+a^\dagger) \tag{17.2.3}$$

it has no *diagonal* matrix elements. The physics of what is happening is more transparent in the coordinate basis, where

$$E_n^1 = -qf \int (\psi_n^0)^* x \psi_n^0 \, dx$$
$$= -qf \int |\psi_n^0|^2 x \, dx \tag{17.2.4}$$

Now $\psi_n^0(x)$, being the unperturbed eigenfunction, has definite parity $(-1)^n$. Consequently $|\psi_n^0|^2$ is an even function, while the external potential is an odd function. Thus the average interaction with the external field is zero, for the particle is as likely to be found in the region of potential ϕ as in the region of potential $-\phi$. Notice that E_n^1 is the energy of interaction of the *unperturbed configuration* $|n^0\rangle$, with the applied field. Consequently this is not the whole story, for the configuration itself will get modified by the external field to $|n^0\rangle + |n^1\rangle + \cdots$, and we should really be considering the energy of interaction of the perturbed configurations and the applied field. But this is a distinction that is at least a second-order effect, for the change $\delta|n\rangle \equiv |n\rangle - |n^0\rangle$ in the configuration is at least a first-order effect and the interaction of $\delta|n\rangle$ with the applied field involves another order of H^1. So let us calculate the perturbed eigenket to first order and then energy levels to second order. From Eq. (17.1.14).

$$|n\rangle = |n^0\rangle + \sum_m{}' \frac{|m^0\rangle\langle m^0| - qf(\hbar/2m\omega)^{1/2}(a+a^\dagger)|n^0\rangle}{E_n^0 - E_m^0}$$
$$= |n^0\rangle + qf\left(\frac{1}{2m\hbar\omega^3}\right)^{1/2} [(n+1)^{1/2}|(n+1)^0\rangle - n^{1/2}|(n-1)^0\rangle] \tag{17.2.5}$$

Thus to first order, the perturbation mixes the state $|n^0\rangle$ with the states immediately above and below it. It was stated earlier that $E_n^0 - E_m^0$ measures the "rigidity" of the system. We find in this example that this quantity is proportional to ω, which in the mass–spring case measures the force constant. How does the wave function of the perturbed state look? This is not transparent from the above equation, but we expect that it will represent a probability distribution that is no longer centered at and symmetric about $x = 0$, but instead is biased toward positive x (for that is the direction of the external field). We will return to confirm this picture quantitatively.

The second-order energy shift (which reflects the fact that the configuration of the system is not $|n^0\rangle$ but $|n^0\rangle + |n^1\rangle$), is

$$E_n^2 = \langle n^0|H^1|n^1\rangle = \sum_m{}' \frac{|\langle m^0|H^1|n^0\rangle|^2}{E_n^0 - E_m^0}$$
$$= q^2 \cdot f^2 \cdot \frac{\hbar}{2m\omega}\left(\frac{n+1}{-\hbar\omega} + \frac{n}{\hbar\omega}\right) = \frac{-q^2 f^2}{2m\omega^2} \tag{17.2.6}$$

The present problem is a nice testing ground for perturbation theory because it may be exactly solved. This is because H may be written as

$$\begin{aligned} H &= \frac{P^2}{2m} + \frac{1}{2} m\omega^2 X^2 - qfX \\ &= \frac{P^2}{2m} + \frac{1}{2} m\omega^2 \left(X - \frac{qf}{m\omega^2} \right)^2 - \frac{1}{2} \frac{q^2 f^2}{m\omega^2} \end{aligned} \tag{17.2.7}$$

This Hamiltonian also describes an oscillator of frequency ω, but is different in that (i) the oscillator is centered at $x = qf/m\omega^2$; (ii) each state has a constant energy $(-q^2f^2/2m\omega^2)$ added to it. Thus the eigenfunctions of H are just the eigenfunctions of H^0 shifted ay $qf/m\omega^2$ and the eigenvalues are $E_n = E_n^0 - q^2f^2/2m\omega^2$. The classical picture that goes with Eq. (17.2.7) is clear: the effect of a constant force qf on a mass coupled to a spring of force constant $m\omega^2$ is to shift the equilibrium point to $x = qf/m\omega^2$. (Imagine a mass hanging from a spring attached to the ceiling and ask what gravity does to its dynamics.) Let us compare these exact results with the perturbation theory. Consider the energy

$$E_n = E_n^0 - q^2f^2/2m\omega^2 \tag{17.2.8}$$

Since H^1 is proportional to qf, the power of qf gives the order of the term. According to Eq. (17.2.8), there is no first-order shift in energy, and the second-order shift is $-q^2f^2/2m\omega^2$, which agrees with Eq. (17.2.6). Had we tried to go to higher orders, we would have found nothing more.

Now consider the state vectors. The exact result is

$$|n\rangle = T(qf/m\omega^2)|n^0\rangle \tag{17.2.9}$$

where $T(a)$ is the operator that translates the system by an amount a. Since we are working to first order in qf,

$$\begin{aligned} T(qf/m\omega^2) &= e^{-i(qf/m\omega^2\hbar)P} \simeq I - i\left(\frac{qf}{m\omega^2\hbar}\right)P \\ &= I - i\left(\frac{qf}{m\omega^2\hbar}\right)\left(\frac{\hbar m\omega}{2}\right)^{1/2} \cdot \frac{a - a^\dagger}{i} \end{aligned} \tag{17.2.10}$$

so that

$$\begin{aligned} |n\rangle &= \left[I - \left(\frac{qf}{m\omega^2\hbar}\right)\left(\frac{\hbar m\omega}{2}\right)^{1/2} (a - a^\dagger) \right] |n^0\rangle \\ &= |n^0\rangle + qf\left(\frac{1}{2m\hbar\omega^3}\right)^{1/2} [(n+1)^{1/2}|(n+1)^0\rangle - n^{1/2}|(n-1)^0\rangle] \end{aligned} \tag{17.2.11}$$

which agrees with Eq. (17.2.5). It is clear that computing $|n\rangle$ to higher order in perturbation theory will be equivalent to expanding T to higher orders in qf.

*Exercise 17.2.1.** Consider $H^1 = \lambda x^4$ for the oscillator problem.

(1) Show that

$$E_n^1 = \frac{3\hbar^2\lambda}{4m^2\omega^2}[1+2n+2n^2]$$

(2) Argue that no matter how small λ is, the perturbation expansion will break down for some large enough n. What is the physical reason?

*Exercise 17.2.2.** Consider a spin-1/2 particle with gyromagnetic ratio γ in a magnetic field $\mathbf{B} = B\mathbf{i} + B_0\mathbf{k}$. Treating B as a perturbation, calculate the first- and second-order shifts in energy and first-order shift in wave function for the ground state. Then compare the exact answers expanded to the corresponding orders.

Exercise 17.2.3. In our study of the H atom, we assumed that the proton is a point charge e. This leads to the familiar Coulomb interaction $(-e^2/r)$ with the electron. (1) Show that if the proton is a uniformly dense charge distribution of radius R, the interaction is

$$\begin{aligned} V(r) &= -\frac{3e^2}{2R} + \frac{e^2r^2}{2R^3}, \qquad r \le R \\ &= -\frac{e^2}{r}, \qquad r > R \end{aligned}$$

(2) Calculate the first-order shift in the ground-state energy of hydrogen due to this modification. You may assume $e^{-R/a_0} \simeq 1$. You should find $E^1 = 2e^2R^2/5a_0^3$.

*Exercise 17.2.4.** (1) Prove the Thomas–Reiche–Kuhn sum rule

$$\sum_{n'} (E_{n'} - E_n)|\langle n'|X|n\rangle|^2 = \sum_{n'} (E_{n'} - E_n)\langle n|X|n'\rangle\langle n'|X|n\rangle = \frac{\hbar^2}{2m}$$

where $|n\rangle$ and $|n'\rangle$ are eigenstates of $H = P^2/2m + V(X)$. (Hint: Eliminate the $E_{n'} - E_n$ factor in favor of H.)

(2) Test the sum rule on the nth state of the oscillator.

Exercise 17.2.5 (Hard). We have seen that if we neglect the repulsion e^2/r_{12} between the two electrons in the ground state of He, the energy is $-8\ \text{Ry} = -108.8$ eV. Treating e^2/r_{12} as a perturbation, show that

$$\langle 100, 100|H^1|100, 100\rangle = \tfrac{5}{2}\ \text{Ry}$$

so that $E_0^0 + E_0^1 = -5.5.\ \text{Ry} = -74.8$ eV. Recall that the measured value is -78.6 eV and the variational estimate is -77.5 eV. [Hint: $\langle H^1\rangle$ can be viewed as the interaction between two concentric, spherically symmetric exponentially falling charge distributions. Find the potential $\phi(r)$ due to one distribution and calculate the interaction energy between this potential and the other charge distribution.]

Selection Rules

The labor involved in perturbation theory calculations is greatly reduced by the use of selection rules, which allow us to conclude that certain matrix elements of H^1 are zero without explicitly calculating them. They are based on the idea that if

$$[\Omega, H^1]=0$$

then

$$\langle \alpha_2\omega_2|H^1|\alpha_1\omega_1\rangle=0 \quad \text{unless} \quad \omega_1=\omega_2 \tag{17.2.12}‡$$

Proof.

$$0=\langle \alpha_2\omega_2|\Omega H^1-H^1\Omega|\alpha_1\omega_1\rangle=(\omega_2-\omega_1)\langle \alpha_2\omega_2|H^1|\alpha_1\omega_1\rangle \qquad \text{Q.E.D.}$$

Consider for example $H^1=\lambda Z$, which is invariant under rotations around the z axis. Then $[L_z, H^1]=0$ and

$$\langle \alpha_2 m_2|H^1|\alpha_1 m_1\rangle=0 \quad \text{unless} \quad m_2=m_1 \tag{17.2.13}$$

(This result also follows from the Wigner–Eckart theorem.) Or if H^1 is parity invariant, say $H^1=\lambda Z^2$, then its matrix element between states of opposite parity is zero.

There is a simple way to understand Eq. (17.2.12). To say that $[\Omega, H^1]=0$ is to say that H^1 "carries no Ω"; in other words, when it acts on a state it imparts no Ω to it. We see this as follows. Consider $|\omega_1\rangle$, which carries a definite amount of the variable Ω, namely, ω_1:

$$\Omega|\omega_1\rangle=\omega_1|\omega_1\rangle \tag{17.2.14}$$

Let us measure Ω in the state after H^1 acts on it:

$$\Omega(H^1|\omega_1\rangle)=H^1\Omega|\omega_1\rangle=H^1\omega_1|\omega_1\rangle=\omega_1(H^1|\omega_1\rangle) \tag{17.2.15}$$

We find it is the same as before, namely, ω_1. The selection rule then merely reflects the orthogonality of eigenstates with different ω.

This discussion paves the way for an extension of the selection rule to a case where H^1 carries a definite amount of Ω. For instance, if H^1 is a tensor operator T_k^q, it carries angular momentum (k, q) and we know from the Wigner–Eckart theorem that

$$\langle \alpha_2 j_2 m_2|T_k^q|\alpha_1 j_1 m_1\rangle=0 \quad \text{unless} \quad \begin{cases} j_1+k\geq j_2\geq |j_1-k| \\ m_2=m_1+q \end{cases} \tag{17.2.16}$$

---

‡ α stands for other quantum numbers that label the state.

i.e., that the matrix element vanishes unless $|\alpha_2 j_2 m_2\rangle$ has the angular momentum that obtains when we add to $(j_1 m_1)$ the angular momentum (kq) imparted by the operator. For instance, if $H^1 = \lambda Z \sim T_1^0$,

$$\langle \alpha_2 j_2 m_2 | Z | \alpha_1 j_1 m_1 \rangle = 0 \quad \text{unless} \quad \begin{cases} j_2 = j_1 + 1, j_1, j_1 - 1 \\ m_2 = m_1 + 0 \end{cases} \tag{17.2.17}$$

while if $H^1 = \lambda X$ or $\lambda Y (\sim T_1^{\pm 1})$, we have

$$\langle \alpha_2 j_2 m_2 | X \text{ or } Y | \alpha_1 j_1 m_1 \rangle = 0 \quad \text{unless} \quad \begin{cases} j_2 = j_1 + 1, j_1, j_1 - 1 \\ m_2 = m_1 \pm 1 \end{cases} \tag{17.2.18}$$

Another example of this type is an operator that is not parity invariant, but parity odd. An example is X, which obeys

$$\Pi^\dagger X \Pi = -X \tag{17.2.19}$$

You can verify that if X acts on a state of definite parity, it changes the parity of the state. Thus the matrix element of X between eigenstates of parity vanishes unless they have opposite parity. More generally, if

$$\Pi^\dagger \Omega \Pi = -\Omega \tag{17.2.20}$$

then the matrix element of Ω between two parity eigenstates vanishes unless they have opposite parity.

We get more selection rules by combining these selection rules. For instance, we can combine the angular momentum and parity selection rules for the vector operators $\mathbf{R}$ to get (in the case of no spin, $\mathbf{J} = \mathbf{L}$),

$$\begin{aligned} \langle \alpha_2 l_2 m_2 | Z | \alpha_1 l_1 m_1 \rangle = 0 \quad &\text{unless} \quad \begin{cases} l_2 = l_1 \pm 1 \\ m_2 = m_1 \end{cases} \\ \langle \alpha_2 l_2 m_2 | X \text{ or } Y | \alpha_1 l_1 m_1 \rangle = 0 \quad &\text{unless} \quad \begin{cases} l_2 = l_1 \pm 1 \\ m_2 = m_1 \pm 1 \end{cases} \end{aligned} \tag{17.2.21}$$

We rule out the possibility $l_2 = l_1$ by the parity selection rule, for states of orbital angular momentum l have definite parity $(-1)^l$. Equation (17.2.21) is called the *dipole selection rule*.

We now consider an example that illustrates the use of these tricks and a few more. The problem is to determine the response of the hydrogen atom in the ground state to a constant external electric field $\mathbf{E} = \mathscr{E}\mathbf{k}$. This is called the *Stark effect*. Let us first calculate H^1. We do this by determining $\mathscr{H}^1$, its classical counterpart and then making the operator substitution. If $\mathbf{r}_1$ and $\mathbf{r}_2$ are the position vectors of the electron and proton, respectively, and $\phi(\mathbf{r})$ is the electrostatic potential due to **E**, then

$$\begin{aligned}\mathscr{H}^1&=-e\phi(\mathbf{r}_1)+e\phi(\mathbf{r}_2)\\&=e[\phi(\mathbf{r}_2)-\phi(\mathbf{r}_1)]\\&=e(\mathbf{r}_1-\mathbf{r}_2)\cdot\mathbf{E}\quad(\text{recall }\mathbf{E}=-\nabla\phi)\\&=e\mathbf{r}\cdot\mathbf{E}\end{aligned}\tag{17.2.22}$$

where **r** is the relative coordinate or equivalently the position vector of the electron in the CM frame in the limit $m/M=0$. $\mathscr{H}^1$ is called the *dipole interaction*, for in terms of

$$\boldsymbol{\mu}_e=e(\mathbf{r}_2-\mathbf{r}_1)=-e\mathbf{r}\tag{17.2.23}$$

the *electric dipole moment* of the system,

$$\mathscr{H}^1=-\boldsymbol{\mu}_e\cdot\mathbf{E}\tag{17.2.24}$$

(This is the electric analog of $\mathscr{H}=-\boldsymbol{\mu}\cdot\mathbf{B}$.)‡ Thus, for the given electric field

$$H^1=eZ\mathscr{E}\tag{17.2.25}$$

Let us now calculate the first-order shift in the energy of the ground state $|100\rangle$§:

$$E^1_{100}=\langle 100|eZ\mathscr{E}|100\rangle\tag{17.2.26}$$

We can argue that $E^1_{100}=0$ either on the grounds of parity or the Wigner–Eckart theorem. More physically, E^1_{100} vanishes because in the unperturbed state, the electron probability distribution is spherically symmetric and the electron samples $\phi(\mathbf{r})$ and $\phi(-\mathbf{r})=-\phi(\mathbf{r})$ equally. Another way to say this is that the unperturbed atom has no mean electric dipole moment $\langle\boldsymbol{\mu}\rangle$ (by parity or the Wigner–Eckart theorem) so that

$$E^1_{100}=\langle 100|-\boldsymbol{\mu}\cdot\mathbf{E}|100\rangle=-\langle 100|\boldsymbol{\mu}|100\rangle\cdot\mathbf{E}=0\tag{17.2.27}$$

But we expect the second-order energy shift to be nonzero, for the external field will shift the electron distribution downward and induce a dipole moment which can interact with **E**. So let us calculate

$$E^2_{100}=\sum_{nlm}{}'\frac{e^2\mathscr{E}^2|\langle nlm|Z|100\rangle|^2}{E^0_{100}-E^0_{nlm}}\tag{17.2.28}$$

---

‡ In the rest of this chapter we will omit the subscript e on $\boldsymbol{\mu}_e$.

§ When we discuss hydrogen, we will use the symbol $|nlm\rangle$, rather than $|(nlm)^0\rangle$ to denote the *unperturbed* state.

where

$$E_{100}^0 - E_{nlm}^0 = -\mathrm{Ry}\left(1 - \frac{1}{n^2}\right) = \mathrm{Ry}\left(\frac{1-n^2}{n^2}\right) \tag{17.2.29}$$

Unlike in the case of the oscillator, the sum now involves an infinite number of terms. Although we can use dipole selection rules to reduce the sum to

$$E_{100}^2 = \sum_{n=2}^{\infty} \frac{e^2\mathscr{E}^2 |\langle n10|Z|100\rangle|^2}{E_1^0 - E_n^0} \tag{17.2.30}$$

let us keep the form in Eq. (17.2.28) for a while. There are several ways to proceed.

Method 1. Since the magnitude of the energy denominator grows with n, we have the inequality

$$|E_{100}^2| \le \frac{e^2\mathscr{E}^2}{|E_1^0 - E_2^0|} \sum_{nlm}{}' |\langle nlm|Z|100\rangle|^2$$

But since

$$\begin{aligned} &\sum_{nlm}{}' |\langle nlm|Z|100\rangle|^2 \\ &\quad = \sum_{nlm}{}' \langle 100|Z|nlm\rangle\langle nlm|Z|100\rangle \\ &\quad = \sum_{nlm} \langle 100|Z|nlm\rangle\langle nlm|Z|100\rangle - \langle 100|Z|100\rangle^2 \\ &\quad = \langle 100|Z^2|100\rangle - \langle 100|Z|100\rangle^2 \\ &\quad = a_0^2 - 0 = a_0^2 \end{aligned} \tag{17.2.31}$$

we get

$$\begin{aligned} |E_{100}^2| &\le \frac{e^2\mathscr{E}^2}{|(e^2/2a_0)(1-\frac{1}{4})|} a_0^2 \\ &\le \frac{8a_0^3\mathscr{E}^2}{3} \end{aligned} \tag{17.2.32}$$

We can also get a lower bound on $|E_{100}^2|$ by keeping just the first term in Eq. (17.2.30) (since all terms have the same sign):

$$|E_{100}^2| \ge \frac{e^2\mathscr{E}^2}{3e^2/8a_0} |\langle 210|Z|100\rangle|^2 \tag{17.2.33}$$

Now,

$$|\langle 210|Z|100\rangle|^2=\frac{2^{15}a_0^2}{3^{10}}\simeq 0.55a_0^2 \tag{17.2.34}$$

so that

$$|E_{100}^2|\geq (0.55)\tfrac{8}{3}\mathscr{E}^2a_0^3 \tag{17.2.35}$$

We thus manage to restrict $|E_{100}^2|$ to the interval

$$\tfrac{8}{3}\mathscr{E}^2a_0^3\geq |E_{100}^2|\geq 0.55(\tfrac{8}{3})\mathscr{E}^2a_0^3 \tag{17.2.36}$$

Method 2. Consider the general problem of evaluating

$$E_n^2=\sum_m{}' \frac{\langle n^0|H^1|m^0\rangle\langle m^0|H^1|n^0\rangle}{E_n^0-E_m^0} \tag{17.2.37}$$

If it weren't for the energy denominator, we could use the completeness relation to eliminate the sum (after adding and subtracting the $m=n$ term). There exists a way to eliminate the energy denominator.‡ Suppose we can find an operator Ω such that

$$H^1=[\Omega, H^0] \tag{17.2.38}$$

then

$$\begin{aligned}E_n^2&=\sum_m{}' \frac{\langle n^0|H^1|m^0\rangle\langle m^0|\Omega H^0-H^0\Omega|n^0\rangle}{E_n^0-E_m^0}\\&=\sum_m{}' \langle n^0|H^1|m^0\rangle\langle m^0|\Omega|n^0\rangle\\&=\langle n^0|H^1\Omega|n^0\rangle-\langle n^0|H^1|n^0\rangle\langle n^0|\Omega|n^0\rangle\end{aligned} \tag{17.2.39}$$

which calls for computing just three matrix elements. But it is not an easy problem to find the Ω that satisfies Eq. (17.2.38). (There are, however, exceptions, see Exercise 17.2.7.) A more modest proposal is to find Ω such that

$$H^1|n^0\rangle=[\Omega, H^0]|n^0\rangle \tag{17.2.40}$$

for a given $|n^0\rangle$. You can verify that this is all it takes to derive Eq. (17.2.39) *for this value of n.* In the problem we are interested in, we need to solve

$$H^1|100\rangle=[\Omega, H^0]|100\rangle \tag{17.2.41}$$

---

‡ See A. Dalgarno and J. T. Lewis, *Proceedings of the Royal Society*, **A233**, 70 (1955).

By writing this equation in the coordinate basis and assuming Ω is a function of coordinates and not momenta, we can show that

$$\Omega \xrightarrow[\text{coordinate basis}]{} -\frac{ma_0 e\mathscr{E}}{\hbar^2}\left(\frac{r^2\cos\theta}{2}+a_0 r\cos\theta\right) \tag{17.2.42}$$

The exact second-order shift is then

$$\begin{aligned}|E^2_{100}| &= |\langle 100|H^1\Omega|100\rangle - 0|\\ &= |\langle 100|eZ\mathscr{E}\Omega|100\rangle|\\ &= \tfrac{9}{4}a_0^3\mathscr{E}^2 = \tfrac{8}{3}a_0^3\mathscr{E}^2\cdot(\tfrac{27}{32})\\ &= (0.84)\tfrac{8}{3}a_0^3\mathscr{E}^2\end{aligned} \tag{17.2.43}$$

which is roughly in the middle of the interval we restricted it to by Method 1.

Exercise 17.2.6. Verify Eq. (17.2.34).

*Exercise 17.2.7.** For the oscillator, consider $H^1=-qfX$. Find an Ω that satisfies Eq. (17.2.38). Feed it into Eq. (17.2.39) for E^2_n and compare with the earlier calculation.

Exercise 17.2.8. Fill in the steps connecting Eqs. (17.2.41) and (17.2.43). Try to use symmetry arguments to reduce the labor involved in evaluating the integrals.

We argued earlier that E^2_{100} represents the interaction of the induced dipole moment with the applied field. How big is the induced moment μ? One way to find out is to calculate $\langle\boldsymbol{\mu}\rangle$ in the perturbed ground state. An easier way to extract it from E^2_{100}. Suppose we take a system that has no intrinsic dipole moment and turn on an external electric field that starts at 0 and grows to the full value of **E**. During this time the dipole moment grows from 0 to μ. If you imagine charges $\pm q$ separated by a distance x along **E**, you can see that the work done on the system as x changes by dx is

$$\begin{aligned}dW &= -q\mathscr{E}\,dx\\ &= -\mathscr{E}\,d\mu\end{aligned} \tag{17.2.44}$$

If we assume that the induced moment is proportional to **E**:

$$\boldsymbol{\mu} = \alpha\mathbf{E} \tag{17.2.45}$$

(where α is called the *polarizability*), then

$$dW = -\alpha\mathscr{E}\,d\mathscr{E}$$

or

$$W=-\tfrac{1}{2}\alpha\mathscr{E}^2 \tag{17.2.46}$$

We identify W with E_{100}^2 and determine the polarizability

$$\alpha=\frac{18}{4}a_0^3\simeq\frac{18}{4}(0.5\text{ Å})^3\simeq 0.56\text{ Å}^3 \tag{17.2.47}$$

If we use a more accurate value $a_0=0.53$ Å, we get $\alpha=0.67$ Å^3, which is in excellent agreement with the measured value of 0.68 Å^3. For a given $\mathbf{E}$, we can get $\boldsymbol{\mu}$ from Eq. (17.2.45).

Finally note that E_{100}^2 is negative. From Eq. (17.1.17) it is clear that the second-order shift in the ground-state energy is always negative (unless it vanishes). Since E_0^2 measures the energy shift due to the first-order change in the ground-state state vector, we conclude that the system changes its configuration so as to lower its energy of interaction with the external field.

17.3. Degenerate Perturbation Theory

In the face of degeneracy ($E_n^0=E_m^0$) the condition for the validity of the perturbation expansion,

$$\left|\frac{\langle m^0|H^1|n^0\rangle}{E_n^0-E_m^0}\right|\ll 1 \tag{17.3.1}$$

is impossible to fulfill. The breakdown of the method may be understood in less formal terms as follows.

Let us consider the case when neither H^0 nor H^0+H^1 is degenerate. For the purposes of this argument imagine that H^1 is due to some external field that can be continuously varied from zero to its full value. As the total Hamiltonian grows from H^0 to H^0+H^1, the corresponding eigenbasis changes continuously from $|n^0\rangle$ to $|n\rangle$. It is this continuous or analytic variation of the eigenbasis with the perturbation that makes it possible to find $|n\rangle$ starting with $|n^0\rangle$, the way one finds the value of some analytic function at the point $x+a$ starting at the point x and using a Taylor series. Consider now the case when H^0 has a degenerate subspace and H^0+H^1 is nondegenerate in this subspace. (More general cases can be handled the same way.) Imagine starting with the basis $|n\rangle$ and slowly turning *off* the perturbation. We will end up with a basis $|\bar{n}^0\rangle$ of H^0. If we now turn on the perturbation, we can retrace our path back to $|n\rangle$. It is clear that if we start with this basis, $|\bar{n}^0\rangle$, we can evaluate $|n\rangle$ perturbatively. But since H^0 is degenerate, we needn't have started with this basis; we could have started with some other basis $|n^0\rangle$, chosen randomly. But if we start with any basis except $|\bar{n}^0\rangle$, and turn on the external field of infinitesimal size, the change in the basis will not be infinitesimal. It is this nonanalytic behavior that is signaled by the divergence in the first-order matrix element. [This can be compared to the divergence of the first derivative in the Taylor series where $f(x)$ is discontinu-

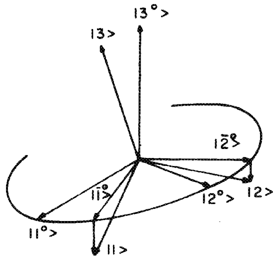

Figure 17.1. An example of the degenerate problem from $\mathbb{V}^3(R)$. In the x–y plane, which is the degenerate subspace, we must start not with some arbitrarily chosen pair of basis vectors $|1^0\rangle$ and $|2^0\rangle$, but with the pair $|\bar{1}^0\rangle$ and $|\bar{2}^0\rangle$ which diagonalizes H^1.

ous.] So, we must start with the right basis in the degenerate space. We have already characterized this basis as one we get if we start with $|n\rangle$ and slowly turn off H^1. A more useful characterization is the following: it is a basis that diagonalizes H^1 within the degenerate space. Why? Because, if we start with this basis, the first-order perturbation coefficient [Eq. (17.1.8)] does not blow up, for the (off-diagonal) matrix element in the numerator vanishes along with energy denominator whenever $|n^0\rangle$ and $|m^0\rangle$ belong to the degenerate space. Figure 17.1 depicts a simple example from $\mathbb{V}^3(R)$, where the x–y plane is the degenerate space and $|1^0\rangle$ and $|2^0\rangle$ are randomly chosen basis vectors in that subspace. The proper starting point is the pair $|\bar{1}^0\rangle$, $|\bar{2}^0\rangle$, which diagonalizes H^1 in the x–y plane.

It is worth noting that to find the proper starting point, we need to find the basis that diagonalizes H^1 only within the degenerate space and not the full Hilbert space. Thus even if we work with infinite-dimensional spaces, the exact diagonalization will usually have to be carried out only in some small, finite-dimensional subspace.

Let us consider, as a concrete example, the Stark effect in the $n=2$ level of hydrogen. (We ignore spin, which is a spectator variable.) Are we to conclude that there is no first-order shift because

$$\langle 2lm|e\mathscr{E}Z|2lm\rangle = 0 \tag{17.3.2}$$

by parity invariance, or equivalently, because the atom in these states has no intrinsic dipole moment? No, because these states need not provide the correct starting points for a perturbative calculation in view of the degeneracy. We must first find the basis in the $n=2$ sector which diagonalizes H^1. Using the selection rules, which tell us that only two of the 16 matrix elements are nonzero, we get

$$H^1 \rightarrow \begin{array}{c|cccc} nlm & 200 & 210 & 211 & 21-1 \\ \hline 200 & 0 & \Delta & 0 & 0 \\ 210 & \Delta & 0 & 0 & 0 \\ 211 & 0 & 0 & 0 & 0 \\ 21-1 & 0 & 0 & 0 & 0 \end{array} \tag{17.3.3}$$

where

$$\Delta = \langle 200|e\mathscr{E}Z|210\rangle = -3e\mathscr{E}a_0 \tag{17.3.4}$$

*Exercise 17.3.1.** Use the dipole selection rules to show that H^1 has the above form and carry out the evaluation of Δ.

Since H^1 is just Δ times the Pauli matrix σ_x in the $m=0$ sector, we infer that its eigenvalues are $\pm\Delta$ and that its eigenstates are $[|200\rangle \pm |210\rangle]/2^{1/2}$. In the $|m| = 1$ sector the old states $|2, 1, \pm 1\rangle$ diagonalize H^1. Our calculation tells us the following.

(1) The zeroth-order states stable under the perturbation are $|2, 1, \pm 1\rangle$ and $[|200\rangle \pm |210\rangle]/2^{1/2}$.

(2) The first-order shift E^1 is zero for the first two states and $\pm\Delta$ for the next two. (Note that Δ is negative.)

Notice that the stable eigenstates for which $E^1 \neq 0$ are mixtures of $l=0$ and $l=1$. Thus they have indefinite parity and can have a nonzero intrinsic dipole moment which can interact with E and produce a first-order energy shift. From the energy shift, we infer that the size of the dipole moment is $3ea_0$.

Degenerate perturbation theory is relevant not only when the levels are exactly degenerate but also when they are close, that is to say, when the inequality (17.3.1) is not respected. In that case one must diagonalize $H^0 + H^1$ exactly in the almost degenerate subspace.

Exercise 17.3.2. Consider a spin-1 particle (with no orbital degrees of freedom). Let $H = AS_z^2 + B(S_x^2 - S_y^2)$, where S_i are 3×3 spin matrices, and $A \gg B$. Treating the B term as a perturbation, find the eigenstates of $H^0 = AS_z^2$ that are stable under the perturbation. Calculate the energy shifts to first order in B. How are these related to the exact answers?

Fine Structure

The Coulomb potential $(-e^2/r)$ does not give the complete interaction between the electron and the proton, though it does provide an excellent first approximation.‡ There are "fine-structure" corrections to this basic interaction, which produce energy shifts of the order of α^2 times the binding energy due to the Coulomb potential. Since the electron velocity (in a semiclassical picture) is typically $\beta = v/c \simeq O(\alpha)$, these are corrections of the order $(v/c)^2$ *relative* to binding energy, which is itself proportional to $(v/c)^2$. Thus these are relativistic in origin. There are two parts to this effect.

The first reflects the fact that to order $(v/c)^4$ the kinetic energy of the electron is not $p^2/2m$ but

$$T = (c^2p^2 + m^2c^4)^{1/2} - mc^2 = \frac{p^2}{2m} - \frac{p^4}{8m^3c^2} + O(p^6 \text{ or } v^6) \qquad (17.3.5)$$

We now wish to calculate the effect of this extra term

$$H_T = -P^4/8m^3c^2 \qquad (17.3.6)$$

---

‡ We consider here just the fine structure of hydrogen. The analysis may be extended directly to hydrogen-like atoms. We also ignore the difference between the reduced mass and the electron mass.

treating it as a perturbation. Since H_T is rotationally invariant, it is diagonal in the $|nlm\rangle$ basis. (In other words, the $|nlm\rangle$ basis is stable under this perturbation.) So we can forget about the fact that the levels at each n are degenerate and determine E_T^1 simply from

$$E_T^1 = -\frac{1}{8m^3c^2}\langle nlm|P^4|nlm\rangle \tag{17.3.7}$$

We evaluate the matrix element by noting that

$$P^4 = 4m^2\left(\frac{P^2}{2m}\right)^2 = 4m^2\left(H^0 + \frac{e^2}{r}\right)^2 \tag{17.3.8}$$

so that

$$E_T^1 = -\frac{1}{2mc^2}\left[(E_n^0)^2 + 2E_n^0 e^2\left\langle\frac{1}{r}\right\rangle_{nlm} + e^4\left\langle\frac{1}{r^2}\right\rangle_{nlm}\right] \tag{17.3.9}$$

From the virial theorem [Eq. (13.1.34)]

$$-\left\langle\frac{e^2}{r}\right\rangle_{nlm} = 2E_n^0 \tag{17.3.10}$$

while from Exercise (17.3.4)

$$\left\langle\frac{e^4}{r^2}\right\rangle_{nlm} = \frac{e^4}{a_0^2 n^3(l+1/2)} = \frac{4E_0^2 n}{l+1/2} \tag{17.3.11}$$

so that

$$\begin{aligned} E_T^1 &= -\frac{(E_n^0)^2}{2mc^2}\left(-3 + \frac{4n}{l+1/2}\right) \\ &= -\frac{1}{2}(mc^2)\alpha^4\left[-\frac{3}{4n^4} + \frac{1}{n^3(l+1/2)}\right] \end{aligned} \tag{17.3.12}$$

The other relativistic effect is called the *spin–orbit interaction*. Its origin may be understood as follows. The Coulomb interaction $(-e^2/r)$ is the whole story only if the electron is at rest. If it moves at a velocity **v**, there is an extra term which we find as follows. In the electron rest frame, the proton will be moving at a velocity

$(-\mathbf{v})$ and will produce a magnetic field

$$\mathbf{B} = -\frac{e}{c}\frac{\mathbf{v}\times\mathbf{r}}{r^3} \tag{17.3.13}$$

The interaction of the magnetic moment of the electron with this field leads to the *spin–orbit energy*

$$\begin{aligned}\mathscr{H}_{\text{s.o.}} &= -\boldsymbol{\mu}\cdot\mathbf{B} = \frac{e}{mcr^3}\boldsymbol{\mu}\cdot(\mathbf{p}\times\mathbf{r}) \\ &= -\frac{e}{mc}\frac{\boldsymbol{\mu}\cdot\mathbf{l}}{r^3}\end{aligned} \tag{17.3.14}$$

So we expect that in the quantum theory there will be a perturbation

$$\begin{aligned}H_{\text{s.o.}} &= \left(-\frac{e}{mc}\right)\left(-\frac{e}{mc}\right)\frac{\mathbf{S}\cdot\mathbf{L}}{r^3} \\ &= \frac{e^2}{m^2c^2r^3}\mathbf{S}\cdot\mathbf{L}\end{aligned} \tag{17.3.15}$$

However, the correct answer is half as big:

$$H_{\text{s.o.}} = \frac{e^2}{2m^2c^2r^3}\mathbf{S}\cdot\mathbf{L} \tag{17.3.16}$$

The reason is that the "rest frame of the electron" doesn't have a fixed velocity relative to the CM of the atom since the motion of the electron is not rectilinear. Thus $\mathscr{H}_{\text{s.o.}}$ deduced in the comoving frame does not directly translate into what must be used in the CM frame. The transformation and the factor of 1/2 were found by Thomas.‡ In Chapter 20 we will derive Eq. (17.3.16) from the Dirac equation, which has relativistic kinematics built into it. The *Thomas factor* of 1/2 will drop out automatically.

Since $H_{\text{s.o.}}$ involves the spin, we must now reinstate it. Since the states at a given n are degenerate, we must start with a basis that diagonalizes $H_{\text{s.o.}}$. Since we can rewrite $H_{\text{s.o.}}$ as

$$H_{\text{s.o.}} = \frac{e^2}{4m^2c^2r^3}[J^2 - L^2 - S^2] \tag{17.3.17}$$

---

‡ L. H. Thomas *Nature* **117**, 574 (1926).

the states of total angular momentum suggest themselves. In this basis,

$$\langle j', m'; l', 1/2 | H_{\text{s.o.}} | j, m; l, 1/2 \rangle$$

$$= \delta_{jj'}\,\delta_{mm'}\,\delta_{ll'}\,\frac{e^2}{4m^2c^2}\left\langle \frac{1}{r^3}\right\rangle_{nl} \hbar^2[j(j+1) - l(l+1) - 3/4] \tag{17.3.18}$$

(Note that two states with the same total jm, but built from different l's, are orthogonal because of the orthogonality of the spherical harmonics. Thus, for example, at $n=2$, we can build $j=1/2$ either from $l=0$ or $l=1$. The states $|j=1/2, m; 0, 1/2\rangle$ and $|j=1/2, m; 1, 1/2\rangle$ are orthogonal.) Feeding $j=l\pm 1/2$ into Eq. (17.3.18) we get

$$E^1_{\text{s.o.}} = \frac{\hbar^2 e^2}{4m^2c^2}\left\langle \frac{1}{r^3}\right\rangle_{nl} \begin{Bmatrix} l \\ -(l+1) \end{Bmatrix} \tag{17.3.19}$$

where the upper and lower values correspond to $j=l\pm 1/2$. Using the result from Exercise 17.3.4

$$\left\langle \frac{1}{r^3}\right\rangle_{nl} = \frac{1}{a_0^3}\,\frac{1}{n^3 l(l+1/2)(l+1)} \tag{17.3.20}$$

we get

$$E^1_{\text{s.o.}} = \frac{1}{4}mc^2\alpha^4\,\frac{\begin{Bmatrix} l \\ -(l+1) \end{Bmatrix}}{n^3(l)(l+1/2)(l+1)} \tag{17.3.21}$$

This formula has been derived for $l\neq 0$. When $l=0$, $\langle 1/r^3\rangle$ diverges and $\langle \mathbf{L}\cdot\mathbf{S}\rangle$ vanishes. But if we set $l=0$ in Eq. (17.3.21) we get a finite limit, which in fact happens to give the correct level shift for $l=0$ states. This will be demonstrated when we study the Dirac equation in Chapter 20. The physical origin behind this shift (which is clearly not the spin–orbit interaction) will be discussed then. Since $E^1_{\text{s.o.}}$ and E^1_T are both α^4 effects, we combine them to get the *total fine-structure energy shift*

$$E^1_{\text{f.s.}} = E^1_T + E^1_{\text{s.o.}} = -\frac{mc^2\alpha^2}{2n^2}\cdot\frac{\alpha^2}{n}\left(\frac{1}{j+1/2} - \frac{3}{4n}\right) \tag{17.3.22}$$

for both $j=l\pm 1/2$.

The fine-structure formula can be extended to other atoms as well, provided we make the following change in Eq. (17.3.19):

$$\left\langle \frac{e^2}{r^3} \right\rangle \to \left\langle \frac{1}{r}\frac{dV}{dr} \right\rangle$$

where V is the potential energy of the electron in question. Consider, for example, the $n=4$ states of potassium. We have seen in the prespin treatment that due to penetration and shielding effects the $4s$ level lies below the $4p$ level. If we add spin to this picture, the s state can only become ${}^2S_{1/2}$ while the p state can generate both ${}^2P_{3/2}$ and ${}^2P_{1/2}$. The last two are split by the fine-structure effect‡ by an amount $(3\hbar^2/4m^2c^2)\langle(1/r)(dV/dr)\rangle$, where V is the potential seen by the $n=4$, $l=1$ electron. In the $4p \to 4s$ transition, the fine-structure interaction generates two lines in the place of one, with wavelengths 7644.9 Å and 7699.0 Å.

Exercise 17.3.3. Consider the case where H^0 includes the Coulomb plus spin–orbit interaction and H^1 is the effect of a weak magnetic field $\mathbf{B}=B\mathbf{k}$. Using the appropriate basis, show that the first-order level shift is related to j_z by

$$E^1 = \left(\frac{eB}{2mc}\right)\left(1 \pm \frac{1}{2l+1}\right) j_z, \qquad j = l \pm 1/2$$

Sketch the levels for the $n=2$ level assuming that $E^1 \ll E^1_{\text{f.s.}}$.

*Exercise 17.3.4.** We discuss here some tricks for evaluating the expectation values of certain operators in the eigenstates of hydrogen.

(1) Suppose we want $\langle 1/r \rangle_{nlm}$. Consider first $\langle \lambda/r \rangle$. We can interpret $\langle \lambda/r \rangle$ as the first-order correction due to a perturbation λ/r. Now this problem can be solved exactly; we just replace e^2 by $e^2-\lambda$ everywhere. (Why?) So the exact energy, from Eq. (13.1.16) is $E(\lambda) = -(e^2-\lambda)^2 m/2n^2\hbar^2$. The first-order correction is the term linear in λ, that is, $E^1 = me^2\lambda/n^2\hbar^2 = \langle \lambda/r \rangle$, from which we get $\langle 1/r \rangle = 1/n^2 a_0$, in agreement with Eq. (13.1.36). For later use, let us observe that as $E(\lambda) = E^0 + E^1 + \cdots = E(\lambda=0) + \lambda(dE/d\lambda)_{\lambda=0} + \cdots$, one way to extract E^1 from the exact answer is to calculate $\lambda(dE/d\lambda)_{\lambda=0}$.

(2) Consider now $\langle \lambda/r^2 \rangle$. In this case, an exact solution is possible since the perturbation just modifies the centrifugal term as follows:

$$\frac{\hbar^2 l(l+1)}{2mr^2} + \frac{\lambda}{r^2} = \frac{\hbar^2 l'(l'+1)}{2mr^2} \tag{17.3.23}$$

where l' is a function of λ. Now the dependence of E on $l'(\lambda)$ is, from Eq. (13.1.14),

$$E(l') = \frac{-me^4}{2\hbar^2(k+l'+1)^2} = E(\lambda) = E^0 + E^1 + \cdots$$

---

‡ Actually the splitting at a given l is solely due to the spin–orbit interaction. The kinetic energy correction depends only on l and does not contribute to the splitting between the $P_{3/2}$ and $P_{1/2}$ levels.

Show that

$$\left\langle\frac{\lambda}{r^3}\right\rangle = E^1 = \lambda\left.\frac{dE}{d\lambda}\right|_{\lambda=0} = \left(\frac{dE}{dl'}\right)_{l'=l}\cdot\left(\frac{dl'}{d\lambda}\right)_{l'=l}\cdot\lambda = \frac{\lambda}{n^3a_0^2(l+\frac{1}{2})}$$

Canceling λ on both sides, we get Eq. (17.3.11).

(3) Consider finally $\langle l/r^3\rangle$. Since there is no such term in the Coulomb Hamiltonian, we resort to another trick. Consider the *radial* momentum operator, $p_r = -i\hbar(\partial/\partial r + 1/r)$, in terms of which we may write the radial part of the Hamiltonian

$$\left(\frac{-\hbar^2}{2m}\right)\left(\frac{1}{r^2}\frac{\partial}{\partial r}r^2\frac{\partial}{\partial r}\right)$$

as $p_r^2/2m$. (Verify this.) Using the fact that $\langle[H, p_r]\rangle = 0$ in the energy eigenstates, and by explictly evaluating the commutator, show that

$$\left\langle\frac{1}{r^3}\right\rangle = \frac{1}{a_0(l)(l+1)}\left\langle\frac{1}{r^2}\right\rangle$$

combining which with the result from part (2) we get Eq. (17.3.20).

(4) Find the mean kinetic energy using the trick from part (1), this time rescaling the mass. Regain the virial theorem.