Quantum Mechanics Part 1

Chapter 8

John Boccio

Chapter 8 **Time-Independent Perturbation Theory**

8.1 **Nondegenerate Case**

For most physically interesting systems, it is not possible to find simple, exact formulas for the energy eigenvalues and state vectors.

In many cases, however, the real system is very *similar* to another system that we can solve exactly in closed form.

Our procedure will then be to approximate the real system by the similar system and approximately calculate corrections to find the corresponding values for the real system. The approximation method that is most often used is called perturbation theory.

8.1.1 **Rayleigh-Schrodinger Perturbation Theory**

Consider the problem of finding the energies (eigenvalues) and state vectors (eigenvectors) for a system with a Hamiltonian $\hat{H}$ that can be written in the form

$$\hat{H} = \hat{H}_0 + \hat{V} \tag{8.1}$$

where we have already solved the system described by $\hat{H}_0$, i.e., we know that

$$\hat{H}_0 \left|n\right\rangle = \varepsilon_n \left|n\right\rangle \tag{8.2}$$

with $\langle m \,|\, n\rangle = \delta_{mn}$ (remember that the eigenvectors of a Hermitian operator always form a complete orthonormal set ... or we can make them so using the Gram-Schmidt process if degeneracy exists).

We call this solvable system the *unperturbed* or *zero-order* system.

We then assume that the extra term $\hat{V}$ is a *small correction* to $\hat{H}_0$ (that is what we mean by *similar* systems).

This says that the real physical system will have a solution given by

$$\hat{H}\,|N\rangle = E_n\,|N\rangle \tag{8.3}$$

where the real physical state vectors $|N\rangle$ are only *slightly different* from the unperturbed state vectors $|n\rangle$ and the real physical energies E_n are only *slightly different* from the unperturbed energies ε_n. Mathematically, we can express this situation by writing the perturbation in the form

$$\hat{V} = g\hat{U} \tag{8.4}$$

where g is some small ($\ll 1$) constant factor pulled out of the correction term $\hat{V}$ that characterizes its strength (or effect on the system described by $\hat{H}_0$) of the perturbation.

As $g \to 0$, each eigenvector $|N\rangle$ of $\hat{H}$ must approach the corresponding eigenvector $|n\rangle$ of $\hat{H}_0$ and each energy eigenvalue E_n of $\hat{H}$ must approach the corresponding energy eigenvalue ε_n of $\hat{H}_0$.

We can guarantee that this property is true by assuming that power series expansions in the small parameter g exist for all physically relevant quantities of the real system, i.e.,

$$\hat{H} = \hat{H}_0 + \hat{V} = \hat{H}_0 + g\hat{U} \tag{8.5}$$

$$|N\rangle = |n\rangle + g\left|N^{(1)}\right\rangle + g^2\left|N^{(2)}\right\rangle + \ldots \tag{8.6}$$

$$E_n = \varepsilon_n + gE_n^{(1)} + g^2E_n^{(2)} + \ldots \tag{8.7}$$

where the terms $\left|N^{(i)}\right\rangle$ and $E_n^{(i)}$ are called the i^{th}-order correction to the unperturbed or zero-order solution. This is a major assumption, that we cannot, in general, prove is true a priori, i.e., we cannot prove that the power series converge and therefore make sense.

The usual normalization condition we might impose would be $\langle N \,|\, N\rangle = 1$. Since the results of any calculation are independent of the choice of normalization (remember the expectation value and density operator definitions all include the norm in the denominator), we choose instead to use the normalization condition

$$\langle n \mid N \rangle = 1 \tag{8.8}$$

which will greatly simplify our derivations and subsequent calculations.

Substituting the power series expansion into the normalization condition we get

$$\langle n \mid N \rangle = 1 = \langle n \mid n \rangle + g \left\langle n \,\middle|\, N^{(1)} \right\rangle + g^2 \left\langle n \,\middle|\, N^{(2)} \right\rangle + \ldots \tag{8.9}$$

But since we already have assumed that $\langle n \,|\, n \rangle = 1$, we must have

$$0 = g \left\langle n \,\middle|\, N^{(1)} \right\rangle + g^2 \left\langle n \,\middle|\, N^{(2)} \right\rangle + ... \tag{8.10}$$

Now the only way for a power series to be *identically zero* is for the coefficient of each power of g to be separately equal to zero. This gives the result

$$\left\langle n \,\middle|\, N^{(i)} \right\rangle = 0 \quad , \quad i = 1, 2, 3, 4, \ldots \tag{8.11}$$

as a direct consequence of the normalization condition, i.e., all corrections to the state vector are orthogonal to the unperturbed state vector.

We now substitute all of these power series into the original energy eigenvalue equation for $\hat{H}$:

$$\hat{H}\,|N\rangle = E_n\,|N\rangle$$

$$\left(\hat{H}_0 + g\hat{U}\right)\left(|n\rangle + g\left|N^{(1)}\right\rangle + g^2 + \left|N^{(2)}\right\rangle ....\right)$$

$$= \left(\varepsilon_n + gE_n^{(1)} + g^2E_n^{(2)} + ...\right)\left(|n\rangle + g\left|N^{(1)}\right\rangle + g^2 + \left|N^{(2)}\right\rangle ....\right)$$

We now multiply everything out and collect terms in a single power series in g. We get

$$
\begin{aligned}
0 = & \left(\hat{H}_0 \left|n\right\rangle - \varepsilon_n \left|n\right\rangle\right) g^0 \\
& + \left(\hat{H}_0 \left|N^{(1)}\right\rangle + \hat{U} \left|n\right\rangle - \varepsilon_n \left|N^{(1)}\right\rangle - E_n^{(1)} \left|n\right\rangle\right) g^1 \\
& + ... \\
& + \left(\hat{H}_0 \left|N^{(k)}\right\rangle + \hat{U} \left|N^{(k-1)}\right\rangle - \varepsilon_n \left|N^{(k)}\right\rangle \right. \\
& \left. - E_n^{(1)} \left|N^{(k-1)}\right\rangle - .... - E_n^{(k)} \left|n\right\rangle\right) g^k \\
& + ...
\end{aligned}
$$

Since the power series is equal to zero, the coefficient of each power of g must be equal to zero. We get (labelling the equation by the corresponding power of g)

$$
0^{th} - order \qquad \hat{H}_0 \left|n\right\rangle = \varepsilon_n \left|n\right\rangle \tag{8.12}
$$

which is just our original assumption = unperturbed solution.

$$1^{st} - order \qquad \hat{H}_0 \left| N^{(1)} \right\rangle + \hat{U} \left| n \right\rangle = \varepsilon_n \left| N^{(1)} \right\rangle + E_n^{(1)} \left| n \right\rangle \quad (8.13)$$

.....

.....

$$k^{th} - order \qquad \hat{H}_0 \left| N^{(k)} \right\rangle + \hat{U} \left| N^{(k-1)} \right\rangle = \varepsilon_n \left| N^{(k)} \right\rangle + E_n^{(1)} \left| N^{(k-1)} \right\rangle + ... + E_n^{(k)} \left| N^{(0)} \right\rangle \quad (8.14)$$

.....

where we have used the notation $\left| n \right\rangle = \left| N^{(0)} \right\rangle$.

Let us consider the $1^{st} - order$ equation. If we apply the linear functional $\langle n|$ we get

$$\langle n|\,\hat{H}_0\left|N^{(1)}\right\rangle + \langle n|\,\hat{U}\,|n\rangle = \langle n|\,\varepsilon_n\left|N^{(1)}\right\rangle + \langle n|\,E_n^{(1)}\,|n\rangle \quad (8.15)$$

$$\varepsilon_n\left\langle n\,\middle|\,N^{(1)}\right\rangle + \langle n|\,\hat{U}\,|n\rangle = \varepsilon_n\left\langle n\,\middle|\,N^{(1)}\right\rangle + E_n^{(1)}\,\langle n\mid n\rangle \quad (8.16)$$

or since $\left\langle n\,\middle|\,N^{(1)}\right\rangle = 0$ we get

$E_n^{(1)} = \langle n|\,\hat{U}\,|n\rangle = 1^{st} - order$ correction to the energy

$=$ diagonal matrix element of $\hat{U}$ in the $|n\rangle$ (unperturbed) basis

or the expectation value of $\hat{U}$ in the state $|n\rangle$

or in the n^{th} unperturbed state

Therefore, to first order in g we have

$$\begin{aligned} E_n &= \varepsilon_n + gE_n^{(1)} = \varepsilon_n + g\,\langle n|\,\hat{U}\,|n\rangle \\ &= \varepsilon_n + \langle n|\,\hat{V}\,|n\rangle \end{aligned} \quad (8.17)$$

where we have reabsorbed the factor g back into the original potential energy function.

In the same manner, if we apply the linear functional $\langle n|$ to the $k^{th} - order$ equation we get

$$E_n^{(k)} = \langle n| \hat{U} \left| N^{(k-1)} \right\rangle \tag{8.18}$$

This says that, if we know the correction to the eigenvector to order $(k-1)$, then we can calculate the correction to the energy eigenvalue to order k (the next order).

Now the $k^{th} - order$ correction to the eigenvector, $\left|N^{(k)}\right\rangle$ is just another vector in the space and, hence, we can expand it as a linear combination of the $|n\rangle$ states (since they are a basis).

$$\left|N^{(k)}\right\rangle = \sum_{m \neq n} |m\rangle \left\langle m \left| N^{(k)} \right.\right\rangle \tag{8.19}$$

The state $|n\rangle$ is not included because $\left\langle n \left| N^{(i)} \right.\right\rangle = 0$ by our choice of normalization.

In order to evaluate this sum, we must find an expression for the coefficients $\left\langle m \left| N^{(k)} \right.\right\rangle$.

This can be done by applying the linear functional $\langle m|$, $m \neq n$ to the $k^{th} - order$ equation. We get

$$\langle m| \hat{H}_0 \left|N^{(k)}\right\rangle + \langle m| \hat{U} \left|N^{(k-1)}\right\rangle$$
$$= \langle m| \varepsilon_n \left|N^{(k)}\right\rangle + \langle m| E_n^{(1)} \left|N^{(k-1)}\right\rangle + ... + \langle m| E_n^{(k)} \left|N^{(0)}\right\rangle$$
$$\varepsilon_m \left\langle m \left| N^{(k)}\right.\right\rangle + \langle m| \hat{U} \left|N^{(k-1)}\right\rangle$$
$$= \varepsilon_n \left\langle m \left| N^{(k)}\right.\right\rangle + E_n^{(1)} \left\langle m \left| N^{(k-1)}\right.\right\rangle + ... + E_n^{(k)} \left\langle m \left| N^{(0)}\right.\right\rangle$$

If we assume that $\varepsilon_m \neq \varepsilon_n$ (we have *nondegenerate levels*) we get

$$\left\langle n \left| M^{(k)}\right.\right\rangle$$
$$= \frac{1}{\varepsilon_n - \varepsilon_m} \left(\langle m| \hat{U} \left|N^{(k-1)}\right\rangle - E_n^{(1)} \left\langle m \left| N^{(k-1)}\right.\right\rangle \right.$$
$$\left. - ... - E_n^{(k)} \left\langle m \left| N^{(0)}\right.\right\rangle \right. \qquad (8.20)$$

This formula allows us to find the $k^{th} - order$ correction to the eigenvector in terms of lower order corrections to $|N\rangle$ and E_n as long as $|n\rangle$ corresponds to a nondegenerate level.

To see how this works, we will calculate the corrections to second order.

For first order, we let $k = 1$ and get

$$\begin{aligned}\left\langle m \left| N^{(1)} \right\rangle &= \frac{1}{\varepsilon_n - \varepsilon_m} \left(\langle m| \hat{U} \left| N^{(0)} \right\rangle - E_n^{(1)} \left\langle m \left| N^{(0)} \right\rangle \right) \right. \\ &= \frac{1}{\varepsilon_n - \varepsilon_m} \left(\langle m| \hat{U} |n\rangle - E_n^{(1)} \langle m \mid n\rangle \right) \\ &= \frac{1}{\varepsilon_n - \varepsilon_m} \langle m| \hat{U} |n\rangle \end{aligned} \tag{8.21}$$

which gives

$$\left| N^{(1)} \right\rangle = \sum_{m \neq n} |m\rangle \left\langle m \left| N^{(1)} \right\rangle = \sum_{m \neq n} |m\rangle \frac{1}{\varepsilon_n - \varepsilon_m} \langle m| \hat{U} |n\rangle \right. \tag{8.22}$$

Therefore, to first order in g we have

$$|N\rangle = |n\rangle + g\left|N^{(1)}\right\rangle = |n\rangle + \sum_{m\neq n} |m\rangle \frac{1}{\varepsilon_n - \varepsilon_m} \langle m|\hat{V}|n\rangle \quad (8.23)$$

We then use $N^{(1)}$ to calculate $E_n^{(2)}$, the second order correction to the energy, using

$$E_n^{(2)} = \langle n|\hat{U}\left|N^{(1)}\right\rangle = \langle n|\hat{U}\left(\sum_{m\neq n} |m\rangle \frac{1}{\varepsilon_n - \varepsilon_m} \langle m|\hat{U}|n\rangle\right)$$

$$= \sum_{m\neq n} \frac{\left|\langle n|\hat{U}|m\rangle\right|^2}{\varepsilon_n - \varepsilon_m} \quad (8.24)$$

Therefore, to second order in g we have

$$E_n = \varepsilon_n + gE_n^{(1)} + g^2 E_n^{(2)}$$

$$= \varepsilon_n + \langle n|\hat{V}|n\rangle + \sum_{m\neq n} \frac{\left|\langle n|\hat{V}|m\rangle\right|^2}{\varepsilon_n - \varepsilon_m} \quad (8.25)$$

We then obtain the second order correction to the state vector in the same way.

$$\begin{aligned}
\left\langle m \,\middle|\, N^{(2)} \right\rangle &= \frac{1}{\varepsilon_n - \varepsilon_m} \left(\langle m| \hat{U} \left| N^{(1)} \right\rangle - E_n^{(1)} \left\langle m \,\middle|\, N^{(1)} \right\rangle \right. \\
&\qquad \left. - E_n^{(2)} \left\langle m \,\middle|\, N^{(0)} \right\rangle \right) \\
&= \frac{1}{\varepsilon_n - \varepsilon_m} \langle m| \hat{U} \left| N^{(1)} \right\rangle - \frac{1}{\varepsilon_n - \varepsilon_m} \langle n| \hat{U} |n\rangle \left\langle m \,\middle|\, N^{(1)} \right\rangle \\
&= \frac{1}{\varepsilon_n - \varepsilon_m} \langle m| \hat{U} \left(\sum_{k \neq n} |k\rangle \frac{1}{\varepsilon_n - \varepsilon_k} \langle k| \hat{U} |n\rangle \right) \\
&\qquad - \frac{1}{\varepsilon_n - \varepsilon_m} \langle n| \hat{U} |n\rangle \frac{1}{\varepsilon_n - \varepsilon_m} \langle m| \hat{U} |n\rangle \\
&= \sum_{k \neq n} \frac{\langle m| \hat{U} |k\rangle \langle k| \hat{U} |n\rangle}{(\varepsilon_n - \varepsilon_m)(\varepsilon_n - \varepsilon_k)} - \frac{\langle n| \hat{U} |n\rangle \langle m| \hat{U} |n\rangle}{(\varepsilon_n - \varepsilon_m)^2}
\end{aligned} \tag{8.26}$$

Therefore,

$$\left|N^{(2)}\right\rangle = \sum_{m\neq n}\sum_{k\neq n}|m\rangle\,\frac{\langle m|\,\hat{U}\,|k\rangle\,\langle k|\,\hat{U}\,|n\rangle}{(\varepsilon_n-\varepsilon_m)(\varepsilon_n-\varepsilon_k)} - \sum_{m\neq n}|m\rangle\,\frac{\langle n|\,\hat{U}\,|n\rangle\,\langle m|\,\hat{U}\,|n\rangle}{(\varepsilon_n-\varepsilon_m)^2} \tag{8.27}$$

and so on.

An Example

We will now do an example where we know the exact answer so that we can compare it to the perturbation results.

We consider a 1-dimensional system represented by a perturbed harmonic oscillator where

$$\hat{H} = \hat{H}_0 + \hat{V} \tag{8.28}$$

with

$$\hat{H}_0 = \hbar\omega(\hat{a}^+\hat{a} + \frac{1}{2}) \rightarrow \text{ harmonic oscillator} \tag{8.29}$$

$$\hat{H}_0 \left|n\right\rangle = \varepsilon_n \left|n\right\rangle = \hbar\omega(n + \frac{1}{2}) \left|n\right\rangle \tag{8.30}$$

In standard operator notation

$$\hat{H}_0 = \frac{\hat{p}^2}{2m} + \frac{1}{2}kx^2 \quad , \quad k = m\omega^2 \tag{8.31}$$

We now perturb the system with the potential energy term

$$\hat{V} = \frac{1}{2}k'x^2 \quad , \quad k' << k \tag{8.32}$$

Therefore,

$$\hat{H} = \frac{\hat{p}^2}{2m} + \frac{1}{2}(k + k')x^2 \tag{8.33}$$

We still have a harmonic oscillator (with a changed spring constant). This says that the new energies are given by

$$E_n = \hbar\tilde{\omega}(n + \frac{1}{2}) \tag{8.34}$$

where

$$\tilde{\omega} = \sqrt{\frac{k + k'}{m}} = \sqrt{\frac{k}{m}}\sqrt{1 + \frac{k'}{k}} = \omega\sqrt{1 + \frac{k'}{k}} \tag{8.35}$$

Therefore, the exact energies for the perturbed system are

$$E_n = \hbar\tilde{\omega}(n + \frac{1}{2}) = \hbar\omega\sqrt{1 + \frac{k'}{k}}(n + \frac{1}{2}) \tag{8.36}$$

For $k' \ll k$, we can expand this as

$$E_n = \hbar\omega(n + \frac{1}{2})\left(1 + \frac{1}{2}\frac{k'}{k} - \frac{1}{8}\left(\frac{k'}{k}\right)^2 + \ldots\right) \tag{8.37}$$

which should correspond to the perturbation calculated energy calculated to 2^{nd} order in perturbation theory. We now do the perturbation calculation.

Our earlier derivation gives

$$E_n = \varepsilon_n + \langle n| \hat{V} |n\rangle + \sum_{m \neq n} \frac{\left| \langle n| \hat{V} |m\rangle \right|^2}{\varepsilon_n - \varepsilon_m} \tag{8.38}$$

$$|N\rangle = |n\rangle + \sum_{m \neq n} |m\rangle \frac{1}{\varepsilon_n - \varepsilon_m} \langle m| \hat{V} |n\rangle \tag{8.39}$$

where

$$\hat{V} = \frac{1}{2} k' x^2 = \frac{1}{4} \frac{k' \hbar}{m\omega} (\hat{a} + \hat{a}^+)^2 \tag{8.40}$$

We need to calculate this matrix element

$$
\begin{aligned}
\langle m|\,\hat{V}\,|n\rangle &= \frac{1}{4}\frac{k'\hbar}{m\omega}\,\langle m|\,(\hat{a}+\hat{a}^{+})^{2}\,|n\rangle \\
&= \frac{1}{4}\frac{k'\hbar}{m\omega}\,\langle m|\,\hat{a}^{2}+\hat{a}\hat{a}^{+}+\hat{a}^{+}\hat{a}+(\hat{a}^{+})^{2}\,|n\rangle \\
&= \frac{1}{4}\frac{k'\hbar}{m\omega}\,\langle m|\,\Big(\sqrt{n(n-1)}\,|n-2\rangle+(n+1)\,|n\rangle \\
&\qquad\qquad +n\,|n\rangle+\sqrt{(n+1)(n+2)}\,|n+2\rangle\Big) \\
&= \frac{1}{4}\frac{k'\hbar}{m\omega}\,\Big(\sqrt{n(n-1)}\delta_{m,n-2}+(2n+1)\delta_{m,n} \\
&\qquad\qquad +\sqrt{(n+1)(n+2)}\delta_{m,n+2}\Big) \qquad (8.41)
\end{aligned}
$$

where we have used

$$\langle m|\,\hat{a}\,|n\rangle = \sqrt{n}\,\langle m\mid n-1\rangle = \sqrt{n}\delta_{m,n-1} \qquad (8.42)$$

$$\langle m|\,\hat{a}^{+}\,|n\rangle = \sqrt{n+1}\,\langle m\mid n+1\rangle = \sqrt{n+1}\delta_{m,n+1} \qquad (8.43)$$

Therefore,

$$\langle n|\,\hat{V}\,|n\rangle = \frac{1}{4}\frac{k'\hbar}{m\omega}(2n+1) = \hbar\omega(n+\frac{1}{2})\left(\frac{1}{2}\frac{k'}{k}\right) \tag{8.44}$$

and

$$\sum_{m\neq n}\frac{\left|\langle n|\,\hat{V}\,|m\rangle\right|^2}{\varepsilon_n-\varepsilon_m} = \left(\frac{1}{4}\frac{k'\hbar}{m\omega}\right)^2\left[\frac{n(n-1)}{2\hbar\omega}-\frac{(n+1)(n+2)}{(-2\hbar\omega)}\right]$$

$$= -\hbar\omega\left(n+\frac{1}{2}\right)\left[\frac{1}{8}\left(\frac{k'}{k}\right)^2\right] \tag{8.45}$$

which then gives

$$E_n = \hbar\omega(n+\frac{1}{2})\left(1+\frac{1}{2}\frac{k'}{k}-\frac{1}{8}\left(\frac{k'}{k}\right)^2+\ldots\right) \tag{8.46}$$

in agreement with the exact result (to 2^{nd} order).

To calculate the new state vector to first order we need

$$\sum_{m\neq n} |m\rangle \frac{1}{\varepsilon_n - \varepsilon_m} \langle m| \hat{V} |n\rangle$$

$$= \frac{1}{4}\frac{k'\hbar}{m\omega}\frac{\sqrt{n(n-1)}}{2\hbar\omega} |n-2\rangle + \frac{1}{4}\frac{k'\hbar}{m\omega}\frac{\sqrt{(n+1)(n+2)}}{(-2\hbar\omega)} |n+2\rangle \tag{8.47}$$

which gives

$$|N\rangle = |n\rangle + \frac{1}{4}\frac{k'\hbar}{m\omega}\frac{\sqrt{n(n-1)}}{2\hbar\omega} |n-2\rangle - \frac{1}{4}\frac{k'\hbar}{m\omega}\frac{\sqrt{(n+1)(n+2)}}{2\hbar\omega} |n+2\rangle \tag{8.48}$$

What does the new ground state wave function look like? We have

$$|N=0\rangle = |0\rangle - \frac{\sqrt{2}}{2\hbar\omega} |2\rangle \tag{8.49}$$

and

$$\langle x \mid N=0\rangle = \langle x \mid 0\rangle - \frac{1}{4}\frac{k'\hbar}{m\omega}\frac{\sqrt{2}}{2\hbar\omega}\langle x \mid 2\rangle \tag{8.50}$$

$$\psi_{N=0}(x) = \psi_0(x) - \frac{1}{4}\frac{k'\hbar}{m\omega}\frac{\sqrt{2}}{2\hbar\omega}\psi_2(x) \tag{8.51}$$

Now we found earlier that

$$\langle x \mid 0\rangle = \psi_0(x) = \left(\frac{m\omega}{\pi\hbar}\right)^{1/4} e^{-\frac{m\omega x^2}{2\hbar}} \tag{8.42}$$

and

$$\langle x \mid 2\rangle = \psi_2(x) = \left(\frac{m\omega}{4\pi\hbar}\right)^{1/4} (2\frac{m\omega}{\hbar}x^2 - 1)e^{-\frac{m\omega x^2}{2\hbar}} \tag{8.53}$$

which gives

$$\begin{aligned}\psi_{N=0}(x) &= \left(\frac{m\omega}{\pi\hbar}\right)^{1/4} e^{-\frac{m\omega x^2}{2\hbar}} \left(1 + \frac{1}{4}\frac{k'\hbar}{m\omega}\frac{\sqrt{2}}{2\hbar\omega}\frac{1}{\sqrt{2}}\left(1 - 2\frac{m\omega}{\hbar}x^2\right)\right) \\ &= \left(\frac{m\omega}{\pi\hbar}\right)^{1/4} e^{-\frac{m\omega x^2}{2\hbar}} \left(1 + \frac{k'}{8k} - \frac{m\omega k'}{4\hbar k}x^2\right)\end{aligned} \tag{8.54}$$

Since we are only changing the spring constant we should have

$$\begin{aligned}
\psi_{N=0}(x) &= \left(\frac{m\tilde{\omega}}{\pi\hbar}\right)^{1/4} e^{-\frac{m\tilde{\omega}x^2}{2\hbar}} = \left(\frac{m\omega}{\pi\hbar}\right)^{1/4} \left(1+\frac{k'}{k}\right)^{1/8} e^{-\frac{m\omega x^2}{2\hbar}\sqrt{1+\frac{k'}{k}}} \\
&= \left(\frac{m\omega}{\pi\hbar}\right)^{1/4} \left(1+\frac{k'}{8k}\right) e^{-\frac{m\omega x^2}{2\hbar}\left(1+\frac{k'}{2k}\right)} \\
&= \left(\frac{m\omega}{\pi\hbar}\right)^{1/4} \left(1+\frac{k'}{8k}\right)\left(1-\frac{m\omega k'}{4\hbar k}x^2\right) e^{-\frac{m\omega x^2}{2\hbar}} \\
&= \left(\frac{m\omega}{\pi\hbar}\right)^{1/4} e^{-\frac{m\omega x^2}{2\hbar}} \left(1+\frac{k'}{8k}-\frac{m\omega k'}{4\hbar k}x^2\right) \qquad (8.55)
\end{aligned}$$

which agrees with the perturbation result to this order.

The perturbation theory we have developed so far breaks down if there are any states where

$$\varepsilon_n = \varepsilon_m \quad \text{but} \quad \langle m|\,\hat{V}\,|n\rangle \neq 0 \qquad (8.56)$$

i.e., degenerate states with nonzero matrix elements of the perturbing potential between them.

8.2 **Degenerate Case**

We handle this case as follows. Suppose we have a group of k states

$$|n_1\rangle \ , \ |n_2\rangle \ , \ |n_3\rangle \ , \ \ldots \ , \ |n_k\rangle \tag{8.57}$$

that are degenerate states of the unperturbed Hamiltonian $\hat{H}_0$, i.e.,

$$\hat{H}_0 \, |n_i\rangle = \varepsilon_{n_1} \, |n_i\rangle \ , \ i = 1, 2, 3, 4, 5, \ldots, k \tag{8.58}$$

If $\langle n_i | \, \hat{V} \, |n_j\rangle \neq 0$ for $i \neq j$ within this set, the previous perturbation formulas will fail because the energy denominators $\varepsilon_{n_i} - \varepsilon_{n_j} \to 0$.

Remember, however, that any linear combination of the degenerate states

$$|n_1\rangle \ , \ |n_2\rangle \ , \ |n_3\rangle \ , \ \ldots \ , \ |n_k\rangle \tag{8.59}$$

is also an eigenstate of $\hat{H}_0$ with the *same* energy ε_{n_1}.

Therefore, if we can choose a different set of basis states (start off with a *new set of zero-order states*) within this degenerate subspace, i.e., choose a new set of k orthogonal states (linear combinations of the old set of degenerate states)

$$|n_\alpha\rangle = \sum_{i=1}^{k} C_{\alpha i} |n_i\rangle \tag{8.60}$$

such that we have

$$\langle n_\alpha | \hat{V} | n_\beta \rangle = 0 \text{ for } \alpha \neq \beta \tag{8.61}$$

then we can use the perturbation formulas as derived earlier.

This procedure will work because the terms with zero denominators will have zero numerators and if one looks at the derivation, this means that these terms do not even appear in the final results, i.e., the zero numerators take effect before the zero denominators appear.

This condition says that the correct choice of zero-order states *within the degenerate subspace (the set of degenerate vectors)* for doing degenerate perturbation theory is that set which diagonalizes the matrix representation of $\hat{V}$ *within each* group of degenerate states.

The problem of diagonalizing $\hat{V}$ within a group of k states

$$|n_1\rangle \ , \ |n_2\rangle \ , \ |n_3\rangle \ , \ \ldots \ , \ |n_k\rangle \tag{8.62}$$

that of finding the eigenvectors and eigenvalues of the $k \times k$ matrix

$$\begin{pmatrix} \langle n_1|\hat{V}|n_1\rangle & \langle n_1|\hat{V}|n_2\rangle & \cdot & \langle n_1|\hat{V}|n_k\rangle \\ \langle n_2|\hat{V}|n_1\rangle & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \langle n_k|\hat{V}|n_1\rangle & \cdot & \cdot & \langle n_k|\hat{V}|n_k\rangle \end{pmatrix} \tag{8.63}$$

We now show that if the coefficients $C_{\alpha i}$ of the new zero-order states are just the components of the eigenvectors of this matrix, then it will be diagonalized in the degenerate subspace.

Suppose that we represent the eigenvector by the column vector

$$|n_\alpha\rangle = \begin{pmatrix} C_{\alpha 1} \\ C_{\alpha 2} \\ . \\ C_{\alpha k} \end{pmatrix} \tag{8.64}$$

then the statement that $|n_\alpha\rangle$ is an eigenvector of the $k \times k$ submatrix of $\hat{V}$ with eigenvalues that we write as $E^{(1)}_{n_\alpha}$ is equivalent to writing

$$\hat{V}\,|n_\alpha\rangle = E^{(1)}_{n_\alpha}\,|n_\alpha\rangle \tag{8.65}$$

or

$$\begin{pmatrix} \langle n_1|\hat{V}|n_1\rangle & \langle n_1|\hat{V}|n_2\rangle & \cdot & \langle n_1|\hat{V}|n_k\rangle \\ \langle n_2|\hat{V}|n_1\rangle & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \langle n_k|\hat{V}|n_1\rangle & \cdot & \cdot & \langle n_k|\hat{V}|n_k\rangle \end{pmatrix} \begin{pmatrix} C_{\alpha 1} \\ C_{\alpha 2} \\ \cdot \\ C_{\alpha k} \end{pmatrix} = E^{(1)}_{n_\alpha} \begin{pmatrix} C_{\alpha 1} \\ C_{\alpha 2} \\ \cdot \\ C_{\alpha k} \end{pmatrix} \tag{8.66}$$

or finally

$$\sum_i \langle n_j|\hat{V}|n_i\rangle\, C_{\alpha i} = E^{(1)}_{n_\alpha} C_{\alpha i} \tag{8.67}$$

All of these calculations take place within the k-dimensional degenerate subspace. We will assume that the eigenvectors are normalized, which implies that

$$\sum_i |C_{\alpha i}|^2 = 1 \tag{8.68}$$

i.e., vectors are normalized to one.

Now consider another of the new vectors given by

$$|n_\beta\rangle = \sum_{j=1}^{k} C_{\beta j} \, |n_j\rangle \tag{8.69}$$

We then have

$$\langle n_\beta| = \sum_{j=1}^{k} C^*_{\beta j} \, \langle n_j| \tag{8.70}$$

Applying the linear functional $\langle n_\beta|$ to the eigenvector/eigenvalue equation we get

$$\sum_j \sum_i \langle n_j | C^*_{\beta j} \hat{V} C_{\alpha i} \, |n_i\rangle = E^{(1)}_{n_\alpha} \sum_j C^*_{\beta j} C_{\alpha i} \tag{8.71}$$

Now, since the eigenvectors of any Hermitian matrix are always a complete orthonormal set (or can always be made so using the Gram-Schmidt process), the orthonormality of the new vectors says that

$$\sum_j C^*_{\beta j} C_{\alpha i} = \langle n_\beta \mid n_\alpha \rangle = \delta_{\beta\alpha} \tag{8.72}$$

Therefore the vectors

$$|n_\alpha\rangle = \sum_{i=1}^{k} C_{\alpha i} |n_i\rangle \tag{8.73}$$

satisfy

$$\left(\sum_j \langle n_j | C^*_{\beta j} \right) \hat{V} \left(\sum_i C_{\alpha i} |n_i\rangle \right) = E^{(1)}_{n_\alpha} \delta_{\alpha\beta} \tag{8.74}$$

$$\langle n_\beta | \hat{V} |n_\alpha\rangle = E^{(1)}_{n_\alpha} \delta_{\alpha\beta} \tag{8.75}$$

This says that the corresponding eigenvalue $E^{(1)}_{n_\alpha}$ of one of the new vectors is the first order energy corrections for the state $|N_\alpha\rangle$. The states

$$|n_\alpha\rangle \ , \ |n_\beta\rangle \ , \ |n_\gamma\rangle \ , \ \ldots \ , \ |n_\kappa\rangle \tag{8.76}$$

are called the *new zeroth-order state vectors.*

Thus the group of states

$$|n_1\rangle \ , \ |n_2\rangle \ , \ |n_3\rangle \ , \ \ldots \ , \ |n_k\rangle \tag{8.77}$$

in the presence of the perturbation $\hat{V}$ split(rearrange) into the k states

$$|n_\alpha\rangle \ , \ |n_\beta\rangle \ , \ |n_\gamma\rangle \ , \ \ldots \ , \ |n_\kappa\rangle \tag{8.78}$$

which are given to first order by

$$|N_\alpha\rangle = |n_\alpha\rangle + \sum_{m \neq \alpha,\beta,..,\kappa} \frac{|m\rangle \langle m| \hat{V} |n_\alpha\rangle}{\varepsilon_{n_1} - \varepsilon_m} \tag{8.79}$$

and the energy shift to second order is

$$E_{n_\alpha} = \varepsilon_{n_1} + \langle n_\alpha | \hat{V} | n_\alpha \rangle + \sum_{m \neq \alpha, \beta, .., \kappa} \frac{\left| \langle m | \hat{V} | n_\alpha \rangle \right|^2}{\varepsilon_{n_1} - \varepsilon_m} \tag{8.80}$$

where

$$\langle n_\alpha | \hat{V} | n_\alpha \rangle = E^{(1)}_{n_\alpha} \tag{8.81}$$

is an eigenvalue of the $\hat{V}$ matrix in the degenerate subspace.

An Example

We now consider a 2−dimensional oscillator that is perturbed by a potential of the form

$$\hat{V} = \lambda \hat{x} \hat{y} \tag{8.82}$$

We then have

$$\hat{H} = \hat{H}_0 + \hat{V} \tag{8.83}$$

where

$$\hat{H}_0 = \frac{\hat{p}_x^2}{2m} + \frac{\hat{p}_y^2}{2m} + \frac{1}{2}k(x^2 + y^2) \tag{8.84}$$

As we showed earlier, using the $\hat{a}_x$ and $\hat{a}_y$ operators we get

$$\hat{H}_0 = \hbar\omega(\hat{a}_x^+ \hat{a}_x + \hat{a}_y^+ \hat{a}_y + 1) \tag{8.85}$$

$$\hat{H}_0 \left|n_x, n_y\right\rangle = \varepsilon_{n_x,n_y} \left|n_x, n_y\right\rangle = \hbar\omega(n_x + n_y + 1) \left|n_x, n_y\right\rangle \tag{8.86}$$

$$\text{degeneracy } = n_x + n_y + 1 \tag{8.87}$$

and

$$\hat{V} = \lambda \frac{\hbar}{2m\omega}(\hat{a}_x + \hat{a}_x^+)(\hat{a}_y + \hat{a}_y^+) \tag{8.88}$$

The unperturbed ground-state is $|0,0\rangle$ with $\varepsilon_{0,0} = \hbar\omega$. It is a nondegenerate level, so we can apply the standard perturbation theory to get

$$E_0 = \varepsilon_{0,0} + \langle 0,0| \hat{V} |0,0\rangle + \sum_{\substack{m\neq 0 \\ n\neq 0}} \frac{\left|\langle m,n| \hat{V} |0,0\rangle\right|}{\varepsilon_{0,0} - \varepsilon_{m,n}} \tag{8.89}$$

Now

$$\langle 0,0|\,\hat{V}\,|0,0\rangle = \frac{\lambda\hbar}{2m\omega}\,\langle 0,0|\,(\hat{a}_x+\hat{a}_x^+)(\hat{a}_y+\hat{a}_y^+)\,|0,0\rangle = 0 \quad (8.90)$$

and

$$\begin{aligned}\langle m,n|\,\hat{V}\,|0,0\rangle &= \frac{\lambda\hbar}{2m\omega}\,\langle m,n|\,(\hat{a}_x+\hat{a}_x^+)(\hat{a}_y+\hat{a}_y^+)\,|0,0\rangle \\ &= \frac{\lambda\hbar}{2m\omega}\,\langle m,n|\,\hat{a}_x^+\hat{a}_y^+\,|0,0\rangle = \frac{\lambda\hbar}{2m\omega}\,\langle m,n\mid 1,1\rangle \\ &= \frac{\lambda\hbar}{2m\omega}\delta_{m,1}\delta_{n,1} \qquad (8.91)\end{aligned}$$

Thus, the correction to first order is zero. Calculating to second order we get

$$E_0 = \hbar\omega - \left(\frac{\lambda\hbar}{2m\omega}\right)^2 \frac{1}{2\hbar\omega} = \hbar\omega\left(1-\frac{\lambda^2}{8m^2\omega^4}\right) \qquad (8.92)$$

The next unperturbed level is 2−fold degenerate, i.e.,

$$n_x = 0, n_y = 1 \rightarrow \varepsilon_{0,1} = 2\hbar\omega$$
$$n_x = 1, n_y = 0 \rightarrow \varepsilon_{1,0} = 2\hbar\omega$$

For ease of notation we will sometimes denote

$$|1,0\rangle = |a\rangle \text{ and } |0,1\rangle = |b\rangle \tag{8.93}$$

We now use degenerate perturbation theory.

The procedure is to evaluate the $\hat{V}$ matrix in the 2×2 degenerate subspace, diagonalize it and obtain the first order corrections.

The 2×2 matrix is

$$V = \begin{pmatrix} \hat{V}_{aa} & \hat{V}_{ab} \\ \hat{V}_{ba} & \hat{V}_{bb} \end{pmatrix} = \begin{pmatrix} \langle a|\hat{V}|a\rangle & \langle a|\hat{V}|b\rangle \\ \langle b|\hat{V}|a\rangle & \langle b|\hat{V}|b\rangle \end{pmatrix} \tag{8.94}$$

Now

$$\hat{V}_{aa} = \langle 1,0|\, \hat{V}\, |1,0\rangle = 0 = \langle 0,1|\, \hat{V}\, |0,1\rangle = \hat{V}_{bb} \tag{8.95}$$

$$\begin{aligned} \hat{V}_{ab} = \hat{V}_{ba} &= \langle 1,0|\, \hat{V}\, |0,1\rangle = \frac{\lambda\hbar}{2m\omega} \langle 1,0|\, \hat{a}_x^+ \hat{a}_y\, |0,1\rangle \\ &= \frac{\lambda\hbar}{2m\omega} \end{aligned} \tag{8.96}$$

Therefore the 2×2 submatrix is

$$V = \frac{\lambda\hbar}{2m\omega} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \tag{8.97}$$

This is simple to diagonalize. We get these results

$$|a'\rangle = \frac{1}{\sqrt{2}}(|a\rangle + |b\rangle) = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \rightarrow \text{ eigenvalue } = +\frac{\lambda\hbar}{2m\omega} \tag{8.98}$$

$$|b'\rangle = \frac{1}{\sqrt{2}}(|a\rangle - |b\rangle) = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix} \rightarrow \text{ eigenvalue } = -\frac{\lambda\hbar}{2m\omega} \tag{8.99}$$

$|a'\rangle$ and $|b'\rangle$ are the *new zeroth order state vectors* (eigenvectors of the 2×2 submatrix) and

$$\pm \frac{\lambda \hbar}{2m\omega} \tag{8.100}$$

are the corresponding first order energy corrections.

Thus, the 2−fold degenerate level splits into two nondegenerate levels as shown in the figure below.

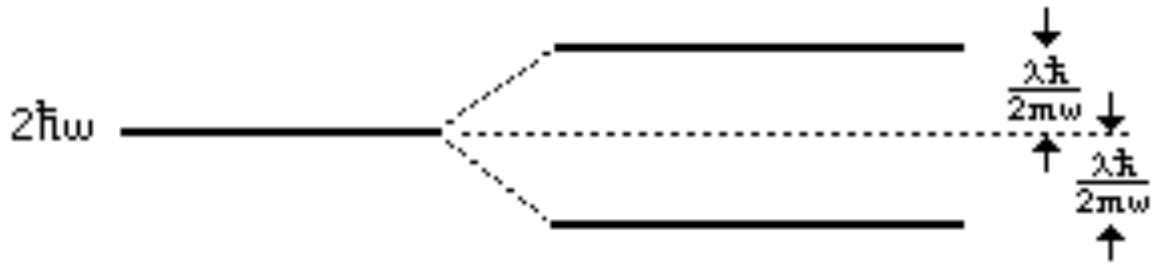

Figure: Splitting of a degenerate Level

where

$$E_{a'} = 2\hbar\omega - \frac{\lambda\hbar}{2m\omega} \tag{8.101}$$

$$E_{b'} = 2\hbar\omega + \frac{\lambda\hbar}{2m\omega} \tag{8.102}$$

$$\Delta E = \text{levelsplitting} = \frac{\lambda\hbar}{m\omega} \tag{8.103}$$

Another Example

Now let us consider a system of two spin$-1/2$ particles in a magnetic field. We also assume that there exists a direct spin-spin interaction so that the Hamiltonian takes the form

$$\hat{H} = \left(\alpha\vec{S}_{1,op} + \beta\vec{S}_{2,op}\right) \cdot \vec{B} + \gamma\vec{S}_{1,op} \cdot \vec{S}_{2,op} \tag{8.104}$$

If we choose $\vec{B} = B\hat{z}$ and $\gamma \gg \alpha, \beta$, then we can write

$$\hat{H} = \hat{H}_0 + \hat{V} \tag{8.105}$$

$$\hat{H}_0 = \gamma \vec{S}_{1,op} \cdot \vec{S}_{2,op} \tag{8.106}$$

$$\hat{V} = \alpha B \hat{S}_{1z} + \beta B \hat{S}_{2z} \tag{8.107}$$

We define

$$\vec{S}_{op} = \vec{S}_{1,op} + \vec{S}_{2,op} = \text{ total spin angular momentum} \tag{8.108}$$

and then we have

$$\begin{aligned} \vec{S}_{op}^2 = \vec{S}_{op} \cdot \vec{S}_{op} &= (\vec{S}_{1,op} + \vec{S}_{2,op}) \cdot (\vec{S}_{1,op} + \vec{S}_{2,op}) \\ &= \vec{S}_{1,op}^2 + \vec{S}_{2,op}^2 + 2\vec{S}_{1,op} \cdot \vec{S}_{2,op} \\ &= \frac{3}{4}\hbar^2 \hat{I} + \frac{3}{4}\hbar^2 \hat{I} + 2\vec{S}_{1,op} \cdot \vec{S}_{2,op} \end{aligned} \tag{8.109}$$

or

$$\hat{H}_0 = \gamma \vec{S}_{1,op} \cdot \vec{S}_{2,op} = \frac{\gamma}{2}\left(\vec{S}_{op}^2 - \frac{3}{2}\hbar^2 \hat{I}\right) \tag{8.110}$$

Our earlier discussion of the addition of angular momentum says that when we add two spin$-1/2$ angular momenta we get the resultant total angular momentum values 0 and 1, i.e.,

$$\frac{1}{2} \otimes \frac{1}{2} = 0 \oplus 1 \tag{8.111}$$

Each separate spin$-1/2$ system has the eigenvectors/eigenvalues

$$\vec{S}^2_{1,op} \left|\pm\right\rangle = \frac{3}{4}\hbar \left|\pm\right\rangle \quad , \quad \vec{S}_{1z} \left|\pm\right\rangle = \pm\frac{\hbar}{2} \left|\pm\right\rangle \tag{8.112}$$

The corresponding direct-product states are

$$\left|++\right\rangle \ , \left|+-\right\rangle \ , \left|-+\right\rangle \ , \left|--\right\rangle \tag{8.113}$$

where the symbols mean

$$\left|+-\right\rangle = \left|+\right\rangle_1 \left|-\right\rangle_2 \tag{8.114}$$

and so on.

The total angular momentum states are (we derived them earlier) labeled as $|s, m\rangle$ where

$$\vec{S}_{op}^{2} |s, m\rangle = \hbar^2 s(s+1) |s, m\rangle \tag{8.115}$$

$$\hat{S}_z |s, m\rangle = (\hat{S}_{1z} + \hat{S}_{2z}) |s, m\rangle = \pm m\hbar |s, m\rangle \tag{8.116}$$

They are given in terms of the direct product states by

$$|1, 1\rangle = |{+}{+}\rangle = |1\rangle \tag{8.117}$$

$$|1, 0\rangle = \frac{1}{\sqrt{2}} |{+}{-}\rangle + \frac{1}{\sqrt{2}} |{-}{+}\rangle = |2\rangle \tag{8.118}$$

$$|1, -1\rangle = |{-}{-}\rangle = |3\rangle \tag{8.119}$$

$$|0, 0\rangle = \frac{1}{\sqrt{2}} |{+}{-}\rangle - \frac{1}{\sqrt{2}} |{-}{+}\rangle = |4\rangle \tag{8.120}$$

The total angular momentum states are eigenstates of $\hat{H}_0$ and we use them as the unperturbed or zero-order states.

$$\hat{H}_0 \left|1,1\right\rangle = \frac{\gamma\hbar^2}{4} \left|1,1\right\rangle = \varepsilon_1 \left|1,1\right\rangle \tag{8.121}$$

$$\hat{H}_0 \left|1,0\right\rangle = \frac{\gamma\hbar^2}{4} \left|1,0\right\rangle = \varepsilon_2 \left|1,0\right\rangle \tag{8.122}$$

$$\hat{H}_0 \left|1,-1\right\rangle = \frac{\gamma\hbar^2}{4} \left|1,-1\right\rangle = \varepsilon_3 \left|1,-1\right\rangle \tag{8.123}$$

$$\hat{H}_0 \left|0,0\right\rangle = -\frac{3\gamma\hbar^2}{4} \left|0,0\right\rangle = \varepsilon_4 \left|0,0\right\rangle \tag{8.124}$$

We thus have one nondegenerate level and one 3−fold degenerate level. Now using

$$\hat{V} = \alpha B \hat{S}_{1z} + \beta B \hat{S}_{2z} \tag{8.125}$$

we do perturbation theory on these levels.

Nondegenerate Level

First order:

$$
\begin{aligned}
E_4^{(1)} &= \langle 4| \hat{V} |4\rangle \\
&= \left(\frac{1}{\sqrt{2}} \langle +-| + \frac{1}{\sqrt{2}} \langle -+| \right) (\alpha B \hat{S}_{1z} + \beta B \hat{S}_{2z}) \left(\frac{1}{\sqrt{2}} |+-\rangle - \frac{1}{\sqrt{2}} |-+\rangle \right) \\
&= 0
\end{aligned}
$$

Second order:

$$E_4^{(2)} = \sum_{m \neq 4} \frac{\left| \langle m | \hat{V} | 4 \rangle \right|^2}{\varepsilon_4 - \varepsilon_m}$$

$$= \frac{\left| \langle ++ | \, \alpha B \hat{S}_{1z} + \beta B \hat{S}_{2z} \left(\frac{1}{\sqrt{2}} |+-\rangle - \frac{1}{\sqrt{2}} |-+\rangle \right) \right|^2}{\varepsilon_4 - \varepsilon_1}$$

$$+ \frac{\left| \left(\frac{1}{\sqrt{2}} \langle +- | + \frac{1}{\sqrt{2}} \langle -+ | \right) \alpha B \hat{S}_{1z} + \beta B \hat{S}_{2z} \left(\frac{1}{\sqrt{2}} |+-\rangle - \frac{1}{\sqrt{2}} |-+\rangle \right) \right|^2}{\varepsilon_4 - \varepsilon_2}$$

$$+ \frac{\left| \langle -- | \, \alpha B \hat{S}_{1z} + \beta B \hat{S}_{2z} \left(\frac{1}{\sqrt{2}} |+-\rangle - \frac{1}{\sqrt{2}} |-+\rangle \right) \right|^2}{\varepsilon_4 - \varepsilon_3}$$

$$
\begin{aligned}
&E_4^{(2)} \\
&= \frac{\left|\left(\frac{1}{\sqrt{2}}\langle +-| + \frac{1}{\sqrt{2}}\langle -+|\right)\alpha B\hat{S}_{1z} + \beta B\hat{S}_{2z}\left(\frac{1}{\sqrt{2}}|+-\rangle - \frac{1}{\sqrt{2}}|-+\rangle\right)\right|^2}{\varepsilon_4 - \varepsilon_2} \\
&= -\frac{B^2(\alpha-\beta)^2}{\gamma}
\end{aligned}
$$

Therefore the energy to second order for the non-degenerate level is

$$
E_4 = \varepsilon_4 + E_4^{(1)} + E_4^{(2)} = -\frac{3\gamma\hbar^2}{4} - \frac{B^2(\alpha-\beta)^2}{\gamma} \qquad (8.126)
$$

Degenerate Level

In this case, the 3×3 degenerate submatrix of $\hat{V}$ is

$$\begin{pmatrix} \langle 1|\hat{V}|1\rangle & \langle 1|\hat{V}|2\rangle & \langle 1|\hat{V}|3\rangle \\ \langle 2|\hat{V}|1\rangle & \langle 2|\hat{V}|2\rangle & \langle 2|\hat{V}|3\rangle \\ \langle 3|\hat{V}|1\rangle & \langle 3|\hat{V}|2\rangle & \langle 3|\hat{V}|3\rangle \end{pmatrix} = \frac{(\alpha+\beta)\hbar B}{2} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix} \tag{8.127}$$

which is already diagonal. Since the diagonal elements are the first order energy corrections, we have (to first order)

$$E_1 = \frac{\gamma\hbar^2}{4} + \frac{(\alpha+\beta)\hbar B}{2} \tag{8.128}$$

$$E_2 = \frac{\gamma\hbar^2}{4} \tag{8.129}$$

$$E_3 = \frac{\gamma\hbar^2}{4} - \frac{(\alpha+\beta)\hbar B}{2} \tag{8.130}$$

Exact Solution

We can, in fact, solve this problem exactly and compare it to the perturbation result. We do this by choosing a new basis set (arbitrary choice made to simplify calculations) and rewriting $\hat{H}$ in terms of operators appropriate to the basis choice(that is what is meant by simplify calculations).

We then use the new basis to construct the 4×4 $\hat{H}$ matrix and then diagonalize the matrix. This method always works for a system with a small number of states.

Choose the direct product states as a basis

$$|++\rangle = |1\rangle \ , \ |+-\rangle = |2\rangle \ , \ |-+\rangle = |3\rangle \ , \ |--\rangle = |4\rangle \quad (8.131)$$

Write $\hat{H}$ as (choose operators appropriate(easy to calculate) to the basis or the HOME space)

$$\begin{aligned}\hat{H} &= \alpha B\hat{S}_{1z} + \beta B\hat{S}_{2z} + \gamma\left(\hat{S}_{1z}\hat{S}_{2z} + \hat{S}_{1x}\hat{S}_{2x} + \hat{S}_{1y}\hat{S}_{2y}\right) \\ &= \alpha B\hat{S}_{1z} + \beta B\hat{S}_{2z} + \gamma\left(\hat{S}_{1z}\hat{S}_{2z} + \frac{1}{2}\left(\hat{S}_{1+}\hat{S}_{2-} + \hat{S}_{1-}\hat{S}_{2+}\right)\right)\end{aligned} \tag{8.132}$$

Construct the 4×4 $\hat{H}$ matrix

$$\begin{pmatrix} \langle 1|\hat{H}|1\rangle & \langle 1|\hat{H}|2\rangle & \langle 1|\hat{H}|3\rangle & \langle 1|\hat{H}|4\rangle \\ \langle 2|\hat{H}|1\rangle & \langle 2|\hat{H}|2\rangle & \langle 2|\hat{H}|3\rangle & \langle 2|\hat{H}|4\rangle \\ \langle 3|\hat{H}|1\rangle & \langle 3|\hat{H}|2\rangle & \langle 3|\hat{H}|3\rangle & \langle 3|\hat{H}|4\rangle \\ \langle 4|\hat{H}|1\rangle & \langle 4|\hat{H}|2\rangle & \langle 4|\hat{H}|3\rangle & \langle 4|\hat{H}|4\rangle \end{pmatrix} \tag{8.133}$$

using

$$\hat{S}_z \left|\pm\right\rangle = \pm\frac{\hbar}{2}\left|\pm\right\rangle \tag{8.134}$$

$$\hat{S}_+ \left|+\right\rangle = 0 = \hat{S}_- \left|-\right\rangle \tag{8.135}$$

$$\hat{S}_+ \left|-\right\rangle = \hbar\left|+\right\rangle \text{ and } \hat{S}_- \left|+\right\rangle = \hbar\left|-\right\rangle \tag{8.136}$$

We get

$$\begin{pmatrix} B\hbar(\alpha+\beta) + \frac{\gamma\hbar^2}{4} & 0 & 0 & 0 \\ 0 & B\hbar(\alpha-\beta) - \frac{\gamma\hbar^2}{4} & \frac{\gamma\hbar^2}{2} & 0 \\ 0 & \frac{\gamma\hbar^2}{2} & -B\hbar(\alpha-\beta) - \frac{\gamma\hbar^2}{4} & 0 \\ 0 & 0 & 0 & -B\hbar(\alpha+\beta) + \frac{\gamma\hbar^2}{4} \end{pmatrix}$$

Diagonalizing to get the eigenvalues we find the exact energies

$$E_1 = \frac{\gamma\hbar^2}{4} + B\hbar(\alpha+\beta) \quad , \quad E_2 = -\frac{\gamma\hbar^2}{4} + \frac{1}{2}\sqrt{\gamma^2\hbar^4 + 4B^2\hbar^2(\alpha-\beta)^2}$$

$$E_3 = -\frac{\gamma\hbar^2}{4} - \frac{1}{2}\sqrt{\gamma^2\hbar^4 + 4B^2\hbar^2(\alpha-\beta)^2} \quad , \quad E_4 = \frac{\gamma\hbar^2}{4} - B\hbar(\alpha+\beta)$$

To compare to the perturbation calculation we let $B \to 0$ and we get the approximation

$$E_1 = \frac{\gamma\hbar^2}{4} + B\hbar(\alpha+\beta) \quad , \quad E_2 = \frac{\gamma\hbar^2}{4} + \frac{B^2(\alpha-\beta)^2}{\gamma}$$

$$E_3 = -\frac{3\gamma\hbar^2}{4} - \frac{B^2(\alpha-\beta)^2}{\gamma} \quad , \quad E_4 = \frac{\gamma\hbar^2}{4} - B\hbar(\alpha+\beta)$$

which agrees with the perturbation results.

8.2.1 More Ideas about Perturbation Methods

The main problem with Rayleigh-Schrodinger perturbation theory ($RSPT$) is that the form of the higher order terms becomes increasingly complex and, hence, the series is difficult to evaluate.

The Brillouin-Wigner Method

This technique allows us to see the higher order structure of the perturbation series more clearly.

Consider the energy eigenvalue equation

$$\hat{H}\left|N\right\rangle = E_n\left|N\right\rangle = \left(\hat{H}_0 + g\hat{U}\right)\left|N\right\rangle \tag{8.137}$$

Applying the linear functional $\langle m|$ we get

$$\langle m|\,\hat{H}\,|N\rangle = E_n\,\langle m\mid N\rangle = \langle m|\,\hat{H}_0\,|N\rangle + g\,\langle m|\,\hat{U}\,|N\rangle$$
$$E_n\,\langle m\mid N\rangle = \varepsilon_m\,\langle m\mid N\rangle + g\,\langle m|\,\hat{U}\,|N\rangle$$
$$(E_n-\varepsilon_m)\,\langle m\mid N\rangle = g\,\langle m|\,\hat{U}\,|N\rangle \tag{8.138}$$

We will use the normalization $\langle n\,|\,N\rangle = 1$ once again.

Now since the $|m\rangle$ states are a complete orthonormal basis we can always write

$$|N\rangle = \sum_m |m\rangle\,\langle m\mid N\rangle = |n\rangle\,\langle n\mid N\rangle + \sum_{m\neq n}|m\rangle\,\langle m\mid N\rangle$$
$$= |n\rangle + \sum_{m\neq n}|m\rangle\,\langle m\mid N\rangle \tag{8.139}$$

Using the results above (10.139) we get

$$|N\rangle = |n\rangle + \sum_{m\neq n}|m\rangle\frac{g}{E_n-\varepsilon_m}\,\langle m|\,\hat{U}\,|N\rangle \tag{8.140}$$

We now develop a series expansion of $|N\rangle$ in powers of g as follows:

$0^{th} - order$:

$$|N\rangle = |n\rangle$$

$1^{st} - order$: (substitute $0^{th} - order$ result for $|N\rangle$ into general formula (8.140))

$$|N\rangle = |n\rangle + \sum_{m \neq n} |m\rangle \frac{g}{E_n - \varepsilon_m} \langle m| \hat{U} |n\rangle$$

This is *not the same result* as in RSPT since the full energy E_n remains in the denominator.

$2^{nd} - order$: (substitute $1^{st} - order$ result for $|N\rangle$ into general formula (8.140))

$$
\begin{aligned}
|N\rangle &= |n\rangle + \sum_{m\neq n} |m\rangle \frac{g}{E_n - \varepsilon_m} \langle m| \hat{U} \left(|n\rangle + \sum_{j\neq n} |j\rangle \frac{g}{E_n - \varepsilon_j} \langle j| \hat{U} |n\rangle \right) \\
&= |n\rangle + \sum_{m\neq n} |m\rangle \frac{g}{E_n - \varepsilon_m} \langle m| \hat{U} |n\rangle \\
&\quad + g^2 \sum_{m\neq n} \sum_{j\neq n} |m\rangle \frac{g}{E_n - \varepsilon_m} \langle m| \hat{U} |j\rangle \frac{g}{E_n - \varepsilon_j} \langle j| \hat{U} |n\rangle
\end{aligned}
$$

and so on.

This is a complex power series in g since the full energy E_n remains in the denominator. If we let $|m\rangle = |n\rangle$ in (8.138) we get

$$E_n \langle n \mid N \rangle = \varepsilon_n \langle n \mid N \rangle + g \langle n | \hat{U} | N \rangle \tag{8.141}$$

$$E_n = \varepsilon_n + g \langle n | \hat{U} | N \rangle \tag{8.142}$$

This can be expanded to give E_n as a series in powers of g, i.e., substituting the $0^{th} - order$ approximation for $|N\rangle$ gives the $1^{st} - order$ approximation for E_n and substituting the $1^{st} - order$ approximation for $|N\rangle$ gives the $2^{nd} - order$ approximation for E_n and so on.

If we substitute the $1^{st} - order$ approximation for $|N\rangle$ we get

$$\begin{aligned} E_n &= \varepsilon_n + g \langle n | \hat{U} | N \rangle \\ &= \varepsilon_n + g \langle n | \hat{U} \left(|n\rangle + \sum_{m \neq n} |m\rangle \frac{g}{E_n - \varepsilon_m} \langle m | \hat{U} | n \rangle \right) \\ &= \varepsilon_n + g \langle n | \hat{U} | n \rangle + g^2 \sum_{m \neq n} \frac{\left| \langle m | \hat{U} | n \rangle \right|^2}{E_n - \varepsilon_m} \end{aligned} \tag{8.143}$$

which is the second-order energy.

So the BWPT and the RSPT agree at each order of perturbation theory, as they must. The structure of the equations, however, is very different.

A simple example shows the very different properties of the two methods.

Consider the Hamiltonian given by $\hat{H} = \hat{H}_0 + \hat{V}$ where

$$\hat{H}_0 = \begin{pmatrix} \varepsilon_1 & 0 \\ 0 & \varepsilon_2 \end{pmatrix} \rightarrow \text{ eigenvectors } |1\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \text{ and } |2\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{8.144}$$

and

$$\hat{V} = \begin{pmatrix} 0 & \alpha \\ \alpha^* & 0 \end{pmatrix} \tag{8.145}$$

The exact energy eigenvalues are obtained by diagonalizing the $\hat{H}$ matrix

$$\hat{H} = \begin{pmatrix} \varepsilon_1 & \alpha \\ \alpha^* & \varepsilon_2 \end{pmatrix} \tag{8.146}$$

to get the characteristic equation

$$\det \begin{bmatrix} \varepsilon_1 - E & \alpha \\ \alpha^* & \varepsilon_2 - E \end{bmatrix} = 0 = E^2 - (\varepsilon_1 + \varepsilon_2)E + (\varepsilon_1\varepsilon_2 - |\alpha|^2) \tag{8.147}$$

which has solutions

$$E_1 = \frac{1}{2}(\varepsilon_1 + \varepsilon_2) + \frac{1}{2}\sqrt{(\varepsilon_1 - \varepsilon_2)^2 + 4\,|\alpha|^2} \tag{8.148}$$

$$E_2 = \frac{1}{2}(\varepsilon_1 + \varepsilon_2) - \frac{1}{2}\sqrt{(\varepsilon_1 - \varepsilon_2)^2 + 4\,|\alpha|^2} \tag{8.149}$$

In the *degenerate limit*, $\varepsilon_1 = \varepsilon_2 = \varepsilon$, we have the *exact* solutions

$$E_1 = \varepsilon + |\alpha| \text{ and } E_2 = \varepsilon - |\alpha| \tag{8.150}$$

Now, BWPT gives

$$E_n = \varepsilon_n + \langle n| \hat{V} |N\rangle$$

$$= \varepsilon_n + \langle n| \hat{V} |n\rangle + \sum_{m \neq n} \frac{\left|\langle m| \hat{V} |n\rangle\right|^2}{E_n - \varepsilon_m} \tag{8.151}$$

or

$$E_1 = \varepsilon_n + \langle 1| \hat{V} |1\rangle + \frac{\left|\langle 2| \hat{V} |1\rangle\right|^2}{E_1 - \varepsilon_2} = \varepsilon_1 + \frac{|\alpha|^2}{E_1 - \varepsilon_2} \tag{8.152}$$

Rearranging we get

$$E_1^2 - (\varepsilon_1 + \varepsilon_2)E_1 + (\varepsilon_1\varepsilon_2 - |\alpha|^2) = 0 \tag{8.153}$$

which is the same eigenvalue equation as the exact solution. In the degenerate limit, $\varepsilon_1 = \varepsilon_2 = \varepsilon$, we have

$$E_1 = \varepsilon + |\alpha| \tag{8.154}$$

or BWPT gives the exact answer for this simple system, *even in the degenerate case.*

On the other hand, RSPT gives to second-order

$$E_1 = \varepsilon_1 + \langle 1| \hat{V} |1\rangle + \frac{|\alpha|^2}{\varepsilon_1 - \varepsilon_2} \tag{8.155}$$

This is equivalent to the exact formula to second order only!

In the degenerate limit we get nonsense since the denominator vanishes. As we saw earlier, RSPT requires an *entirely different procedure in the degenerate case.*

Notice that RSPT is in trouble even if $\varepsilon_1 \approx \varepsilon_2$ which implies that

$$\frac{|\alpha|^2}{\varepsilon_1 - \varepsilon_2} \text{ is very large} \tag{8.156}$$

and thus, that the perturbation expansion makes no sense(the terms are supposed to get smaller!). A clever trick for handling these *almost degenerate* cases using RSPT goes as follows.

Almost Degenerate Perturbation Theory

Given $\hat{H} = \hat{H}_0 + \hat{V}$, suppose, as in the last example, we have two states $|n\rangle$ and $|m\rangle$ of the unperturbed (zero order) Hamiltonian $\hat{H}_0$ that have energies that are approximately equal.

This is a troublesome situation for RSPT because it is an expansion that includes increasing numbers of

$$\frac{1}{\varepsilon_n - \varepsilon_m} \tag{8.157}$$

terms. This implies that successive terms in the perturbation series might decrease slowly or not at all.

To develop a more rapidly converging perturbation expansion we rearrange the calculation as follows. We use the definition of the identity operator in terms of projection operators to write

$$\hat{V} = \hat{I}\hat{V}\hat{I} = \sum_{i,j} |i\rangle \langle i| \hat{V} |j\rangle \langle j| \tag{8.158}$$

We then break up $\hat{V}$ into two parts

$$\hat{V} = \hat{V}_1 + \hat{V}_2 \tag{8.159}$$

where we separate out the m and n terms into $\hat{v}_1$

$$\begin{aligned}\hat{V}_1 = &|m\rangle\langle m|\hat{V}|m\rangle\langle m| + |m\rangle\langle m|\hat{V}|n\rangle\langle n| \\ &+ |n\rangle\langle n|\hat{V}|m\rangle\langle m| + |n\rangle\langle n|\hat{V}|n\rangle\langle n|\end{aligned} \tag{8.160}$$

and $\hat{v}_2 =$ the rest of the terms. We then write

$$\hat{H} = \hat{H}_0 + \hat{V} = \hat{H}_0 + \hat{V}_1 + \hat{V}_2 = \hat{H}_0' + \hat{V}_2 \tag{8.161}$$

This new procedure then finds exact eigenvectors/eigenvalues of $\hat{H}_0'$ and treats $\hat{v}_2$ by ordinary perturbation theory.

Since the basis is orthonormal, we have from the definition of $\hat{v}_2$, i.e.,

$$\hat{V}_2 = \sum_{i,j\neq m,n} |i\rangle\langle i|\hat{V}|j\rangle\langle j| \tag{8.162}$$

which gives

$$0 = \langle n| \hat{V}_2 |n\rangle = \langle n| \hat{V}_2 |m\rangle = \langle m| \hat{V}_2 |n\rangle = \langle m| \hat{V}_2 |m\rangle \quad (8.163)$$

Thus, the closeness of the levels ε_n and ε_m will not prevent us from applying standard perturbation theory to $\hat{v}_2$, i.e., the numerators of terms with very small energy denominators, which might cause the series to diverge, all vanish identically!

Now if $|i\rangle$ is an eigenvector of $\hat{H}_0$ (not $|m\rangle$ or $|n\rangle$), then it is also an eigenvector of $\hat{H}'_0$ since, by the orthonormality condition,

$$\hat{V}_1 |i\rangle = 0 \quad (8.164)$$

Neither $|m\rangle$ nor $|n\rangle$ is an eigenvector of $\hat{H}'_0$ however.

Now, the $\hat{H}_0$ matrix is diagonal since we are using its eigenvectors as a basis. The $\hat{H}'_0$ matrix is diagonal also except for the 2×2 submatrix

$$\begin{pmatrix} \langle m| \hat{H}_0' |m\rangle & \langle m| \hat{H}_0' |n\rangle \\ \langle n| \hat{H}_0' |m\rangle & \langle n| \hat{H}_0' |n\rangle \end{pmatrix} \tag{8.165}$$

Therefore, we can finish the solution of the problem of $\hat{H}_0'$ by diagonalizing this 2×2 matrix.

Diagonalizing the 2×2 matrix is equivalent to finding the linear combinations (or new zero order eigenvectors)

$$\alpha |n\rangle + \beta |m\rangle \tag{8.166}$$

that diagonalize the 2×2 matrix.

We must have

$$\hat{H}_0' (\alpha |n\rangle + \beta |m\rangle) = \left(\hat{H}_0 + \hat{V}_1\right) (\alpha |n\rangle + \beta |m\rangle) = E' (\alpha |n\rangle + \beta |m\rangle) \tag{8.167}$$

Now

$$
\begin{aligned}
\hat{H}_0' \left|n\right\rangle &= \hat{H}_0 \left|n\right\rangle + \hat{V}_1 \left|n\right\rangle \\
&= \varepsilon_n \left|n\right\rangle + \left|m\right\rangle \left\langle m\right| \hat{V} \left|m\right\rangle \left\langle m \mid n\right\rangle + \left|m\right\rangle \left\langle m\right| \hat{V} \left|n\right\rangle \left\langle n \mid n\right\rangle \\
&\qquad + \left|n\right\rangle \left\langle n\right| \hat{V} \left|m\right\rangle \left\langle m \mid n\right\rangle + \left|n\right\rangle \left\langle n\right| \hat{V} \left|n\right\rangle \left\langle n \mid n\right\rangle \\
&= \varepsilon_n \left|n\right\rangle + \left|m\right\rangle \left\langle m\right| \hat{V} \left|n\right\rangle + \left|n\right\rangle \left\langle n\right| \hat{V} \left|n\right\rangle \\
&= \left(\varepsilon_n + \left\langle n\right| \hat{V} \left|n\right\rangle\right) \left|n\right\rangle + \left|m\right\rangle \left\langle m\right| \hat{V} \left|n\right\rangle \\
&= E_n^{(1)} \left|n\right\rangle + \left\langle m\right| \hat{V} \left|n\right\rangle \left|m\right\rangle
\end{aligned}
\tag{8.168}
$$

and similarly

$$
\hat{H}_0' \left|m\right\rangle = E_m^{(1)} \left|m\right\rangle + \left\langle n\right| \hat{V} \left|m\right\rangle \left|n\right\rangle \tag{8.169}
$$

Therefore, we get

$$
\begin{aligned}
\alpha \left(E_n^{(1)} \left|n\right\rangle + \left\langle m\right| \hat{V} \left|n\right\rangle \left|m\right\rangle\right) + \beta \left(E_m^{(1)} \left|m\right\rangle + \left\langle n\right| \hat{V} \left|m\right\rangle \left|n\right\rangle\right) \\
= E' \left(\alpha \left|n\right\rangle + \beta \left|m\right\rangle\right)
\end{aligned}
\tag{8.170}
$$

Since the state vectors $\left|m\right\rangle$ and $\left|n\right\rangle$ are orthogonal, we must then have

$$E_n^{(1)}\alpha + \langle n|\,\hat{V}\,|m\rangle\,\beta = E'\alpha \tag{8.171}$$

$$\langle m|\,\hat{V}\,|n\rangle\,\alpha + E_m^{(1)}\beta = E'\beta \tag{8.172}$$

These equations have two solutions, namely,

$$\alpha = \langle n|\,\hat{V}\,|m\rangle \tag{8.173}$$

$$\beta_\pm = \frac{E_m^{(1)} - E_n^{(1)}}{2} \pm \sqrt{\left(\frac{E_m^{(1)} - E_n^{(1)}}{2}\right)^2 + \left|\langle n|\,\hat{V}\,|m\rangle\right|^2} \tag{8.174}$$

which then give the results

$$E_\pm = \frac{E_m^{(1)} + E_n^{(1)}}{2} \pm \sqrt{\left(\frac{E_m^{(1)} - E_n^{(1)}}{2}\right)^2 + \left|\langle n|\,\hat{V}\,|m\rangle\right|^2} \tag{8.175}$$

We then know all the eigenvectors/eigenvalues of $\hat{H}_0'$ (we know all of the unperturbed states) and we can deal with $\hat{V}_2$ by perturbation theory.

Finally, let us introduce another interesting idea.

Fake Degenerate Perturbation Theory

Consider the problem of finding the energy eigenvalues and state vectors for a system with a Hamiltonian $\hat{H} = \hat{H}_0 + \hat{V}$ where we know the solution to the zero-order system

$$\hat{H}_0 |n\rangle = \varepsilon_n |n\rangle \tag{8.176}$$

We will assume that the unperturbed states are nondegenerate.

Now define

$$E_{average} = E_{av} = \frac{1}{n} \sum_n \varepsilon_n \tag{8.177}$$

and redefine

$$\hat{H} = E_{av} \hat{I} + \hat{U} \tag{8.178}$$

where

$$\hat{U} = \hat{H}_0 - E_{av} \hat{I} + \hat{V} \tag{8.179}$$

If the energies associated with $\hat{U}$ are small corrections to E_{av}, then we can use degenerate perturbation theory to solve this problem, i.e., the new unperturbed Hamiltonian is

$$\hat{H}_0' = E_{av}\hat{I} \tag{8.180}$$

and all of its levels are degenerate in zero order.

The problem is then solved by diagonalizing the $\hat{U}$ matrix in the basis of $\hat{H}_0$ states.

8.2.2 **Thoughts on Degeneracy and Position Representation**

When we derived the energy spectrum of the hydrogen atom we found that the states were labeled by three quantum numbers

$$|\psi\rangle = |n\ell m\rangle \tag{8.181}$$

where

$n =$ the radial quantum number

$\ell =$ orbital angular momentum quantum number

$m = z -$ component of orbital angular momentum quantum number

and we found that

$$n = 1, 2, 3, ...$$
$$\ell = 0, 1, 2, ..., n-1 \text{ for a given value of } n$$
$$m = -\ell, -\ell+1, ...\ell-1, \ell \text{ for a given value of } \ell$$

The energy eigenvalues, however, did not depend on ℓ or m. We found that

$$E_{n\ell m} = E_n = -\frac{e^2}{2a_0 n^2} \tag{8.182}$$

Therefore, each energy level had a degeneracy given by

$$g = \sum_{\ell=0}^{n-1} \sum_{m=-\ell}^{\ell} 1 = \sum_{\ell=0}^{n-1} (2\ell + 1) = 2 \sum_{\ell=0}^{n-1} \ell + \sum_{\ell=0}^{n-1} 1$$
$$= 2\frac{n(n-1)}{2} + n = n^2 \tag{8.183}$$

The degeneracy with respect to m is understandable since no direction is explicitly preferred in the Hamiltonian. We expect that this degeneracy will disappear as soon as a preferred direction is added to the Hamiltonian, as in the case of external electric(Stark effect) or magnetic(Zeeman effect) fields.

The degeneracy with respect to ℓ is a property peculiar to the pure $1/r$ Coulomb potential. Since no other atom except hydrogen has a pure Coulomb potential, we expect this degeneracy to vanish in other atoms.

Such a degeneracy is called an *accidental degeneracy.*

Now the electron and proton making up the hydrogen atom also have spin angular momentum. The presence of these extra(internal) degrees of freedom should change the Hamiltonian.

The Schrodinger equation was derived from the eigenvalue equation for the Hamiltonian

$$\hat{H}\,|\psi\rangle = E\,|\psi\rangle \tag{8.184}$$

by re-expressing that equation in the position representation. The associated Schrodinger wave functions were given by the scalar product(linear functional) relation

$$\psi(\vec{r}) = \langle \vec{r}\,|\,\psi\rangle \tag{8.185}$$

The single particle Schrodinger equation is relevant for problems where the Hamiltonian contains terms dependent on ordinary 3−dimensional space(for many-particle systems we must use a multi-dimensional configuration space which bears no simple relationship to ordinary three-dimensional space). Spin is an internal degree of freedom that has no representation in the 3−dimensional space of the Schrodinger wave equation.

The Schrodinger picture, however, does not choose a particular representation and, therefore, we can include spin within the context of solving the Schrodinger equation in the following ad hoc manner. A more rigorous treatment requires relativity.

If there are spin-dependent terms in the Hamiltonian, then we expand the Hilbert space used to solved the problem by constructing a new basis that is made up of direct product states of the following type

$$|\psi_{new}\rangle = |\psi\rangle \otimes |s, m_s\rangle \tag{8.186}$$

where $|\psi\rangle$ depends on only ordinary 3−dimensional space and $|s, m_s\rangle$ is an eigenvector of $\vec{S}^2_{op}$ and $\hat{S}_z$.

The energy eigenvalue equation becomes

$$\widehat{H}\,|\psi_{new}\rangle$$
$$= (((\text{3-space operators}))\,|\psi\rangle) \otimes (((\text{spin-dependent operators}))\,|s, m_s\rangle)$$

and the corresponding *wave function* is

$$\langle \vec{r}\,|\,\psi_{new}\rangle = \langle \vec{r}\,|\,\psi\rangle\,|s, m_s\rangle = \psi(\vec{r})\,|s, m_s\rangle \tag{8.187}$$

where abstract spin vector is stuck onto the wave function in some way (maybe with superglue).

Let us now investigate what happens in atoms when we add in spin, some aspects of relativity and external fields. We restrict our attention to one-electron atoms like hydrogen at this point.

8.3 **Spin-Orbit Interaction - Fine Structure**

The proton in hydrogen generates an electric field

$$\vec{\mathcal{E}} = \frac{e}{r^2}\hat{r} = \frac{e}{r^3}\vec{r} \tag{8.188}$$

that acts on the moving electron. This result is approximately true (to first order) in most atoms. Now special relativity says that an electron moving with a velocity $\vec{v}$ through an electric field $\vec{\mathcal{E}}$ also behaves as if it is interacting with a magnetic field given by

$$\vec{B} = -\frac{1}{c}\vec{v} \times \vec{\mathcal{E}} \tag{8.189}$$

to first order in v/c.

This magnetic field interacts with the spin (actually with its associated magnetic moment) to produce an additional contribution to the energy of the form

$$E = -\vec{M}_{spin} \cdot \vec{B} \tag{8.190}$$

where

$$\vec{M}_{spin} = -\frac{e}{mc}\vec{S} \tag{8.191}$$

Substituting everything in we get

$$\begin{aligned} E &= -\frac{e}{mc^2}\vec{S}\cdot\left(\vec{v}\times\vec{\mathcal{E}}\right) = -\frac{e}{mc^2}\vec{S}\cdot\left(\vec{v}\times\frac{e}{r^3}\vec{r}\right) \\ &= \frac{1}{m^2c^2}\vec{S}\cdot\vec{L}\frac{e^2}{r^3} \end{aligned} \tag{8.192}$$

Now

$$\frac{e^2}{r^3} = \frac{1}{r}\frac{dV}{dr} \text{ for } V(r) = -\frac{e^2}{r} = \text{ potential energy of the electron} \tag{8.193}$$

so that we finally obtain the so-called spin-orbit energy contribution

$$E = \left[\frac{1}{m^2c^2}\frac{1}{r}\frac{dV}{dr}\right]\vec{S}\cdot\vec{L} = E_{spin-orbit} = E_{so} \tag{8.194}$$

This corresponds to an additional term in the Hamiltonian of the form

$$\hat{H}_{so} = \left[\frac{1}{m^2c^2}\frac{1}{r}\frac{dV}{dr}\right]\vec{S}_{op}\cdot\vec{L}_{op} \tag{8.195}$$

This term couples the orbital and spin angular momentum degrees of freedom (hence the label spin-orbit energy) and mixes 3−dimensional space with *spin* space. That is why we had to expand the Hilbert space as we discussed earlier.

Another way to think about this interaction is that the electron spin magnetic moment vector (or spin vector) is precessing about the direction of the magnetic field. The equations for such a precessional motion are

$$\vec{M}_{spin}\times\vec{B} = \frac{d\vec{S}}{dt} = \vec{\Omega}_{L(armor)}\times\vec{S} \tag{8.196}$$

where

$$\left|\vec{\Omega}_L\right| = \frac{eB}{m} \tag{8.197}$$

Now

$$\vec{B} = -\frac{1}{c}\vec{v} \times \vec{\mathcal{E}} = \frac{1}{emc^2}\frac{1}{r}\frac{dV}{dr}\vec{L} \tag{8.198}$$

which implies that

$$\left|\vec{\Omega}_L\right| = \frac{1}{m^2c^2}\frac{1}{r}\frac{dV}{dr}\left|\vec{L}\right| \tag{8.199}$$

It turns out that this is exactly a factor of 2 too large. There is another relativistic effect, which gives another precession (called Thomas precession) effect, that cancels exactly one-half of this spin-orbit effect.

8.3.1 **Thomas Precession**

This is a relativistic kinematic effect. It results from the time dilation between the rest frames of the electron and the proton. This causes observers in these two frames to disagree on the time required for one of the particles to a make a complete revolution about the other particle.

If an observer on the electron measures a time interval T, then the observer on the proton measures

$$T' = \gamma T \text{ where } \gamma = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} \quad , \quad v = \text{ speed of the electron} \tag{8.200}$$

We assume uniform circular motion for simplicity.

The orbital angular velocities measured by the observers are

$$\frac{2\pi}{T} \text{ and } \frac{2\pi}{T'} \tag{8.201}$$

respectively.

In the rest frame of the electron, the spin angular momentum vector maintains its direction in space. This implies that an observer on the proton sees this spin vector precessing at a rate equal to the *difference* of the two angular velocities, i.e., the precessional frequency is

$$\begin{aligned}\Omega_{Thomas} &= \frac{2\pi}{T} - \frac{2\pi}{T'} = \frac{2\pi}{T'}\left(\frac{T'}{T} - 1\right) \\ &= \frac{2\pi}{T'}\left(\frac{1}{\sqrt{1-\frac{v^2}{c^2}}} - 1\right) \\ &\approx \frac{2\pi}{T'}\left(\frac{v^2}{2c^2}\right) \end{aligned} \tag{8.202}$$

But we also have

$$\frac{2\pi}{T'} = \omega = \frac{\left|\vec{L}\right|}{mr^2} \text{ and } \frac{mv^2}{r} = -\frac{dV}{dr} \tag{8.203}$$

for circular motion.

Thus, we get

$$\left|\vec{\Omega}_T\right| = -\frac{1}{2}\frac{1}{m^2c^2}\frac{1}{r}\frac{dV}{dr}\left|\vec{L}\right| = -\frac{1}{2}\left|\vec{\Omega}_L\right| \tag{8.204}$$

Therefore, the combined precession is reduced by a factor of two and we get the result

$$\hat{H}_{so} = \left[\frac{1}{2m^2c^2}\frac{1}{r}\frac{dV}{dr}\right]\vec{S}_{op}\cdot\vec{L}_{op} \tag{8.205}$$

The energy levels arising from this correction are called the *atomic fine structure.*

8.4 **Another Relativity Correction**

The correct relativistic kinetic energy term is

$$\begin{aligned} K = (\gamma-1)mc^2 &= \left(\frac{1}{\sqrt{1-\frac{v^2}{c^2}}}-1\right)mc^2 \\ &= \left(\left(1+\frac{1}{2}\frac{v^2}{c^2}-\frac{1}{8}\frac{v^4}{c^4}+...\right)-1\right)mc^2 \\ &= \left(\frac{1}{2}\frac{v^2}{c^2}-\frac{1}{8}\frac{v^4}{c^4}\right)mc^2 = \frac{\vec{p}_{op}^{\,2}}{2m}-\frac{\vec{p}_{op}^{\,4}}{8m^3c^2} \end{aligned} \tag{8.206}$$

Therefore, we must correct the $\vec{p}_{op}^2/2m$ we have already included in the Hamiltonian by adding a terms of the form

$$\hat{H}_{relativity} = -\frac{\vec{p}_{op}^4}{8m^3c^2} \tag{8.207}$$

Thus, if no external field are present we have the Hamiltonian

$$\hat{H} = \hat{H}_0 + \hat{H}_{relativity} + \hat{H}_{so} \tag{8.208}$$

where

$$\hat{H}_0 = \frac{\vec{p}_{op}^2}{2m} - e^2 \left(\frac{1}{r}\right)_{op} \tag{8.209}$$

$$\hat{H}_{relativity} = -\frac{\vec{p}_{op}^4}{8m^3c^2} \tag{8.210}$$

$$\hat{H}_{so} = \left[\frac{1}{2m^2c^2}\frac{1}{r}\frac{dV}{dr}\right]\vec{S}_{op}\cdot\vec{L}_{op} \tag{8.211}$$

8.5 **External Fields - Zeeman and Stark Effects; Hyperfine Structure**

8.5.1 **Zeeman Effect**

If an external magnetic field exists, then it interacts with the total magnetic moment of the electron, where

$$\vec{M}_{total} = \vec{M}_{orbital} + \vec{M}_{spin} = -\frac{e}{2mc}\left(g_\ell \vec{L} + g_s \vec{S}\right) \tag{8.212}$$

as we derived earlier. If we define

$$\mu_B = \text{ Bohr magneton } = \frac{e\hbar}{mc} \tag{8.213}$$

and let $\vec{B}_{ext} = B\hat{z}$, then we have, using $g_\ell = 1$ and $g_s = 2$, the result

$$E_{Zeeman} = -\vec{M}_{total} \cdot \vec{B}_{ext} = \frac{\mu_B B}{\hbar}(L_z + 2S_z) \tag{8.214}$$

Thus, we must add a term of the form

$$\hat{H}_{Zeeman} = \frac{\mu_B B}{\hbar}(\hat{L}_z + 2\hat{S}_z) \tag{8.215}$$

to the Hamiltonian when an external magnetic field is present.

We can see directly how the orbital angular momentum part of this energy arises. We saw earlier that if we had a Hamiltonian

$$\hat{H}_0 = \frac{\vec{p}_{op}^{\,2}}{2m} + V(\vec{r}_{op}) \tag{8.216}$$

and we add an electromagnetic field characterized by a vector potential $\vec{A}$, where $\vec{B} = \nabla \times \vec{A}$, then the momentum operator changes to

$$\vec{p}_{em} = \vec{p} - \frac{e}{c}\vec{A}(\vec{r}) \tag{8.217}$$

and the Hamiltonian changes to

$$\hat{H} = \frac{\vec{p}^2_{em,op}}{2m} + V(\vec{r}_{op}) = \frac{\left(\vec{p}_{op} - \frac{e}{c}\vec{A}(\vec{r}_{op})\right)^2}{2m} + V(\vec{r}_{op})$$
$$= \hat{H}_0 - \frac{e}{2mc}\left(\vec{p}_{op}\cdot\vec{A}(\vec{r}_{op}) + \vec{A}(\vec{r}_{op})\cdot\vec{p}_{op}\right) + \frac{e^2}{2mc^2}\vec{A}^2(\vec{r}_{op}) \tag{8.218}$$

The magnetic field has to be enormous or the radial quantum number n very large for the $\vec{A}^2$ term to have any effect, so we will neglect it for now. Let us look at the term

$$\vec{p}_{op}\cdot\vec{A}(\vec{r}_{op}) + \vec{A}(\vec{r}_{op})\cdot\vec{p}_{op} \tag{8.219}$$

For a uniform (constant in magnitude and direction) external field $\vec{B}$, we have

$$\vec{A} = -\frac{1}{2}\vec{r}\times\vec{B} \tag{8.220}$$

I will prove this so we get a chance to see the use of ϵ_{ijk} in vector algebra.

$$
\begin{aligned}
\nabla \times \vec{A} &= -\frac{1}{2} \nabla \times \left(\vec{r} \times \vec{B} \right) = -\frac{1}{2} \sum_{ijk} \varepsilon_{ijk} \frac{\partial}{\partial x_j} \left(\vec{r} \times \vec{B} \right)_k \hat{e}_i \\
&= -\frac{1}{2} \sum_{ijk} \varepsilon_{ijk} \frac{\partial}{\partial x_j} \left(\sum_{mn} \varepsilon_{kmn} x_m B_n \right) \hat{e}_i \\
&= -\frac{1}{2} \sum_{ij} \sum_{mn} \left(\sum_{k} \varepsilon_{ijk} \varepsilon_{mnk} \right) \frac{\partial}{\partial x_j} \left(x_m B_n \right) \hat{e}_i \\
&= -\frac{1}{2} \sum_{ij} \sum_{mn} \left(\delta_{im} \delta_{jn} - \delta_{in} \delta_{jm} \right) \frac{\partial}{\partial x_j} \left(x_m B_n \right) \hat{e}_i \\
&= -\frac{1}{2} \sum_{ij} \left[\frac{\partial}{\partial x_j} \left(x_i B_j \right) \hat{e}_i - \frac{\partial}{\partial x_j} \left(x_j B_i \right) \hat{e}_i \right] \\
&= -\frac{1}{2} \sum_{ij} \left[\frac{\partial x_i}{\partial x_j} B_j + x_i \frac{\partial B_j}{\partial x_j} - \frac{\partial x_j}{\partial x_j} B_i - x_j \frac{\partial B_i}{\partial x_j} \right] \hat{e}_i
\end{aligned}
$$

Now

$$\frac{\partial x_i}{\partial x_j} = \delta_{ij} \text{ and } \frac{\partial B_i}{\partial x_j} = 0 \tag{8.221}$$

so we get

$$\begin{aligned}
\nabla \times \vec{A} &= -\frac{1}{2} \sum_{ij} [\delta_{ij} B_j - \delta_{jj} B_i] \hat{e}_i \\
&= -\frac{1}{2} \left[\sum_i B_i \hat{e}_i - 3 \sum_i B_i \hat{e}_i \right] \\
&= -\frac{1}{2} \left[\vec{B} - 3\vec{B} \right] = \vec{B}
\end{aligned}$$

Therefore, we have

$$\begin{aligned}
\vec{p}_{op} \cdot \vec{A}(\vec{r}_{op}) + \vec{A}(\vec{r}_{op}) \cdot \vec{p}_{op} &= -\frac{1}{2} \left[\vec{p}_{op} \cdot \left(\vec{r}_{op} \times \vec{B} \right) + \left(\vec{r}_{op} \times \vec{B} \right) \cdot \vec{p}_{op} \right] \\
= -\frac{1}{2} \left[\sum_{ijk} \varepsilon_{ijk} B_k \left[\hat{p}_i \hat{x}_j + \hat{x}_j \hat{p}_i \right] \right] &= -\frac{1}{2} \left[\sum_{ijk} \varepsilon_{ijk} B_k \left[2\hat{x}_j \hat{p}_i - i\hbar \delta_{ij} \right] \right] \\
= \sum_{ijk} \varepsilon_{kji} \hat{x}_j \hat{p}_i B_k = (\vec{r}_{op} \times \vec{p}_{op}) \cdot \vec{B} &= \vec{L}_{op} \cdot \vec{B}
\end{aligned} \tag{8.222}$$

which then gives

$$\hat{H} = \hat{H}_0 - \frac{e}{2mc}\vec{L}_{op} \cdot \vec{B} + \frac{e^2}{2mc^2}\vec{A}^2(\vec{r}_{op}) \tag{8.223}$$

which accounts for the orbital angular momentum part of the Zeeman energy.

The spin angular momentum part of the Zeeman energy cannot be derived from the non-relativistic Schrodinger equation. When one derives the Dirac *relativistic* equation for the electron, the $\vec{S}_{op} \cdot \vec{B}$ term appears naturally.

8.5.2 **Stark Effect**

If a hydrogen atom is placed in an external electric field $\vec{\mathcal{E}}$ which is constant is space and time (uniform and static), then an additional energy appears. It corresponds to an interaction between and electric dipole made up of the electron and proton separated by a distance and the external electric field. We introduce the electric dipole moment operator

$$\vec{d}_{op} = -e\vec{r}_{op} \tag{8.224}$$

where $\vec{r}$ is the position vector of the electron relative to the proton. We then write the extra energy term as

$$\hat{H}_{dipole} = -\vec{d}_{op} \cdot \vec{\mathcal{E}} \tag{8.225}$$

If we choose $\vec{\mathcal{E}} = \mathcal{E}\hat{z}$, then we have $\hat{H}_{dipole} = -ez\mathcal{E}$. The full Hamiltonian is then

$$\hat{H} = \hat{H}_0 + \hat{H}_{relativity} + \hat{H}_{so} + \hat{H}_{Zeeman} + \hat{H}_{dipole} \tag{8.226}$$

where

$$\hat{H}_0 = \frac{\vec{p}_{op}^{\,2}}{2m} - e^2 \left(\frac{1}{r}\right)_{op} \tag{8.227}$$

$$\hat{H}_{relativity} = -\frac{\vec{p}_{op}^{\,4}}{8m^3c^2} \tag{8.228}$$

$$\hat{H}_{so} = \left[\frac{1}{2m^2c^2}\frac{1}{r}\frac{dV}{dr}\right]\vec{S}_{op} \cdot \vec{L}_{op} \tag{8.229}$$

$$\hat{H}_{Zeeman} = \frac{\mu_B}{\hbar}(\vec{L}_{op} + 2\vec{S}_{op}) \cdot \vec{B} \tag{8.230}$$

$$\hat{H}_{dipole} = -e\vec{r}_{op} \cdot \vec{\mathcal{E}} \tag{8.231}$$

8.5.3 **Hyperfine Structure**

The nuclear magnetic dipole moment also generates a magnetic field. If we assume that it is a point dipole $\vec{M}_N$, then the magnetic field is given by

$$\vec{B}(\vec{r}) = \left(\frac{3\left(\vec{M}_N \cdot \vec{r}\right)\vec{r}}{r^5} - \frac{\vec{M}_N}{r^3} \right) + \frac{8\pi}{3}\vec{M}_N \delta(\vec{r}) \tag{8.232}$$

where the first two terms are the standard result of the magnetic field due to a loop of current as seen from very far away (approximates dipole as a point) and the last term is peculiar to a point dipole. The last term will give a contribution only for spherically symmetric states ($\ell = 0$). The extra energy is then

$$\hat{H}_{hyperfine} = -\vec{M}_e \cdot \vec{B}$$
$$= -\left(\frac{3\left(\vec{M}_N \cdot \vec{r}\right)\left(\vec{M}_e \cdot \vec{r}\right)}{r^5} - \frac{\vec{M}_N \cdot \vec{M}_e}{r^3} \right) - \frac{8\pi}{3}\vec{M}_N \cdot \vec{M}_e \delta(\vec{r}) \tag{8.233}$$

where

$$\vec{M}_N = g_N \frac{Ze}{2m_N c}\vec{S}_{N,op} \text{ and } \vec{M}_e = \frac{e}{mc}\vec{S}_{e,op} \tag{8.234}$$

This is clearly due to spin-spin interactions between the electron and the nucleus and gives rise to the so-called hyperfine level splitting.

8.6 **Examples**

Now that we have identified all of the relevant corrections to the Hamiltonian for atoms, let us illustrate the procedures for calculation of the new energy levels via perturbation theory. We look at the simplest atom first.

8.6.1 **Spin-Orbit, Relativity, Zeeman Effect in Hydrogen Atom**

The Hamiltonian is

$$\hat{H} = \hat{H}_0 + \hat{H}_{relativity} + \hat{H}_{so} + \hat{H}_{Zeeman} \tag{8.235}$$

where

$$\hat{H}_0 = \frac{\vec{p}^2_{op}}{2m} - e^2 \left(\frac{1}{r}\right)_{op} \tag{8.236}$$

$$\hat{H}_{relativity} = -\frac{\vec{p}^4_{op}}{8m^3c^2} \tag{8.237}$$

$$\hat{H}_{so} = \left[\frac{1}{2m^2c^2}\frac{1}{r}\frac{dV}{dr}\right]\vec{S}_{op}\cdot\vec{L}_{op} \tag{8.238}$$

$$\hat{H}_{Zeeman} = \frac{\mu_B}{\hbar}(\vec{L}_{op} + 2\vec{S}_{op})\cdot\vec{B} \tag{8.239}$$

The first step is to calculate all relevant commutators so that we can find those operators that have a common eigenbasis.

$$\left[\vec{J}^2_{op}, \hat{H}_0\right] = 0 = \left[\hat{J}_z, \hat{H}_0\right] = \left[\vec{L}^2_{op}, \hat{H}_0\right] = \left[\vec{S}^2_{op}, \hat{H}_0\right] \tag{8.240}$$

$$\left[\vec{L}^2_{op}, \vec{S}^2_{op}\right] = 0 = \left[\vec{L}^2_{op}, \vec{J}^2_{op}\right] = \left[\vec{S}^2_{op}, \vec{J}^2_{op}\right] \tag{8.241}$$

$$\left[\vec{J}^2_{op}, \hat{J}_z\right] = 0 = \left[\vec{L}^2_{op}, \hat{J}_z\right] = \left[\vec{S}^2_{op}, \hat{J}_z\right] \tag{8.242}$$

This says that there exists a common set of eigenvectors in the unperturbed system for the set of operators

$$\hat{H}_0, \vec{J}^2_{op}, \vec{L}^2_{op}, \vec{S}^2_{op}, \hat{J}_z \tag{8.243}$$

We label these states by the corresponding eigenvalues of the commuting set of observables (these are called *good quantum numbers*)

$$|n, \ell, s, j, m_j\rangle \tag{8.244}$$

We also have

$$\left[\hat{S}_z, \hat{H}_0\right] = 0 = \left[\hat{L}_z, \hat{H}_0\right] = \left[\vec{L}^2_{op}, \hat{H}_0\right] = \left[\vec{S}^2_{op}, \hat{H}_0\right] \tag{8.245}$$

$$\left[\vec{L}^2_{op}, \vec{S}^2_{op}\right] = 0 = \left[\vec{L}^2_{op}, \hat{S}_z\right] = \left[\vec{S}^2_{op}, \hat{S}_z\right] = \left[\vec{L}^2_{op}, \hat{L}_z\right] = \left[\vec{S}^2_{op}, \hat{L}_z\right] \tag{8.246}$$

which says that there exists another common set of eigenvectors in the unperturbed system for the operators

$$\hat{H}_0, \vec{L}^2_{op}, \vec{S}^2_{op}, \hat{L}_z, \hat{S}_z \tag{8.247}$$

We label these states by the corresponding eigenvalues of this commuting set of observables (again these are called *good quantum numbers*)

$$|n, \ell, s, m_\ell, m_s\rangle \tag{8.248}$$

In this latter basis, the unperturbed or zero-order Hamiltonian has solutions represented by

$$\hat{H}_0 |n, \ell, m_\ell, s, m_s\rangle = E_n^{(0)} |n, \ell, m_\ell, s, m_s\rangle \quad , \quad E_n^{(0)} = -\frac{Ze^2}{2a_0 n^2}$$

$Ze =$ nucleus charge ($Z = 1$ for hydrogen)

$$a_0 = \text{ Bohr radius } = \frac{\hbar^2}{me^2}$$

$$\psi_{n\ell m_\ell s m_s}(r, \theta, \phi) = \langle \vec{r}| \left(|nlm_\ell s m_s\rangle\right) = \langle \vec{r}| \left(|nlm_\ell\rangle |sm_s\rangle\right)$$
$$= \langle \vec{r} \,|\, nlm_\ell\rangle |sm_s\rangle = \psi_{n\ell m_\ell}(r, \theta, \phi) |sm_s\rangle$$

and first few unperturbed wave functions are

$$\psi_{100}(r,\theta,\phi) = \frac{1}{\sqrt{\pi}}\left(\frac{Z}{a_0}\right)^{3/2} e^{-\frac{Zr}{a_0}} \tag{8.249}$$

$$\psi_{200}(r,\theta,\phi) = \frac{1}{\sqrt{32\pi}}\left(\frac{Z}{a_0}\right)^{3/2}\left(2-\frac{Zr}{a_0}\right) e^{-\frac{Zr}{2a_0}} \tag{8.250}$$

$$\psi_{210}(r,\theta,\phi) = \frac{1}{\sqrt{32\pi}}\left(\frac{Z}{a_0}\right)^{3/2}\frac{Zr}{a_0} e^{-\frac{Zr}{2a_0}}\cos\theta \tag{8.251}$$

$$\psi_{21\pm1}(r,\theta,\phi) = \frac{1}{\sqrt{64\pi}}\left(\frac{Z}{a_0}\right)^{3/2}\frac{Zr}{a_0} e^{-\frac{Zr}{2a_0}}\sin\theta e^{\pm i\phi} \tag{8.252}$$

We also have the relations below for the unperturbed states.

$$\vec{L}^2_{op}\,|n,\ell,s,j,m_j\rangle = \hbar^2\ell(\ell+1)\,|n,\ell,s,j,m_j\rangle \tag{8.253}$$

$$\vec{S}^2_{op}\,|n,\ell,s,j,m_j\rangle = \hbar^2 s(s+1)\,|n,\ell,s,j,m_j\rangle \tag{8.254}$$

$$\vec{J}^2_{op}\,|n,\ell,s,j,m_j\rangle = \hbar j(j+1)\,|n,\ell,s,j,m_j\rangle \tag{8.255}$$

$$\hat{J}_z\,|n,\ell,s,j,m_j\rangle = \hbar^2 m_j\,|n,\ell,s,j,m_j\rangle \tag{8.256}$$

$$\vec{L}^2_{op}\,|n,\ell,s,m_\ell,m_s\rangle = \hbar^2\ell(\ell+1)\,|n,\ell,s,m_\ell,m_s\rangle \tag{8.257}$$

$$\vec{S}^2_{op}\,|n,\ell,s,m_\ell,m_s\rangle = \hbar^2 s(s+1)\,|n,\ell,s,m_\ell,m_s\rangle \tag{8.258}$$

$$\hat{L}_z\,|n,\ell,s,m_\ell,m_s\rangle = \hbar m_\ell\,|n,\ell,s,m_\ell,m_s\rangle \tag{8.259}$$

$$\hat{S}_z\,|n,\ell,s,m_\ell,m_s\rangle = \hbar m_s\,|n,\ell,s,m_\ell,m_s\rangle \tag{8.260}$$

$$\hat{J}_z\,|n,\ell,s,m_\ell,m_s\rangle = (\hat{L}_z+\hat{S}_z)\,|n,\ell,s,m_\ell,m_s\rangle \tag{8.261}$$

$$= \hbar(m_\ell+m_s)\,|n,\ell,s,m_\ell,m_s\rangle \tag{8.262}$$

$$= \hbar m_j\,|n,\ell,s,m_\ell,m_s\rangle \tag{8.263}$$

Since the total angular momentum is given by

$$\vec{J}_{op} = \vec{L}_{op} + \vec{S}_{op} \tag{8.264}$$

the rules we developed for the addition of angular momentum say that

$$j = \ell + s, \ell + s - 1, ..., |\ell - s| + 1, |\ell - s| \tag{8.265}$$

and

$$m_j = j, j - 1, j - 2, ..., -j + 1, -j \tag{8.266}$$

In the case of hydrogen, where $s = 1/2$, we have only two allowed total j values for each ℓ value, namely,

$$j = \ell \pm \frac{1}{2} \tag{8.267}$$

We can use either of the two sets of basis states (both are an orthonormal basis)

$$|n, \ell, s, j, m_j\rangle \text{ or } |nlm_\ell sm_s\rangle \tag{8.268}$$

as the zero-order states for a perturbation theory development of the energies. The choice depends on the specific perturbations we are trying to calculate.

Let us start off by using the $|n, \ell, s, j, m_j\rangle$ states.

If we use the potential energy function

$$V(r) = -\frac{e^2}{r} \tag{8.269}$$

for hydrogen, then the spin-orbit correction to the Hamiltonian becomes

$$\hat{H}_{so} = \left[\frac{1}{2m^2c^2}\frac{1}{r}\frac{dV}{dr}\right]\vec{S}_{op}\cdot\vec{L}_{op} = \frac{e^2}{2m^2c^2}\left(\frac{1}{r^3}\right)\vec{S}_{op}\cdot\vec{L}_{op} \tag{8.270}$$

Now $\vec{J}_{op} = \vec{L}_{op} + \vec{S}_{op}$ implies that

$$\vec{J}_{op}^2 = \vec{L}_{op}^2 + \vec{S}_{op}^2 + 2\vec{L}_{op} \cdot \vec{S}_{op} \tag{8.271}$$

$$\rightarrow \vec{L}_{op} \cdot \vec{S}_{op} = \frac{1}{2}\left(\vec{J}_{op}^2 - \vec{L}_{op}^2 - \vec{S}_{op}^2\right) \tag{8.272}$$

and therefore

$$\hat{H}_{so} = \frac{e^2}{4m^2c^2}\left(\frac{1}{r^3}\right)\left(\vec{J}_{op}^2 - \vec{L}_{op}^2 - \vec{S}_{op}^2\right) \tag{8.273}$$

Therefore,

$$\left[\vec{J}_{op}^2, \hat{H}_{so}\right] = 0 = \left[\hat{J}_z, \hat{H}_{so}\right] = \left[\vec{L}_{op}^2, \hat{H}_{so}\right] = \left[\vec{S}_{op}^2, \hat{H}_{so}\right] = \left[\hat{H}_0, \hat{H}_{so}\right] \tag{8.274}$$

which implies that the state vectors $|n, \ell, s, j, m_j\rangle$ are also eigenvectors of $\hat{H}_{so}$. This means that the matrix representation of $\hat{H}_{so}$ in this basis will be diagonal and we can apply standard non-degenerate perturbation theory.

Applying our rules for first order perturbation theory we have

$$E_{n\ell sjm_j} = E_n^{(0)} + \langle n\ell sjm_j | \hat{H}_{so} | n\ell sjm_j \rangle$$
$$= E_n^{(0)} + \frac{e^2\hbar^2}{4m^2c^2} \left(j(j+1) - \ell(\ell+1) - s(s+1) \right) \langle n\ell sjm_j | \frac{1}{r_{op}^3} | n\ell sjm_j \rangle \tag{8.275}$$

We now evaluate

$$\langle n\ell sjm_j | \frac{1}{r_{op}^3} | n\ell sjm_j \rangle$$
$$= \int d^3\vec{r}' \int d^3\vec{r} \langle n\ell sjm_j \mid \vec{r}' \rangle \langle \vec{r}' | \frac{1}{r_{op}^3} | \vec{r} \rangle \langle \vec{r} \mid n\ell sjm_j \rangle \tag{8.276}$$

Now

$$\langle \vec{r}' | \frac{1}{r_{op}^3} | \vec{r} \rangle = \frac{1}{r^3} \langle \vec{r}' \mid \vec{r} \rangle = \frac{1}{r^3} \delta(\vec{r}' - \vec{r}) \tag{8.277}$$

which gives

$$\langle n\ell sjm_j| \frac{1}{r_{op}^3} |n\ell sjm_j\rangle = \int d^3\vec{r} \frac{1}{r^3} |\langle \vec{r} \mid n\ell sjm_j\rangle|^2$$
$$= \int d^3\vec{r} \frac{1}{r^3} \left|\psi_{n\ell sjm_j}(\vec{r})\right|^2$$

Therefore, we can calculate the energy corrections once we know $|n, \ell, s, j, m_j\rangle$.

We first consider the trivial case of the $n = 1$ level in hydrogen. We have

$$E_1^{(0)} = -\frac{e^2}{2a_0} \tag{8.278}$$

and the corresponding states are shown in Table 8.1 below.

n	ℓ	s	m_ℓ	m_s	j	m_j
1	0	1/2	0	+1/2	1/2	+1/2
1	0	1/2	0	-1/2	1/2	-1/2

Table: $n = 1$ level quantum numbers

or

$$\left|1,0,\frac{1}{2},\frac{1}{2},\frac{1}{2}\right\rangle_{jm} \quad \text{and} \quad \left|1,0,\frac{1}{2},\frac{1}{2},-\frac{1}{2}\right\rangle_{jm} \tag{8.279}$$

where we have added the label jm to distinguish them from the $|n,\ell,s,m_\ell,m_s\rangle$ states which we label with $m_\ell m_s = mm$. We are able to specify m_ℓ and m_s also in this case because when $\ell = 0$ we must have $m_\ell = 0$ and $j = s$ which says that $m_j = m_s$.

This is a *two-fold degenerate* ground state for the atom in *zeroth order*.

Since $\ell = 0$, which implies that $j = s = 1/2$, the expectation value $\langle \hat{H}_{so} \rangle = 0$. Thus, there is no spin orbit correction for this state to first order, In fact, there is no spin orbit correction to any order for an $\ell = 0$ state.

Now in general, we can write

$$|n,\ell,s,j,m_j\rangle = \sum_{\substack{m_\ell,m_s \\ m_\ell+m_s=m_j}} a_{n\ell sjm_jm_\ell m_s} \, |n,\ell,s,m_\ell,m_s\rangle \tag{8.280}$$

where the $a_{n\ell sjm_j m_\ell m_s}$ are the relevant Clebsch-Gordon(CG) coefficients.

For the $n = 1$ level we have the simple cases where

$$\left|1, 0, \frac{1}{2}, \frac{1}{2}, \pm\frac{1}{2}\right\rangle_{jm} = \left|1, 0, \frac{1}{2}, 0, \pm\frac{1}{2}\right\rangle_{mm} \tag{8.281}$$

i.e., the CG coefficients are equal to 1,

$$a_{1,0,\frac{1}{2},\frac{1}{2},\pm\frac{1}{2},0,\pm\frac{1}{2}} = 1 \tag{8.282}$$

which is always true for the (maximum j, maximum(minimum) m_j) states. There is always only one such state.

The next level is the $n = 2$ level of hydrogen and the complexity of the calculation increases fast. We have

$$E_2^{(0)} = -\frac{e^2}{8a_0} \tag{8.283}$$

It is always the case that the direct-product states $|n, \ell, s, m_\ell, m_s\rangle$ are easier to write down. For this level the $|n, \ell, s, j, m_j\rangle$ states need to be constructed from the $|n, \ell, s, m_\ell, m_s\rangle$ states. Before we proceed, we can enumerate the states in both schemes. The degeneracy is given by

$$\begin{aligned} \text{degeneracy} &= 2\sum_{\ell=0}^{n-1}(2\ell+1) = 2n^2 = 8 \\ &= \sum_{\ell=0}^{n-1}\sum_{j=|\ell-s|}^{\ell+s}(2j+1) = 2\sum_{\ell=0}^{n-1}(2\ell+1) \end{aligned}$$

The states are shown in the tables below.

n	ℓ	s	m_ℓ	m_s	ket
2	1	1/2	1	+1/2	$\|2,1,1/2,1,1/2\rangle_{mm}$
2	1	1/2	0	+1/2	$\|2,1,1/2,0,1/2\rangle_{mm}$
2	1	1/2	-1	+1/2	$\|2,1,1/2,-1,1/2\rangle_{mm}$
2	1	1/2	1	-1/2	$\|2,1,1/2,1,-1/2\rangle_{mm}$
2	1	1/2	0	-1/2	$\|2,1,1/2,0,-1/2\rangle_{mm}$
2	1	1/2	-1	-1/2	$\|2,1,1/2,-1,-1/2\rangle_{mm}$
2	0	1/2	0	1/2	$\|2,0,1/2,0,1/2\rangle_{mm}$
2	0	1/2	0	-1/2	$\|2,0,1/2,0,-1/2\rangle_{mm}$

Table: $n = 2$ level quantum numbers $m_\ell m_s$ states

n	ℓ	s	j	m_j	ket
2	1	1/2	3/2	3/2	$\|2,1,1/2,1,1/2\rangle_{mm}$
2	1	1/2	3/2	1/2	$\|2,1,1/2,0,1/2\rangle_{mm}$
2	1	1/2	3/2	-1/2	$\|2,1,1/2,-1,1/2\rangle_{mm}$
2	1	1/2	3/2	-3/2	$\|2,1,1/2,1,-1/2\rangle_{mm}$
2	1	1/2	1/2	1/2	$\|2,1,1/2,0,-1/2\rangle_{mm}$
2	1	1/2	1/2	-1/2	$\|2,1,1/2,-1,-1/2\rangle_{mm}$
2	0	1/2	1/2	1/2	$\|2,0,1/2,0,1/2\rangle_{mm}$
2	0	1/2	1/2	-1/2	$\|2,0,1/2,0,-1/2\rangle_{mm}$

Table: $n = 2$ level quantum numbers jm_j states

In the first set of states, we could have also included the m_j label since we must have $m_j = m_\ell + m_s$.

In order to learn all the intricate details of this type of calculation, we shall proceed in two ways using the spin-orbit correction as an example.

In method #1, we will construct the $|n, \ell, s, j, m_j\rangle$ states (the zero-order state vectors) from the $|n, \ell, s, m_\ell, m_s\rangle$ and then calculate the first-order energy corrections. In this basis, the $\langle \hat{H}_{so} \rangle$ matrix will be diagonal.

In method #2, we will construct the $\langle \hat{H}_{so} \rangle$ matrix using the easiest states to write down, namely the $|n, \ell, s, m_\ell, m_s\rangle$ states, and then diagonalize it to find the correct first order energies and new zero-order state vectors, which should be the $|n, \ell, s, j, m_j\rangle$ states.

Method #1

We start with the state with maximum j and m_j values. This state always has a CG coefficient equal to 1, i.e., there is only one way to construct it from the other angular momenta.

$$\left|n=2,\ell=1,s=\frac{1}{2},j=\frac{3}{2},m_j=\frac{3}{2}\right\rangle_{jm}$$
$$=\left|n=2,\ell=1,s=\frac{1}{2},m_\ell=1,m_s=\frac{1}{2}\right\rangle_{mm}$$

where we have shown all the labels explicitly. From now on we will write such equations as

$$\left|2,1,\frac{1}{2},\frac{3}{2},\frac{3}{2}\right\rangle_{jm}=\left|2,1,\frac{1}{2},1,\frac{1}{2}\right\rangle_{mm}$$

We then use the lowering operators to obtain

$$\hat{J}_- \left|2,1,\frac{1}{2},\frac{3}{2},\frac{3}{2}\right\rangle_{jm} = \hbar\sqrt{\frac{3}{2}\left(\frac{3}{2}+1\right)-\frac{3}{2}\left(\frac{3}{2}-1\right)}\left|2,1,\frac{1}{2},\frac{3}{2},\frac{1}{2}\right\rangle_{jm}$$

$$= \hbar\sqrt{3}\left|2,1,\frac{1}{2},\frac{3}{2},\frac{1}{2}\right\rangle_{jm} = \left(\hat{L}_- + \hat{S}_-\right)\left|2,1,\frac{1}{2},1,\frac{1}{2}\right\rangle_{mm}$$

$$= \hbar\sqrt{1\,(1+1)-1\,(1-1)}\left|2,1,\frac{1}{2},0,\frac{1}{2}\right\rangle_{mm}$$

$$+\hbar\sqrt{\frac{1}{2}\left(\frac{1}{2}+1\right)-\frac{1}{2}\left(\frac{1}{2}-1\right)}\left|2,1,\frac{1}{2},1,-\frac{1}{2}\right\rangle_{mm}$$

$$= \hbar\sqrt{2}\left|2,1,\frac{1}{2},0,\frac{1}{2}\right\rangle_{mm} + \hbar\sqrt{1}\left|2,1,\frac{1}{2},1,-\frac{1}{2}\right\rangle_{mm}$$

or

$$\left|2,1,\frac{1}{2},\frac{3}{2},\frac{1}{2}\right\rangle_{jm} = \sqrt{\frac{2}{3}}\left|2,1,\frac{1}{2},0,\frac{1}{2}\right\rangle_{mm} + \sqrt{\frac{1}{3}}\left|2,1,\frac{1}{2},1,-\frac{1}{2}\right\rangle_{mm} \tag{8.284}$$

Notice that we use the total J operators on the left and the L and S operators on the right.

The result is a linear combination of the $|n, \ell, s, m_\ell, m_s\rangle$ states all with $m_j = 1/2$ as we expected.

Continuing this process we have

$$\begin{aligned}
\hat{J}_- \left|2, 1, \frac{1}{2}, \frac{3}{2}, \frac{1}{2}\right\rangle_{jm} &= \hbar\sqrt{\frac{3}{2}\left(\frac{3}{2}+1\right) - \frac{1}{2}\left(\frac{1}{2}-1\right)}\left|2, 1, \frac{1}{2}, \frac{3}{2}, -\frac{1}{2}\right\rangle_{jm} \\
&= 2\hbar\left|2, 1, \frac{1}{2}, \frac{3}{2}, -\frac{1}{2}\right\rangle_{jm} \\
&= \left(\hat{L}_- + \hat{S}_-\right)\left(\sqrt{\frac{2}{3}}\left|2, 1, \frac{1}{2}, 0, \frac{1}{2}\right\rangle_{mm} + \sqrt{\frac{1}{3}}\left|2, 1, \frac{1}{2}, 1, -\frac{1}{2}\right\rangle_{mm}\right) \\
&= \hbar\sqrt{\frac{4}{3}}\left|2, 1, \frac{1}{2}, -1, \frac{1}{2}\right\rangle_{mm} + \hbar\sqrt{\frac{8}{3}}\left|2, 1, \frac{1}{2}, 0, -\frac{1}{2}\right\rangle_{mm}
\end{aligned}$$

or

$$\left|2,1,\frac{1}{2},\frac{3}{2},-\frac{1}{2}\right\rangle_{jm}$$
$$=\sqrt{\frac{1}{3}}\left|2,1,\frac{1}{2},-1,\frac{1}{2}\right\rangle_{mm}+\sqrt{\frac{2}{3}}\left|2,1,\frac{1}{2},0,-\frac{1}{2}\right\rangle_{mm} \tag{8.285}$$

and finally

$$\left|2,1,\frac{1}{2},\frac{3}{2},-\frac{3}{2}\right\rangle_{jm}=\left|2,1,\frac{1}{2},-1,-\frac{1}{2}\right\rangle_{mm} \tag{8.286}$$

We now need to construct the maximum state for then next lowest value of j, namely,

$$\left|2,1,\frac{1}{2},\frac{1}{2},\frac{1}{2}\right\rangle_{jm} \tag{8.287}$$

This state has $m_j = 1/2$ so it must be constructed out of the same states that make up

$$\left|2,1,\frac{1}{2},\frac{3}{2},\frac{1}{2}\right\rangle_{jm} \tag{8.288}$$

or it can be written as

$$\left|2,1,\frac{1}{2},\frac{1}{2},\frac{1}{2}\right\rangle_{jm} = a\left|2,1,\frac{1}{2},0,\frac{1}{2}\right\rangle_{mm} + b\left|2,1,\frac{1}{2},1,-\frac{1}{2}\right\rangle_{mm} \tag{8.289}$$

Now we must have

$$\left\langle 2,1,\frac{1}{2},\frac{3}{2},\frac{1}{2}\right|\left.2,1,\frac{1}{2},\frac{1}{2},\frac{1}{2}\right\rangle_{jm} = 0 \text{ orthogonality} \tag{8.290}$$

$$\left\langle 2,1,\frac{1}{2},\frac{1}{2},\frac{1}{2}\right|\left.2,1,\frac{1}{2},\frac{1}{2},\frac{1}{2}\right\rangle_{jm} = 1 \text{ normalization} \tag{8.291}$$

Using the orthonormality of the $|n,\ell,s,m_\ell,m_s\rangle$ states we get

$$\sqrt{\frac{2}{3}}a + \sqrt{\frac{1}{3}}b = 0 \text{ and } a^2 + b^2 = 1 \tag{8.292}$$

The solution is

$$a = \sqrt{\frac{1}{3}} \text{and} b = -\sqrt{\frac{2}{3}} \tag{8.293}$$

and therefore

$$\left|2,1,\frac{1}{2},\frac{1}{2},\frac{1}{2}\right\rangle_{jm} = \sqrt{\frac{1}{3}}\left|2,1,\frac{1}{2},0,\frac{1}{2}\right\rangle_{mm} - \sqrt{\frac{2}{3}}\left|2,1,\frac{1}{2},1,-\frac{1}{2}\right\rangle_{mm} \tag{8.294}$$

In a similar manner, we find

$$\begin{aligned}\left|2,1,\frac{1}{2},\frac{1}{2},-\frac{1}{2}\right\rangle_{jm} \\ = \sqrt{\frac{2}{3}}\left|2,1,\frac{1}{2},-1,\frac{1}{2}\right\rangle_{mm} - \sqrt{\frac{1}{3}}\left|2,1,\frac{1}{2},0,-\frac{1}{2}\right\rangle_{mm}\end{aligned} \tag{8.295}$$

Finally, we construct the other $j = 1/2$ states with $\ell = 0$. They are

$$\left|2,0,\frac{1}{2},0,\frac{1}{2}\right\rangle_{jm} = \left|2,0,\frac{1}{2},0,\frac{1}{2}\right\rangle_{mm} \tag{8.296}$$

$$\left|2,0,\frac{1}{2},0,-\frac{1}{2}\right\rangle_{jm} = \left|2,0,\frac{1}{2},0,-\frac{1}{2}\right\rangle_{mm} \tag{8.297}$$

We can now calculate the first-order energy corrections. We do not actually need the detailed construction of the states to do this, but we will need these states to compare with the results of Method #2 later. We found earlier (10.276) that in the $|n,\ell,s,j,m_j\rangle$ basis

$$E_{n\ell sjm_j} = E_n^{(0)} + \frac{e^2\hbar^2 A_{n\ell sjm_j}}{4m^2c^2}\left(j(j+1) - \ell(\ell+1) - s(s+1)\right) \tag{8.298}$$

$$A_{n\ell sjm_j} = \int d^3\vec{r}\frac{1}{r^3}\left|\psi_{n\ell sjm_j}(\vec{r})\right|^2 \tag{8.299}$$

or

$$\Delta E_{n,\ell,s,j,m_j} = E_{n,\ell,s,j,m_j} - E_n^{(0)} = \frac{e^2\hbar^2 A_\pm}{4m^2c^2} \begin{cases} \ell & j = \ell + 1/2 \\ -(\ell+1) & j = \ell - 1/2 \\ 0 & \ell = 0 \end{cases} \tag{8.300}$$

$$A_\pm = \int d^3\vec{r} \frac{1}{r^3} \left| \psi_{n,\ell,s,j=\ell\pm\frac{1}{2},m_j}(\vec{r}) \right|^2 \tag{8.301}$$

Evaluating the integrals we get

$$\Delta E_{n,\ell,s,j,m_j} = \frac{Z^2 \left| E_n^{(0)} \right| \alpha^2}{n(2\ell+1)(\ell+1)} \qquad j = \ell + \frac{1}{2} \tag{8.302}$$

$$\Delta E_{n,\ell,s,j,m_j} = -\frac{Z^2 \left| E_n^{(0)} \right| \alpha^2}{n\ell(2\ell+1)} \qquad j = \ell - \frac{1}{2} \tag{8.303}$$

$$\Delta E_{n,\ell,s,j,m_j} = 0 \qquad \ell = 0 \tag{8.304}$$

where

$$\alpha = \frac{e^2}{\hbar c} = \text{ fine structure constant} \tag{8.305}$$

Therefore, for the $n = 2$ level we have

$$\Delta E_{2,1,\frac{1}{2},\frac{3}{2},m_j} = \frac{Z^2 \left|E_2^{(0)}\right| \alpha^2}{12} \qquad j = \ell + \frac{1}{2} = \frac{3}{2} \quad (8.306)$$

$$\Delta E_{2,1,\frac{1}{2},\frac{1}{2},m_j} = -\frac{Z^2 \left|E_2^{(0)}\right| \alpha^2}{6} \qquad j = \ell - \frac{1}{2} = \frac{1}{2} \quad (8.307)$$

$$\Delta E_{2,0,\frac{1}{2},\frac{1}{2},m_j} = 0 \qquad j = \ell + \frac{1}{2} = \frac{1}{2} \quad (8.3058)$$

We note that for hydrogen $Z^2\alpha^2 \approx 10^{-4}$ and thus, the fine structure splitting is significantly smaller than the zero-order energies.

The relativity correction is the same order of magnitude as the spin-orbit correction. We found

$$\hat{H}_{relativity} = -\frac{\vec{p}_{op}^{\,4}}{8m^3c^2} \quad (8.309)$$

This gives the correction

$$\Delta E_{rel} = -\frac{\hbar^4}{8m^3c^2}\int_0^\infty r^2 dr \psi^*_{n\ell m}(r)\nabla^4\psi_{n\ell m}(r)$$
$$= \frac{Z^2\left|E_n^{(0)}\right|\alpha^2}{4n^2}\left(3-\frac{4n}{\ell+\frac{1}{2}}\right) \qquad (8.310)$$

Combining these two correction terms(spin-orbit and relativity) gives

$$\Delta E_{fine\ structure} = \frac{Z^2\left|E_2^{(0)}\right|\alpha^2}{4n^2}\left(3-\frac{4n}{j+\frac{1}{2}}\right) \qquad (8.311)$$

The result is independent of ℓ. It turns out that this result is valid for $\ell = 0$ also. There is an additional term that must be added to the Hamiltonian which contributes only in $\ell = 0$ states. It is called the Darwin term.

The Darwin term comes from the relativistic equation for the electron and takes the form

$$\hat{H}_{Darwin} = \frac{\hbar^2}{8m^2c^2}\nabla^2 V = -\frac{\hbar^2}{8m^2c^2}\left(4\pi e Q_{nuclear}(\vec{r})\right) = \frac{\pi\hbar^2 Ze^2}{2m^2c^2}\delta(\vec{r}) \tag{8.312}$$

where

$$Q_{nuclear}(\vec{r}) = \text{ the nuclear charge density} \tag{8.313}$$

Because of the delta function, a contribution

$$\left\langle \hat{H}_{Darwin} \right\rangle_{nj\ell} = \frac{\pi\hbar^2 Ze^2}{2m^2c^2}\left|\psi_{n\ell}(0)\right|^2 = \frac{mc^2 Z^4\alpha^4}{2n^3}\delta_{\ell,0} \tag{8.314}$$

arises for $\ell = 0$ states *only*. This is identical to the contribution $\langle \hat{H}_{so} + \hat{H}_{rel} \rangle$ for $\ell = 0, j = 1/2$.

Method #2

We use the $|n, \ell, s, m_\ell, m_s\rangle$ basis. In this case, the best form of the operators to use (means we know how to evaluate them with these states) are

$$\begin{aligned}\hat{H}_{so} &= \frac{e^2}{2m^2c^2}\left(\frac{1}{r^3}\right)\vec{S}_{op}\cdot\vec{L}_{op} \\ &= \frac{e^2}{2m^2c^2}\left(\frac{1}{r^3}\right)\left(\hat{L}_z\hat{S}_z + \frac{1}{2}\left(\hat{L}_+\hat{S}_- + \hat{L}_-\hat{S}_+\right)\right)\end{aligned} \tag{8.315}$$

If we label the rows and columns of the matrix representation by

$$\begin{aligned}|1\rangle &= \left|2, 1, \frac{1}{2}, 1, \frac{1}{2}\right\rangle_{mm} \qquad m_j = \frac{3}{2} \\ |2\rangle &= \left|2, 1, \frac{1}{2}, 1, -\frac{1}{2}\right\rangle_{mm} \qquad m_j = \frac{1}{2} \\ |3\rangle &= \left|2, 1, \frac{1}{2}, 0, \frac{1}{2}\right\rangle_{mm} \qquad m_j = \frac{1}{2}\end{aligned} \tag{1}$$

$$
\begin{aligned}
|4\rangle &= \left|2, 1, \frac{1}{2}, 0, -\frac{1}{2}\right\rangle_{mm} & m_j &= -\frac{1}{2} \\
|5\rangle &= \left|2, 1, \frac{1}{2}, -1, \frac{1}{2}\right\rangle_{mm} & m_j &= -\frac{1}{2} \\
|6\rangle &= \left|2, 1, \frac{1}{2}, -1, -\frac{1}{2}\right\rangle_{mm} & m_j &= -\frac{3}{2} \\
|7\rangle &= \left|2, 0, \frac{1}{2}, 0, \frac{1}{2}\right\rangle_{mm} & m_j &= \frac{1}{2} \\
|8\rangle &= \left|2, 0, \frac{1}{2}, 0, -\frac{1}{2}\right\rangle_{mm} & m_j &= -\frac{1}{2}
\end{aligned}
$$

then we get the matrix for $\langle \hat{H}_{so} \rangle$ as shown in the table below.

I have used a table rather than an equation format so that I could clearly label the rows and columns by the state index.

	1	2	3	4	5	6	7	8
1	a	0	0	0	0	0	0	0
2	0	b	c	0	0	0	0	0
3	0	c	d	0	0	0	0	0
4	0	0	0	e	f	0	0	0
5	0	0	0	f	g	0	0	0
6	0	0	0	0	0	h	0	0
7	0	0	0	0	0	0	p	0
8	0	0	0	0	0	0	0	q

Table: $\langle \hat{H}_{so} \rangle$ matrix

We have marked the non-zero elements. Using the operator properties derived earlier we get

$$
\begin{aligned}
a &= \langle 1|\, \hat{H}_{so}\, |1\rangle = \left\langle 2,1,\frac{1}{2},1,\frac{1}{2}\right|\, \hat{H}_{so}\, \left|2,1,\frac{1}{2},1,\frac{1}{2}\right\rangle_{mm} \\
&= \left\langle 2,1,\frac{1}{2},1,\frac{1}{2}\right| \frac{e^2}{2m^2c^2}\left(\frac{1}{r^3}\right) \\
&\quad \times \left(\hat{L}_z\hat{S}_z + \frac{1}{2}\left(\hat{L}_+\hat{S}_- + \hat{L}_-\hat{S}_+\right)\right)\left|2,1,\frac{1}{2},1,\frac{1}{2}\right\rangle_{mm} \\
&= \frac{e^2}{2m^2c^2}\,\langle 2,1|\,\frac{1}{r^3}\,|2,1\rangle_{mm} \\
&\quad \times \left\langle 2,1,\frac{1}{2},1,\frac{1}{2}\right|\left(\hat{L}_z\hat{S}_z + \frac{1}{2}\left(\hat{L}_+\hat{S}_- + \hat{L}_-\hat{S}_+\right)\right)\left|2,1,\frac{1}{2},1,\frac{1}{2}\right\rangle_{mm} \\
&= \frac{e^2}{2m^2c^2}\,\langle 2,1|\,\frac{1}{r^3}\,|2,1\rangle_{mm}\left\langle 2,1,\frac{1}{2},1,\frac{1}{2}\right|\hat{L}_z\hat{S}_z\left|2,1,\frac{1}{2},1,\frac{1}{2}\right\rangle \\
&= \frac{e^2\hbar^2}{4m^2c^2}\int_0^\infty \frac{1}{r}R_{n\ell}^2(r)dr = \frac{e^2\hbar^2}{4m^2c^2}\left[\frac{Z^3}{a_0^3 n^3 \ell(\ell+\frac{1}{2})(\ell+1)}\right]
\end{aligned}
\tag{8.316}
$$

Similar calculations give $\langle \hat{H}_{so} \rangle$ as shown in the table below.

	1	2	3	4	5	6	7	8
1	a	0	0	0	0	0	0	0
2	0	-a	$\sqrt{2}$a	0	0	0	0	0
3	0	$\sqrt{2}$a	0	0	0	0	0	0
4	0	0	0	0	$\sqrt{2}$a	0	0	0
5	0	0	0	$\sqrt{2}$a	-a	0	0	0
6	0	0	0	0	0	a	0	0
7	0	0	0	0	0	0	0	0
8	0	0	0	0	0	0	0	0

Table: $\langle \hat{H}_{so} \rangle$ matrix - revised

This says that (only diagonal elements)

$$E_1^{(1)} = a = E_6^{(1)} \quad , \quad E_7^{(1)} = 0 = E_8^{(1)} \tag{8.317}$$

and

$$|1\rangle = \left|2, 1, \frac{1}{2}, \frac{3}{2}, \frac{3}{2}\right\rangle_{jm} \text{ and } |6\rangle = \left|2, 1, \frac{1}{2}, \frac{3}{2}, -\frac{3}{2}\right\rangle_{jm}$$

$$|7\rangle = \left|2, 0, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}\right\rangle_{jm} \text{ and } |8\rangle = \left|2, 0, \frac{1}{2}, \frac{1}{2}, -\frac{1}{2}\right\rangle_{jm}$$

In order to find $E_{2'}^{(1)}, E_{3'}^{(1)}, E_{4'}^{(1)}, E_{5'}^{(1)}$, corresponding to the new zero-order state vectors $|2'\rangle, |3'\rangle, |4'\rangle, |5'\rangle$, we must diagonalize the two 2×2 submatrices.

We begin with the submatrix involving states $|2\rangle$ and $|3\rangle$ as shown in the table below, namely,

	2	3
2	-a	$\sqrt{2}$a
3	$\sqrt{2}$a	0

Table: $\langle \hat{H}_{so} \rangle$ 2-3 submatrix

The characteristic equation is

$$(-a - E)(-E) - 2a^2 = 0 = E^2 + aE - 2a^2 \tag{8.318}$$

or

$$E^{(1)}_{2'} = a \text{ and } E^{(1)}_{3'} = -2a \tag{8.319}$$

Notice that these energies are

$$E^{(1)}_{2'} = \ell a \text{ and } E^{(1)}_{3'} = -(\ell + 1)a \tag{8.320}$$

as expected for the

$$j = \ell + \frac{1}{2} \text{ and } j = \ell - \frac{1}{2} \tag{8.321}$$

states respectively.

We find the eigenvectors using the eigenvalue equations. For $|2'\rangle$ we have

$$\begin{pmatrix} -a & a\sqrt{2} \\ a\sqrt{2} & 0 \end{pmatrix} |2'\rangle = \begin{pmatrix} -a & a\sqrt{2} \\ a\sqrt{2} & 0 \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix}$$
$$= E^{(1)}_{2'} \begin{pmatrix} u \\ v \end{pmatrix} = a \begin{pmatrix} u \\ v \end{pmatrix}$$

or

$$-u + \sqrt{2}v = u \text{ and } \sqrt{2}u = v \tag{8.322}$$

Using the normalization condition $u^2 + v^2 = 1$ we get

$$u = \sqrt{\frac{1}{3}} \text{ and } v = \sqrt{\frac{2}{3}} \tag{8.323}$$

or

$$|2'\rangle = \sqrt{\frac{1}{3}}\,|2\rangle + \sqrt{\frac{2}{3}}\,|3\rangle = \left|2, 1, \frac{1}{2}, \frac{3}{2}, \frac{1}{2}\right\rangle_{jm} \tag{8.324}$$

and similarly

$$|3'\rangle = \sqrt{\frac{2}{3}}\,|2\rangle - \sqrt{\frac{1}{3}}\,|3\rangle = \left|2, 1, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}\right\rangle_{jm} \tag{8.325}$$

We then deal with the submatrix involving states $|4\rangle$ and $|5\rangle$ as shown in table below, namely,

	4	5
4	0	$\sqrt{2}$a
5	$\sqrt{2}$a	-a

Table: $\langle \hat{H}_{so} \rangle$ 4-5 submatrix

The characteristic equation is

$$(-a - E)(-E) - 2a^2 = 0 = E^2 + aE - 2a^2 \tag{8.326}$$

or

$$E^{(1)}_{4'} = a \text{ and } E^{(1)}_{5'} = -2a \tag{8.327}$$

and the eigenvectors are

$$\left|4'\right\rangle = \sqrt{\frac{2}{3}}\left|4\right\rangle + \sqrt{\frac{1}{3}}\left|5\right\rangle = \left|2,1,\frac{1}{2},\frac{3}{2},-\frac{1}{2}\right\rangle_{jm} \tag{8.328}$$

$$\left|5'\right\rangle = \sqrt{\frac{1}{3}}\left|4\right\rangle - \sqrt{\frac{2}{3}}\left|5\right\rangle = \left|2,1,\frac{1}{2},\frac{1}{2},-\frac{1}{2}\right\rangle_{jm} \tag{8.329}$$

So including the spin-orbit correction we end up with the energy levels

$$E^{(0)}_2 + a \text{ for } 1, 2', 4', 6 \rightarrow \text{ 4 - fold degenerate}$$

$$E^{(0)}_2 \text{ for } 7, 8 \rightarrow \text{ 2 - fold degenerate}$$

$$E^{(0)}_2 - 2a \text{ for } 3', 5' \rightarrow \text{ 2 - fold degenerate}$$

8.6.2 **Spin-Orbit and Arbitrary Magnetic Field**

Now let us add on the Zeeman correction (for $\vec{B} = B\hat{z}$)

$$\hat{H}_{Zeeman} = \frac{\mu_B}{\hbar}(\vec{L}_{op} + 2\vec{S}_{op}) \cdot \vec{B} = \frac{\mu_B B}{\hbar}(\hat{L}_z + 2S_z) \qquad (8.330)$$

We can solve this problem for *arbitrary magnetic field* by repeating Method #2 using the correction term as the sum of spin-orbit and Zeeman effects.

The zero-order Hamiltonian is $\hat{H}_0$ and the zero-order state vectors are the $|n, \ell, s, m_\ell, m_s\rangle$ states. The eight zero-order $n = 2$ states are all degenerate with energy

$$E_2^{(0)} = -\frac{e^2}{8a_0} \qquad (8.331)$$

so we must use degenerate perturbation theory.

We have already calculated the $\langle \hat{H}_{so} \rangle$ in this basis. It is shown in the table below.

	1	2	3	4	5	6	7	8
1	a	0	0	0	0	0	0	0
2	0	-a	$\sqrt{2}$a	0	0	0	0	0
3	0	$\sqrt{2}$a	0	0	0	0	0	0
4	0	0	0	0	$\sqrt{2}$a	0	0	0
5	0	0	0	$\sqrt{2}$a	-a	0	0	0
6	0	0	0	0	0	a	0	0
7	0	0	0	0	0	0	0	0
8	0	0	0	0	0	0	0	0

Table: $\langle \hat{H}_{so} \rangle$ matrix

where

$$a = \frac{e^2 \hbar^2}{96 m^2 a_0^3 c^2} \tag{8.332}$$

The $\langle \hat{L}_z + 2\hat{S}_z \rangle$ matrix is diagonal in this representation and its diagonal elements are given by $m_\ell + 2m_s$ and so the Zeeman contribution $\langle \hat{H}_{Zeeman} \rangle$ is shown in the table below.

	1	2	3	4	5	6	7	8
1	2b	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0	0
3	0	0	b	0	0	0	0	0
4	0	0	0	-b	0	0	0	0
5	0	0	0	0	0	0	0	0
6	0	0	0	0	0	-2b	0	0
7	0	0	0	0	0	0	b	0
8	0	0	0	0	0	0	0	-b

Table: $\langle \hat{H}_{Zeeman} \rangle$ matrix

where $b = \mu_B B$. The combined perturbation matrix $\langle \hat{V} \rangle$ is then given in the table below.

	1	2	3	4	5	6	7	8
1	a+2b	0	0	0	0	0	0	0
2	0	-a	$a\sqrt{2}$	0	0	0	0	0
3	0	$a\sqrt{2}$	b	0	0	0	0	0
4	0	0	0	-b	$a\sqrt{2}$	0	0	0
5	0	0	0	$a\sqrt{2}$	-a	0	0	0
6	0	0	0	0	0	a-2b	0	0
7	0	0	0	0	0	0	b	0
8	0	0	0	0	0	0	0	-b

Table: $\langle \hat{V} \rangle$ matrix

After diagonalizing, the new energies are

$$E_{1'} = -\frac{e^2}{8a_0} + \frac{e^2}{8a_0}\frac{\alpha^2}{12} + 2\mu_B B$$

$$E_{2'} = -\frac{e^2}{8a_0} + \frac{1}{2}(-(a-b) + \sqrt{9a^2 + 2ab + b^2})$$

$$E_{3'} = -\frac{e^2}{8a_0} + \frac{1}{2}(-(a-b) - \sqrt{9a^2 + 2ab + b^2})$$

$$E_{4'} = -\frac{e^2}{8a_0} + \frac{1}{2}(-(a+b) + \sqrt{9a^2 - 2ab + b^2})$$

$$E_{5'} = -\frac{e^2}{8a_0} + \frac{1}{2}(-(a+b) - \sqrt{9a^2 - 2ab + b^2})$$

$$E_{6'} = -\frac{e^2}{8a_0} + \frac{e^2}{8a_0}\frac{\alpha^2}{12} - 2\mu_B B$$

$$E_{7'} = -\frac{e^2}{8a_0} + \mu_B B$$

$$E_{8'} = -\frac{e^2}{8a_0} - \mu_B B$$

If we let B be small so that $b \ll a$, we then get the approximate energies

$$E_{1'} = -\frac{e^2}{8a_0} + \frac{e^2}{8a_0}\frac{\alpha^2}{12} + 2\mu_B B$$

$$E_{2'} = -\frac{e^2}{8a_0} + \frac{1}{2}(-(a-b) + 3a(1+\frac{b}{9a}))$$

$$= -\frac{e^2}{8a_0} + a + \frac{2}{3}b = -\frac{e^2}{8a_0} + \frac{e^2}{8a_0}\frac{\alpha^2}{12} + \frac{2}{3}\mu_B B$$

$$E_{3'} = -\frac{e^2}{8a_0} + \frac{1}{2}(-(a-b) - 3a(1+\frac{b}{9a}))$$

$$= -\frac{e^2}{8a_0} - 2a + \frac{1}{3}b = -\frac{e^2}{8a_0} - \frac{e^2}{8a_0}\frac{\alpha^2}{6} + \frac{1}{3}\mu_B B$$

$$E_{4'} = -\frac{e^2}{8a_0} + \frac{1}{2}(-(a+b) + 3a(1-\frac{b}{9a}))$$

$$= -\frac{e^2}{8a_0} + a - \frac{2}{3}b = -\frac{e^2}{8a_0} + \frac{e^2}{8a_0}\frac{\alpha^2}{12} - \frac{2}{3}\mu_B B$$

$$E_{5'} = -\frac{e^2}{8a_0} + \frac{1}{2}(-(a+b) - 3a(1 - \frac{b}{9a}))$$

$$= -\frac{e^2}{8a_0} - 2a - \frac{1}{3}b = -\frac{e^2}{8a_0} - \frac{e^2}{8a_0}\frac{\alpha^2}{6} - \frac{1}{3}\mu_B B$$

$$E_{6'} = -\frac{e^2}{8a_0} + \frac{e^2}{8a_0}\frac{\alpha^2}{12} - 2\mu_B B$$

$$E_{7'} = -\frac{e^2}{8a_0} + \mu_B B$$

$$E_{8'} = -\frac{e^2}{8a_0} - \mu_B B$$

This is clearly a perturbation of the spin-orbit energy levels. We assume that the new state vectors become the zero-order vectors in the spin-orbit case for low fields. In this case the Zeeman effect corrections(to the fine structure energies) are given by the table below.

ℓ	s	j	m_j	ΔE_{Zeeman}	State
1	1/2	3/2	3/2	$2\mu_B B$	$1'$
1	1/2	3/2	1/2	$2\mu_B B/3$	$2'$
1	1/2	1/2	1/2	$\mu_B B/3$	$3'$
1	1/2	3/2	-1/2	$-2\mu_B B/3$	$4'$
1	1/2	1/2	-1/2	$-\mu_B B/3$	$5'$
1	1/2	3/2	-3/2	$-2\mu_B B$	$6'$
0	1/2	1/2	1/2	$\mu_B B$	$7'$
0	1/2	1/2	-1/2	$-\mu_B B$	$8'$

Table: $n = 2$ energy corrections for small B

A little bit of study shows the general relation

$$\Delta E_{Zeeman} = g\mu_B B m_j \tag{8.333}$$

where

$$g = \text{ Lande g - factor } = 1 + \frac{j(j+1) - \ell(\ell+1) + s(s+1)}{2j(j+1)} \tag{8.334}$$

This is called the *Zeeman effect.*

We can prove this result in general. The general method uses the *Wigner-Eckart Theorem.*

8.6.3 **Wigner-Eckart Theorem**

Consider a vector operator $\vec{A}_{op}$. We have already shown that the Cartesian components of any vector operator has the following commutation relations with the Cartesian components of the angular momentum operator

$$\left[\hat{A}_i, \hat{J}_j\right] = i\hbar\varepsilon_{ijk}\hat{A}_k \tag{8.334}$$

We will now prove the following powerful theorem:

In a basis that diagonalizes $\vec{J}^2_{op}$ and $\hat{J}_z$ (i.e., the $|\lambda, \ell, s, j, m_j\rangle$ states, where λ signifies other operators that commute with $\vec{J}^2_{op}$ and $\hat{J}_z$), the matrix elements of $\vec{A}_{op}$ between states with the *same* $j-$value are proportional to the matrix elements of $\vec{J}_{op}$ and the proportionality factor is independent of m_j.

The algebra involved in the proof is simpler if we work in the so-called *spherical basis* instead of the Cartesian basis. The spherical basis uses

$$\hat{J}_\pm = \hat{J}_x \pm i\hat{J}_y \quad , \quad \hat{J}_0 = \hat{J}_z \tag{8.335}$$

$$\hat{A}_\pm = \hat{A}_x \pm i\hat{A}_y \quad , \quad \hat{A}_0 = \hat{A}_z \tag{8.336}$$

The corresponding commutators are

$$\left[\hat{A}_\pm, \hat{J}_0\right] = \mp\hbar\hat{A}_\pm \quad , \quad \left[\hat{A}_\pm, \hat{J}_\pm\right] = 0 \tag{8.337}$$

$$\left[\hat{A}_\pm, \hat{J}_\mp\right] = \pm 2\hbar\hat{A}_0 \quad , \quad \left[\hat{A}_0, \hat{J}_0\right] = 0 \tag{8.338}$$

$$\left[\hat{A}_0, \hat{J}_\pm\right] = \pm\hbar\hat{A}_\pm \tag{8.339}$$

which all follow from the original commutator for the arbitrary vector operator. Now, by definition of the operators, we have

$$\hat{J}_0 \left|\lambda, j, m_j\right\rangle = \hbar m_j \left|\lambda, j, m_j\right\rangle = \left\langle j, m_j\right| \hat{J}_0 \left|j, m_j\right\rangle \left|\lambda, j, m_j\right\rangle \tag{8.340}$$

$$\begin{aligned} \hat{J}_\pm \left|\lambda, j, m_j\right\rangle &= \hbar\sqrt{j(j+1) - m_j(m_j \pm 1)} \left|\lambda, j, m_j \pm 1\right\rangle \\ &= \left\langle j, m_j \pm 1\right| \hat{J}_\pm \left|j, m_j\right\rangle \left|\lambda, j, m_j \pm 1\right\rangle \end{aligned} \tag{8.341}$$

$$\begin{aligned} \left\langle \lambda, j, m_j\right| \hat{J}_\pm &= \left(\hat{J}_\mp \left|\lambda, j, m_j\right\rangle\right)^+ \\ &= \left\langle \lambda, j, m_j \mp 1\right| \hbar\sqrt{j(j+1) - m_j(m_j \mp 1)} \\ &= \left\langle \lambda, j, m_j \mp 1\right| \left\langle j, m_j\right| \hat{J}_\pm \left|j, m_j \mp 1\right\rangle \end{aligned} \tag{8.343}$$

We now work with the matrix elements of some of the commutators and use the defining relations above to prove the theorem.

First, we have

$$\begin{aligned}\mp\hbar\left\langle\lambda', j, m_j'\right| \hat{A}_\pm \left|\lambda, j, m_j\right\rangle &= \left\langle\lambda', j, m_j'\right| \left[\hat{A}_\pm, \hat{J}_0\right] \left|\lambda, j, m_j\right\rangle \\ &= \left\langle\lambda', j, m_j'\right| \left[\hat{A}_\pm \hat{J}_0 - \hat{J}_0 \hat{A}_\pm\right] \left|\lambda, j, m_j\right\rangle \\ &= (m_j - m_j')\hbar \left\langle\lambda', j, m_j'\right| \hat{A}_\pm \left|\lambda, j, m_j\right\rangle\end{aligned} \tag{8.344}$$

or

$$0 = (m_j - m_j' \pm 1)\hbar \left\langle\lambda', j, m_j'\right| \hat{A}_\pm \left|\lambda, j, m_j\right\rangle \tag{8.345}$$

This says that either $m_j' = m_j \pm 1$ or the matrix element

$$\left\langle\lambda', j, m_j'\right| \hat{A}_\pm \left|\lambda, j, m_j\right\rangle = 0 \tag{8.346}$$

Since we have an identical property for the matrix elements of $\hat{J}_\pm$ this implies that the matrix elements of $\hat{A}_\pm$ are proportional to those of $\hat{J}_\pm$ and we can write the proportionality constant as

$$\frac{\left\langle\lambda', j, m_j \pm 1\right| \hat{A}_\pm \left|\lambda, j, m_j\right\rangle}{\left\langle j, m_j \pm 1\right| \hat{J}_\pm \left|j, m_j\right\rangle} \tag{8.347}$$

Second, we have

$$\left\langle \lambda', j, m_j' \right| \left[\hat{A}_\pm, \hat{J}_\pm \right] \left| \lambda, j, m_j \right\rangle = 0 \tag{8.348}$$

$$\left\langle \lambda', j, m_j' \right| \hat{A}_\pm \hat{J}_\pm \left| \lambda, j, m_j \right\rangle = \left\langle \lambda', j, m_j' \right| \hat{J}_\pm \hat{A}_\pm \left| \lambda, j, m_j \right\rangle \tag{8.349}$$

$$\begin{aligned} &\left\langle \lambda', j, m_j' \right| \hat{A}_\pm \left| \lambda, j, m_j \pm 1 \right\rangle \left\langle \lambda', j, m_j \pm 1 \right| \hat{J}_\pm \left| \lambda, j, m_j \right\rangle \\ &\quad = \left\langle \lambda', j, m_j' \mp 1 \right| \hat{A}_\pm \left| \lambda, j, m_j \right\rangle \left\langle \lambda, j, m_j \pm 1 \right| \hat{J}_\pm \left| \lambda, j, m_j' \mp 1 \right\rangle \end{aligned} \tag{8.350}$$

Using the result from the first commutator this says that $m_j' = m_j \pm 2$, which, in turn, implies that

$$\frac{\left\langle \lambda', j, m_j \pm 2 \right| \hat{A}_\pm \left| \lambda, j, m_j \pm 1 \right\rangle}{\left\langle j, m_j \pm 2 \right| \hat{J}_\pm \left| j, m_j \pm 1 \right\rangle} = \frac{\left\langle \lambda', j, m_j \pm 1 \right| \hat{A}_\pm \left| \lambda, j, m_j \right\rangle}{\left\langle j, m_j \pm 1 \right| \hat{J}_\pm \left| j, m_j \right\rangle} \tag{8.351}$$

This says that the proportionality constant is independent of m_j.

We define a new symbol for the proportionality constant

$$\left\langle \lambda', j \right| \left| A \right| \left| \lambda, j \right\rangle_\pm = \text{ the reduced matrix element} \tag{8.352}$$

which gives the relation

$$\left\langle \lambda', j, m'_j \right| \hat{A}_{\pm} \left| \lambda, j, m_j \right\rangle = \left\langle \lambda', j \right| \left| A \right| \left| \lambda, j \right\rangle_{\pm} \left\langle j, m'_j \right| \hat{J}_{\pm} \left| j, m_j \right\rangle \tag{8.353}$$

To complete the proof we need to show that the same result holds for $\hat{A}_0$ and that

$$\left\langle \lambda', j \right| \left| A \right| \left| \lambda, j \right\rangle_{+} = \left\langle \lambda', j \right| \left| A \right| \left| \lambda, j \right\rangle_{-} \tag{8.354}$$

We have

$$\begin{aligned} \pm 2\hbar \left\langle \lambda', j, m'_j \right| \hat{A}_0 \left| \lambda, j, m_j \right\rangle &= \left\langle \lambda', j, m'_j \right| \left[\hat{A}_{\pm}, \hat{J}_{\mp} \right] \left| \lambda, j, m_j \right\rangle \\ &= \left\langle \lambda', j, m'_j \right| \left[\hat{A}_{\pm} \hat{J}_{\mp} - \hat{J}_{\mp} \hat{A}_{\pm} \right] \left| \lambda, j, m_j \right\rangle \\ &= \left\langle \lambda', j, m'_j \right| \hat{A}_{\pm} \left| \lambda, j, m_j \mp 1 \right\rangle \left\langle j, m_j \mp 1 \right| \hat{J}_{\mp} \left| j, m_j \right\rangle \\ &\quad - \left\langle \lambda', j, m'_j \pm 1 \right| \hat{A}_{\pm} \left| \lambda, j, m_j \right\rangle \left\langle j, m'_j \right| \hat{J}_{\mp} \left| j, m'_j \pm 1 \right\rangle \end{aligned} \tag{8.355}$$

Now substituting in the matrix element of $\hat{A}_{\pm}$ we get

$$\pm 2\hbar \left\langle \lambda', j, m_j' \right| \hat{A}_0 \left| \lambda, j, m_j \right\rangle$$
$$= \left\langle \lambda', j \right| |A| \left| \lambda, j \right\rangle_{\pm} \left[\left\langle j, m_j' \right| \hat{J}_{\pm} \left| j, m_j \mp 1 \right\rangle \left\langle j, m_j \mp 1 \right| \hat{J}_{\mp} \left| j, m_j \right\rangle \right.$$
$$\left. - \left\langle j, m_j' \pm 1 \right| \hat{J}_{\pm} \left| j, m_j \right\rangle \left\langle j, m_j' \right| \hat{J}_{\mp} \left| j, m_j' \pm 1 \right\rangle \right] \tag{8.356}$$

This says that $\hat{A}_0$ has non-vanishing matrix elements only when $m_j' = m_j$. We then get

$$\pm 2\hbar \left\langle \lambda', j, m_j' \right| \hat{A}_0 \left| \lambda, j, m_j \right\rangle$$
$$= \left\langle \lambda', j \right| |A| \left| \lambda, j \right\rangle_{\pm} \left[\left| \left\langle j, m_j \mp 1 \right| \hat{J}_{\mp} \left| j, m_j \right\rangle \right|^2 - \left| \left\langle j, m_j \pm 1 \right| \hat{J}_{\mp} \left| j, m \right. \right. \right.$$
$$= \pm 2\hbar m_j = \pm 2\hbar \left\langle \lambda', j, m_j' \right| \hat{J}_0 \left| \lambda, j, m_j \right\rangle \tag{8.357}$$

Putting it all together we get

$$\left\langle \lambda', j, m_j' \right| \hat{A}_0 \left| \lambda, j, m_j \right\rangle = \left\langle \lambda', j \right| |A| \left| \lambda, j \right\rangle_{\pm} \left\langle j, m_j' \right| \hat{J}_0 \left| j, m_j \right\rangle \tag{8.358}$$

Since no operator has a $\pm$ subscript, this also says that

$$\left\langle \lambda', j \right| |A| \left| \lambda, j \right\rangle_{+} = \left\langle \lambda', j \right| |A| \left| \lambda, j \right\rangle_{-} = \left\langle \lambda', j \right| |A| \left| \lambda, j \right\rangle \tag{8.359}$$

and we finally have

$$\left\langle \lambda', j, m'_j \right| \vec{A}_{op} \left| \lambda, j, m_j \right\rangle = \left\langle \lambda', j \right| \left| A \right| \left| \lambda, j \right\rangle_{\pm} \left\langle j, m'_j \right| \vec{J}_{op} \left| j, m_j \right\rangle \tag{8.360}$$

This completes the proof of the *Wigner-Eckart theorem.*

A very important extension of this theorem is the following result:

$$\begin{aligned} \left\langle \lambda', j, m'_j \right| \vec{A}_{op} \cdot \vec{J}_{op} \left| \lambda, j, m_j \right\rangle &= \left\langle \lambda', j \right| \left| A \right| \left| \lambda, j \right\rangle \left\langle j, m'_j \right| \vec{J}^2_{op} \left| j, m_j \right\rangle \\ &= \delta_{m'_j, m_j} \hbar^2 j(j+1) \left\langle \lambda', j \right| \left| A \right| \left| \lambda, j \right\rangle \end{aligned} \tag{8.361}$$

which says that the scalar product is diagonal in m_j. This result follows directly from the Wigner-Eckart theorem

$$\left\langle \lambda', j, m'_j \right| \vec{A}_{op} \cdot \vec{J}_{op} \left| \lambda, j, m_j \right\rangle = \sum_k \left\langle \lambda', j, m'_j \right| A_{op,k} J_{op,k} \left| \lambda, j, m_j \right\rangle$$

$$= \sum_k \left\langle \lambda', j, m'_j \right| A_{op,k} \left(\sum_{m''_j} \left| \lambda, j, m''_j \right\rangle \left\langle \lambda, j, m''_j \right| \right) J_{op,k} \left| \lambda, j, m_j \right\rangle$$

$$= \sum_{\lambda'', j'', m''_j} \sum_k \left\langle \lambda', j, m'_j \right| A_{op,k} \left| \lambda, j, m''_j \right\rangle \left\langle \lambda, j, m''_j \right| J_{op,k} \left| \lambda, j, m_j \right\rangle$$

$$= \sum_{m''_j} \sum_k \left\langle \lambda', j \right| \left| A \right| \left| \lambda, j \right\rangle \left\langle \lambda', j, m'_j \right| J_{op,k} \left| \lambda, j, m''_j \right\rangle$$

$$\times \left\langle \lambda, j, m''_j \right| J_{op,k} \left| \lambda, j, m_j \right\rangle$$

$$= \left\langle \lambda', j \right| \left| A \right| \left| \lambda, j \right\rangle \sum_k \left\langle \lambda', j, m'_j \right| J_{op,k} \left(\sum_{m''_j} \left| \lambda, j, m''_j \right\rangle \left\langle \lambda, j, m''_j \right| \right)$$

$$\times J_{op,k} \left| \lambda, j, m_j \right\rangle$$

$$= \left\langle \lambda', j \right| \left| A \right| \left| \lambda, j \right\rangle \left\langle j, m'_j \right| \vec{J}^2_{op} \left| j, m_j \right\rangle = \delta_{m'_j, m_j} \hbar^2 j(j+1) \left\langle \lambda', j \right| \left| A \right| \left| \lambda, j \right\rangle$$

Now back to the Zeeman effect. In the low field limit, we need to evaluate the diagonal matrix elements

$$\begin{aligned}\langle \ell s j m_j | \left(\hat{L}_z + 2\hat{S}_z\right) | \ell s j m_j \rangle &= \langle \ell s j m_j | \left(\hat{J}_z + \hat{S}_z\right) | \ell s j m_j \rangle \\ &= \hbar m_j + \langle \ell s j m_j | \hat{S}_z | \ell s j m_j \rangle \end{aligned} \quad (8.362)$$

Now the Wigner-Eckart theorem says that

$$\begin{aligned}\langle \ell s j m_j | \hat{S}_z | \ell s j m_j \rangle &= \langle \ell s j | |S| | \ell s j \rangle \langle \ell s j m_j | \hat{J}_z | \ell s j m_j \rangle \\ &= \hbar m_j \langle \ell s j | |S| | \ell s j \rangle \end{aligned} \quad (8.363)$$

The scalar product matrix element formula gives

$$\begin{aligned}\langle \ell s j m_j | \vec{S}_{op} \cdot \vec{J}_{op} | \ell s j m_j \rangle &= \langle \ell s j | |S| | \ell s j \rangle \langle j m_j | \vec{J}^2_{op} | j m_j \rangle \\ &= \hbar^2 j(j+1) \langle \ell s j | |S| | \ell s j \rangle \end{aligned} \quad (8.364)$$

But we also have

$$(\vec{J}_{op} - \vec{S}_{op})^2 = \vec{L}^2_{op} = \vec{J}^2_{op} + \vec{S}^2_{op} - 2\vec{S}_{op} \cdot \vec{J}_{op} \quad (8.365)$$

$$\langle \ell s j m_j | \, \vec{S}_{op} \cdot \vec{J}_{op} \, | \ell s j m_j \rangle = \frac{1}{2} \langle \ell s j m_j | \, (\vec{J}^2_{op} + \vec{S}^2_{op} - \vec{L}^2_{op}) \, | \ell s j m_j \rangle$$
$$= \frac{1}{2} \hbar^2 (j(j+1) + s(s+1) - \ell(\ell+1)) \tag{8.366}$$

or

$$\langle \ell s j | \, |S| \, | \ell s j \rangle = \frac{j(j+1) + s(s+1) - \ell(\ell+1)}{2j(j+1)} \tag{8.367}$$

and thus

$$\langle \ell s j m_j | \, (\hat{L}_z + 2\hat{S}_z) \, | \ell s j m_j \rangle$$
$$= \hbar m_j + \hbar m_j \frac{j(j+1) + s(s+1) - \ell(\ell+1)}{2j(j+1)}$$
$$= \hbar m_j g_{j\ell s} \tag{8.368}$$

$$g_{j\ell s} = 1 + \frac{j(j+1) + s(s+1) - \ell(\ell+1)}{2j(j+1)} = \text{ Lande g - factor} \tag{8.369}$$

Finally, we have

$$\langle \ell s j m_j | \, \hat{H}_{Zeeman} \, | \ell s j m_j \rangle = \mu_B B m_j g_{j\ell s} \tag{8.370}$$

and the result we found earlier in the special example case is now proved in general.

8.6.4 **Paschen-Bach Effect**

When B is large enough such that $\Delta E_{Zeeman} \gg \Delta E_{so}$, but not large enough so that the $\vec{B}^2$ term we neglected earlier is important. We have the so-called *Paschen-Bach* effect. If the $\vec{B}^2$ term is dominant we have the so-called *quadratic Zeeman effect.*

The best way to see what is happening for all magnetic field values is a plot. In CGS Gaussian units

$$\mu_B = 5.7884 \times 10^{-9} \frac{eV}{gauss}, a_0 = 5.2918 \times 10^{-8} cm, e = 4.80 \times 10^{-10} esu$$

$$\frac{e^2}{a_0} = 27.2\, eV a = 1.509 \times 10^{-5}\, eV b = 5.7884 \times 10^{-9} B\, eV$$

Using our earlier results we then have

$$E_{1'} = -\frac{e^2}{8a_0} + \frac{e^2}{8a_0}\frac{\alpha^2}{12} + 2\mu_B B$$

$$E_{2'} = -\frac{e^2}{8a_0} + \frac{1}{2}(-(a-b) + \sqrt{9a^2 + 2ab + b^2})$$

$$E_{3'} = -\frac{e^2}{8a_0} + \frac{1}{2}(-(a-b) - \sqrt{9a^2 + 2ab + b^2})$$

$$E_{4'} = -\frac{e^2}{8a_0} + \frac{1}{2}(-(a+b) + \sqrt{9a^2 - 2ab + b^2})$$

$$E_{5'} = -\frac{e^2}{8a_0} + \frac{1}{2}(-(a+b) - \sqrt{9a^2 - 2ab + b^2})$$

$$E_{6'} = -\frac{e^2}{8a_0} + \frac{e^2}{8a_0}\frac{\alpha^2}{12} - 2\mu_B B$$

$$E_{7'} = -\frac{e^2}{8a_0} + \mu_B B$$

$$E_{8'} = -\frac{e^2}{8a_0} - \mu_B B$$

A plot of

$$\left(E + \frac{e^2}{8a_0}\right) \times 10^5\, eV \text{ versus } \log_e(B(gauss))$$

looks like the figure below.

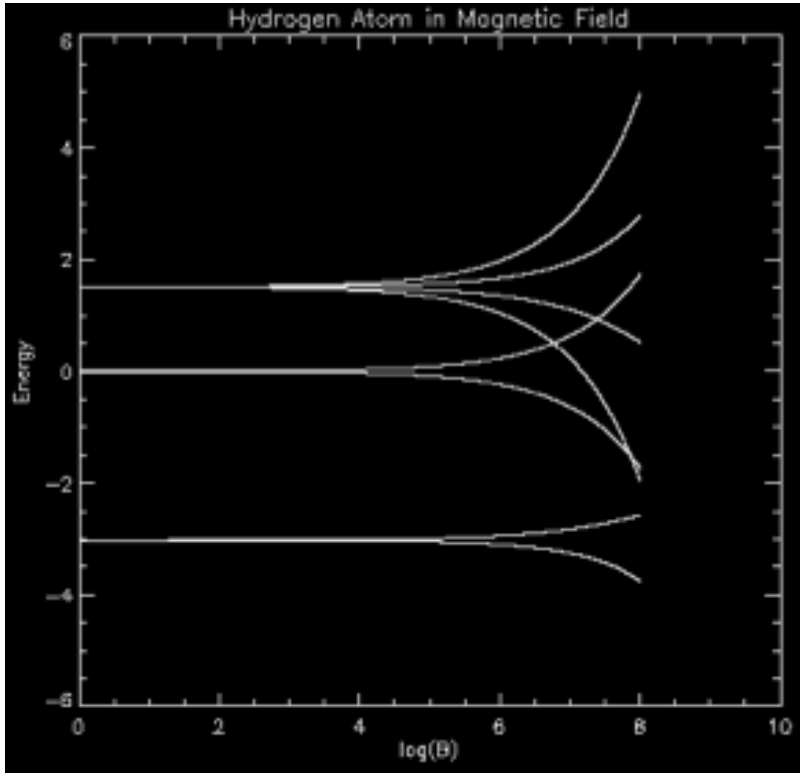

Figure: Hydrogen Atom In a Magnetic Field - Zeeman Effect

This plot for fields below 400 gauss ($\log_e(B) \approx 6$) shows the characteristic level structure of the Zeeman effect.

The very large magnetic field Paschen-Bach effect is illustrated in Figure 8.3 below.

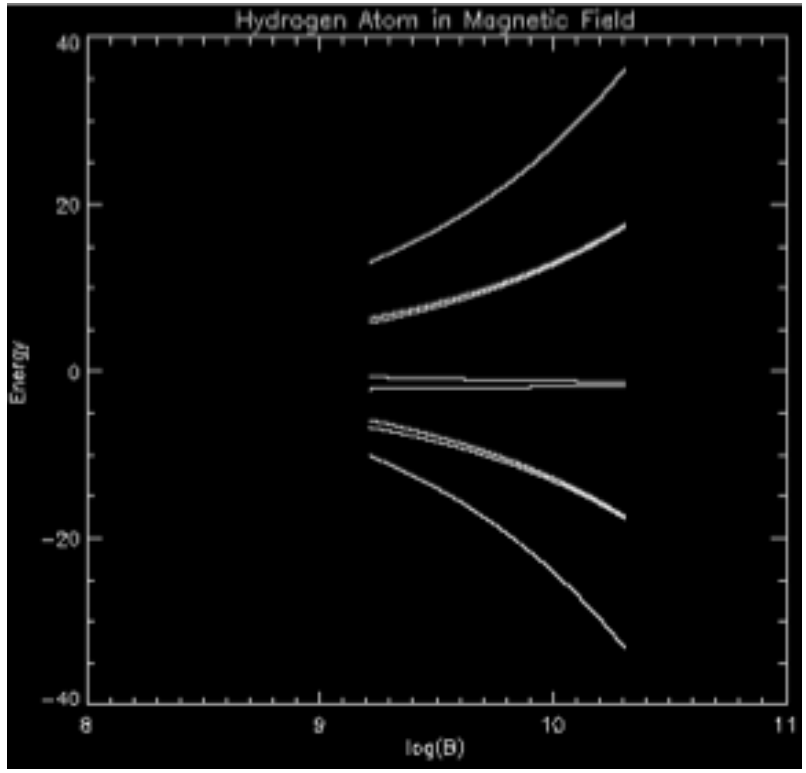

Figure: Hydrogen Atom In a Magnetic Field - Paschen-Bach Effect

Notice the equally-spaced level signature of the Paschen-Bach effect.

We now define some notation that will be important later as we study atomic spectra. For small magnetic fields we found that the approximate state vectors are the $|n\ell sjm_j\rangle$ states. The energy levels including spin-orbit effects are

$$\begin{aligned} E_n &= E_n^{(0)} + \Delta E_{so} \\ &= -\frac{Ze^2}{2a_0 n^2} \\ &\quad + \frac{Z^2\alpha^2 \left|E_n^{(0)}\right|}{n\ell(2\ell+1)(\ell+1)}(1-\delta_{\ell,0})\left(\ell\delta_{j,\ell+1/2} - (\ell+1)\delta_{j,\ell-1/2}\right) \end{aligned} \tag{8.371}$$

We define a spectroscopic notation to label the energy levels using the scheme shown below:

$$|n\ell sjm_j\rangle \to n^{2S+1}L(symbol)_J \tag{8.372}$$

so that

$$\left|21\frac{1}{2}\frac{3}{2}m_j\right\rangle \to 2^2P_{\frac{3}{2}}\ ,\ \left|21\frac{1}{2}\frac{1}{2}m_j\right\rangle \to 2^2P_{\frac{1}{2}}\ ,\ \left|20\frac{1}{2}\frac{1}{2}m_j\right\rangle \to 2^2S_{\frac{1}{2}}$$

The $L(symbols)$ are defined by the table below.

L	0	1	2	3
Symbol	S	P	D	F

Table: Spectroscopic Labels

The energy level diagram for $n = 1$ and $n = 2$ is shown in the figure below.

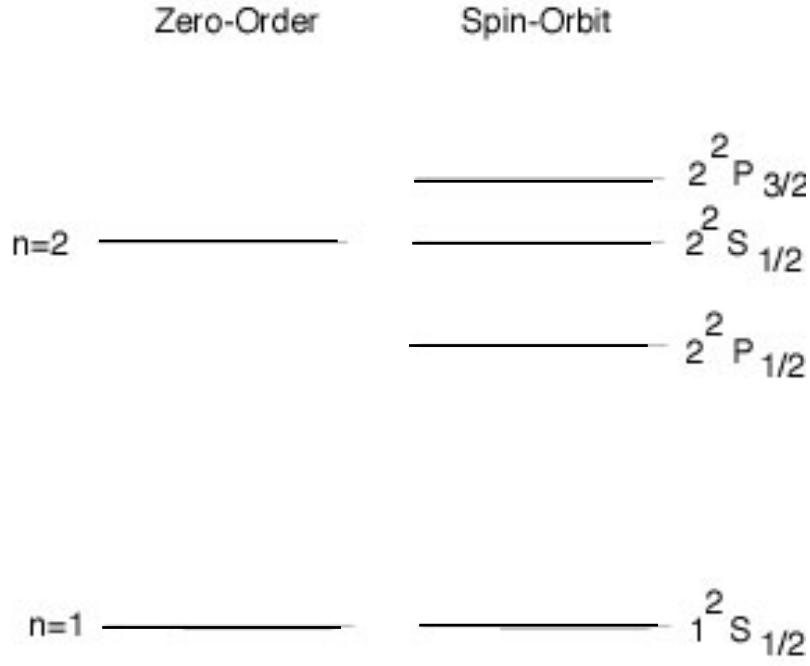

Figure: Spin-Orbit Energy Levels

Earlier we calculated the relativistic correction and found that it was the same order of magnitude as the spin-orbit correction for hydrogen. We found

$$\Delta E_{rel} = -\frac{Z^2\alpha^2 \left|E_n^{(0)}\right|}{n}\left(\frac{2}{2\ell+1} - \frac{3}{4n}\right) \tag{8.373}$$

Combining these two corrections we have

$$\Delta E_{fs} = \Delta E_{so} + \Delta E_{rel} = -\frac{Z^2\alpha^2 \left|E_n^{(0)}\right|}{n}\left(\frac{1}{j+\frac{1}{2}} - \frac{3}{4n}\right) \quad j = \ell \pm \frac{1}{2} \tag{8.374}$$

which is independent of ℓ. This changes the energy level structure to that shown in the figure below.

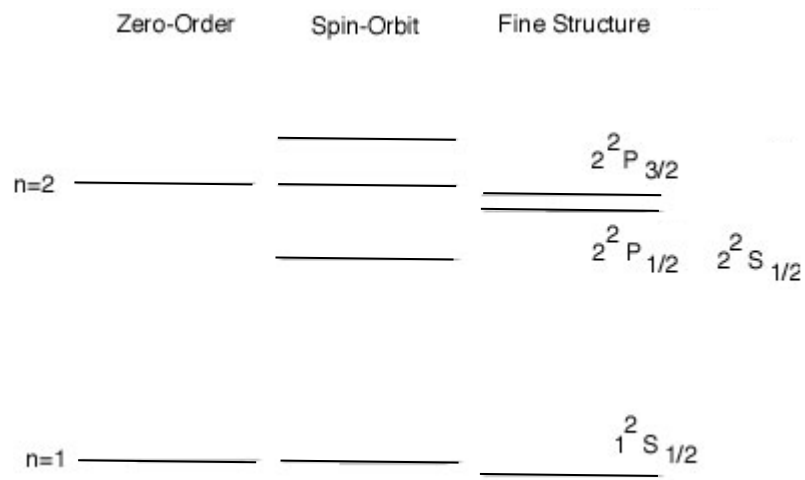

Figure: Fine Structure Energy Levels

The observed spectral lines result from an electron making a transition between these levels. We will discuss this topic later.

8.6.5 **Stark Effect**

When a hydrogen atom is placed in an external electric field $\vec{\mathcal{E}}_0$, the potential energy of the proton and the electron is given by

$$V_{dipole}(\vec{r}_e, \vec{r}_p) = -e\vec{\mathcal{E}}_0 \cdot \vec{r}_p + e\vec{\mathcal{E}}_0 \cdot \vec{r}_e$$
$$= e\mathcal{E}_0(z_e - z_p) = e\mathcal{E}_0 z \tag{8.375}$$

where

$$z = z_e - z_p = z_{relative} \tag{8.376}$$

Therefore, we consider the Hamiltonian

$$\hat{H} = \hat{H}_0 + \hat{H}_{dipole} \tag{8.377}$$

where

$$\hat{H}_0 = \frac{\vec{p}^2_{op}}{2m} - e^2 \left(\frac{1}{r}\right)_{op} \tag{8.378}$$

$$\hat{H}_{dipole} = e\mathcal{E}_0 z_{op} \tag{8.379}$$

For weak electric fields, we can apply perturbation theory (we ignore spin in this calculation). First, we apply perturbation theory to the $n = 1$ ground state of the hydrogen atom.

For the ground state, the wave function is $\psi_{100}(\vec{r})$ and the first-order correction to the energy is

$$
\begin{aligned}
E_1^{(1)} &= \langle 100| \, e\mathcal{E}_0 z_{op} \, |100\rangle \\
&= e\mathcal{E}_0 \int d^3\vec{r} d^3\vec{r}' \, \langle 100 \mid \vec{r}\rangle \, \langle \vec{r}| \, z_{op} \, \left|\vec{r}'\right\rangle \left\langle \vec{r}' \mid 100\right\rangle \\
&= e\mathcal{E}_0 \int d^3\vec{r} d^3\vec{r}' \, z\psi_{100}^*(\vec{r}) \left\langle \vec{r} \mid \vec{r}'\right\rangle \psi_{100}(\vec{r}') \\
&= e\mathcal{E}_0 \int d^3\vec{r} d^3\vec{r}' \, z\psi_{100}^*(\vec{r}) \delta(\vec{r} - \vec{r}') \psi_{100}(\vec{r}') \\
&= e\mathcal{E}_0 \int d^3\vec{r} z \, |\psi_{100}(\vec{r})|^2
\end{aligned} \tag{8.380}
$$

This equals zero since the integrand is the product of an even and odd functions. Thus, the first-order correction is zero for the ground state.

The second-order correction is given by non-degenerate perturbation theory as

$$
E_1^{(2)} = \sum_{n=2}^{\infty} \sum_{\ell=0}^{n-1} \sum_{m=-\ell}^{\ell} \frac{\left|\langle n\ell m| \, e\mathcal{E}_0 z_{op} \, |100\rangle\right|^2}{E_1^{(0)} - E_m^{(0)}} \tag{8.381}
$$

Using $z = r\cos\theta$ we have

$$\langle n\ell m|\, z_{op}\, |100\rangle = \int d^3\vec{r}\,[R_{n\ell}(r)Y^*_{\ell m}(\theta,\phi)]\,[r\cos\theta]\,R_{10}(r)Y_{00}(\theta,\phi) \tag{8.382}$$

Now

$$Y_{00} = \frac{1}{\sqrt{4\pi}} \text{ and } z = \sqrt{\frac{4\pi}{3}}Y_{10} \tag{8.383}$$

Therefore,

$$\langle n\ell m|\, z_{op}\, |100\rangle = \int r^3 dr R_{n\ell}(r)R_{10}(r)\frac{1}{\sqrt{3}}\int d\Omega Y^*_{\ell m}(\theta,\phi)Y_{10}(\theta,\phi) \tag{8.384}$$

Now

$$\int d\Omega Y^*_{\ell m}(\theta,\phi)Y_{10}(\theta,\phi) = \delta_{\ell,1}\delta_{m,0} \tag{8.385}$$

by the orthonormality of the $(\vec{L}^2_{op}, \hat{L}_z)$ eigenfunctions. Therefore,

$$\langle n\ell m|\, z_{op}\, |100\rangle = \frac{1}{\sqrt{3}}\delta_{\ell,1}\delta_{m,0}\int_0^\infty r^3 dr R_{n1}(r)R_{10}(r) \qquad (8.386)$$

and

$$|\langle n10|\, z_{op}\, |100\rangle|^2 = \frac{1}{3}\frac{2^8 n^7 (n-1)^{2n-5}}{(n+1)^{2n+5}}a_0^2 = \beta(n)a_0^2 \qquad (8.387)$$

Finally,

$$E_1^{(2)} = (e\mathcal{E}_0 a_0)^2 \sum_{n=2}^{\infty}\frac{\beta(n)}{\frac{e^2}{2a_0}\left(1-\frac{1}{n^2}\right)} = -2F\mathcal{E}_0^2 a_0^3 \qquad (8.388)$$

where

$$F = \sum_{n=2}^{\infty}\frac{n^2\beta(n)}{(n^2-1)} \approx 1.125 \qquad (8.389)$$

Therefore, the ground state exhibits a *quadratic* Stark effect.

The $n = 2$ level, which is the first excited state of hydrogen, has 4 degenerate states.

$$n = 2 \to \ell = 0 \to \psi_{200} = \psi_1$$
$$\to \ell = 1 \to m = \begin{cases} 1 \to \psi_{211} = \psi_2 \\ 0 \to \psi_{210} = \psi_3 \\ -1 \to \psi_{21-1} = \psi_4 \end{cases}$$

We must use degenerate perturbation theory. We construct the 4×4 $\langle e\mathcal{E}_0 z_{op} \rangle$ matrix and then diagonalize it. We have

$$\langle e\mathcal{E}_0 z_{op} \rangle = e\mathcal{E}_0 \begin{pmatrix} \langle 1| z_{op} |1\rangle & \langle 1| z_{op} |2\rangle & \langle 1| z_{op} |3\rangle & \langle 1| z_{op} |4\rangle \\ \langle 2| z_{op} |1\rangle & \langle 2| z_{op} |2\rangle & \langle 2| z_{op} |3\rangle & \langle 2| z_{op} |4\rangle \\ \langle 3| z_{op} |1\rangle & \langle 3| z_{op} |2\rangle & \langle 3| z_{op} |3\rangle & \langle 3| z_{op} |4\rangle \\ \langle 4| z_{op} |1\rangle & \langle 4| z_{op} |2\rangle & \langle 4| z_{op} |3\rangle & \langle 4| z_{op} |4\rangle \end{pmatrix} \tag{8.390}$$

Now z has no φ dependence and therefore,

$$\langle j| z_{op} |k\rangle = 0 \text{ if } m_j \neq m_k \tag{8.391}$$

Thus,

$$\begin{aligned}
&\langle 1|\, z_{op}\, |2\rangle = 0 = \langle 1|\, z_{op}\, |4\rangle \\
&\langle 2|\, z_{op}\, |1\rangle = 0 = \langle 2|\, z_{op}\, |3\rangle = \langle 2|\, z_{op}\, |4\rangle \\
&\langle 3|\, z_{op}\, |2\rangle = 0 = \langle 3|\, z_{op}\, |4\rangle \\
&\langle 4|\, z_{op}\, |1\rangle = 0 = \langle 4|\, z_{op}\, |2\rangle = \langle 4|\, z_{op}\, |3\rangle
\end{aligned}$$

and the matrix becomes

$$\langle e\mathcal{E}_0 z_{op}\rangle = e\mathcal{E}_0 \begin{pmatrix} \langle 1|\, z_{op}\, |1\rangle & 0 & \langle 1|\, z_{op}\, |3\rangle & 0 \\ 0 & \langle 2|\, z_{op}\, |2\rangle & 0 & 0 \\ \langle 3|\, z_{op}\, |1\rangle & 0 & \langle 3|\, z_{op}\, |3\rangle & 0 \\ 0 & 0 & 0 & \langle 4|\, z_{op}\, |4\rangle \end{pmatrix} \tag{8.392}$$

We also have

$$\langle 1|\, z_{op}\, |1\rangle = 0 = \langle 2|\, z_{op}\, |2\rangle = \langle 3|\, z_{op}\, |3\rangle = \langle 4|\, z_{op}\, |4\rangle \tag{8.393}$$

since these integrands involve the product of even and odd functions.

Finding out which matrix elements are equal to zero without actually evaluating the integrals corresponds to finding what are called *selection rules*. We will elaborate on the idea of selection rules in the next section on the Van der Waal's interaction.

Thus, the matrix finally becomes (after relabeling the rows and columns)

$$\langle e\mathcal{E}_0 z_{op}\rangle = e\mathcal{E}_0 \begin{pmatrix} 0 & \langle 1|\, z_{op}\, |3\rangle & 0 & 0 \\ \langle 3|\, z_{op}\, |1\rangle & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \tag{8.394}$$

where

$$\langle 1|\, z_{op}\, |3\rangle = \langle 3|\, z_{op}\, |1\rangle = \int \psi_{200}(\vec{r}) z \psi_{210}(\vec{r}) d^3\vec{r} = -3e\mathcal{E}_0 a_0 \tag{8.395}$$

Diagonalizing the 2×2 submatrix gives eigenvalues $\pm 3e\mathcal{E}_0 a_0$. The first-order energies and new zero-order wave functions are

$$\psi_{211}(\vec{r}) \rightarrow E_{211} = E_2^{(0)} \text{ remains degenerate} \tag{8.396}$$

$$\psi_{21-1}(\vec{r}) \rightarrow E_{21-1} = E_2^{(0)} \text{ remains degenerate} \tag{8.397}$$

$$\psi_{(+)}(\vec{r}) = \frac{1}{\sqrt{2}} \left(\psi_{200}(\vec{r}) - \psi_{210}(\vec{r})\right) \rightarrow E_+ = E_2^{(0)} + 3e\mathcal{E}_0 a_0$$

$$\psi_{(-)}(\vec{r}) = \frac{1}{\sqrt{2}} \left(\psi_{200}(\vec{r}) + \psi_{210}(\vec{r})\right) \rightarrow E_- = E_2^{(0)} - 3e\mathcal{E}_0 a_0$$

The degeneracy is broken for the $m = 0$ levels and we see a *linear* Stark effect. The linear Stark effect only appears for degenerate levels.

8.6.6 **Van der Waal's Interaction**

We now consider a system consisting of two widely separated atoms. In particular, we consider the interaction between two hydrogen atoms, where we treat the two protons as fixed point charges separated by a vector $\vec{R}$ and we define

$$\vec{r}_1 = \text{ vector from first proton to its electron}$$
$$\vec{r}_2 = \text{ vector from second proton to its electron}$$

as shown in Figure 8.6 below.

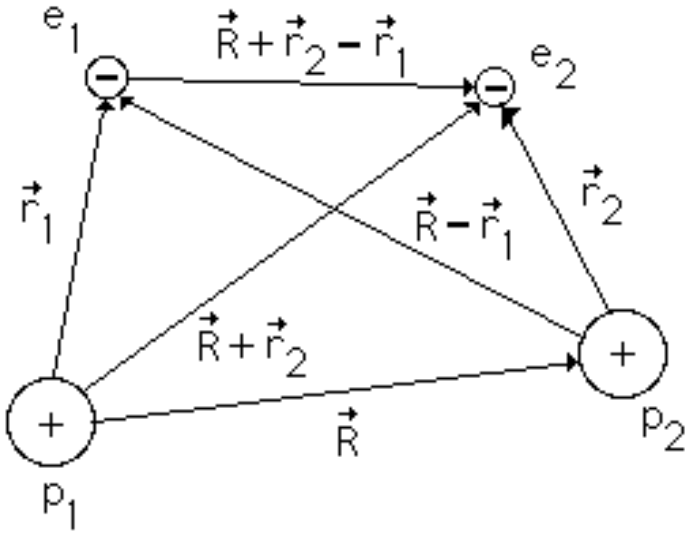

Figure: Van der Wall's System

The Hamiltonian is given by

$$\hat{H} = \hat{H}_0 + V \tag{8.398}$$

where

$$\hat{H}_0 = \frac{\vec{p}_{1,op}^2}{2\mu} - \frac{e^2}{r_1} + \frac{\vec{p}_{2,op}^2}{2\mu} - \frac{e^2}{r_2} = \text{ 2 non-interacting hydrogen atoms} \tag{8.399}$$

and

$$\begin{aligned} \hat{V} &= \text{ rest of the Coulomb interactions} \\ &= V_{p_1p_2} + V_{e_1e_2} + V_{e_1p_2} + V_{e_2p_1} \\ &= e^2 \left(\frac{1}{R} + \frac{1}{\left|\vec{R} + \vec{r}_2 - \vec{r}_1\right|} - \frac{1}{\left|\vec{R} + \vec{r}_2\right|} - \frac{1}{\left|\vec{R} - \vec{r}_1\right|} \right) \end{aligned} \tag{8.400}$$

This is the perturbation potential.

We know the zero-order solution for two non-interacting hydrogen atoms. It is

$$\text{zero-order states}: \quad |n_1\ell_1 m_1\rangle\,|n_2\ell_2 m_2\rangle \tag{8.401}$$

with

$$\text{zero-order energies}: \quad E^{(0)}_{n_1 n_2} = -\frac{e^2}{2a_0}\left(\frac{1}{n_1^2}+\frac{1}{n_2^2}\right) \tag{8.402}$$

where

$$\hat{H}_0\,|n_1\ell_1 m_1\rangle\,|n_2\ell_2 m_2\rangle = -\frac{e^2}{2a_0}\left(\frac{1}{n_1^2}+\frac{1}{n_2^2}\right)|n_1\ell_1 m_1\rangle\,|n_2\ell_2 m_2\rangle \tag{8.403}$$

This expression for the perturbation potential is too complicated to calculate. We will need to make an approximation. We make the reasonable assumption that

$$R >> r_2 \text{ and } R >> r_1 \tag{8.404}$$

We have two useful mathematical results that we can apply. In general, we can write

$$\frac{1}{\left|\vec{R}+\vec{a}\right|} = \frac{1}{\left[R^2+2\vec{R}\cdot\vec{a}+a^2\right]^{1/2}} = \frac{1}{R}\left[1+\frac{2\vec{R}\cdot\vec{a}}{R^2}+\frac{a^2}{R^2}\right]^{-1/2} \tag{8.405}$$

and for small x we have

$$[1+x]^{-1/2} \approx 1-\frac{1}{2}x+\frac{3}{8}x^2-\frac{5}{16}x^3+..... \tag{8.406}$$

Using

$$x = \frac{2\vec{R}\cdot\vec{a}}{R^2}+\frac{a^2}{R^2} \tag{8.407}$$

we get the general result

$$\frac{1}{\left|\vec{R}+\vec{a}\right|} = \frac{1}{R}\left[1-\frac{1}{2}\left(\frac{2\vec{R}\cdot\vec{a}}{R^2}+\frac{a^2}{R^2}\right)+\frac{3}{8}\left(\frac{2\vec{R}\cdot\vec{a}}{R^2}+\frac{a^2}{R^2}\right)^2-.....\right]$$

$$= \frac{1}{R}\left[1-\frac{\vec{a}\cdot\vec{R}}{R^2}-\frac{1}{2}\frac{a^2}{R^2}+\frac{3}{2}\frac{\left(\vec{a}\cdot\vec{R}\right)^2}{R^4}+.....\right] \qquad (8.408)$$

Therefore, we have

$$\frac{1}{\left|\vec{R}+\vec{r}_2\right|} = \frac{1}{R}-\frac{\vec{r}_2\cdot\vec{R}}{R^3}-\frac{1}{2}\frac{r_2^2}{R^3}+\frac{3}{2}\frac{\left(\vec{r}_2\cdot\vec{R}\right)^2}{R^5} \qquad (8.409)$$

$$\frac{1}{\left|\vec{R}-\vec{r}_1\right|} = \frac{1}{R}+\frac{\vec{r}_1\cdot\vec{R}}{R^3}-\frac{1}{2}\frac{r_1^2}{R^3}+\frac{3}{2}\frac{\left(\vec{r}_1\cdot\vec{R}\right)^2}{R^5} \qquad (8.410)$$

$$\frac{1}{\left|\vec{R}+\vec{r}_2-\vec{r}_1\right|} = \frac{1}{R} - \frac{(\vec{r}_2-\vec{r}_1)\cdot\vec{R}}{R^2} - \frac{1}{2}\frac{(\vec{r}_2-\vec{r}_1)^2}{R^2} + \frac{3}{2}\frac{\left((\vec{r}_2-\vec{r}_1)\cdot\vec{R}\right)^2}{R^4}$$

$$= \frac{1}{R} - \frac{\vec{r}_2\cdot\vec{R}}{R^3} + \frac{\vec{r}_1\cdot\vec{R}}{R^3} - \frac{1}{2}\frac{r_1^2}{R^3} - \frac{1}{2}\frac{r_2^2}{R^3} + \frac{\vec{r}_1\cdot\vec{r}_2}{R^3}$$

$$+ \frac{3}{2}\frac{\left(\vec{r}_2\cdot\vec{R}\right)^2}{R^5} + \frac{3}{2}\frac{\left(\vec{r}_1\cdot\vec{R}\right)^2}{R^5} - 3\frac{\left((\vec{r}_1\cdot\vec{R})(\vec{r}_2\cdot\vec{R}\right)}{R^5} \tag{8.411}$$

Putting all this together we get

$$V = \frac{e^2}{R^3}\left[\vec{r}_1\cdot\vec{r}_2 - 3\frac{\left((\vec{r}_1\cdot\vec{R})(\vec{r}_2\cdot\vec{R}\right)}{R^2}\right] \tag{8.412}$$

Physically, this says that for large separations, the interaction between the atoms is the same as that between two dipoles $e\vec{r}_1$ and $e\vec{r}_2$ separated by $\vec{R}$.

To simplify the algebra, we now choose the vector $\vec{R}$ to lie along the $z-$axis

$$\vec{R} = R\hat{z} \tag{8.413}$$

which gives

$$\begin{aligned} V &= \frac{e^2}{R^3}\left[(x_1x_2 + y_1y_2 + z_1z_2) - 3\frac{z_1z_2R^2}{R^2}\right] \\ &= \frac{e^2}{R^3}(x_1x_2 + y_1y_2 - 2z_1z_2) \end{aligned} \tag{8.414}$$

We now specialize to consider the case $n_1 = n_2 = 2$. When $n = 2$, there are 4 electron states for each atom

$$\ell = 0, m = 0$$
$$\ell = 1, m = 1, 0, -1$$

Therefore, there are $16 = (4 \times 4)$ degenerate unperturbed zero-order states, each with energy

$$E_0 = -\frac{e^2}{8a_0} - \frac{e^2}{8a_0} = -\frac{e^2}{4a_0} \tag{8.415}$$

We use degenerate perturbation theory. To carry out degenerate perturbation theory, we must construct the 16×16 matrix representation of $\langle \hat{V} \rangle$ and diagonalize it to find the energies corrected to first-order.

The typical matrix element is (leaving off the n labels)

$$\begin{aligned}
&\langle \ell_1 m_1 \ell_2 m_2 | \, \hat{V} \, | \ell_1 m_1 \ell_2 m_2 \rangle \\
&\quad = \frac{e^2}{R^3} \langle \ell_1 m_1 | \, \hat{x}_1 \, | \ell_1 m_1 \rangle \langle \ell_2 m_2 | \, \hat{x}_2 \, | \ell_2 m_2 \rangle \\
&\qquad + \frac{e^2}{R^3} \langle \ell_1 m_1 | \, \hat{y}_1 \, | \ell_1 m_1 \rangle \langle \ell_2 m_2 | \, \hat{y}_2 \, | \ell_2 m_2 \rangle \\
&\qquad - 2\frac{e^2}{R^3} \langle \ell_1 m_1 | \, \hat{z}_1 \, | \ell_1 m_1 \rangle \langle \ell_2 m_2 | \, \hat{z}_2 \, | \ell_2 m_2 \rangle
\end{aligned} \tag{8.416}$$

We have

$$x = r\sin\theta\cos\phi = -r\sqrt{\frac{2\pi}{3}}\left(Y_{1,1} - Y_{1,-1}\right) \tag{8.417}$$

$$y = r\sin\theta\sin\phi = +ir\sqrt{\frac{2\pi}{3}}\left(Y_{1,1} + Y_{1,-1}\right) \tag{8.416}$$

$$z = r\cos\theta = r\sqrt{\frac{4\pi}{3}}Y_{10} \tag{8.419}$$

and

$$\langle n\ell m|\, x\, |n\ell' m'\rangle$$
$$= -\sqrt{\frac{2\pi}{3}} \left[\int_0^\infty r^3 R_{n\ell}(r) R_{n\ell'}(r) dr\right] \left[\int d\Omega Y^*_{\ell m}(Y_{1,1} - Y_{1,-1}) Y_{\ell' m'}\right] \tag{8.420}$$

$$\langle n\ell m|\, y\, |n\ell' m'\rangle$$
$$= i\sqrt{\frac{2\pi}{3}} \left[\int_0^\infty r^3 R_{n\ell}(r) R_{n\ell'}(r) dr\right] \left[\int d\Omega Y^*_{\ell m}(Y_{1,1} + Y_{1,-1}) Y_{\ell' m'}\right] \tag{8.421}$$

$$\langle n\ell m|\, z\, |n\ell' m'\rangle$$
$$= \sqrt{\frac{4\pi}{3}} \left[\int_0^\infty r^3 R_{n\ell}(r) R_{n\ell'}(r) dr\right] \left[\int d\Omega Y^*_{\ell m} Y_{10} Y_{\ell' m'}\right] \tag{8.422}$$

Now let us return to the subject of *selection rules*.

We will just begin the discussion here and then elaborate and finish it later when we cover the topic of time-dependent perturbation theory.

Consider the integrals involving the spherical harmonics above. We have

$$\int d\Omega Y^*_{\ell m} Y_{1m''} Y_{\ell' m'} = 0 \text{unless} \begin{cases} \ell + \ell' + 1 = even \\ m = m' + m'' \end{cases} \tag{8.423}$$

These rules follow from doing the integrations over the θ and φ variables.

In particular, when the perturbation involves x, y, or z we have

$$\begin{array}{cc} \text{for } x \text{ and } y & m = m' \pm 1 \\ \text{for } z & m = m' \end{array}$$

which is the so-called

$$\Delta m = \pm 1, 0 \tag{8.424}$$

selection rule for this type of perturbation.

In addition, we have the

$$\Delta \ell = \pm 1 \tag{8.425}$$

selection rule for this type of perturbation.

These two rules will enable us to say many matrix elements are equal to zero by inspection.

We can derive two more very useful selection rules as follows. We know that

$$\left[\hat{L}_i, r_j\right] = i\hbar\varepsilon_{ijk}r_k \tag{8.426}$$

This allows us to write (after much algebra)

$$\left[\hat{L}_z, \hat{V}\right] = \left[(\hat{L}_{1z} + \hat{L}_{2z}), \hat{V}\right] = \left[\hat{L}_{1z}, \hat{V}\right] + \left[\hat{L}_{2z}, \hat{V}\right]$$
$$= \frac{e^2}{R^3}\left[\hat{L}_{1z}, (x_1x_2 + y_1y_2 - 2z_1z_2)\right]$$
$$+ \frac{e^2}{R^3}\left[\hat{L}_{2z}, (x_1x_2 + y_1y_2 - 2z_1z_2)\right] = 0$$

This implies that $\left[\hat{L}_z, \hat{H}\right] = 0$ or that the $z-$component of the total angular momentum of the electrons is not changed by this perturbation (it is conserved).

This gives the selection rule

$$m_1 + m_2 = m_1' + m_2' \tag{8.427}$$

Summarizing the selection rules we have

$$\ell_1 + \ell_1' + 1 = even \qquad (= \text{reason b for a zero})$$
$$\ell_2 + \ell_2' + 1 = even \qquad (= \text{reason c for a zero})$$
$$m_1 - m_1' = \pm 1, 0 \qquad (= \text{reason d for a zero})$$
$$m_2 - m_2' = \pm 1, 0 \qquad (= \text{reason d for a zero})$$
$$m_1 + m_2 = m_1' + m_2' \qquad (= \text{reason a for a zero})$$

and we have also given reason labels for each.

The unperturbed states are (using the format $|\ell_1 m_1\rangle\,|\ell_2 m_2\rangle$ are

$$|1\rangle = |00\rangle\,|00\rangle\,,\ |2\rangle = |00\rangle\,|11\rangle\,,\ |3\rangle = |00\rangle\,|10\rangle\,,\ |4\rangle = |00\rangle\,|1,-1\rangle$$
$$|5\rangle = |11\rangle\,|00\rangle\,,\ |6\rangle = |11\rangle\,|11\rangle\,,\ |7\rangle = |11\rangle\,|10\rangle\,,\ |8\rangle = |11\rangle\,|1,-1\rangle$$
$$|9\rangle = |10\rangle\,|00\rangle\,,\ |10\rangle = |10\rangle\,|11\rangle\,,\ |11\rangle = |10\rangle\,|10\rangle\,,\ |12\rangle = |10\rangle\,|1,-1\rangle$$
$$|13\rangle = |1,-1\rangle\,|00\rangle\,,\ |14\rangle = |1,-1\rangle\,|11\rangle\,,\ |15\rangle = |1,-1\rangle\,|10\rangle\,,$$
$$|16\rangle = |1,-1\rangle\,|1,-1\rangle$$

The $\langle \hat{V} \rangle$ matrix looks like (using labels (VALUE) or (0=reason)) and labeling the rows/columns in order as

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16

0bc	0ba	0a	0a	0a	0a	0a	A	0a	0a	B	0a	0a	C	0a	0a
0ba	0bc	0a	0a	D	0a	0a	0a	0a	0a	0a	0a	0a	0a	0a	0a
0a	0a	0a	0bc	0a	0a	0a	0a	E	0a	0a	0a	0a	0a	0a	0a
0a	0a	0bc	0a	0a	0a	0a	0a	0a	0a	0a	0a	F	0a	0a	0a
0a	D	0a	0a	0bc	0a	0a	0a	0a	0a	0a	0a	0a	0a	0a	0a
0a	0a	0a	0a	0a	0bc	0a	0a	0a	0a	0a	0a	0a	0a	0a	0a
0a	0a	0a	0a	0a	0a	0bc	0a	0a	0a	0a	0a	0a	0a	0a	0a
A	0a	0a	0a	0a	0a	0a	0bc	0a	0a	0a	0a	0a	0a	0a	0a
0a	0a	E	0a	0a	0a	0a	0a	0bc	0a	0a	0a	0a	0a	0a	0a
0a	0a	0a	0a	0a	0a	0a	0a	0a	0bc	0a	0a	0a	0a	0a	0a
B	0a	0a	0a	0a	0a	0a	0a	0a	0a	0bc	0a	0a	0bc	0a	0a
0a	0a	0a	0a	0a	0a	0a	0a	0a	0a	0a	0bc	0a	0a	0a	0a
0a	0a	0a	F	0a	0a	0a	0a	0a	0a	0a	0a	0bc	0a	0a	0a
C	0a	0a	0a	0a	0a	0a	0a	0a	0a	0bc	0a	0a	0bc	0a	0a
0a	0a	0a	0a	0a	0a	0a	0a	0a	0a	0a	0a	0a	0a	0bc	0a
0a	0a	0a	0a	0a	0a	0a	0a	0a	0a	0a	0a	0a	0a	0a	0bc

Figure: $\langle \hat{V} \rangle$ matrix entries

There are only 12 nonzero elements(out of 256) and because the matrix is Hermitian we only have 6 to calculate (one side of the diagonal). It should now be clear why finding the relevant selection rules is so important!!!

The 10 nonzero elements are given by the expressions

$$A = C = \frac{e^2}{R^3} \langle 200| \langle 200| (x_1 x_2 + y_1 y_2) |211\rangle |21, -1\rangle \tag{8.428}$$

$$B = E = -2\frac{e^2}{R^3} \langle 200| \langle 200| z_1 z_2 |210\rangle |21, 0\rangle \tag{8.429}$$

$$D = \frac{e^2}{R^3} \langle 200| \langle 211| (x_1 x_2 + y_1 y_2) |211\rangle |200\rangle \tag{8.430}$$

$$F = \frac{e^2}{R^3} \langle 200| \langle 21, -1| (x_1 x_2 + y_1 y_2) |21, -1\rangle |200\rangle \tag{8.431}$$

If we define

$$\alpha = \sqrt{\frac{8\pi}{3}} \int_0^\infty r^3 R_{20} R_{21} dr \tag{8.432}$$

we have

we have

$$\langle 200|\, x\, |211\rangle = \frac{\alpha}{2} = -\langle 200|\, x\, |21,-1\rangle \tag{8.433}$$

$$\langle 200|\, y\, |211\rangle = i\frac{\alpha}{2} = \langle 200|\, x\, |21,-1\rangle \tag{8.434}$$

$$\langle 200|\, z\, |211\rangle = \frac{\sqrt{2}}{2}\alpha \tag{8.435}$$

and

$$A = C = \frac{E}{2} = \frac{B}{2} = -D = -F = -\frac{1}{2}\alpha^2\frac{e^2}{R^3} \tag{8.436}$$

Now we rearrange the row/column labels(the original choice was arbitrary) to create a *Jordan canonical form* with blocks on the diagonal. We choose

1 8 11 14 2 5 3 9 4 13 6 7 10 12 15 16

0	A	2A	A	0	0	0	0	0	0	0	0	0	0	0	0
A	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2A	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
A	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	-A	0	0	0	0	0	0	0	0	0	0
0	0	0	0	-A	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	2A	0	0	0	0	0	0	0	0
0	0	0	0	0	0	2A	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	-A	0	0	0	0	0
0	0	0	0	0	0	0	0	0	-A	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Table: Jordan Canonical Form

The is called the *block-diagonalized form.* We have one 4×4 and three 2×2 matrices to diagonalize. We get the eigenvalues

$$\begin{aligned} &4 \times 4 \to 0, 0, \pm\sqrt{6}A \qquad , \qquad 2 \times 2 \to \pm A \\ &2 \times 2 \to \pm 2A \qquad , \qquad 2 \times 2 \to \pm A \end{aligned}$$

Therefore, the energies correct to first-order are

$$E = \begin{cases} E_0 + \sqrt{6}A & \text{degeneracy} = 1 \\ E_0 + 2A & \text{degeneracy} = 1 \\ E_0 + A & \text{degeneracy} = 2 \\ E_0 & \text{degeneracy} = 8 \\ E_0 - A & \text{degeneracy} = 2 \\ E_0 - 2A & \text{degeneracy} = 1 \\ E_0 - \sqrt{6}A & \text{degeneracy} = 1 \end{cases} \tag{8.437}$$

That was a *real* problem!

8.7 **Variational Methods**

All perturbation methods rely on our ability to make the separation $\hat{H} = \hat{H}_0 + \hat{V}$ where $\hat{H}_0$ is solvable exactly and $\hat{V}$ is a small correction. The Rayleigh-Ritz variational method is not subject to any such restrictions. This method is based on the following mathematical results.

We can always write

$$\hat{H} = \hat{I}\hat{H} = \sum_N |N\rangle \langle N| \hat{H} = \sum_N E_n |N\rangle \langle N| \tag{8.438}$$

where

$$\hat{H} |N\rangle = E_n |N\rangle \tag{8.439}$$

This is just the spectral decomposition of $\hat{H}$ in terms of its eigenvectors and eigenvalues. Now, if we choose some arbitrary state vector $|\psi\rangle$ (called a *trial* vector), then have

$$\begin{aligned}
\langle\psi|\,\hat{H}\,|\psi\rangle &= \sum_N E_n \langle\psi \mid N\rangle \langle N \mid \psi\rangle \\
&\geq \sum_N E_0 \langle\psi \mid N\rangle \langle N \mid \psi\rangle = E_0 \sum_N \langle\psi \mid N\rangle \langle N \mid \psi\rangle \\
&\geq E_0 \langle\psi| \left(\sum_N |N\rangle \langle N| \right) |\psi\rangle = E_0 \langle\psi|\,\hat{I}\,|\psi\rangle = E_0 \langle\psi \mid \psi\rangle
\end{aligned} \tag{8.440}$$

or

$$\frac{\langle\psi|\,\hat{H}\,|\psi\rangle}{\langle\psi \mid \psi\rangle} \geq E_0 \tag{8.441}$$

for *any choice* of the trial vector $|\psi\rangle$, where E_0 is the ground state energy (the lowest energy). Equality holds only if $|\psi\rangle$ is the true ground state vector. This result says that

$$\frac{\langle\psi|\,\hat{H}\,|\psi\rangle}{\langle\psi \mid \psi\rangle} \text{ is an upper bound for } E_0 \tag{8.442}$$

Procedure

1. Pick a trial vector $|\psi\rangle$ that contains unknown parameters $\{\alpha_k\}$
2. Calculate
$$\frac{\langle\psi|\,\hat{H}\,|\psi\rangle}{\langle\psi\mid\psi\rangle} = E_0\left(\{\alpha_k\}\right) \tag{8.443}$$
3. Since $E_0\left(\{\alpha_k\}\right)$ is an upper bound, we then minimize it with respect to all of the parameters $\{\alpha_k\}$. This gives a least upper bound for that choice of the functional form for the trial vector.
4. We perform the minimization by setting
$$\frac{\partial E_0}{\partial \alpha_k} = 0 \text{ for all } k \tag{8.444}$$
5. The more complex the trial vector, i.e., the more parameters we incorporate allows us to better approximate the true functional form of the ground state vector and we will get closer and closer to the true energy value.

What about states other than the ground state? If the ground state has different symmetry properties than the first excited state, i.e.,

$$\begin{aligned} \text{ground state} &\rightarrow \ell = 0 \rightarrow \text{ contains } Y_{00} \\ 1^{st} \text{ excited state} &\rightarrow \ell = 1 \rightarrow \text{ contains } Y_{1m} \end{aligned}$$

then if we choose a trial vector with the symmetry of an $\ell = 0$ state we obtain an approximation to the ground state energy. If, however, we choose a trial vector with the symmetry of an $\ell = 1$ state, then we obtain an approximation to the first-excited state energy and so on.

In other words, the variational method always gives the least upper bound for the energy of the state with the *same symmetry* as the trial vector.

Example

Let us choose the harmonic oscillator Hamiltonian

$$\hat{H} = -\frac{\hbar^2}{2m}\frac{d^2}{dx^2} + \frac{1}{2}kx^2 \tag{8.445}$$

and a trial wave function

$$\psi(x,a) = \begin{cases} (a^2 - x^2)^2 & |x| < a \\ 0 & |x| \geq a \end{cases} \tag{8.4446}$$

where a is an unknown parameter. The variational principle says that

$$\frac{\langle\psi(a)|\,\hat{H}\,|\psi(a)\rangle}{\langle\psi(a)\mid\psi(a)\rangle} = E_0(a) \geq E_0 = \text{ true ground state energy}$$

We get a *best* value for this choice of trial function by minimizing with respect to a using

$$\frac{dE_0(a)}{da} = 0 \tag{8.447}$$

Now we need to calculate the integrals. We have for the denominator (the normalization integral)

$$\langle \psi(a) \mid \psi(a) \rangle = \int_{-a}^{a} \psi^2(x,a) dx = \int_{-a}^{a} (a^2 - x^2)^4 dx$$
$$= 2 \int_{0}^{a} (a^2 - x^2)^4 dx = \frac{336}{315} a^9 \qquad (8.448)$$

and for the numerator

$$\langle \psi(a) | \hat{H} | \psi(a) \rangle = -\frac{\hbar^2}{2m} 2 \int_{0}^{a} \psi(x,a) \frac{d^2 \psi(x,a)}{dx^2} dx + \frac{1}{2} k 2 \int_{0}^{a} \psi^2(x,a) x^2 dx$$
$$= -\frac{\hbar^2}{2m} \left(-\frac{80}{21} a^7 \right) + \frac{1}{2} k \left(\frac{336}{3465} a^{11} \right) \qquad (8.449)$$

Therefore,

$$E_0(a) = 1.786\frac{\hbar^2}{ma^2} + 0.045ka^2 \tag{8.450}$$

The minimum condition gives

$$a^2 = 6.300\left(\frac{\hbar^2}{mk}\right)^{1/2} \tag{8.451}$$

which then says that

$$0.566\hbar\omega \geq E_0 \tag{8.452}$$

The true value is $0.500\hbar\omega$. This is an excellent result considering that the trial function does not look at all like the correct ground-state wave function (it is a Gaussian function). This points out clearly how powerful the variational technique can be for many problems.

8.8, 2^{nd}-Order Degenerate Perturbation Theory

Suppose that the first order correction in perturbation theory is zero for some degenerate states so that the states remain degenerate. In this case, second-order degenerate perturbation theory must be applied. This is complex. We follow the derivation in Schiff (using our notation).

We assume that

$$\varepsilon_m = \varepsilon_k \quad , \quad V_{km} = 0 \text{ and } V_{kk} = V_{mm} \tag{8.453}$$

so that the degeneracy is not removed in first order.

We assume the equations (to second order)

$$\hat{H} = \hat{H}_0 + \hat{V} = \hat{H}_0 + g\hat{U} \tag{8.454}$$

$$|M\rangle = a_m |m\rangle + a_k |k\rangle + g \sum_{l \neq m,k} a_l^{(1)} |l\rangle + g^2 \sum_{l \neq m,k} a_l^{(2)} |l\rangle \tag{8.455}$$

$$|K\rangle = b_m |m\rangle + b_k |k\rangle + g \sum_{l \neq m,k} b_l^{(1)} |l\rangle + g^2 \sum_{l \neq m,k} b_l^{(2)} |l\rangle \tag{8.456}$$

$$|N\rangle = |n\rangle + g \sum_{l \neq m,k} a_{nl}^{(1)} |l\rangle + g^2 \sum_{l \neq m,k} a_{nl}^{(2)} |l\rangle \quad , \quad n \neq m, k \tag{8.457}$$

$$E_p = \varepsilon_p + g E_p^{(1)} + g^2 E_p^{(2)} \tag{8.458}$$

$$\hat{H} |M\rangle = (\hat{H}_0 + \hat{V}) |M\rangle = E_m |M\rangle = (\varepsilon_m + g E_m^{(1)} + g^2 E_m^{(2)}) |M\rangle \tag{8.459}$$

where the degenerate states are labeled by k, m and we assume the degenerate zero-order states are linear combinations of the two zero order degenerate states as shown.

A very important point is being made here.

If the system remains degenerate after first-order correction, then one must redevelop the equations for degenerate perturbation theory, using the correct zero order state vectors, i.e., linear combinations of the degenerate states. Even in the special case where the degenerate basis is uncoupled, i.e.,

$$\langle k| V |m\rangle = V_{km} = 0 \tag{8.460}$$

we *must not use* non-degenerate perturbation theory, *as one might do* if many textbooks are to be *believed*.

Remember that the primary object of degenerate perturbation theory is not only to eliminate the divergent terms, but to determine the appropriate linear combinations to use for zero-order state vectors. If we start with the wrong linear combinations, then we would have a *discontinuous change* of states in the limit of vanishing interactions, which says that the perturbation expansion is not valid.

We then have

$$
\begin{aligned}
& g a_m U \left|m\right\rangle + g a_k U \left|k\right\rangle + g \sum_{l \neq m,k} a_l^{(1)} \varepsilon_l \left|l\right\rangle \\
& \qquad + g^2 \sum_{l \neq m,k} a_l^{(2)} \varepsilon_l \left|l\right\rangle + g^2 \sum_{l \neq m,k} a_l^{(1)} U \left|l\right\rangle \\
& = \left(g E_m^{(1)} + g^2 E_m^{(2)} \right) \left(a_m \left|m\right\rangle + a_k \left|k\right\rangle \right) + g \sum_{l \neq m,k} a_l^{(1)} \varepsilon_m \left|l\right\rangle \\
& \qquad + g^2 \sum_{l \neq m,k} a_l^{(2)} \varepsilon_m \left|l\right\rangle + g^2 \sum_{l \neq m,k} a_l^{(1)} E_m^{(1)} \left|l\right\rangle
\end{aligned}
$$

Applying the linear functional $\left\langle m\right|$ we get

$$
g a_m U_{mm} + g^2 \sum_{l \neq m,k} a_l^{(1)} U_{ml} = g E_m^{(1)} a_m + g^2 E_m^{(2)} a_m \qquad (8.461)
$$

Applying the linear functional $\langle k|$ we get

$$g a_k U_{kk} + g^2 \sum_{l \neq m,k} a_l^{(1)} U_{kl} = g E_m^{(1)} a_k + g^2 E_m^{(2)} a_k \qquad (8.462)$$

Applying the linear functional $\langle n|$ we get

$$g a_m U_{nm} + g a_k U_{nk} + g \varepsilon_n a_n^{(1)} + g^2 \varepsilon_n a_n^{(2)} + g^2 \sum_{l \neq m,k} a_l^{(1)} U_{nl}$$

$$= g \varepsilon_m a_m^{(1)} + g^2 \varepsilon_m a_m^{(2)} + g^2 E_m^{(1)} a_m^{(2)} \qquad (8.463)$$

The first-order terms in (10.462) and (10.463) give the expected result

$$E_m^{(1)} = U_{mm} = U_{kk} \qquad (8.464)$$

The second-order terms in (8.462) and (8.463) give (equation (8.465))

$$\sum_{l \neq m,k} a_l^{(1)} U_{ml} = E_m^{(2)} a_m \quad , \quad \sum_{l \neq m,k} a_l^{(1)} U_{kl} = E_m^{(2)} a_k \qquad (8.465)$$

The first-order terms in (8.464) give an expression for $a_l^{(1)}$ when $n = l \neq m, k$ (8.467)

$$a_l^{(1)}(\varepsilon_m - \varepsilon_l) = a_m U_{lm} + a_k U_{lk} \tag{8.466}$$

Substituting (8.467) into (8.466) we get a pair of homogeneous algebraic equations for a_m and a_k.

These equations have a nonvanishing solution if and only if the determinant of the coefficients of a_m and a_k is zero or

$$\det \begin{bmatrix} \sum\limits_{l\neq m,k} \frac{U_{ml}U_{lm}}{\varepsilon_m - \varepsilon_l} - E_m^{(2)} & \sum\limits_{l\neq m,k} \frac{U_{ml}U_{lk}}{\varepsilon_m - \varepsilon_l} \\ \sum\limits_{l\neq m,k} \frac{U_{kl}U_{lm}}{\varepsilon_m - \varepsilon_l} & \sum\limits_{l\neq m,k} \frac{U_{kl}U_{lk}}{\varepsilon_m - \varepsilon_l} - E_m^{(2)} \end{bmatrix} = 0 \tag{8.467}$$

or

$$\det \begin{bmatrix} \sum\limits_{l\neq m,k} \frac{V_{ml}V_{lm}}{\varepsilon_m-\varepsilon_l} - g^2E_m^{(2)} & \sum\limits_{l\neq m,k} \frac{V_{ml}V_{lk}}{\varepsilon_m-\varepsilon_l} \\ \sum\limits_{l\neq m,k} \frac{V_{kl}V_{lm}}{\varepsilon_m-\varepsilon_l} & \sum\limits_{l\neq m,k} \frac{V_{kl}V_{lk}}{\varepsilon_m-\varepsilon_l} - g^2E_m^{(2)} \end{bmatrix} = 0 \tag{8.468}$$

The two roots of this equation are $g^2E_m^{(2)}$ and $g^2E_k^{(2)}$ and the two pairs of solutions of (10.466) are a_m,a_k and b_m,b_k. We thus obtain perturbed energy levels in which the degeneracy has been removed in second order and we also find the correct linear combinations of the unperturbed degenerate state vectors $|m\rangle$ and $|k\rangle$.

Example

This is a tricky problem because the degeneracy between the first and second state is not removed in first order degenerate perturbation theory.

A system that has three unperturbed states can be represented by the perturbed Hamiltonian matrix

$$\begin{aligned}\hat{H} &= \hat{H}_0 + \hat{V} \\ &= \begin{pmatrix} E_1 & 0 & 0 \\ 0 & E_1 & 0 \\ 0 & 0 & E_2 \end{pmatrix} + \begin{pmatrix} 0 & 0 & a \\ 0 & 0 & b \\ a^* & b^* & 0 \end{pmatrix} = \begin{pmatrix} E_1 & 0 & a \\ 0 & E_1 & b \\ a^* & b^* & E_2 \end{pmatrix}\end{aligned} \tag{8.469}$$

where $E_2 > E_1$. The quantities a and b are to be regarded as perturbations that are of same order and are small compared to $E_2 - E_1$. The procedure is:

1. Diagonalize the matrix to find the exact eigenvalues.
2. Use second-order nondegenerate perturbation theory to calculate the perturbed energies. Is this procedure correct?
3. Use second-order degenerate perturbation theory to calculate the perturbed energies.
4. Compare the three results obtained.

Solution - We have

$$\hat{H} = \hat{H}_0 + \hat{V}$$
$$= \begin{pmatrix} E_1 & 0 & 0 \\ 0 & E_1 & 0 \\ 0 & 0 & E_2 \end{pmatrix} + \begin{pmatrix} 0 & 0 & a \\ 0 & 0 & b \\ a^* & b^* & 0 \end{pmatrix} = \begin{pmatrix} E_1 & 0 & a \\ 0 & E_1 & b \\ a^* & b^* & E_2 \end{pmatrix}$$

with $E_2 > E_1$ and $E_2 - E_1 \gg a = b$.

1. For an exact solution we need to find the eigenvalues of

$$\begin{pmatrix} E_1 & 0 & a \\ 0 & E_1 & b \\ a^* & b^* & E_2 \end{pmatrix} \tag{8.470}$$

This leads to the characteristic equation

$$(E_1 - E)(E_1 - E)(E_2 - E) - (E_1 - E)\,|b|^2 - (E_1 - E)\,|a|^2 = 0 \tag{8.471}$$

This says that one of the eigenvalues is $E = E_1$ and the remaining quadratic equation is

$$E^2 - (E_1 + E_2)E + (E_1 E_2 - |b|^2 - |a|^2) = 0 \tag{8.472}$$

or the other two eigenvalues are

$$E = \frac{1}{2}\left((E_1 + E_2) \pm \sqrt{(E_1 + E_2)^2 - 4(E_1 E_2 - |b|^2 - |a|^2)}\right) \tag{8.473}$$

The exact energy values are

$$E_1$$

$$\frac{1}{2}\left((E_1+E_2)+\sqrt{(E_1+E_2)^2-4(E_1E_2-|b|^2-|a|^2)}\right)$$

$$\approx E_1+\frac{|a|^2+|b|^2}{E_1-E_2}$$

$$E=\frac{1}{2}\left((E_1+E_2)-\sqrt{(E_1+E_2)^2-4(E_1E_2-|b|^2-|a|^2)}\right)$$

$$\approx E_2-\frac{|a|^2+|b|^2}{E_1-E_2}$$

2. Apply non-degenerate second-order perturbation theory. The unperturbed system has

$$\hat{H}_0 = \begin{pmatrix} E_1 & 0 & 0 \\ 0 & E_1 & 0 \\ 0 & 0 & E_2 \end{pmatrix} \tag{8.474}$$

Since this is diagonal we have

$E_1^{(0)} = E_1 = E_2^{(0)}$, $E_3^{(0)} = E_2$ (levels 1 and 2 are degenerate)

and unperturbed eigenvectors

$$|1\rangle = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \quad , \quad |2\rangle = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \quad , \quad |3\rangle = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \tag{8.475}$$

The perturbation is (in the unperturbed basis)

$$\hat{V} = \begin{pmatrix} 0 & 0 & a \\ 0 & 0 & b \\ a^* & b^* & 0 \end{pmatrix} \tag{8.476}$$

Since the diagonal matrix elements of the perturbation are zero we have

$$E_1^{(1)} = E_2^{(1)} = E_3^{(1)} = 0 \text{ or no first - order corrections} \quad (8.477)$$

Thus, levels 1 and 2 remain degenerate.

If we formally (and incorrectly) apply non-degenerate second-order perturbation theory to this system we get

$$E_n^{(2)} = \sum_{m \neq n} \frac{|V_{mn}|^2}{E_n^{(0)} - E_m^{(0)}} \quad (8.478)$$

Now $V_{12} = 0, V_{13} = a, V_{23} = b$ and so we get

$$E_1^{(2)} = \sum_{m \neq 1} \frac{|V_{m1}|^2}{E_1^{(0)} - E_m^{(0)}}$$

$$= \frac{0}{0} + \frac{|V_{13}|^2}{E_1^{(0)} - E_3^{(0)}} \stackrel{?}{=} \frac{|a|^2}{E_1 - E_2} \text{ incorrect because of } 0/0 \text{ term}$$

$$E_2^{(2)} = \sum_{m \neq 2} \frac{|V_{m2}|^2}{E_2^{(0)} - E_m^{(0)}}$$

$$= \frac{0}{0} + \frac{|V_{23}|^2}{E_2^{(0)} - E_3^{(0)}} \stackrel{?}{=} \frac{|b|^2}{E_1 - E_2} \text{ incorrect because of } 0/0 \text{ term}$$

$$E_3^{(2)} = \sum_{m \neq 3} \frac{|V_{m3}|^2}{E_3^{(0)} - E_m^{(0)}}$$

$$= \frac{|V_{13}|^2}{E_3^{(0)} - E_1^{(0)}} + \frac{|V_{23}|^2}{E_3^{(0)} - E_2^{(0)}} = \frac{|a|^2 + |b|^2}{E_2 - E_1}$$

which agrees with the exact solution.

3. Now we apply second-order degenerate perturbation theory for the two degenerate levels. We have

$$\det \begin{vmatrix} \frac{|V_{13}|^2}{E_1^{(0)}-E_3^{(0)}} - E^{(2)} & \frac{V_{13}V_{32}}{E_1^{(0)}-E_3^{(0)}} \\ \frac{V_{23}V_{31}}{E_2^{(0)}-E_3^{(0)}} & \frac{|V_{23}|^2}{E_2^{(0)}-E_3^{(0)}} - E^{(2)} \end{vmatrix} \tag{8.479}$$

$$= \det \begin{vmatrix} \frac{|a|^2}{E_1-E_2} - E^{(2)} & \frac{ab^*}{E_1-E_2} \\ \frac{ba^*}{E_1-E_2} & \frac{|b|^2}{E_1-E_2} - E^{(2)} \end{vmatrix} = 0$$

$$(E^{(2)})^2 - \frac{|a|^2+|b|^2}{E_1-E_2}E^{(2)} + \frac{|a|^2|b|^2}{(E_1-E_2)^2} - \frac{|a|^2|b|^2}{(E_1-E_2)^2}$$

$$= E^{(2)}\left(E^{(2)} - \frac{|a|^2+|b|^2}{E_1-E_2}\right) = 0 \tag{8.480}$$

corresponding to

$$E^{(2)} = 0$$

$$E^{(2)} = \frac{|a|^2+|b|^2}{E_1-E_2}$$

so that to second-order we have

$$
\begin{aligned}
& E_1 \\
& E_1 + \frac{|a|^2 + |b|^2}{E_1 - E_2} \\
& E_2 - \frac{|a|^2 + |b|^2}{E_1 - E_2}
\end{aligned}
$$

which agrees with the exact solution.