Chapter 8

Time-Independent Perturbation Theory

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.

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

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

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

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

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

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

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

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

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 exin the small parameter g exist for all physically relevant quantities of the real system, i.e..

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

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

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

where the terms $|N^{(i)}\rangle$ and $E^{(i)}$ 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 $\langle N \mid N \rangle = 1$.

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

$$\langle n \mid N \rangle = 1$$

which will greatly simplify our derivations and subsequent calculations.

Substituting the power series expansion into the normalization condition we get

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

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

$$0 = g \left\langle n \mid N^{(1)} \right\rangle + g^2 \left\langle n \mid N^{(2)} \right\rangle + \ldots.$$

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

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

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

$$0 = \left(\hat{H}_0\,|n\rangle - \varepsilon_n\,|n\rangle\right)g^0 + \left(\hat{H}_0\left|N^{(1)}\right\rangle + \hat{U}\,|n\rangle - \varepsilon_n\left|N^{(1)}\right\rangle - E_n^{(1)}\,|n\rangle\right)g^1$$

$$+ .....$$

$$+ \left(\hat{H}_0\left|N^{(k)}\right\rangle + \hat{U}\left|N^{(k-1)}\right\rangle - \varepsilon_n\left|N^{(k)}\right\rangle - E_n^{(1)}\left|N^{(k-1)}\right\rangle - .... - E_n^{(k)}\,|n\rangle\right)g^k$$

$$+ .....$$

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)

$$0^{th} - order \qquad \hat{H}_0\,|n\rangle = \varepsilon_n\,|n\rangle$$

which is just our original assumption = unperturbed solution.

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

.....

.....

$$k^{th} - order \qquad \hat{H}_0\left|N^{(k)}\right\rangle + \hat{U}\left|N^{(k-1)}\right\rangle$$

$$= \varepsilon_n\left|N^{(k)}\right\rangle + E_n^{(1)}\left|N^{(k-1)}\right\rangle + ... + E_n^{(k)}\left|N^{(0)}\right\rangle$$

.....

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

Let us consider the 1^{st} – order equation.

If we apply the linear functional $\langle n|$ we get

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

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

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

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

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

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

$$\text{or in the } n^{th} \text{ unperturbed state}$$

Therefore, to first order in g we have

$$E_n = \varepsilon_n + gE_n^{(1)} = \varepsilon_n + g\langle n|\hat{U}|n\rangle$$

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

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

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

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, $|N^{(k)}\rangle$ 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).

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

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

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

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

We get

$$\begin{aligned}
&\langle m|\,\hat{H}_0\left|N^{(k)}\right\rangle + \langle m|\,\hat{U}\left|N^{(k-1)}\right\rangle \\
&\quad = \langle m|\,\varepsilon_n\left|N^{(k)}\right\rangle + \langle m|\,E_n^{(1)}\left|N^{(k-1)}\right\rangle + \ldots. + \langle m|\,E_n^{(k)}\left|N^{(0)}\right\rangle \\
&\varepsilon_m \left\langle m \mid N^{(k)}\right\rangle + \langle m|\,\hat{U}\left|N^{(k-1)}\right\rangle \\
&\quad = \varepsilon_n \left\langle m \mid N^{(k)}\right\rangle + E_n^{(1)} \left\langle m \mid N^{(k-1)}\right\rangle + \ldots\ldots + E_n^{(k)} \left\langle m \mid N^{(0)}\right\rangle
\end{aligned}$$

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

$$\begin{aligned}
&\left\langle m \mid N^{(k)}\right\rangle \\
&\quad = \frac{1}{\varepsilon_n - \varepsilon_m}\left(\langle m|\,\hat{U}\left|N^{(k-1)}\right\rangle - E_n^{(1)}\left\langle m \mid N^{(k-1)}\right\rangle - \ldots\ldots - E_n^{(k)}\left\langle m \mid N^{(0)}\right\rangle\right)
\end{aligned}$$

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

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

which gives

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

Therefore, to first order in g we have

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

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

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

Therefore, to second order ın g we have

$$\begin{aligned}E_n &= \varepsilon_n + gE_n^{(1)} + g^2E_n^{(2)} \\ &= \varepsilon_n + \langle n|\hat{V}|n\rangle + \sum_{m\neq n}\frac{\left|\langle n|\hat{V}|m\rangle\right|^2}{\varepsilon_n - \varepsilon_m}\end{aligned}$$

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

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

Therefore,

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

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

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

with

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

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

In standard operator notation

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

We now perturb the system with the potential energy term

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

Therefore

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

We still have a harmonic oscillator (with a changed spring constant).

This says that the new energies are given by

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

where

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

Therefore, the exact energies for the perturbed system are

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

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

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

which should correspond to the perturbation calculated energy calculated to 2nd order in perturbation theory.

We now do the perturbation calculation.

Our earlier derivation gives

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

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

where

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

We need to calculate this matrix element

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

where we have used

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

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

Therefore

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

and

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

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

which then gives

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

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

To calculate the new state vector to first order we need

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

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

which gives

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

What does the new ground state wave function look like?

We have

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

and

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

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

Now we found earlier that

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

and

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

which gives

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

$$= \left(\frac{m\omega}{\pi\hbar}\right)^{1/4} e^{-\frac{m\omega x^2}{2\hbar}}\left(1 + \frac{k'}{8k} - \frac{m\omega k'}{4\hbar k}x^2\right)$$

Since we are only changing the spring constant we should have

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

$$= \left(\frac{m\omega}{\pi\hbar}\right)^{1/4} \left(1+\frac{k'}{8k}\right) e^{-\frac{m\omega x^2}{2\hbar}\left(1+\frac{k'}{2k}\right)}$$

$$= \left(\frac{m\omega}{\pi\hbar}\right)^{1/4} \left(1+\frac{k'}{8k}\right)\left(1-\frac{m\omega k'}{4\hbar k}x^2\right) e^{-\frac{m\omega x^2}{2\hbar}}$$

$$= \left(\frac{m\omega}{\pi\hbar}\right)^{1/4} e^{-\frac{m\omega x^2}{2\hbar}} \left(1+\frac{k'}{8k}-\frac{m\omega k'}{4\hbar k}x^2\right)$$

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

$$\varepsilon_n = \varepsilon_m \text{ but } \langle m|V|n\rangle \neq 0$$

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

Degenerate Case

We handle this case as follows.

Suppose we have a group of *k* states

$$|n_1\rangle \ , \ |n_2\rangle \ , \ |n_3\rangle \ , \ \ldots \ , \ |n_k\rangle$$

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

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

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 $\varepsilon_{ni} - \varepsilon_{nj} \rightarrow 0$.

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

$$|n_1\rangle \ , \ |n_2\rangle \ , \ |n_3\rangle \ , \ \ldots \ , \ |n_k\rangle$$

is also an eigenstate of $\hat{H}_0$ with the same energy ε_{n1} .

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)

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

such that we have

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

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

$$|n_1\rangle \ , \ |n_2\rangle \ , \ |n_3\rangle \ , \ \ldots \ , \ |n_k\rangle$$

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

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

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

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

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 $E^{(1)}$ is equivalent to writing

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

or

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

or finally

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

All of these calculations take place within the k-dimensional degenerate subspace.

We will assume that the eigenvectors are normalized, which implies that

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

i.e., vectors are normalized to one.

Now consider another of the new vectors given by

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

We then have

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

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

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

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

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

Therefore the vectors

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

satisfy

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

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

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

$$|n_\alpha\rangle \ , \ |n_\beta\rangle \ , \ |n_\gamma\rangle \ , \ \ldots \ , \ |n_\kappa\rangle$$

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

Thus the group of states

$$|n_1\rangle \ , \ |n_2\rangle \ , \ |n_3\rangle \ , \ \ldots \ , \ |n_k\rangle$$

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

$$|n_\alpha\rangle \ , \ |n_\beta\rangle \ , \ |n_\gamma\rangle \ , \ \ldots \ , \ |n_\kappa\rangle$$

which are given to first order by

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

and the energy shift to second order is

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

where

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

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

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

We then have

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

where

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

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

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

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

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

and

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

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

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

Now

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

and

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

$$= \frac{\lambda\hbar}{2m\omega}\langle m,n|\hat{a}_x^+\hat{a}_y^+|0,0\rangle = \frac{\lambda\hbar}{2m\omega}\langle m,n \mid 1,1\rangle$$

$$= \frac{\lambda\hbar}{2m\omega}\delta_{m,1}\delta_{n,1}$$

Thus, the correction to first order is zero.

Calculating to second order we get

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

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

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

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

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

Now

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

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

$$= \frac{\lambda\hbar}{2m\omega}$$

Therefore the 2×2 submatrix is

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

This is simple to diagonalize.

We get these results

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

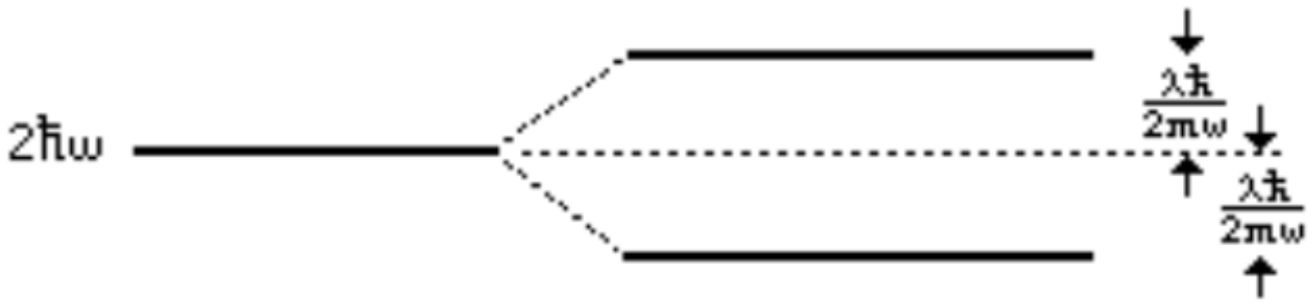
$$|b'\rangle = \frac{1}{\sqrt{2}}(|a\rangle - |b\rangle) = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \end{pmatrix} \rightarrow \text{ eigenvalue } = -\frac{\lambda\hbar}{2m\omega}$$

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

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

are the corresponding first order energy corrections.

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



Splitting of a degenerate Level

where

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

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

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

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

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

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

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

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

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

We define

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

and then we have

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

or

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

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.,

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

Each separate spin−1/2 system has the eigenvectors/eigenvalues

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

The corresponding direct-product states are

$$|{+}{+}\rangle \ , \ |{+}{-}\rangle \ , \ |{-}{+}\rangle \ , \ |{-}{-}\rangle$$

where the symbols mean

$$|{+}{-}\rangle = |+\rangle_1 |-\rangle_2$$

and so on.

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

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

They are given in terms of the direct product states by

$$|1,1\rangle = |{+}{+}\rangle = |1\rangle$$
$$|1,0\rangle = \frac{1}{\sqrt{2}} |{+}{-}\rangle + \frac{1}{\sqrt{2}} |{-}{+}\rangle = |2\rangle$$
$$|1,-1\rangle = |{-}{-}\rangle = |3\rangle$$
$$|0,0\rangle = \frac{1}{\sqrt{2}} |{+}{-}\rangle - \frac{1}{\sqrt{2}} |{-}{+}\rangle = |4\rangle$$

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

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

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

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

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

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

Now using

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

we do perturbation theory on these levels.

Nondegenerate Level

First order:

$$E_4^{(1)} = \langle 4| \hat{V} |4\rangle$$

$$= \left(\frac{1}{\sqrt{2}} \langle +-| + \frac{1}{\sqrt{2}} \langle -+| \right) (\alpha B \hat{S}_{1z} + \beta B \hat{S}_{2z}) \left(\frac{1}{\sqrt{2}} |+-\rangle - \frac{1}{\sqrt{2}} |-+\rangle \right) = 0$$

Second order:

$$E_4^{(2)} = \sum_{m\neq 4} \frac{\left|\langle m|\hat{V}|4\rangle\right|^2}{\varepsilon_4-\varepsilon_m} = \frac{\left|\langle ++|\,\alpha B\hat{S}_{1z}+\beta B\hat{S}_{2z}\left(\frac{1}{\sqrt{2}}|+-\rangle-\frac{1}{\sqrt{2}}|-+\rangle\right)\right|^2}{\varepsilon_4-\varepsilon_1}$$
$$+\frac{\left|\left(\frac{1}{\sqrt{2}}\langle +-|+\frac{1}{\sqrt{2}}\langle -+|\right)\alpha B\hat{S}_{1z}+\beta B\hat{S}_{2z}\left(\frac{1}{\sqrt{2}}|+-\rangle-\frac{1}{\sqrt{2}}|-+\rangle\right)\right|^2}{\varepsilon_4-\varepsilon_2}$$
$$+\frac{\left|\langle --|\,\alpha B\hat{S}_{1z}+\beta B\hat{S}_{2z}\left(\frac{1}{\sqrt{2}}|+-\rangle-\frac{1}{\sqrt{2}}|-+\rangle\right)\right|^2}{\varepsilon_4-\varepsilon_3}$$

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

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

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

Degenerate Level

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

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

which is already diagonal.

Since the diagonal elements are the first order energy corrections, we have (to first order)

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

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

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

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

$$|{+}{+}\rangle = |1\rangle \ , \ |{+}{-}\rangle = |2\rangle \ , \ |{-}{+}\rangle = |3\rangle \ , \ |{-}{-}\rangle = |4\rangle$$

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

$$\hat{H} = \alpha B\hat{S}_{1z} + \beta B\hat{S}_{2z} + \gamma\left(\hat{S}_{1z}\hat{S}_{2z} + \hat{S}_{1x}\hat{S}_{2x} + \hat{S}_{1y}\hat{S}_{2y}\right)$$

$$= \alpha B\hat{S}_{1z} + \beta B\hat{S}_{2z} + \gamma\left(\hat{S}_{1z}\hat{S}_{2z} + \frac{1}{2}\left(\hat{S}_{1+}\hat{S}_{2-} + \hat{S}_{1-}\hat{S}_{2+}\right)\right)$$

Construct the 4×4 $\hat{H}$ matrix

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

Using

$$\hat{S}_z|\pm\rangle = \pm\frac{\hbar}{2}|\pm\rangle$$

$$\hat{S}_+|+\rangle = 0 = \hat{S}_-|-\rangle$$

$$\hat{S}_+|-\rangle = \hbar|+\rangle \text{ and } \hat{S}_-|+\rangle = \hbar|-\rangle$$

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.

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

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

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

$$\langle m|\,\hat{H}\,|N\rangle = E_n\,\langle m \mid N\rangle = \langle m|\,\hat{H}_0\,|N\rangle + g\,\langle m|\,\hat{U}\,|N\rangle$$

$$E_n\,\langle m \mid N\rangle = \varepsilon_m\,\langle m \mid N\rangle + g\,\langle m|\,\hat{U}\,|N\rangle$$

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

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

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

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

Using these results we get

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

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)

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

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)

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

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 the general state expression we get

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

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

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

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

$$= \varepsilon_n + g \langle n | \hat{U} \left(|n\rangle + \sum_{m \neq n} |m\rangle \frac{g}{E_n - \varepsilon_m} \langle m | \hat{U} | n \rangle \right)$$

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

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

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

and

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

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

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

to get the characteristic equation

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

which has solutions

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

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

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

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

Now, BWPT gives

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

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

or

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

Rearranging we get

$$E_1^2 - (\varepsilon_1 + \varepsilon_2)E_1 + (\varepsilon_1\varepsilon_2 - |\alpha|^2) = 0$$

which is the same eigenvalue equation as the exact solution.

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

$$E_1 = \varepsilon + |\alpha|$$

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

On the other hand, RSPT gives to second-order

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

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

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

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

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

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

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

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

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

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

$$\hat{V}_1 = |m\rangle\langle m|\hat{V}|m\rangle\langle m| + |m\rangle\langle m|\hat{V}|n\rangle\langle n| + |n\rangle\langle n|\hat{V}|m\rangle\langle m| + |n\rangle\langle n|\hat{V}|n\rangle\langle n|$$

and $\hat{V}_2$ = the rest of the terms.

We then write

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

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.,

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

which gives

$$0 = \langle n|\hat{V}_2|n\rangle = \langle n|\hat{V}_2|m\rangle = \langle m|\hat{V}_2|n\rangle = \langle m|\hat{V}_2|m\rangle$$

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,

$$\hat{V}_1|i\rangle = 0$$

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

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

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)

$$\alpha|n\rangle + \beta|m\rangle$$

that diagonalize the 2×2 matrix.

We must have

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

Now

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

and similarly

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

Therefore, we get

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

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

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

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

These equations have two solutions, namely,

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

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

which then give the results

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

We then know all the eigenvectors/eigenvalues of $\hat{H}_0'$ (we know all of the unpeturbed 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

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

We will assume that the unperturbed states are nondegenerate.

Now define

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

and redefine

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

where

$$U = H_0 - E_{av} I + V$$

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

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

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.

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

$$|\psi\rangle = |n\ell m\rangle$$

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

$n = 1, 2, 3, \ldots\ldots\ldots$

$\ell = 0, 1, 2, \ldots\ldots, n-1$ for a given value of n

$m = -\ell, -\ell+1, \ldots\ldots \ell-1, \ell$ for a given value of ℓ

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

We found that

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

Therefore, each energy level had a degeneracy given by

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

$$= 2\frac{n(n-1)}{2} + n = n^2$$

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

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

by re-expressing that equation in the position representation.

The associated Schrodinger wave functions were given by the scalar product(linear functional) relation

$$\psi(\vec{r}) = \langle \vec{r} \mid \psi \rangle$$

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

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

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

The energy eigenvalue equation becomes

$$\begin{aligned} \hat{H}\,|\psi_{new}\rangle &= E\,|\psi_{new}\rangle \\ &= (((3-\text{space operators}))\,|\psi\rangle) \otimes (((\text{spin - dependent operators}))\,|s, m_s\rangle) \end{aligned}$$

and the corresponding wave function is

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

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.

stop

Spin-Orbit Interaction - Fine Structure

The proton in hydrogen generates an electric field

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

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

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

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

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

where

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

Substituting everything in we get

$$E = -\frac{e}{mc^2}\vec{S} \cdot \left(\vec{v} \times \vec{\mathcal{E}}\right) = -\frac{e}{mc^2}\vec{S} \cdot \left(\vec{v} \times \frac{e}{r^3}\vec{r}\right)$$

$$= \frac{1}{m^2c^2}\vec{S} \cdot \vec{L}\frac{e^2}{r^3}$$

Now

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

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

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

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

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

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

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

where

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

Now

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

which implies that

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

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.

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

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

We assume uniform circular motion for simplicity.

The orbital angular velocities measured by the observers are

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

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

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

But we also have

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

for circular motion.

Thus, we get

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

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

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

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

Another Relativity Correction

The correct relativistic kinetic energy term is

$$K = (\gamma - 1)mc^2 = \left(\frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} - 1 \right) mc^2$$

$$= \left(\left(1 + \frac{1}{2}\frac{v^2}{c^2} - \frac{1}{8}\frac{v^4}{c^4} + \ldots \right) - 1 \right) mc^2$$

$$= \left(\frac{1}{2}\frac{v^2}{c^2} - \frac{1}{8}\frac{v^4}{c^4} \right) mc^2 = \frac{\vec{p}_{op}^{\,2}}{2m} - \frac{\vec{p}_{op}^{\,4}}{8m^3c^2}$$

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

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

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

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

where

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

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

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

External Fields - Zeeman and Stark Effects; Hyperfine Structure

Zeeman Effect

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

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

as we derived earlier.

If we define

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

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

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

Thus, we must add a term of the form

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

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

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

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

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

and the Hamiltonian changes to

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

$$= \hat{H}_0 - \frac{e}{2mc}\left(\vec{p}_{op}\cdot\vec{A}(\vec{r}_{op}) + \vec{A}(\vec{r}_{op})\cdot\vec{p}_{op}\right) + \frac{e^2}{2mc^2}\vec{A}^2(\vec{r}_{op})$$

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

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

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

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

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

$$\nabla \times \vec{A} = -\frac{1}{2}\nabla \times \left(\vec{r} \times \vec{B}\right) = -\frac{1}{2}\sum_{ijk} \varepsilon_{ijk} \frac{\partial}{\partial x_j} \left(\vec{r} \times \vec{B}\right)_k \hat{e}_i$$

$$= -\frac{1}{2}\sum_{ijk} \varepsilon_{ijk} \frac{\partial}{\partial x_j} \left(\sum_{mn} \varepsilon_{kmn} x_m B_n\right) \hat{e}_i$$

$$= -\frac{1}{2}\sum_{ij}\sum_{mn} \left(\sum_{k} \varepsilon_{ijk}\varepsilon_{mnk}\right) \frac{\partial}{\partial x_j} \left(x_m B_n\right) \hat{e}_i$$

$$= -\frac{1}{2}\sum_{ij}\sum_{mn} \left(\delta_{im}\delta_{jn} - \delta_{in}\delta_{jm}\right) \frac{\partial}{\partial x_j} \left(x_m B_n\right) \hat{e}_i$$

$$= -\frac{1}{2}\sum_{ij} \left[\frac{\partial}{\partial x_j} \left(x_i B_j\right) \hat{e}_i - \frac{\partial}{\partial x_j} \left(x_j B_i\right) \hat{e}_i\right]$$

$$= -\frac{1}{2}\sum_{ij} \left[\frac{\partial x_i}{\partial x_j} B_j + x_i \frac{\partial B_j}{\partial x_j} - \frac{\partial x_j}{\partial x_j} B_i - x_j \frac{\partial B_i}{\partial x_j}\right] \hat{e}_i$$

Now

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

so we get

$$\nabla \times \vec{A} = -\frac{1}{2}\sum_{ij} \left[\delta_{ij} B_j - \delta_{jj} B_i\right] \hat{e}_i$$

$$= -\frac{1}{2}\left[\sum_{i} B_i \hat{e}_i - 3\sum_{i} B_i \hat{e}_i\right]$$

$$= -\frac{1}{2}\left[\vec{B} - 3\vec{B}\right] = \vec{B}$$

Therefore, we have

$$\vec{p}_{op}\cdot\vec{A}(\vec{r}_{op})+\vec{A}(\vec{r}_{op})\cdot\vec{p}_{op}=-\frac{1}{2}\left[\vec{p}_{op}\cdot\left(\vec{r}_{op}\times\vec{B}\right)+\left(\vec{r}_{op}\times\vec{B}\right)\cdot\vec{p}_{op}\right]$$

$$=-\frac{1}{2}\left[\sum_{ijk}\varepsilon_{ijk}B_k\left[\hat{p}_i\hat{x}_j+\hat{x}_j\hat{p}_i\right]\right]=-\frac{1}{2}\left[\sum_{ijk}\varepsilon_{ijk}B_k\left[2\hat{x}_j\hat{p}_i-i\hbar\delta_{ij}\right]\right]$$

$$=\sum_{ijk}\varepsilon_{kji}\hat{x}_j\hat{p}_iB_k=(\vec{r}_{op}\times\vec{p}_{op})\cdot\vec{B}=\vec{L}_{op}\cdot\vec{B}$$

which then gives

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

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.

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

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

where $\vec{r}$ is the position vector of the electron relative to the proton.

We then write the extra energy term as

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

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

The full Hamiltonian is then

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

where

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

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

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

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

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

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

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

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 ($\ell = 0$).

extra energy is then

$$\hat{H}_{hyperfine} = -\vec{M}_e \cdot \vec{B}$$

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

where

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

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

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.

Spin-Orbit, Relativity, Zeeman Effect in Hydrogen Atom

The Hamiltonian is

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

where

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

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

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

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

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

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

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

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

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

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

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

$$|n, \ell, s, j, m_j\rangle$$

We also have

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

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

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

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

$$|n, \ell, s, m_\ell, m_s\rangle$$

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

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

$$Ze = \text{ nucleus charge } (Z = 1 \text{ for hydrogen }) \, , \; a_0 = \text{ Bohr radius } = \frac{\hbar^2}{me^2}$$

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

and first few unperturbed wave functions are

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

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

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

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

We also have the relations below for the unperturbed states.

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

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

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

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

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

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

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

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

$$\begin{aligned}\hat{J}_z\left|n,\ell,s,m_\ell,m_s\right\rangle &= (\hat{L}_z+\hat{S}_z)\left|n,\ell,s,m_\ell,m_s\right\rangle \\ &= \hbar(m_\ell+m_s)\left|n,\ell,s,m_\ell,m_s\right\rangle \\ &= \hbar m_j\left|n,\ell,s,m_\ell,m_s\right\rangle\end{aligned}$$

Since the total angular momentum is given by

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

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

$$j = \ell + s, \ell + s - 1, ..., |\ell - s| + 1, |\ell - s|$$

and

$$m_j = j, j - 1, j - 2, ..., -j + 1, -j$$

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

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

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

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

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

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

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

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

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

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

and therefore

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

Therefore

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

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

$$E_{n\ell sjm_j} = E_n^{(0)} + \langle n\ell sjm_j | \hat{H}_{so} | n\ell sjm_j \rangle$$

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

We now evaluate

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

Now

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

which gives

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

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

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

and the corresponding states are shown in the Table 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

***n* = 1 level quantum numbers**

or

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

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

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

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

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

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

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

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

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

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

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

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

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	$\|2,0,1/2,0,-1/2\rangle_{mm}$

: $n = 2$ 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	$\|2,0,1/2,0,-1/2\rangle_{mm}$

$n = 2$ 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 $m_\text{j} = m_\ell + m_\text{s}$.

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

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

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

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

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

$$= \left(\hat{L}_- + \hat{S}_-\right)\left(\sqrt{\frac{2}{3}}\left|2,1,\frac{1}{2},0,\frac{1}{2}\right\rangle_{mm} + \sqrt{\frac{1}{3}}\left|2,1,\frac{1}{2},1,-\frac{1}{2}\right\rangle_{mm}\right)$$

$$= \hbar\sqrt{\frac{4}{3}}\left|2,1,\frac{1}{2},-1,\frac{1}{2}\right\rangle_{mm} + \hbar\sqrt{\frac{8}{3}}\left|2,1,\frac{1}{2},0,-\frac{1}{2}\right\rangle_{mm}$$

or

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

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

and finally

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

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

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

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

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

or it can be written as

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

Now we must have

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

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

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

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

The solution is

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

and therefore

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

In a similar manner, we find

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

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

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

They are

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

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

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 that in the $|n,\ell,s,j,m_j\rangle$ basis

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

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

or

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

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

Evaluating the integrals we get

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

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

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

where

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

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

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

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

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

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

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

This gives the correction

$$\Delta E_{rel} = -\frac{\hbar^4}{8m^3c^2}\int_0^\infty r^2 dr \psi^*_{n\ell m}(r)\nabla^4\psi_{n\ell m}(r)$$

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

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

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

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

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

where

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

Because of the delta function, a contribution

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

arises for $\ell = 0$ states only.

This is identical to the contribution $\langle \hat{H}_{\text{so}} + \hat{H}_{\text{rel}} \rangle$ 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

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

$$= \frac{e^2}{2m^2c^2}\left(\frac{1}{r^3}\right)\left(\hat{L}_z\hat{S}_z + \frac{1}{2}\left(\hat{L}_+\hat{S}_- + \hat{L}_-\hat{S}_+\right)\right)$$

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

$$|1\rangle = \left|2, 1, \frac{1}{2}, 1, \frac{1}{2}\right\rangle_{mm} \qquad m_j = \frac{3}{2}$$

$$|2\rangle = \left|2, 1, \frac{1}{2}, 1, -\frac{1}{2}\right\rangle_{mm} \qquad m_j = \frac{1}{2}$$

$$|3\rangle = \left|2, 1, \frac{1}{2}, 0, \frac{1}{2}\right\rangle_{mm} \qquad m_j = \frac{1}{2}$$

$$|4\rangle = \left|2, 1, \frac{1}{2}, 0, -\frac{1}{2}\right\rangle_{mm} \qquad m_j = -\frac{1}{2}$$

$$|5\rangle = \left|2, 1, \frac{1}{2}, -1, \frac{1}{2}\right\rangle_{mm} \qquad m_j = -\frac{1}{2}$$

$$|6\rangle = \left|2, 1, \frac{1}{2}, -1, -\frac{1}{2}\right\rangle_{mm} \qquad m_j = -\frac{3}{2}$$

$$|7\rangle = \left|2, 0, \frac{1}{2}, 0, \frac{1}{2}\right\rangle_{mm} \qquad m_j = \frac{1}{2}$$

$$|8\rangle = \left|2, 0, \frac{1}{2}, 0, -\frac{1}{2}\right\rangle_{mm} \qquad m_j = -\frac{1}{2}$$

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

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

We have marked the non-zero elements.

Using the operator properties derived earlier we get

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

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

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

This says that (only diagonal elements)

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

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 $E^{(1)}{}_{2'}$, $E^{(1)}{}_{3'}$, $E^{(1)}{}_{4'}$, $E^{(1)}{}_{5'}$, 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, namely,

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

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

The characteristic equation is

$$(-a - E)(-E) - 2a^2 = 0 = E^2 + aE - 2a^2$$

or

$$E_{2'}^{(1)} = a \text{ and } E_{3'}^{(1)} = -2a$$

Notice that these energies are

$$E_{2'}^{(1)} = \ell a \text{ and } E_{3'}^{(1)} = -(\ell + 1)a$$

as expected for the

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

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

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

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

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

or

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

and similarly

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

We then deal with the submatrix involving state $|4\rangle$ and $|5\rangle$ in the Table, namely

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

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

The characteristic equation is

$$(-a-E)(-E)-2a^2=0=E^2+aE-2a^2$$

or

$$E_{4'}^{(1)}=a \text{ and } E_{5'}^{(1)}=-2a$$

and the eigenvectors are

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

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

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

$$E_2^{(0)}+a \text{ for } 1,2',4',6 \rightarrow \text{ 4 - fold degenerate}$$

$$E_2^{(0)} \text{ for } 7,8 \rightarrow \text{ 2 - fold degenerate}$$

$$E_2^{(0)}-2a \text{ for } 3',5' \rightarrow \text{ 2 - fold degenerate}$$

Spin-Orbit and Arbitrary Magnetic Field

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

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

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

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

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

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

where

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

The $\langle\ \hat{L}_z + 2\hat{S}_z\rangle$ matrix is diagonal in this representation and its diagonal elements are given by m_ℓ + $2m_s$ 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

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

where $b = \mu_B B$.

The combined perturbation matrix $\langle \quad \hat{V} \quad \rangle$ 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

$\langle \hat{V} \rangle$ 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	$-\mu_B B$	$8'$

n = 2 energy corrections for small B

A little bit of study shows the general relation

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

where

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

This is called the *Zeeman effect*.

We can prove this result in general.

The general method uses the *Wigner-Eckart Theorem*.

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

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

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

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

The corresponding commutators are

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

which all follow from the original commutator for the arbitrary vector operator.

Now, by definition of the operators, we have

$$\hat{J}_0 |\lambda, j, m_j\rangle = \hbar m_j |\lambda, j, m_j\rangle = \langle j, m_j| \hat{J}_0 |j, m_j\rangle |\lambda, j, m_j\rangle$$
$$\hat{J}_\pm |\lambda, j, m_j\rangle = \hbar\sqrt{j(j+1) - m_j(m_j \pm 1)} |\lambda, j, m_j \pm 1\rangle$$
$$= \langle j, m_j \pm 1| \hat{J}_\pm |j, m_j\rangle |\lambda, j, m_j \pm 1\rangle$$
$$\langle \lambda, j, m_j| \hat{J}_\pm = \left(\hat{J}_\mp |\lambda, j, m_j\rangle\right)^+$$
$$= \langle \lambda, j, m_j \mp 1| \hbar\sqrt{j(j+1) - m_j(m_j \mp 1)}$$
$$= \langle \lambda, j, m_j \mp 1| \langle j, m_j| \hat{J}_\pm |j, m_j \mp 1\rangle$$

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

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

or

$$0=(m_j-m_j'\pm1)\hbar\left\langle\lambda',j,m_j'\right|\hat{A}_\pm\left|\lambda,j,m_j\right\rangle$$

This says that either $m'_{\mathrm{j}}=m_{\mathrm{j}}\pm1$ or the matrix element

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

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

$$\frac{\left\langle\lambda',j,m_j\pm1\right|\hat{A}_\pm\left|\lambda,j,m_j\right\rangle}{\left\langle j,m_j\pm1\right|\hat{J}_\pm\left|j,m_j\right\rangle}$$

Second, we have

$$\left\langle\lambda',j,m_j'\right|\left[\hat{A}_\pm,\hat{J}_\pm\right]\left|\lambda,j,m_j\right\rangle=0$$
$$\left\langle\lambda',j,m_j'\right|\hat{A}_\pm\hat{J}_\pm\left|\lambda,j,m_j\right\rangle=\left\langle\lambda',j,m_j'\right|\hat{J}_\pm\hat{A}_\pm\left|\lambda,j,m_j\right\rangle$$
$$\left\langle\lambda',j,m_j'\right|\hat{A}_\pm\left|\lambda,j,m_j\pm1\right\rangle\left\langle\lambda',j,m_j\pm1\right|\hat{J}_\pm\left|\lambda,j,m_j\right\rangle$$
$$=\left\langle\lambda',j,m_j'\mp1\right|\hat{A}_\pm\left|\lambda,j,m_j\right\rangle\left\langle\lambda,j,m_j\pm1\right|\hat{J}_\pm\left|\lambda,j,m_j'\mp1\right\rangle$$

Using the result from the first commutator this says that $m'_{\mathrm{j}}=m_{\mathrm{j}}\pm2$, which, in turn, implies that

$$\frac{\left\langle\lambda',j,m_j\pm2\right|\hat{A}_\pm\left|\lambda,j,m_j\pm1\right\rangle}{\left\langle j,m_j\pm2\right|\hat{J}_\pm\left|j,m_j\pm1\right\rangle}=\frac{\left\langle\lambda',j,m_j\pm1\right|\hat{A}_\pm\left|\lambda,j,m_j\right\rangle}{\left\langle j,m_j\pm1\right|\hat{J}_\pm\left|j,m_j\right\rangle}$$

This says that the proportionality constant is independent of m_j.

We define a new symbol for the proportionality constant

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

which gives the relation

$$\langle \lambda', j, m_j' | \hat{A}_{\pm} | \lambda, j, m_j \rangle = \langle \lambda', j || A || \lambda, j \rangle_{\pm} \langle j, m_j' | \hat{J}_{\pm} | j, m_j \rangle$$

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

$$\langle \lambda', j || A || \lambda, j \rangle_{+} = \langle \lambda', j || A || \lambda, j \rangle_{-}$$

We have

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

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

$$\begin{aligned} &\pm 2\hbar \langle \lambda', j, m_j' | \hat{A}_0 | \lambda, j, m_j \rangle \\ &\quad = \langle \lambda', j || A || \lambda, j \rangle_{\pm} [\langle j, m_j' | \hat{J}_{\pm} | j, m_j \mp 1 \rangle \langle j, m_j \mp 1 | \hat{J}_{\mp} | j, m_j \rangle \\ &\qquad - \langle j, m_j' \pm 1 | \hat{J}_{\pm} | j, m_j \rangle \langle j, m_j' | \hat{J}_{\mp} | j, m_j' \pm 1 \rangle] \end{aligned}$$

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

$$\pm 2\hbar \left\langle \lambda', j, m_j' \right| \hat{A}_0 \left| \lambda, j, m_j \right\rangle$$

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

$$= \pm 2\hbar m_j = \pm 2\hbar \left\langle \lambda', j, m_j' \right| \hat{J}_0 \left| \lambda, j, m_j \right\rangle$$

Putting it all together we get

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

Since no operator has a ± subscript, this also says that

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

and we finally have

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

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

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

$$\left\langle \lambda', j, m_j' \right| \vec{A}_{op} \cdot \vec{J}_{op} \left| \lambda, j, m_j \right\rangle = \left\langle \lambda', j \right| |A| \left| \lambda, j \right\rangle \left\langle j, m_j' \right| \vec{J}_{op}^2 \left| j, m_j \right\rangle$$

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

which says that the scalar product is diagonal in m_{j} .

This result follows directly from the Wigner-Eckart theorem

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

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

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

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

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

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

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

$$\left\langle \ell s j m_j \right| (\hat{L}_z + 2\hat{S}_z) \left| \ell s j m_j \right\rangle = \left\langle \ell s j m_j \right| (\hat{J}_z + \hat{S}_z) \left| \ell s j m_j \right\rangle$$

$$= \hbar m_j + \left\langle \ell s j m_j \right| \hat{S}_z \left| \ell s j m_j \right\rangle$$

Now the Wigner-Eckart theorem says that

$$\left\langle \ell s j m_j \right| \hat{S}_z \left| \ell s j m_j \right\rangle = \left\langle \ell s j \right| |S| \left| \ell s j \right\rangle \left\langle \ell s j m_j \right| \hat{J}_z \left| \ell s j m_j \right\rangle$$

$$= \hbar m_j \left\langle \ell s j \right| |S| \left| \ell s j \right\rangle$$

The scalar product matrix element formula gives

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

But we also have

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

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

or

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

and thus

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

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

Finally, we have

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

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

Paschen-Bach Effect

When B is large enough such that $\Delta E_{\text{Zeeman}} \gg \Delta E_{\text{so}}$, 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 $\mu_B = 5.7884 \times 10^{-9} \dfrac{eV}{gauss}$, $a_0 = 5.2918 \times 10^{-8}\, cm$, $e = 4.80 \times 10^{-10}\, esu$

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

Using our earlier results we then have

$$E_{1'} = -\frac{e^2}{8a_0} + \frac{e^2}{8a_0}\frac{\alpha^2}{12} + 2\mu_B B \qquad E_{5'} = -\frac{e^2}{8a_0} + \frac{1}{2}(-(a+b) - \sqrt{9a^2 - 2ab + b^2})$$

$$E_{2'} = -\frac{e^2}{8a_0} + \frac{1}{2}(-(a-b) + \sqrt{9a^2 + 2ab + b^2}) \qquad E_{6'} = -\frac{e^2}{8a_0} + \frac{e^2}{8a_0}\frac{\alpha^2}{12} - 2\mu_B B$$

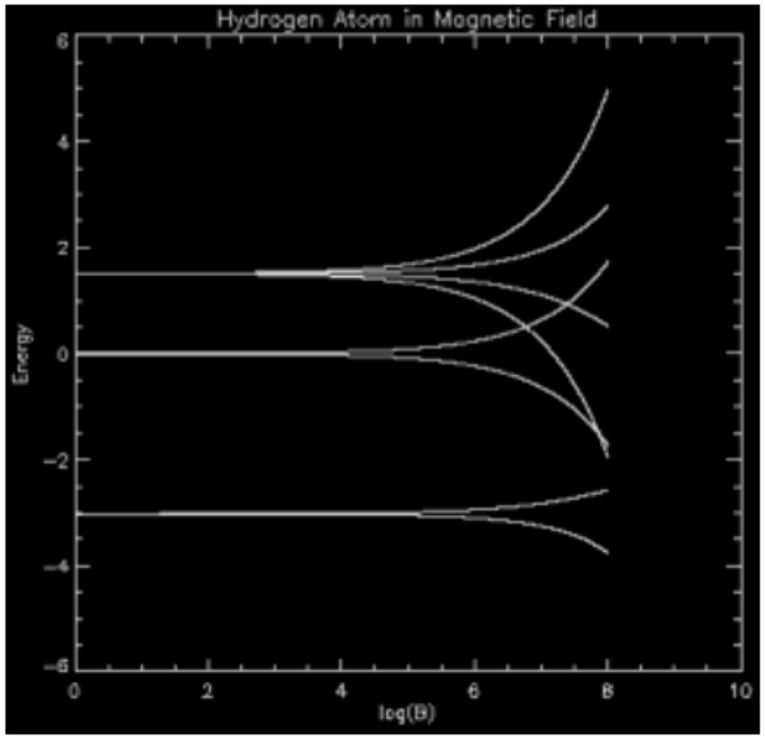
$$E_{3'} = -\frac{e^2}{8a_0} + \frac{1}{2}(-(a-b) - \sqrt{9a^2 + 2ab + b^2}) \qquad E_{7'} = -\frac{e^2}{8a_0} + \mu_B B$$

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

A plot of

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

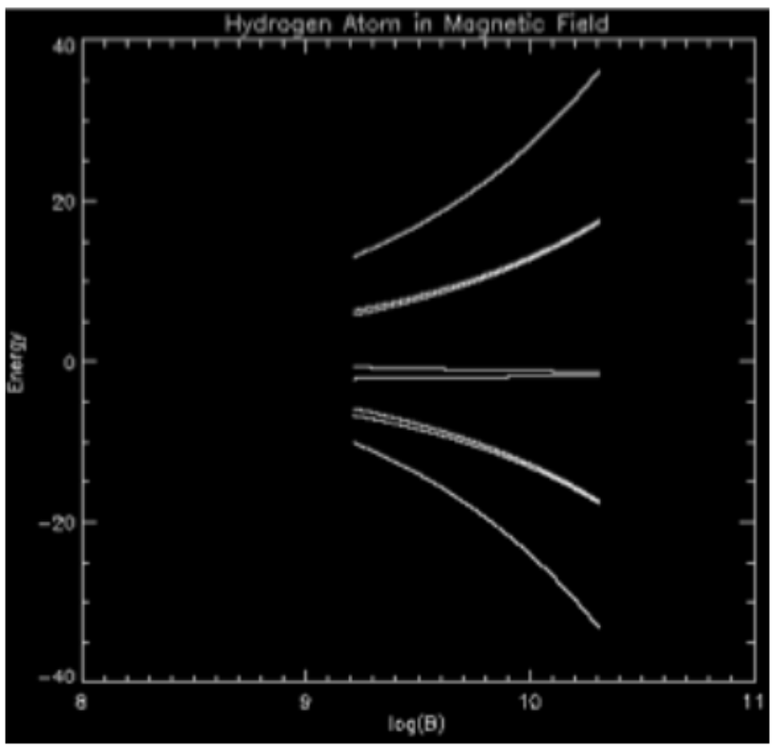
looks like the Figure below.



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 the Figure below.



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 in the 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

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

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

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

so that

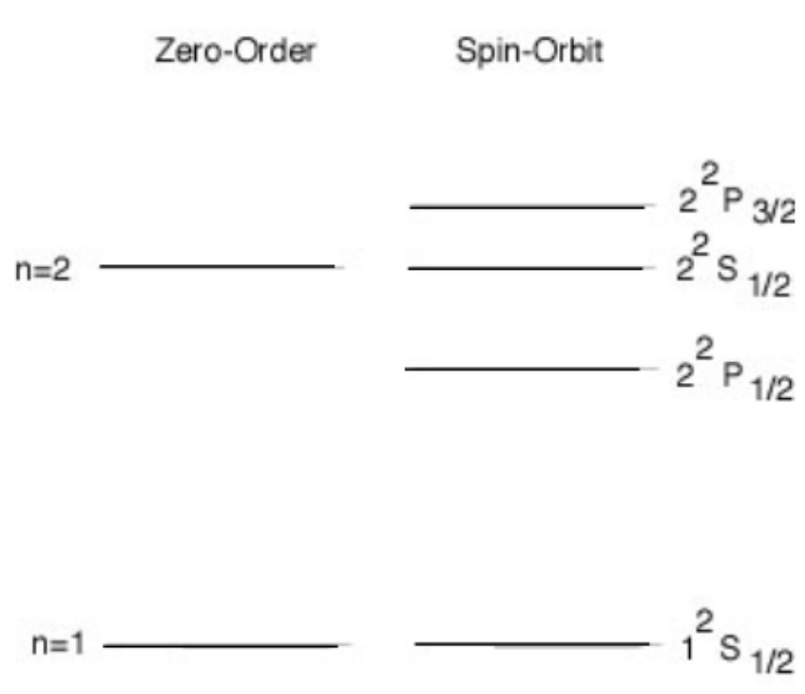
$$\left|21\frac{1}{2}\frac{3}{2}m_j\right\rangle \rightarrow 2^2P_{\frac{3}{2}}\ ,\ \left|21\frac{1}{2}\frac{1}{2}m_j\right\rangle \rightarrow 2^2P_{\frac{1}{2}}\ ,\ \left|20\frac{1}{2}\frac{1}{2}m_j\right\rangle \rightarrow 2^2S_{\frac{1}{2}}$$

The *L*(symbols) are defined by the Table.

L	0	1	2	3
Symbol	S	P	D	F

Spectroscopic Labels

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



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

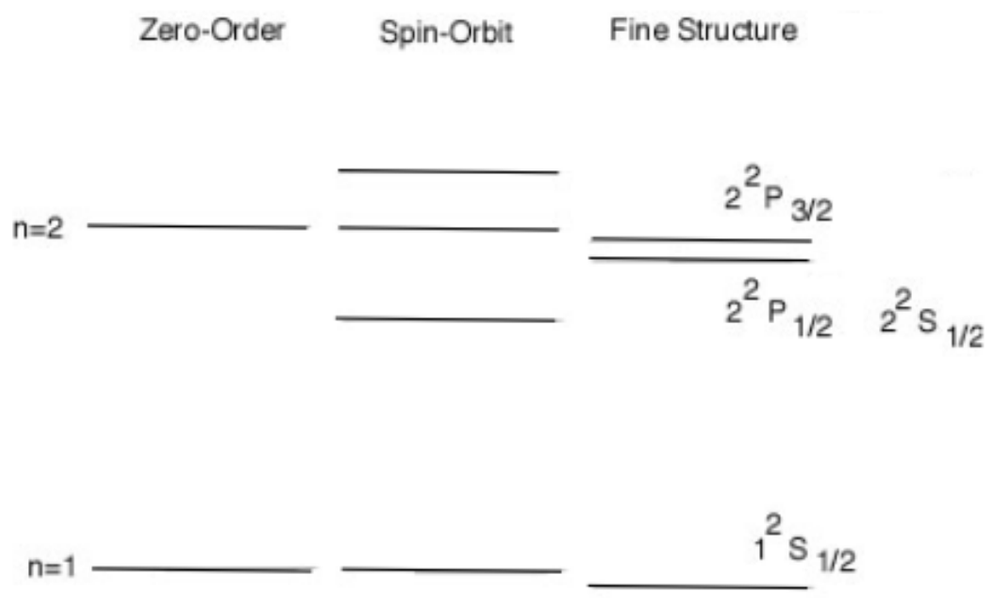
$$\Delta E_{rel} = -\frac{Z^2\alpha^2\left|E_n^{(0)}\right|}{n}\left(\frac{2}{2\ell+1}-\frac{3}{4n}\right)$$

Combining these two corrections we have

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

which is independent of ℓ .

This changes the energy level structure to that shown in the Figure below.



Fine Structure Energy Levels

The observed spectral lines result from an electron making a transition between these levels. stop

Stark Effect

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

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

where

$$z = z_e - z_p = z_{relative}$$

Therefore, we consider the Hamiltonian

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

where

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

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

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

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

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

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

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

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

Now

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

Therefore

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

Now

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

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

Therefore,

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

and

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

Finally

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

where

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

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.

$$n = 2 \rightarrow \ell = 0 \rightarrow \psi_{200} = \psi_1$$

$$\rightarrow \ell = 1 \rightarrow m = \begin{cases} 1 \rightarrow \psi_{211} = \psi_2 \\ 0 \rightarrow \psi_{210} = \psi_3 \\ -1 \rightarrow \psi_{21-1} = \psi_4 \end{cases}$$

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

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

Now z has no ϕ dependence and therefore,

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

Thus,

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

and the matrix becomes

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

We also have

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

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)

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

where

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

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

$$\psi_{211}(\vec{r}) \to E_{211} = E_2^{(0)} \text{ remains degenerate}$$

$$\psi_{21-1}(\vec{r}) \to E_{21-1} = E_2^{(0)} \text{ remains degenerate}$$

$$\psi_{(+)}(\vec{r}) = \frac{1}{\sqrt{2}}\left(\psi_{200}(\vec{r}) - \psi_{210}(\vec{r})\right) \to E_+ = E_2^{(0)} + 3e\mathcal{E}_0 a_0$$

$$\psi_{(-)}(\vec{r}) = \frac{1}{\sqrt{2}}\left(\psi_{200}(\vec{r}) + \psi_{210}(\vec{r})\right) \to E_- = E_2^{(0)} - 3e\mathcal{E}_0 a_0$$

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.

Van der Waal's Interaction

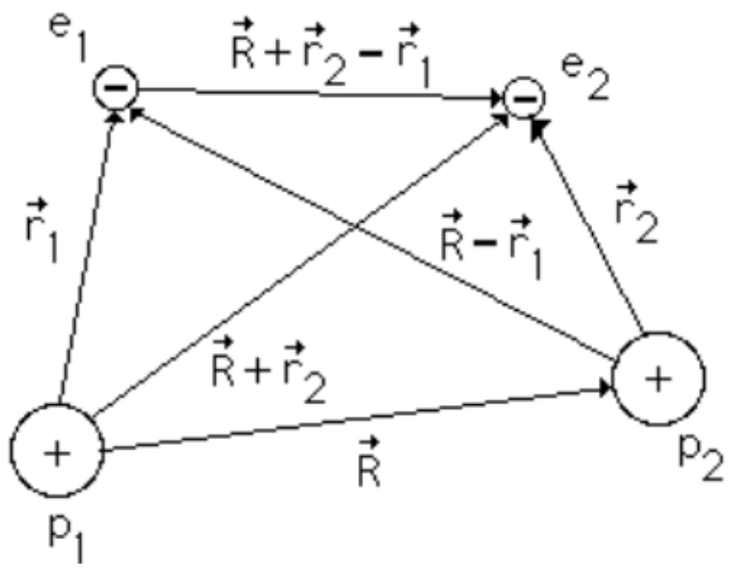
We now consider a system consisting of two widely separated atoms.

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

$\vec{r}_1 =$ vector from first proton to its electron

$\vec{r}_2 =$ vector from second proton to its electron

as shown in the Figure.



Van der Wall's System

The Hamiltonian is given by

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

where

$$\hat{H}_0 = \frac{\vec{p}_{1,op}^{\,2}}{2\mu} - \frac{e^2}{r_1} + \frac{\vec{p}_{2,op}^{\,2}}{2\mu} - \frac{e^2}{r_2} = \text{ 2 non-interacting hydrogen atoms}$$

and

$$\hat{V} = \text{ rest of the Coulomb interactions}$$

$$= V_{p_1p_2} + V_{e_1e_2} + V_{e_1p_2} + V_{e_2p_1}$$

$$= e^2\left(\frac{1}{R} + \frac{1}{\left|\vec{R}+\vec{r}_2-\vec{r}_1\right|} - \frac{1}{\left|\vec{R}+\vec{r}_2\right|} - \frac{1}{\left|\vec{R}-\vec{r}_1\right|}\right)$$

This is the perturbation potential.

We know the zero-order solution for two non-interacting hydrogen atoms.

It is

$$\text{zero-order states}: \ |n_1\ell_1 m_1\rangle\,|n_2\ell_2 m_2\rangle$$

with

$$\text{zero-order energies}: \ E^{(0)}_{n_1 n_2} = -\frac{e^2}{2a_0}\left(\frac{1}{n_1^2}+\frac{1}{n_2^2}\right)$$

where

$$\hat{H}_0\,|n_1\ell_1 m_1\rangle\,|n_2\ell_2 m_2\rangle = -\frac{e^2}{2a_0}\left(\frac{1}{n_1^2}+\frac{1}{n_2^2}\right)|n_1\ell_1 m_1\rangle\,|n_2\ell_2 m_2\rangle$$

This expression for the perturbation potential is too complicated to calculate.

We will need to make an approximation.

We make the reasonable assumption that

$$R >> r_2 \text{ and } R >> r_1$$

We have two useful mathematical results that we can apply.

In general, we can write

$$\frac{1}{\left|\vec{R}+\vec{a}\right|} = \frac{1}{\left[R^2+2\vec{R}\cdot\vec{a}+a^2\right]^{1/2}} = \frac{1}{R}\left[1+\frac{2\vec{R}\cdot\vec{a}}{R^2}+\frac{a^2}{R^2}\right]^{-1/2}$$

and for small x we have

$$[1+x]^{-1/2} \approx 1-\frac{1}{2}x+\frac{3}{8}x^2-\frac{5}{16}x^3+\ldots\ldots$$

Using

$$x = \frac{2\vec{R}\cdot\vec{a}}{R^2} + \frac{a^2}{R^2}$$

we get the general result

$$\frac{1}{\left|\vec{R}+\vec{a}\right|} = \frac{1}{R}\left[1 - \frac{1}{2}\left(\frac{2\vec{R}\cdot\vec{a}}{R^2} + \frac{a^2}{R^2}\right) + \frac{3}{8}\left(\frac{2\vec{R}\cdot\vec{a}}{R^2} + \frac{a^2}{R^2}\right)^2 - .....\right]$$

$$= \frac{1}{R}\left[1 - \frac{\vec{a}\cdot\vec{R}}{R^2} - \frac{1}{2}\frac{a^2}{R^2} + \frac{3}{2}\frac{\left(\vec{a}\cdot\vec{R}\right)^2}{R^4} + ...\right]$$

Therefore, we have

$$\frac{1}{\left|\vec{R}+\vec{r}_2\right|} = \frac{1}{R} - \frac{\vec{r}_2\cdot\vec{R}}{R^3} - \frac{1}{2}\frac{r_2^2}{R^3} + \frac{3}{2}\frac{\left(\vec{r}_2\cdot\vec{R}\right)^2}{R^5}$$

$$\frac{1}{\left|\vec{R}-\vec{r}_1\right|} = \frac{1}{R} + \frac{\vec{r}_1\cdot\vec{R}}{R^3} - \frac{1}{2}\frac{r_1^2}{R^3} + \frac{3}{2}\frac{\left(\vec{r}_1\cdot\vec{R}\right)^2}{R^5}$$

$$\frac{1}{\left|\vec{R}+\vec{r}_2-\vec{r}_1\right|} = \frac{1}{R} - \frac{\left(\vec{r}_2-\vec{r}_1\right)\cdot\vec{R}}{R^2} - \frac{1}{2}\frac{\left(\vec{r}_2-\vec{r}_1\right)^2}{R^2} + \frac{3}{2}\frac{\left(\left(\vec{r}_2-\vec{r}_1\right)\cdot\vec{R}\right)^2}{R^4}$$

$$= \frac{1}{R} - \frac{\vec{r}_2\cdot\vec{R}}{R^3} + \frac{\vec{r}_1\cdot\vec{R}}{R^3} - \frac{1}{2}\frac{r_1^2}{R^3} - \frac{1}{2}\frac{r_2^2}{R^3} + \frac{\vec{r}_1\cdot\vec{r}_2}{R^3}$$

$$+ \frac{3}{2}\frac{\left(\vec{r}_2\cdot\vec{R}\right)^2}{R^5} + \frac{3}{2}\frac{\left(\vec{r}_1\cdot\vec{R}\right)^2}{R^5} - 3\frac{\left(\left(\vec{r}_1\cdot\vec{R}\right)\left(\vec{r}_2\cdot\vec{R}\right)\right)}{R^5}$$

Putting all this together we get

$$V = \frac{e^2}{R^3}\left[\vec{r}_1 \cdot \vec{r}_2 - 3\frac{\left((\vec{r}_1 \cdot \vec{R})(\vec{r}_2 \cdot \vec{R})\right)}{R^2}\right]$$

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

$$\vec{R} = R\hat{z}$$

which gives

$$V = \frac{e^2}{R^3}\left[(x_1x_2 + y_1y_2 + z_1z_2) - 3\frac{z_1z_2R^2}{R^2}\right]$$

$$= \frac{e^2}{R^3}(x_1x_2 + y_1y_2 - 2z_1z_2)$$

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

$$E_0 = -\frac{e^2}{8a_0} - \frac{e^2}{8a_0} = -\frac{e^2}{4a_0}$$

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)

$$\begin{aligned}
&\langle \ell_1 m_1 \ell_2 m_2 | \hat{V} | \ell_1 m_1 \ell_2 m_2 \rangle \\
&\quad = \frac{e^2}{R^3} \langle \ell_1 m_1 | \hat{x}_1 | \ell_1 m_1 \rangle \langle \ell_2 m_2 | \hat{x}_2 | \ell_2 m_2 \rangle \\
&\qquad + \frac{e^2}{R^3} \langle \ell_1 m_1 | \hat{y}_1 | \ell_1 m_1 \rangle \langle \ell_2 m_2 | \hat{y}_2 | \ell_2 m_2 \rangle \\
&\qquad - 2\frac{e^2}{R^3} \langle \ell_1 m_1 | \hat{z}_1 | \ell_1 m_1 \rangle \langle \ell_2 m_2 | \hat{z}_2 | \ell_2 m_2 \rangle
\end{aligned}$$

We have

$$x = r\sin\theta\cos\phi = -r\sqrt{\frac{2\pi}{3}}\left(Y_{1,1} - Y_{1,-1}\right)$$

$$y = r\sin\theta\sin\phi = +ir\sqrt{\frac{2\pi}{3}}\left(Y_{1,1} + Y_{1,-1}\right)$$

$$z = r\cos\theta = r\sqrt{\frac{4\pi}{3}}Y_{10}$$

and

$$\begin{aligned}
&\langle n\ell m | x | n\ell' m' \rangle \\
&\quad = -\sqrt{\frac{2\pi}{3}}\left[\int_0^\infty r^3 R_{n\ell}(r) R_{n\ell'}(r) dr\right]\left[\int d\Omega Y^*_{\ell m}(Y_{1,1} - Y_{1,-1}) Y_{\ell' m'}\right] \\
&\langle n\ell m | y | n\ell' m' \rangle \\
&\quad = i\sqrt{\frac{2\pi}{3}}\left[\int_0^\infty r^3 R_{n\ell}(r) R_{n\ell'}(r) dr\right]\left[\int d\Omega Y^*_{\ell m}(Y_{1,1} + Y_{1,-1}) Y_{\ell' m'}\right] \\
&\langle n\ell m | z | n\ell' m' \rangle \\
&\quad = \sqrt{\frac{4\pi}{3}}\left[\int_0^\infty r^3 R_{n\ell}(r) R_{n\ell'}(r) dr\right]\left[\int d\Omega Y^*_{\ell m} Y_{10} Y_{\ell' m'}\right]
\end{aligned}$$

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

$$\int d\Omega Y^*_{\ell m} Y_{1m''} Y_{\ell' m'} = 0 \quad \text{unless} \quad \begin{cases} \ell + \ell' + 1 = even \\ m = m' + m'' \end{cases}$$

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

$$\text{for } x \text{ and } y \qquad m = m' \pm 1$$

$$\text{for } z \qquad m = m'$$

which is the so-called

$$\Delta m = \pm 1, 0$$

selection rule for this type of perturbation.

In addition, we have the

$$\Delta \ell = \pm 1$$

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

$$\left[\hat{L}_i, r_j\right] = i\hbar\varepsilon_{ijk} r_k$$

This allows us to write (after much algebra)

$$\begin{aligned}\left[\hat{L}_z, \hat{V}\right] &= \left[(\hat{L}_{1z} + \hat{L}_{2z}), \hat{V}\right] = \left[\hat{L}_{1z}, \hat{V}\right] + \left[\hat{L}_{2z}, \hat{V}\right] \\ &= \frac{e^2}{R^3}\left[\hat{L}_{1z}, (x_1x_2 + y_1y_2 - 2z_1z_2)\right] + \frac{e^2}{R^3}\left[\hat{L}_{2z}, (x_1x_2 + y_1y_2 - 2z_1z_2)\right] \\ &= 0\end{aligned}$$

This implies that $[\,\hat{L}_z, \hat{H}\,] = 0$ 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

$$m_1 + m_2 = m_1' + m_2'$$

Summarizing the selection rules we have

$$\begin{aligned}&\ell_1 + \ell_1' + 1 = even \qquad (= \text{reason b for a zero}) \\ &\ell_2 + \ell_2' + 1 = even \qquad (= \text{reason c for a zero}) \\ &m_1 - m_1' = \pm 1, 0 \qquad (= \text{reason d for a zero}) \\ &m_2 - m_2' = \pm 1, 0 \qquad (= \text{reason d for a zero}) \\ &m_1 + m_2 = m_1' + m_2' \qquad (= \text{reason a for a zero})\end{aligned}$$

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

$\langle \hat{V} \rangle$ 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 nonzero elements are given by the expressions

$$A = C = \frac{e^2}{R^3} \langle 200| \langle 200| (x_1 x_2 + y_1 y_2) |211\rangle |21, -1\rangle$$

$$B = E = -2\frac{e^2}{R^3} \langle 200| \langle 200| z_1 z_2 |210\rangle |21, 0\rangle$$

$$D = \frac{e^2}{R^3} \langle 200| \langle 211| (x_1 x_2 + y_1 y_2) |211\rangle |200\rangle$$

$$F = \frac{e^2}{R^3} \langle 200| \langle 21, -1| (x_1 x_2 + y_1 y_2) |21, -1\rangle |200\rangle$$

If we define

$$\alpha = \sqrt{\frac{8\pi}{3}} \int_0^\infty r^3 R_{20} R_{21} dr$$

we have

$$\langle 200| x |211\rangle = \frac{\alpha}{2} = -\langle 200| x |21, -1\rangle$$

$$\langle 200| y |211\rangle = i\frac{\alpha}{2} = \langle 200| x |21, -1\rangle$$

$$\langle 200| z |211\rangle = \frac{\sqrt{2}}{2}\alpha$$

and

$$A = C = \frac{E}{2} = \frac{B}{2} = -D = -F = -\frac{1}{2}\alpha^2 \frac{e^2}{R^3}$$

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

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

$$4 \times 4 \rightarrow 0, 0, \pm\sqrt{6}A \quad , \quad 2 \times 2 \rightarrow \pm A$$
$$2 \times 2 \rightarrow \pm 2A \quad , \quad 2 \times 2 \rightarrow \pm A$$

Therefore, the energies correct to first-order are

$$E = \begin{cases} E_0 + \sqrt{6}A & \text{degeneracy} = 1 \\ E_0 + 2A & \text{degeneracy} = 1 \\ E_0 + A & \text{degeneracy} = 2 \\ E_0 & \text{degeneracy} = 8 \\ E_0 - A & \text{degeneracy} = 2 \\ E_0 - 2A & \text{degeneracy} = 1 \\ E_0 - \sqrt{6}A & \text{degeneracy} = 1 \end{cases}$$

That was a **real** problem!

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

$$\hat{H} = \hat{I}\hat{H} = \sum_N |N\rangle \langle N| \hat{H} = \sum_N E_n |N\rangle \langle N|$$

where

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

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

$$\begin{aligned}
\langle\psi| \hat{H} |\psi\rangle &= \sum_N E_n \langle\psi \mid N\rangle \langle N \mid \psi\rangle \\
&\geq \sum_N E_0 \langle\psi \mid N\rangle \langle N \mid \psi\rangle = E_0 \sum_N \langle\psi \mid N\rangle \langle N \mid \psi\rangle \\
&\geq E_0 \langle\psi| \left(\sum_N |N\rangle \langle N| \right) |\psi\rangle = E_0 \langle\psi| \hat{I} |\psi\rangle = E_0 \langle\psi \mid \psi\rangle
\end{aligned}$$

or

$$\frac{\langle\psi| \hat{H} |\psi\rangle}{\langle\psi \mid \psi\rangle} \geq E_0$$

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

$$\frac{\langle\psi|\hat{H}|\psi\rangle}{\langle\psi\,|\,\psi\rangle} \text{ is an upper bound for } E_0$$

Procedure

1. Pick a trial vector $|\psi\rangle$ that contains unknown parameters $\{\alpha_k\}$
2. Calculate
$$\frac{\langle\psi|\hat{H}|\psi\rangle}{\langle\psi\,|\,\psi\rangle} = E_0\left(\{\alpha_k\}\right)$$
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
$$\frac{\partial E_0}{\partial \alpha_k} = 0 \text{ for all } k$$
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.,

$$\text{ground state} \rightarrow \ell = 0 \rightarrow \text{ contains } Y_{00}$$

$$1^{st} \text{ excited state} \rightarrow \ell = 1 \rightarrow \text{ contains } Y_{1m}$$

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

$$\hat{H} = -\frac{\hbar^2}{2m}\frac{d^2}{dx^2} + \frac{1}{2}kx^2$$

and a trial wave function

$$\psi(x,a) = \begin{cases} (a^2 - x^2)^2 & |x| < a \\ 0 & |x| \geq a \end{cases}$$

where *a* is an unknown parameter.

The variational principle says that

$$\frac{\langle\psi(a)|\,\hat{H}\,|\psi(a)\rangle}{\langle\psi(a)\,|\,\psi(a)\rangle} = E_0(a) \geq E_0 = \text{ true ground state energy}$$

We get a *best* value for this choice of trial function by minimizing with respect to a using

$$\frac{dE_0(a)}{da} = 0$$

Now we need to calculate the integrals.

We have for the denominator (the normalization integral)

$$\langle\psi(a)\,|\,\psi(a)\rangle = \int_{-a}^{a} \psi^2(x,a)dx = \int_{-a}^{a} (a^2 - x^2)^4 dx$$

$$= 2\int_{0}^{a} (a^2 - x^2)^4 dx = \frac{336}{315}a^9$$

and for the numerator

$$\langle\psi(a)|\,\hat{H}\,|\psi(a)\rangle = -\frac{\hbar^2}{2m}2\int_{0}^{a} \psi(x,a)\frac{d^2\psi(x,a)}{dx^2}dx + \frac{1}{2}k2\int_{0}^{a} \psi^2(x,a)x^2dx$$

$$= -\frac{\hbar^2}{2m}\left(-\frac{80}{21}a^7\right) + \frac{1}{2}k\left(\frac{336}{3465}a^{11}\right)$$

Therefore

$$E_0(a) = 1.786\frac{\hbar^2}{ma^2} + 0.045ka^2$$

The minimum condition gives

$$a^2 = 6.300\left(\frac{\hbar^2}{mk}\right)^{1/2}$$

which then says that

$$0.566\hbar\omega \geq E_0$$

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.

Chapter 9

Time-Dependent Perturbation Theory

Theory

Time-independent or stationary-state perturbation theory, as we developed earlier, allows us to find approximations for the energy eigenvalues and eigenvectors in complex physical systems that are not solvable in closed form and where we could write $\hat{H}$ in two parts as

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

For these perturbation methods to work, $\hat{V}$ must be *weak* and *time-independent*.

We now turn our attention to the case

$$H = H_0 + V_t$$

where $\hat{V}_t$ is *weak* and *time-dependent*.

Examples might be the decays of an atomic system by photon emission or the ionization of an atom by shining light on it.

We assume that at some time t_0 the system has evolved into the state $\left|\psi_t^{(0)}\right\rangle$, i.e., the state $\left|\psi_t^{(0)}\right\rangle$ satisfies the time evolution equation

$$i\hbar\frac{\partial}{\partial t}\left|\psi_t^{(0)}\right\rangle = \hat{H}_0\left|\psi_t^{(0)}\right\rangle \qquad t \le t_0$$

It is a solution of the time-dependent Schrodinger equation with no perturbing interactions before t_0 where

$$\hat{H} = \hat{H}_0 \qquad t \le t_0$$

At time t_0 we *turn on* the interaction potential (or perturbation) so that

$$\hat{H} = \hat{H}_0 + \hat{V}_t \qquad t \ge t_0$$

The new state of the system then satisfies

$$i\hbar \frac{\partial}{\partial t} |\psi_t\rangle = \hat{H} |\psi_t\rangle = (\hat{H}_0 + \hat{V}_t) |\psi_t\rangle \qquad t \ge t_0$$

with the *boundary condition*(*initial value*)

$$|\psi_t\rangle = \left|\psi_t^{(0)}\right\rangle \text{ at } t{=}t_0$$

As we said, we assume that the full time-dependent Schrodinger equation cannot be solved in closed form and so we look for approximate solutions.

We let $\hat{V}_t$ be a small perturbation, i.e., we assume there is a natural small parameter in $\hat{V}_t$ (as we saw in time-independent perturbation theory) and we make an expansion of the solution in powers of $\hat{V}_t$ or this small parameter.

Since the effect of $\hat{H}_0$ will be much greater than the effect of $\hat{V}_t$, most of the time dependence comes from $\hat{H}_0$.

If we could neglect $\hat{V}_t$, then since $\hat{H}_0$ is independent of time, we would have the simple time dependence

$$|\psi_t\rangle = e^{-\frac{i}{\hbar}\hat{H}_0 t} \left|\psi_t^{(0)}\right\rangle$$

Let us assume that this is still approximately true and remove this known time dependence from the solution.

This should remove the major portion of the total time dependence from the problem.

We do this by assuming a solution of the form

$$|\psi_t\rangle = e^{-\frac{i}{\hbar}\hat{H}_0 t}\,|\psi(t)\rangle$$

and then determining and solving the equation for the new state vector $|\psi(t)\rangle$.

Substituting this assumption in our original equation, the equation for $|\psi(t)\rangle$ is then given by

$$i\hbar\frac{\partial}{\partial t}\left(e^{-\frac{i}{\hbar}\hat{H}_0 t}\,|\psi(t)\rangle\right) = (\hat{H}_0 + \hat{V}_t)\left(e^{-\frac{i}{\hbar}\hat{H}_0 t}\,|\psi(t)\rangle\right)$$

$$\hat{H}_0 e^{-\frac{i}{\hbar}\hat{H}_0 t}\,|\psi(t)\rangle + i\hbar e^{-\frac{i}{\hbar}\hat{H}_0 t}\frac{\partial}{\partial t}\,|\psi(t)\rangle$$

$$= \hat{H}_0 e^{-\frac{i}{\hbar}\hat{H}_0 t}\,|\psi(t)\rangle + \hat{V}_t e^{-\frac{i}{\hbar}\hat{H}_0 t}\,|\psi(t)\rangle$$

$$i\hbar\frac{\partial}{\partial t}\,|\psi(t)\rangle = \hat{V}(t)\,|\psi(t)\rangle$$

where

$$\hat{V}(t) = e^{\frac{i}{\hbar}\hat{H}_0 t}\hat{V}_t e^{-\frac{i}{\hbar}\hat{H}_0 t}$$

The substitution has removed $\hat{H}_0$ from the equation and changed the time dependence of the perturbing potential.

We are in the so-called *interaction picture* or *representation* where *both* the state vectors and the operators depend on time.

We develop a *formal solution* by integrating this equation of motion for the state vector to get

$$i\hbar \int_{t_0}^{t} \frac{\partial}{\partial t'} |\psi(t')\rangle dt' = \int_{t_0}^{t} \hat{V}(t') |\psi(t')\rangle dt'$$
$$= i\hbar(|\psi(t)\rangle - |\psi(t_0)\rangle)$$

so that the formal solution is given by

$$|\psi(t)\rangle = |\psi(t_0)\rangle + \frac{1}{i\hbar} \int_{t_0}^{t} \hat{V}(t') |\psi(t')\rangle dt'$$

This is an *integral equation* for $|\psi(t)\rangle$.

We solve it as a power series in $\hat{V}_t$ by the *method of iteration*.

The 0^{th}–order approximation is found by neglecting the perturbing potential.

We get $$|\psi(t)\rangle = |\psi(t_0)\rangle$$

The 1^{st}–order approximation is obtained by inserting the 0^{th}–order approximation into the full equation.

We get

$$|\psi(t)\rangle = |\psi(t_0)\rangle + \frac{1}{i\hbar} \int_{t_0}^{t} \hat{V}(t') |\psi(t_0)\rangle dt'$$
$$= \left(1 + \frac{1}{i\hbar} \int_{t_0}^{t} \hat{V}(t') dt'\right) |\psi(t_0)\rangle$$

The 2nd–order approximation is obtained by inserting the 1st–order approximation into the full equation.

We get

$$|\psi(t)\rangle = |\psi(t_0)\rangle + \frac{1}{i\hbar}\int_{t_0}^{t} \hat{V}(t')\,|\psi(t_0)\rangle dt'$$

$$+ \frac{1}{(i\hbar)^2}\int_{t_0}^{t} dt' \int_{t_0}^{t'} dt'' \hat{V}(t')\hat{V}(t'')\,|\psi(t_0)\rangle$$

Notice that in all subsequent iterations the operators $\hat{V}(t'), \hat{V}(t'')$,. . ., etc, always occur in order of increasing time from right to left.

We can write the general result as

$$|\psi(t)\rangle = \hat{U}(t, t_0)\,|\psi(t_0)\rangle$$

where

$$\hat{U}(t, t_0) = \hat{I} + \sum_{n=1}^{\infty} \frac{1}{(i\hbar)^n} \int_{t_0}^{t} dt_1 \int_{t_0}^{t_1} dt_2 \ldots\ldots\ldots \int_{t_0}^{t_{n-1}} dt_n \hat{V}(t_1)\hat{V}(t_2)\ldots\ldots\hat{V}(t_n)$$

The complete, formal solution to the problem is then given by

$$|\psi_t\rangle = e^{-\frac{i}{\hbar}\hat{H}_0 t}\,|\psi(t)\rangle = e^{-\frac{i}{\hbar}\hat{H}_0 t}\hat{U}(t, t_0)\,|\psi(t_0)\rangle$$

so that

$$e^{-\frac{i}{\hbar}\hat{H}_0 t}\hat{U}(t, t_0) = \text{ the total time development operator}$$

Before developing the detailed techniques of time-dependent perturbation theory, let us spend some time with the operator $\hat{U}(t, t_0)$ and discuss some of its properties.

We first introduce the idea of a *time-ordered product* of operators as follows.

The symbol

$$\left(\hat{A}(t)\hat{B}(t_1)\hat{C}(t_2).....\hat{X}(t_n)\right)_+ \equiv T\left(\hat{A}(t)\hat{B}(t_1)\hat{C}(t_2).....\hat{X}(t_n)\right)$$

means the product of the operators where the operators are written from right to left in order of increasing times, i.e.,

$$\left(\hat{A}(t)\hat{B}(t')\right)_+ = \begin{cases} \hat{A}(t)\hat{B}(t') & t \geq t' \\ \hat{B}(t')\hat{A}(t) & t' \geq t \end{cases}$$

Now, we have using the time-ordered product definition

$$\left(\left[\int_{t_0}^{t} \hat{V}(t')dt'\right]^2\right)_+ = \left(\int_{t_0}^{t} \hat{V}(t')dt' \int_{t_0}^{t} \hat{V}(t'')dt''\right)_+$$

$$= \int_{t_0}^{t} dt' \int_{t_0}^{t} dt''(\hat{V}(t')\hat{V}(t''))_+$$

$$= \int_{t_0}^{t} dt' \int_{t_0}^{t'} dt''\hat{V}(t')\hat{V}(t'') + \int_{t_0}^{t} dt'' \int_{t_0}^{t''} dt'\hat{V}(t'')\hat{V}(t')$$

$$= 2\int_{t_0}^{t} dt' \int_{t_0}^{t'} dt''\hat{V}(t')\hat{V}(t'')$$

and in general

$$\left(\left[\int_{t_0}^{t}\hat{V}(t')dt'\right]^n\right)_+ = \left(\int_{t_0}^{t}\hat{V}(t_1)dt_1\int_{t_0}^{t}\hat{V}(t_2)dt_2.....\int_{t_0}^{t}\hat{V}(t_n)dt_n\right)_+$$

$$= \int_{t_0}^{t}dt_1\int_{t_0}^{t}dt_2.....\int_{t_0}^{t}dt_n(\hat{V}(t_1)\hat{V}(t_2).....\hat{V}(t_n))_+$$

$$= n!\int_{t_0}^{t}dt_1\int_{t_0}^{t}dt_2.....\int_{t_0}^{t}dt_n\hat{V}(t_1)\hat{V}(t_2).....\hat{V}(t_n)$$

because there are $n!$ possible orderings of the n terms involved.

This last form is identical to the expression for $\hat{U}(t,t_0)$ and thus we have

$$\hat{U}(t,t_0) = \sum_{n=0}^{\infty}\frac{1}{(i\hbar)^n}\frac{1}{n!}\left(\left[\int_{t_0}^{t}\hat{V}(t')dt'\right]^n\right)_+$$

$$= \left(\exp\left[-\frac{i}{\hbar}\int_{t_0}^{t}\hat{V}(t')dt'\right]\right)_+$$

The last expression is just a *convenient shorthand for the infinite sum*.

In order to verify that this is in fact a solution of

$$i\hbar\frac{\partial}{\partial t}|\psi(t)\rangle = \hat{V}(t)|\psi(t)\rangle$$

we must prove that

$$i\hbar\frac{\partial}{\partial t}\hat{U}(t,t_0)\,|\psi(t_0)\rangle = \hat{V}(t)\hat{U}(t,t_0)\,|\psi(t_0)\rangle$$

$$i\hbar\frac{\partial}{\partial t}\hat{U}(t,t_0) = \hat{V}(t)\hat{U}(t,t_0)$$

Substituting, we have

$$i\hbar\frac{\partial}{\partial t}\hat{U}(t,t_0) = i\hbar\frac{\partial}{\partial t}\left(\exp\left[-\frac{i}{\hbar}\int_{t_0}^{t}\hat{V}(t')dt'\right]\right)_+$$

$$= \left(\hat{V}(t)\exp\left[-\frac{i}{\hbar}\int_{t_0}^{t}\hat{V}(t')dt'\right]\right)_+$$

In the differentiation we do not have to worry about the non-commutation of the operators inside the time-ordered product since the order is already specified.

Since t is certainly the latest time in the time-ordered product and therefore all the other operators will be on the right of $\hat{V}(t)$ we can pull it outside the time-ordered product and write

$$i\hbar\frac{\partial}{\partial t}\hat{U}(t,t_0) = \hat{V}(t)\left(\exp\left[-\frac{i}{\hbar}\int_{t_0}^{t}\hat{V}(t')dt'\right]\right)_+ = \hat{V}(t)\hat{U}(t,t_0)$$

as required.

The *most important question* (really the only question) that is usually asked in quantum mechanics is the following:

> Suppose that the system is initially in an eigenstate$|n\rangle$ of $\hat{H}_0$, i.e., $\hat{H}_0\,|n\rangle = \varepsilon_n|n\rangle$.
> What is the probability that the system will be observed, after the perturbation has had time to act, in a different (and thus orthogonal) eigenstate of $\hat{H}_0$, say $|m\rangle$?

Alternatively, the question is sometimes posed this way:

> What is the probability that the interaction causes the system to make a *transition* from the state $|n\rangle$ to the state $|m\rangle$?

The probability amplitude for observing the system in the state $|m\rangle$ at time t is given by

$$\begin{aligned}\langle m \mid \psi_t \rangle &= \langle m| e^{-\frac{i}{\hbar}\hat{H}_0 t}\hat{U}(t,t_0)\,|\psi(t_0)\rangle \\ &= \langle m| e^{-\frac{i}{\hbar}\hat{H}_0 t}\hat{U}(t,t_0)\,|n\rangle\end{aligned}$$

where

$$|\psi(t_0)\rangle = |n\rangle$$

is the initial state.

Setting $t_0 = 0$ for simplicity and using the 1st–order approximation for $\hat{U}(t,0)$ and also using

$$\langle m| e^{-\frac{i}{\hbar}\hat{H}_0 t} = \left(e^{\frac{i}{\hbar}\hat{H}_0 t}\,|m\rangle\right)^{\dagger} = \left(e^{\frac{i}{\hbar}\varepsilon_m t}\,|m\rangle\right)^{\dagger} = \langle m| e^{-\frac{i}{\hbar}\varepsilon_m t}$$

we get

$$\begin{aligned}\langle m \mid \psi_t \rangle &= \frac{1}{i\hbar} e^{-\frac{i}{\hbar}\varepsilon_m t}\int_0^t dt'\,\langle m|\hat{V}(t')\,|n\rangle \\ &= \frac{1}{i\hbar} e^{-\frac{i}{\hbar}\varepsilon_m t}\int_0^t dt'\,\langle m| e^{\frac{i}{\hbar}\hat{H}_0 t}\hat{V}_{t'} e^{-\frac{i}{\hbar}\hat{H}_0 t}\,|n\rangle \\ &= \frac{1}{i\hbar} e^{-\frac{i}{\hbar}\varepsilon_m t}\int_0^t dt'\, e^{\frac{i}{\hbar}(\varepsilon_m-\varepsilon_n)t}\,\langle m|\hat{V}_{t'}\,|n\rangle\end{aligned}$$

The probability of the transition is then

$$P_{n\to m}(t) = |\langle m \mid \psi_t\rangle|^2 = \frac{1}{\hbar^2}\left|\int_0^t dt' e^{\frac{i}{\hbar}(\varepsilon_m-\varepsilon_n)t}\,\langle m|\,\hat{V}_{t'}\,|n\rangle\right|^2$$

The simplest example is when $\hat{V}_t$ is not a function of t, or $\hat{V}_t = \hat{V}$.

We then have

$$P_{n\to m}(t) = |\langle m \mid \psi_t\rangle|^2 = \frac{\left|\langle m|\,\hat{V}_{t'}\,|n\rangle\right|^2}{\hbar^2}\left|\int_0^t dt' e^{\frac{i}{\hbar}(\varepsilon_m-\varepsilon_n)t}\right|^2$$

If we define $\Delta = \varepsilon_m - \varepsilon_n$, then we have

$$P_{n\to m}(t) = \left|\langle m|\,\hat{V}_{t'}\,|n\rangle\right|^2\left|\frac{1-e^{-\frac{i}{\hbar}\Delta t}}{\Delta}\right|^2 = \left|\langle m|\,\hat{V}_{t'}\,|n\rangle\right|^2\left(\frac{\sin\frac{\Delta t}{2\hbar}}{\Delta/2}\right)^2$$

for the transition probability.

What is the physical meaning of this result?

We must be *very careful* when we use the words

> the perturbation causes a transition
> between eigenstates of $\hat{H}_0$

What this means physically is that the system has absorbed from the perturbing field (or emitted to it) the energy difference $\Delta = \varepsilon_m - \varepsilon_n$ and therefore the system has changed its energy.

Does the statement also mean that the state vector has changed from an initial value $|\psi(0)\rangle = |n\rangle$ to a final value $|\psi(t)\rangle = |m\rangle$?

We can get a better feeling for the correct answer to this question by deriving the result in a different manner.

We have

$$i\hbar\frac{\partial}{\partial t}|\psi_t\rangle = \hat{H}|\psi_t\rangle = (\hat{H}_0 + \hat{V}_t)|\psi_t\rangle$$

and

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

As in our development of time-independent perturbation theory, we let

$$\hat{V}_t = g\hat{U}_t$$

where g is a small parameter.

The set of eigenvectors $\{|n\rangle\}$ is a complete set and therefore we can use it as a basis for the space and, in particular, we can write

$$|\psi_t\rangle = \sum_n a_n(t)e^{-\frac{i}{\hbar}\varepsilon_n t}|n\rangle$$

The reason for pulling out the phase factors will be clear shortly.

It is clear that if $g = 0$, then this is the correct general solution with

$$a_n(t) = a_n(0) = \text{constant}$$

The phase factors we pulled out represent the time dependence due to $\hat{H}_0$ and this is, by assumption, the major time dependence in the system.

If g is small we expect the time dependence of $a_n(t)$, which is due to the perturbation to be weak or that

$$\frac{da_n(t)}{dt} \text{ is small}$$

It is in this sense that we can propose to use perturbation theory on the system.

Using the eigenbasis expansion we have

$$\sum_n \left(i\hbar \frac{da_n(t)}{dt} + \varepsilon_n a_n(t) \right) e^{-\frac{i}{\hbar}\varepsilon_n t} |n\rangle$$
$$= \sum_n \left(\varepsilon_n a_n(t) + g\hat{U}_t a_n(t) \right) e^{-\frac{i}{\hbar}\varepsilon_n t} |n\rangle$$

Applying the linear functional $\langle m|$ from the left and using the orthonormality relation

$$\langle m \mid n \rangle = \delta_{mn}$$

we get

$$i\hbar \frac{da_m(t)}{dt} = g \sum_n \langle m| \hat{U}_t |n\rangle e^{i\omega_{mn}t} a_n(t)$$
$$= \sum_n V_{mn}(t) e^{i\omega_{mn}t} a_n(t)$$

where

$$\omega_{mn} = \frac{\varepsilon_m - \varepsilon_n}{\hbar}$$

This is an exact equation.

It implies that the time dependence of $a_n(t)$ is due entirely to $\hat{V}_t$ (because we explicitly extracted out the dependence due to $\hat{H}_0$).

This is the interaction picture that we had earlier.

Exactly Solvable 2-State Example

Consider a 2-state system with

$$H_0 = \begin{pmatrix} E_1 & 0 \\ 0 & E_2 \end{pmatrix} \quad , \quad V(t) = \begin{pmatrix} 0 & \delta e^{i\omega t} \\ \delta e^{-i\omega t} & 0 \end{pmatrix} = \begin{pmatrix} V_{11} & V_{12} \\ V_{21} & V_{22} \end{pmatrix}$$

In the interaction picture, as derived above, we have

$$i\hbar \frac{dc_m(t)}{dt} = \sum_n V_{mn}(t) e^{i\omega_{mn} t} c_n(t) \quad , \quad |\psi_t\rangle = \sum_n c_n(t) e^{-\frac{i}{\hbar}\varepsilon_n t} |n\rangle$$

or

$$i\hbar \frac{dc_1(t)}{dt} = \delta e^{i\left[\omega + \frac{E_1 - E_2}{\hbar}\right]t} c_2(t)$$

$$i\hbar \frac{dc_2(t)}{dt} = \delta e^{i\left[-\omega - \frac{E_1 - E_2}{\hbar}\right]t} c_1(t)$$

$$|\psi_t\rangle = c_1(t) e^{-\frac{i}{\hbar}E_1 t} |1\rangle + c_2(t) e^{-\frac{i}{\hbar}E_2 t} |2\rangle$$

We can write these equations as

$$\frac{dC(t)}{dt} = -\frac{i\delta}{\hbar} \begin{pmatrix} 0 & e^{i[\omega - \omega_{21}]t} \\ e^{-i[\omega - \omega_{21}]t} & 0 \end{pmatrix} C(t)$$

where

$$C(t) = \begin{pmatrix} c_1(t) \\ c_2(t) \end{pmatrix} \quad , \quad \omega_{21} = \frac{E_2 - E_1}{\hbar}$$

We can find an exact solution.

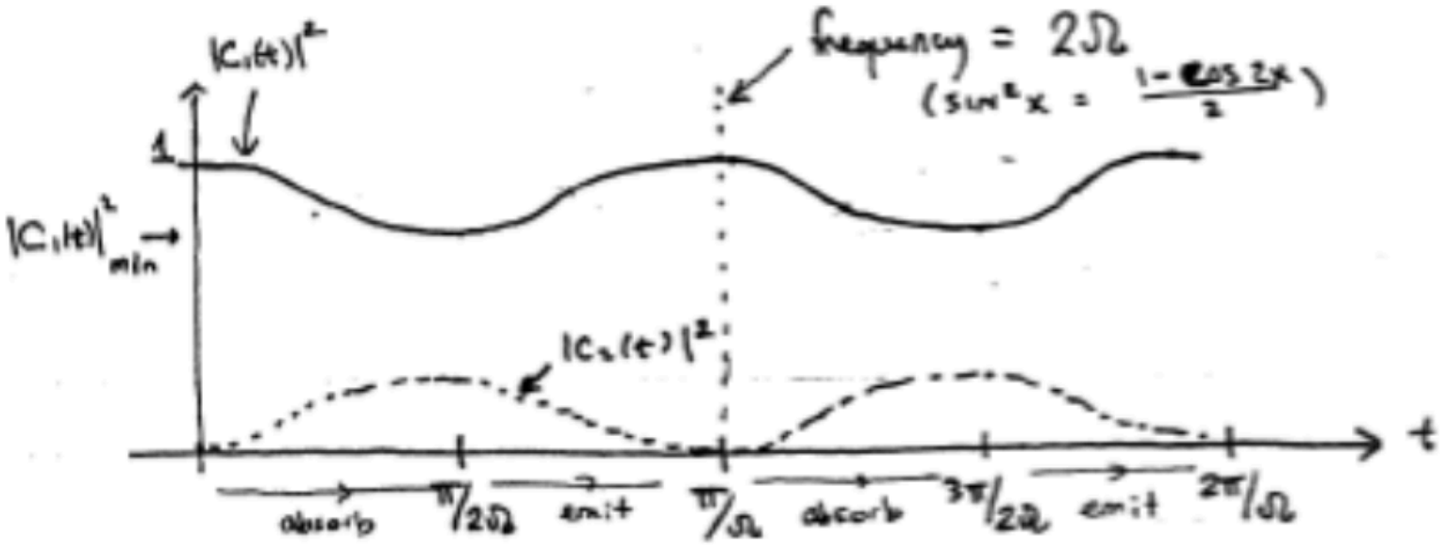
With initial conditions $c_1(0) = 1$ and $c_2(0) = 0$ we get

$$|c_1(t)|^2 = \frac{\delta^2}{\delta^2 + \frac{\hbar^2(\omega-\omega_{21})^2}{4}} \sin^2 \Omega t = 1 - |c_2(t)|^2$$

where

$$\Omega^2 = \frac{\delta^2}{\hbar^2} + \frac{(\omega - \omega_{21})^2}{4}$$

A graph of these functions is shown in the Figure below.



Exact Solution

A straightforward calculation gives

$$|c_1(t)|^2_{\min} = \frac{(\omega - \omega_{21})^2}{(\omega - \omega_{21})^2 + \frac{4\delta^2}{\hbar^2}}$$

At resonance, $\omega = \omega_{21}$, we have

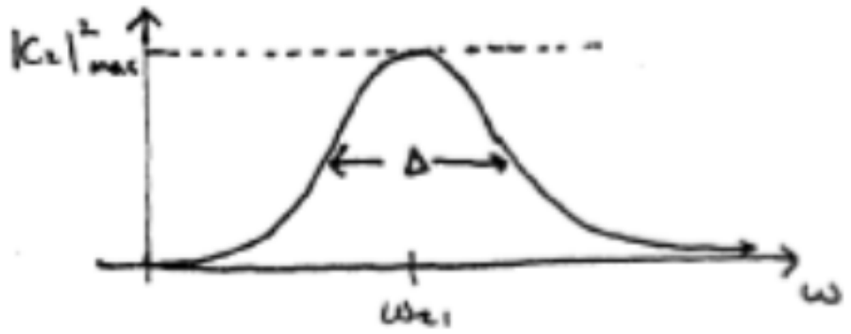
$$\Omega = \frac{\delta}{\hbar} \quad , \quad |c_1(t)|^2_{\min} = 0$$

as shown in the Figure.



At Resonance

The amplitude as a function of ω is shown in the Figure below.



Amplitude versus ω

where Δ = full width at half maximum = $4\delta/\hbar$.

The amplitude is peaked at resonance and the width is proportional to δ (the strength of the perturbation).

This periodically forced 2–state system is a basic problem - it demonstrates the fundamental features of absorption and emission.

We now return to the full, general equations and look for a perturbation solution.

Now we assume (power series)

$$a_n(t) = a_n^{(0)} + g a_n^{(1)} + g^2 a_n^{(2)} + \ldots\ldots$$

Substituting and arranging the terms in a power series in g we have

$$\left(\frac{da_n^{(0)}}{dt}\right) g^0 + \left(i\hbar \frac{da_m^{(1)}}{dt} - \sum_n \langle m|\hat{U}_t|n\rangle e^{\frac{i}{\hbar}\omega_{mn}t} a_n^{(0)}\right) g^1 + \ldots\ldots +$$

$$\left(i\hbar \frac{da_m^{(r+1)}}{dt} - \sum_n \langle m|\hat{U}_t|n\rangle e^{\frac{i}{\hbar}\omega_{mn}t} a_n^{(r)}\right) g^r + \ldots\ldots = 0$$

or looking at each order separately we have

$$0^{th}-\text{order} \quad \frac{da_n^{(0)}}{dt}=0$$

$$1^{st}-\text{order} \quad i\hbar\frac{da_m^{(1)}}{dt}=\sum_n \langle m|\hat{U}_t|n\rangle e^{\frac{i}{\hbar}\omega_{mn}t}a_n^{(0)}$$

.....

.....

$$(r+1)^{st}-\text{order} \quad i\hbar\frac{da_m^{(r+1)}}{dt}=\sum_n \langle m|\hat{U}_t|n\rangle e^{\frac{i}{\hbar}\omega_{mn}t}a_n^{(r)}$$

.....

Note that the coefficients $a_n^{(0)}$ follow from the initial condition

$$|\psi(0)\rangle=\sum_n a_n^{(0)}|n\rangle$$

The solution proceeds as follows:

$$\text{initial condition} \rightarrow a_n^{(0)}$$

$$a_n^{(0)} \rightarrow a_n^{(1)} \text{ using the } 1^{st}-\text{order equation}$$

.....

$$a_n^{(r)} \rightarrow a_n^{(r+1)} \text{ using the } (r+1)^{st}-\text{order equation}$$

Now consider the following example.

We assume that

$$\hat{H}=\hat{H}_0 \qquad t\le 0$$

where

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

and during the time interval $0 \le t \le T$ a perturbation $\hat{V}_t$ is applied to the system and the $a_n(t)$ change with time.

Finally, for $t \ge T$ the perturbation is turned off and $a_n(t) = a_n(T)$.

The probability that, as a result of the perturbation, the energy of the system becomes ε_r, is given by

$$|\langle r \mid \psi_t \rangle|^2 = \left| \sum_i a_i(t) e^{-\frac{i}{\hbar}\varepsilon_i t} \langle r \mid i \rangle \right|^2 = |a_r(t)|^2$$

and as $t \to \infty$ we get

$$|\langle r \mid \psi_t \rangle|^2 = |a_r(T)|^2$$

Now to 1st–order we have

$$i\hbar \frac{da_r^{(1)}}{dt} = \sum_n \langle r| \hat{U}_t |n\rangle e^{\frac{i}{\hbar}\omega_{rn} t} a_n^{(0)}$$

If $|\psi(0)\rangle = |i\rangle$, then

$$a_n^{(0)} = \begin{cases} 1 & n = i \\ 0 & n \neq i \end{cases}$$

This gives

$$i\hbar \frac{da_r^{(1)}}{dt} = \langle r| \hat{U}_t |i\rangle e^{\frac{i}{\hbar}\omega_{ri} t}$$

Integrating we have

$$a_r^{(1)}(T) = \frac{1}{i\hbar} \int_0^T \langle r| \hat{U}_t |i\rangle e^{\frac{i}{\hbar}(\varepsilon_r - \varepsilon_i)t} dt$$

and

$$a_r(T) = a_r^{(0)}(T) + g a_r^{(1)}(T)$$

which is identical to our earlier result as $t \to \infty$.

Now let return to our question.

Has the state changed also?

In the example we found that the perturbation produces a final state $|\psi_t\rangle$ for $t \geq T$ which to 1st–order is

$$|\psi_t\rangle = \sum_n a_n(t) e^{-\frac{i}{\hbar}\varepsilon_n t} |n\rangle$$

This is a coherent (definite relative phases) superposition of eigenvectors of $\hat{H}_0$.

This is NOT a stationary state.

Interference effects between the terms in the sum are detectable.

They do not, however, affect

$$|a_r(T)|^2 = \text{ probability that the energy changes to } \varepsilon_r$$

Thus, the perturbation does not cause a jump from one stationary state $|i\rangle$ of $\hat{H}_0$ to another $|r\rangle$, but instead it produces a non-stationary state.

The conventional language of quantum mechanics produces this ambiguity between the two statements

the energy is ϵ_r and *the state is* $|r\rangle$

For the state

$$|\psi_t\rangle = \sum_n a_n(t) e^{-\frac{i}{\hbar}\varepsilon_n t} |n\rangle$$

it is correct to say

the probability of the energy being ϵ_r is $|a_r(T)|^2$

or

$$Prob(E = \epsilon_r | |\psi\rangle) = |a_r(T)|^2$$

The state, however, is $|\psi_t\rangle$ and NOT $|r\rangle$.

An example

Suppose we perturb an oscillator with a decaying electric field of the form

$$\hat{V}_t = -q\mathcal{E}_0\hat{x}e^{-\frac{t}{\tau}} \qquad t \geq 0$$

To 1st–order, starting with the initial state $|n\rangle$ with energy

$$\varepsilon_n = \hbar\omega(n + \frac{1}{2})$$

we have

$$|\psi(t)\rangle = |n\rangle + \frac{1}{i\hbar}\int_0^t dt'\hat{V}(t)\,|n\rangle$$

where

$$\hat{V}(t) = e^{\frac{i}{\hbar}\hat{H}_0 t}\hat{V}_t e^{-\frac{i}{\hbar}\hat{H}_0 t}$$

We let $n = 0$ (the ground state) for this example.

We then have

$$|\psi(t)\rangle = |0\rangle + \frac{1}{i\hbar}\int_0^t dt' e^{\frac{i}{\hbar}\hat{H}_0 t}(-q\mathcal{E}_0\hat{x})e^{-\frac{t}{\tau}}e^{-\frac{i}{\hbar}\hat{H}_0 t}\,|0\rangle$$

Using

$$e^{-\frac{i}{\hbar}\hat{H}_0 t}\,|0\rangle = e^{-i\frac{\omega}{2}t}\,|0\rangle$$

$$\hat{x}\,|0\rangle = \sqrt{\frac{\hbar}{2m\omega}}\,|1\rangle$$

$$e^{\frac{i}{\hbar}\hat{H}_0 t}\,|1\rangle = e^{i\frac{3\omega}{2}t}\,|0\rangle$$

we get (letting $t \to \infty$)

$$|\psi(t)\rangle = |0\rangle + \frac{1}{i\hbar}(-q\mathcal{E}_0)\sqrt{\frac{\hbar}{2m\omega}}\int_0^\infty dt' e^{i\omega t' - \frac{t}{\tau}}|1\rangle$$

and finally,

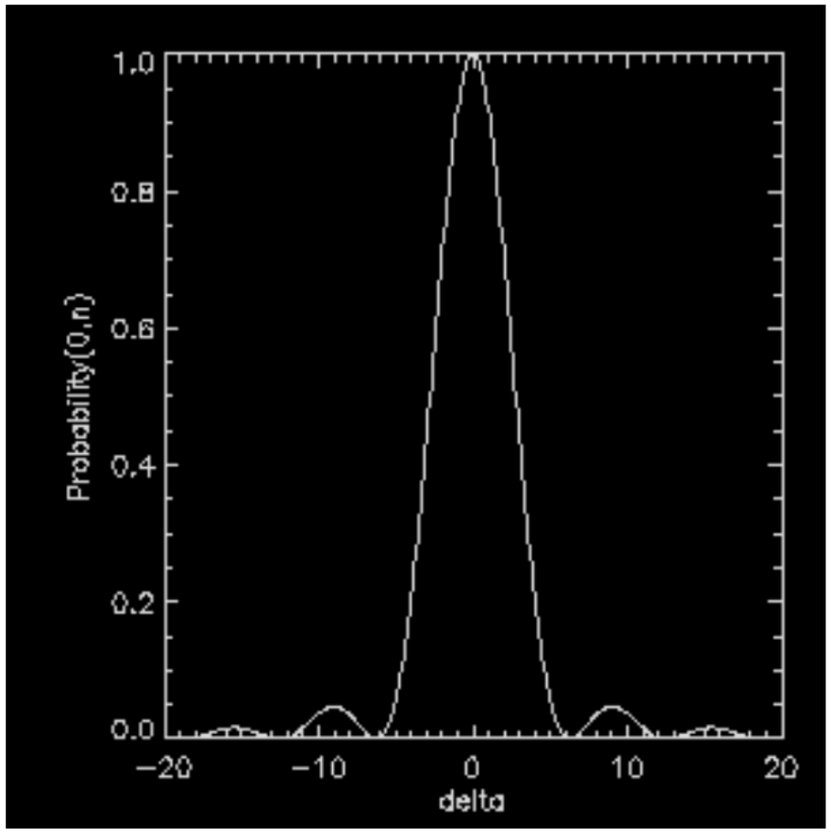
$$P_{0\to 1} = |\langle 1 \mid \psi(t)\rangle|^2 = \frac{q^2\mathcal{E}_0^2}{2m\hbar\omega}\left|\int_0^\infty dt' e^{i\omega t' - \frac{t}{\tau}}\right|^2$$

$$= \frac{q^2\mathcal{E}_0^2}{2m\hbar\omega}\frac{\tau^2}{\tau^2\omega^2 + 1}$$

We now return to the earlier general result we derived for the probability, namely,

$$P_{0\to n}(t) = \left|\langle n|\hat{V}_{t'}|0\rangle\right|^2\left(\frac{\sin\frac{\Delta t}{2\hbar}}{\Delta/2}\right)^2$$

In the Figure we plot this function.



Probability(0,n) versus Delta

The height of the central peak is proportional to t^2 and the location of the first zero is at

$$\Delta = \frac{2\pi\hbar}{t}$$

so that the width of the peak decreases as $1/t$.

The formula implies that for very short times

$$P_{0\to n} \propto t^2 \text{ for all } \varepsilon_n$$

As $t \to \infty$, however, the probability is largest for those states whose energy lies under the sharp bump near $\Delta = 0$ or those states with whose energy lies under the peak around ε_0.

Now the energy $\varepsilon_n \approx \varepsilon_0$ lies under the sharp bump when

$$|\Delta| = |\varepsilon_n - \varepsilon_0| < \frac{2\pi\hbar}{t}$$

The area under the bump is proportional to t and the rest of the area oscillates in time around zero.

This latter feature means that if $\varepsilon_n \neq \varepsilon_0$, the transition probability oscillates in time with a repetition time of

$$\frac{2\pi\hbar}{|\varepsilon_n - \varepsilon_0|}$$

The case, where we are looking for a transition to a single state, is, thus, only valid in perturbation theory for very small time t.

Otherwise the condition that the

$$P_{0\to 0}(t) \approx 1$$

will not be true and perturbation theory breaks down.

We also note that the probability cannot grow larger than one or that, after a while, the higher-order effects of the perturbation which we have neglected so far must become important and prevent the probability from exceeding one.

The condition that tells us whether a transition probability to a state with an energy appreciably different than the original energy is the same condition in time-independent perturbation theory that tells whether the state vector changes appreciably from the unperturbed state, namely

$$\left| \frac{\langle n | \hat{V} | 0 \rangle}{\varepsilon_n - \varepsilon_0} \right| << 1$$

Physically, a more interesting case occurs when the state $|n\rangle$ is one of a continuum of energy states, or it lies in a group of very closely spaced levels.

In this case we ask a different experimental question, namely,

> What is the probability that the system makes a transition to a small group of states near $|n\rangle$ (or has energy near ϵ_n)?

Since the area under the bump near $\Delta = 0$ or $\varepsilon_n \approx \varepsilon_0$ is proportional to t, we expect that the transition probability to a small group of states near ε_0 will grow linearly with t and thus

$$\frac{P_{0\to n}(t)}{t} = \text{ transition rate } = \Gamma = \text{ constant as } t \to \infty$$

Quantities that we measure are related to the transition rate and this result says that these measurements will make sense.

Let us now carry out this derivation in detail.

To calculate this transition rate we must sum $P_{0\to n}$ over the group of final states.

We assume that $|\langle b|\ \hat{V}_{t'}\ |0\rangle|^2$ is relatively constant over the small group of states near $|n\rangle$ (has a weak energy dependence).

We then have

$$\sum_{\substack{n \\ in group}} P_{0\to n}(t) = \left|\langle n|\,\hat{V}_{t'}\,|0\rangle\right|^2 \int_{group} d\varepsilon_n \rho(\varepsilon_n) \left[\frac{\sin\left[\frac{(\varepsilon_n-\varepsilon_0)t}{\hbar}\right]}{\left[\frac{(\varepsilon_n-\varepsilon_0)}{2}\right]}\right]^2$$

where

$\rho(\varepsilon_n)$ = number of states per unit energy

$\rho(\varepsilon_n)d\varepsilon_n$ = number of states in the interval $d\varepsilon_n$

Now in the limit $t\to\infty$

$$\left[\frac{\sin\left[\frac{(\varepsilon_n-\varepsilon_0)t}{\hbar}\right]}{\left[\frac{(\varepsilon_n-\varepsilon_0)}{2}\right]}\right]^2 \to \frac{2\pi t}{\hbar}\delta(\varepsilon_0-\varepsilon_n)$$

i.e., in general, for a sequence of functions

$$\delta_t(\alpha) = \frac{\sin^2 \alpha t}{\pi\alpha^2 t}$$

we have that

$$\delta_t(\alpha) = \begin{cases} \frac{t}{\pi} & \alpha = 0 \\ \leq \frac{1}{\pi\alpha^2 t} & \alpha \neq 0 \end{cases}$$

and

$$\lim_{t\to\infty} \int_{-\infty}^{\infty} d\alpha \delta_t(\alpha) F(\alpha) = F(0)$$

Therefore,

$$\lim_{t\to\infty} \delta_t(\alpha) = \delta(\alpha)$$

Using this result, we have

$$\sum_n P_{0\to n}(t) = \Gamma t$$

and thus

$$\Gamma = \text{ transition rate } = \frac{2\pi}{\hbar} \left|\langle n|\, \hat{V}\, |0\rangle\right|^2 \rho(\varepsilon_n)_{\varepsilon_n=\varepsilon_0}$$

which is called *Fermi's Golden Rule*.

We now consider a perturbation that depends explicitly on time.

In particular, suppose we have a harmonic perturbation of the form

$$\hat{V}_t = \hat{V} e^{\eta t} \cos\omega t = \frac{\hat{V}}{2} e^{\eta t} \left(e^{-i\omega t} + e^{i\omega t}\right)$$

and $|\psi(t_0)\rangle = |0\rangle$, where we let $t_0 \to -\infty$.

The $e^{\eta t}$ factor is necessary to make the mathematical operations valid in the limit.

It is equivalent for small η to turning the perturbation on slowly.

In the end we will let $\eta \to 0$.

We have

$$\langle n \mid \psi(t)\rangle = \langle n \mid 0\rangle + \frac{1}{i\hbar} e^{-\frac{i}{\hbar}\varepsilon_n t} \int_{-\infty}^{0} dt' \, \langle n|\hat{V}(t')\,|0\rangle$$

$$= \frac{e^{\eta t}}{2} \left[\frac{e^{i(\varepsilon_n - \varepsilon_0 - \hbar\omega)\frac{t}{\hbar}}}{\varepsilon_0 - \varepsilon_n + \hbar\omega + i\hbar\eta} + \frac{e^{i(\varepsilon_n - \varepsilon_0 + \hbar\omega)\frac{t}{\hbar}}}{\varepsilon_0 - \varepsilon_n - \hbar\omega + i\hbar\eta} \right] \langle n|\, \hat{V}\, |0\rangle$$

Thus, the probability of the energy being ε_n at the time t is

$$|\langle n \mid \psi(t)\rangle|^2$$
$$= \frac{e^{2\eta t}}{4} \left|\langle n| \hat{V} |0\rangle\right|^2 \left\{ \begin{array}{c} \frac{1}{(\varepsilon_0-\varepsilon_n+\hbar\omega)^2+(\hbar\eta)^2} + \frac{1}{(\varepsilon_0-\varepsilon_n-\hbar\omega)^2+(\hbar\eta)^2} \\ +2\mathrm{Re}\frac{e^{-2i\omega t}}{(\varepsilon_0-\varepsilon_n+\hbar\omega+i\hbar\eta)(\varepsilon_0-\varepsilon_n-\hbar\omega+i\hbar\eta)} \end{array} \right\}$$

The first term comes from the $e^{-i\omega t}$ part of $\hat{V}_t$ (positive frequency) and the second term comes from the $e^{i\omega t}$ part of $\hat{V}_t$ (negative frequency).

The last term represents *interference effects*.

Since

$$P_{0\to n}(t) = \Gamma_{0\to n} t$$

we have

$$\Gamma_{0\to n} = \frac{dP_{0\to n}(t)}{dt}$$

and thus

$$\Gamma_{0\to n}$$
$$= \frac{e^{2\eta t}}{4} \left|\langle n| \hat{V} |0\rangle\right|^2 \left\{ \begin{array}{c} \left[\frac{1}{(\varepsilon_0-\varepsilon_n+\hbar\omega)^2+(\hbar\eta)^2} + \frac{1}{(\varepsilon_0-\varepsilon_n-\hbar\omega)^2+(\hbar\eta)^2}\right](1-\cos 2\omega t) \\ +2\sin 2\omega t\left[\frac{e^{-2i\omega t}}{(\varepsilon_0-\varepsilon_n+\hbar\omega+i\hbar\eta)(\varepsilon_0-\varepsilon_n-\hbar\omega+i\hbar\eta)}\right] \end{array} \right\}$$

The sin ωt and cos ωt terms arise from the interference term.

In the limit $\eta \to 0$ and assuming that $|n\rangle$ is in the continuum part of spectrum, we have

1. the first two terms are not equal to zero only if $\varepsilon_n - \varepsilon_0 = \pm\hbar\omega$

2. the $\sin\omega t$ and $\cos\omega t$ terms average to zero if we assume that $\Gamma_{0\to n}$ is $dP_{0\to n}(t)/dt$ averaged over a few cycles of $\hat{V}_t$

which gives the result

$$\Gamma_{0\to n} = \frac{2\pi}{\hbar}\frac{\left|\langle n|\hat{V}|0\rangle\right|^2}{4}\left[\delta(\varepsilon_n - \varepsilon_0 - \hbar\omega) + \delta(\varepsilon_n - \varepsilon_0 + \hbar\omega)\right]$$

The positive and negative frequency parts act independently and the interference averages to zero.

Thus, the $e^{-i\omega t}$ part produced a $\Delta E > 0$ process (absorption), while the $e^{i\omega t}$ part produced a $\Delta E < 0$ process (emission).

To enhance our understanding of time-dependent perturbation theory, we look at a variation of this harmonic perturbation. Suppose we have a harmonic perturbation of the form

$$\hat{V}_t = \hat{V}e^{-i\omega t} + \hat{V}^+ e^{i\omega t}$$

which is only applied for a finite time interval $0 \le t \le T$.

If we start with energy ε_i in the state $|i\rangle$, then at any time $t \ge T$

$$a_f^{(1)}(t) = 1^{st} - \text{order amplitude for the state } |\psi(t)\rangle$$

$$\text{to have energy } \varepsilon_f \text{ (be in state } |f\rangle ?) \quad f \neq i$$

is given by

$$a_f^{(1)}(T) = \frac{1}{i\hbar}\langle f|\hat{V}|i\rangle\int_0^T e^{i(\omega_{fi}-\omega)t}dt + \frac{1}{i\hbar}\langle f|\hat{V}^+|i\rangle\int_0^T e^{i(\omega_{fi}+\omega)t}dt$$

and

$$\left|a_f^{(1)}(T)\right|^2 = \text{ probability that the final energy will be } \epsilon_f$$

As an example we consider spin resonance (we solved this problem exactly earlier).

We consider a spin = 1/2 particle in a static magnetic field B_0 (in the z–direction).

This says that the unperturbed Hamiltonian is

$$\hat{H}_0 = -\frac{1}{2}\hbar\gamma B_0 \hat{\sigma}_z$$

This operator has the eigenvectors and eigenvalues

$$|+\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \; |-\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad , \quad \epsilon_\pm = \mp\frac{1}{2}\hbar\gamma B_0$$

We now perturb the system with another magnetic field B_1, which is rotating in the $x-y$ plane with angular velocity ω.

This implies that

$$\begin{aligned} \hat{V}_t &= -\frac{1}{2}\hbar\gamma B_1 (\cos\omega t \hat{i} + \sin\omega t \hat{j}) \cdot \hat{\sigma} \\ &= -\frac{1}{2}\hbar\gamma B_1 \left[\hat{\sigma}_x \cos\omega t + \hat{\sigma}_y \sin\omega t\right] = -\frac{1}{2}\hbar\gamma B_1 \begin{pmatrix} 0 & e^{-i\omega t} \\ e^{i\omega t} & 0 \end{pmatrix} \\ &= -\frac{1}{2}\hbar\gamma B_1 \left(\hat{V} e^{-i\omega t} + \hat{V}^+ e^{i\omega t}\right) \end{aligned}$$

where

$$\hat{V} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \quad , \quad \hat{V}^+ = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$$

We choose the initial state to be

$$|i\rangle = |n\rangle \;\; \text{(spin up in the z-direction)}$$

The first order perturbation theory approximation for this result is

$$\left|a_f^{(1)}(T)\right|^2 = \left(\frac{\omega_1}{\omega_0+\omega}\right)^2 \sin^2\frac{1}{2}(\omega_0+\omega)T$$

where

$$\omega_0 = \gamma B_0 \quad , \quad \omega_1 = \gamma B_1$$

and we have used

$$\langle -|\,\hat{V}\,|+\rangle = 0,\ \langle -|\,\hat{V}^+\,|+\rangle = -\frac{1}{2}\hbar\gamma B_1$$

$$\omega_{fi} = \frac{\varepsilon_f - \varepsilon_i}{\hbar} = \omega_0$$

When is the first order perturbation theory result valid?

If we compare the exact result with perturbation theory by expanding the exact result in a power series, we find that the two results agree exactly if

$$\left|\frac{\omega_1}{\omega_0+\omega}\right| << 1$$

which corresponds to $B_1 \ll B_0$ (as long as $\omega_0 + \omega \neq 0$).

When $\omega_0 + \omega = 0$ we have a phenomenon called resonance.

The exact solution gives

$$|a_f(T)|^2 = \sin^2\frac{1}{2}\alpha T$$

where $\alpha^2 = (\omega_0 + \omega)^2 + \omega_1{}^2$ and perturbation theory gives

$$\left|a_f^{(1)}(T)\right|^2 = \left(\frac{\omega_1 T}{2}\right)^2$$

Thus, the results agree only if $|\omega_1 T| \ll 1$ or if the perturbation only acts for a short time.

stop

Atomic Radiation and Selection Rules

We now apply time-independent perturbation theory to the absorption and emission electromagnetic radiation by matter.

The Hamiltonian for an electron in an atom interacting with and electromagnetic field is

$$\hat{H} = \frac{\left(\vec{p}_{op} - \frac{q}{c}\vec{A}\right)^2}{2m_e} + q\phi + U$$

where $q = -e$, $U =$ the potential energy function that binds the electrons in the atom, and $\vec{A}$ and ϕ are the vector and scalar potentials associated with the electromagnetic field.

These potentials imply the electric and magnetic fields

$$\vec{E} = -\nabla\phi - \frac{1}{c}\frac{\partial \vec{A}}{\partial t}, \vec{B} = \nabla \times \vec{A}$$

Note that if $\vec{A} = 0$, the $\vec{B} = 0$ and

$$\phi = -\int_0^{\vec{r}} \vec{E}(\vec{r}, t) \cdot d\vec{r}$$

We rewrite $\hat{H}$ as

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

where

$$\hat{H}_0 = \frac{\vec{p}_{op}^2}{2m_e} + U \; = \; \text{Hamiltonian for the atom with no electromagnetic field}$$

and

$$\hat{V} = \frac{q}{2m_e c}\left(\vec{p}_{op} \cdot \vec{A} + \vec{A} \cdot \vec{p}_{op}\right) + \frac{q^2}{2m_e c^2}\left(\vec{A} \cdot \vec{A}\right) + q\phi$$

is the perturbation due to the presence of the electromagnetic field, i.e., the term $\hat{V}$ tells us how the atom interacts with the electromagnetic field.

The Electric Dipole Approximation

The typical wavelength of visible electromagnetic radiation is ≈ 5000 angstroms and the typical dimension of an atom is ≈ a few .

This implies that the electromagnetic fields are approximately constant over the volume of the atom.

In Gaussian units $|\vec{E}| \approx |\vec{B}|$, but the force due to $\vec{B} \approx (v/c) \times$ the force due to $\vec{E}$.

Thus, magnetic effects are negligible in most atoms compared to electric effects.

We therefore assume

1. $\vec{E} \approx$ constant over the volume of the atom
2. $\vec{B}$ can be neglected

This is the so-called electric dipole approximation.

In this approximation, we have A = 0 and as we said above, B = 0 and

$$\phi = -\int_0^{\vec{r}} \vec{E}(\vec{r}, t) \cdot d\vec{r}$$

This last integral is independent of path since

$$\nabla \times \vec{E} = -\frac{1}{c}\frac{\partial \vec{B}}{\partial t} = 0$$

Since we are assuming that $\vec{E} \approx$ constant over the volume of the atom, we get

$$\phi = -\vec{r} \cdot \vec{E}(t)$$

This adds a perturbation of the form

$$\hat{V} = -q\vec{r} \cdot \vec{E}(t)$$

to the Hamiltonian.

In some derivations the quantity

$$\hat{V} = \frac{q}{2m_e c}\left(\vec{p}_{op} \cdot \vec{A} + \vec{A} \cdot \vec{p}_{op}\right) + \frac{q^2}{2m_e c^2}\left(\vec{A} \cdot \vec{A}\right) + q\phi$$

is chosen as the perturbation.

The next step would be to expand in powers of the potentials.

Since $\nabla \cdot \vec{E} = 0$ for the radiation field, we can choose $\phi = 0$ and $\nabla \cdot \vec{A} = 0$ (if these relations were not true we could make them so with a gauge transformation).

Therefore, we end up with an expansion in powers of $\vec{A}$.

The first term in the expansion is

$$\hat{V}_{Ap} = -\frac{q}{2m_e c}\left(\vec{p}_{op} \cdot \vec{A} + \vec{A} \cdot \vec{p}_{op}\right) = -\frac{q}{m_e c}\vec{A} \cdot \vec{p}_{op}$$

since

$$\vec{p}_{op} \cdot \vec{A} - \vec{A} \cdot \vec{p}_{op} \propto \nabla \cdot \vec{A} = 0$$

In this case, in the electric dipole approximation

$$\vec{E} \approx \text{ constant} \quad , \quad \vec{B} = 0$$

$$\vec{A}(\vec{r}, t) = -c \int_0^{\vec{r}} \vec{E}(\vec{r}, t) dt$$

$$\phi(\vec{r}, t) = 0$$

If $\vec{E}(\vec{r},t) = \vec{E}(\vec{r})e^{-i\omega t}$, then $\vec{A} = e\vec{E}/i\omega$.

Using

$$\frac{\vec{p}_{op}}{m} = \frac{i}{\hbar}\left[\hat{H}_0, \vec{r}_{op}\right]$$

we get

$$\hat{V}_{Ap} = -\frac{q}{\hbar\omega}\left[\hat{H}_0, \vec{r}_{op}\right] \cdot \vec{E}$$

We can then calculate matrix elements

$$\begin{aligned}
\langle m|\, \hat{V}_{Ap}\, |n\rangle &= -\frac{q}{\hbar\omega}\, \langle m|\, \hat{H}_0 \vec{r}_{op} - \vec{r}_{op}\hat{H}_0\, |n\rangle \cdot \vec{E} \\
&= -\frac{q}{\hbar\omega}\, (\varepsilon_m - \varepsilon_n)\, \langle m|\, \vec{r}_{op} \cdot \vec{E}\, |n\rangle \\
&= \frac{\omega_{mn}}{\omega}\, \langle m|\, \hat{V}\, |n\rangle
\end{aligned}$$

where

$$\omega_{mn} = \frac{\varepsilon_m - \varepsilon_n}{\hbar}$$

Thus, the matrix elements of $\hat{V}_{A_p}$ (from $\vec{A}\cdot\vec{p}_{op}$) differ from the matrix elements of $\hat{V}$ (from ϕ) by the factor ω_{mn}/ω.

This implies different transition probabilities in first order except at resonance where $\omega_{mn}/\omega = 1$.

The reason for the differences is as follows:

1. we assumed that the perturbation = 0 for $t \le 0$, $t \ge T$
2. there is no physical problem with $V = -q\vec{r} \cdot \vec{E}$ changing discontinuously
3. however, if $\vec{A} \cdot \vec{p}_{op}$ changes discontinuously, then the relation

$$\vec{E} = -\frac{1}{c}\frac{\partial \vec{A}}{\partial t}$$

generates spurious δ–function type $\vec{E}$ fields

It is clear that one must exercise great care in choosing a starting point for perturbation theory.

Induced Emission and Absorption

We now look at the physics connected with the harmonic perturbation.

Physically, it represents an electromagnetic wave interacting with the atom.

We consider the perturbing potential

$$\hat{V} = \begin{cases} -q\vec{r} \cdot \vec{E}_0 (e^{-i\omega t} + e^{i\omega t}) & 0 < t < T \\ 0 & t < 0, t > T \end{cases}$$

where $\vec{E}_0$ = a constant vector which tells us the strength and polarization of the electromagnetic field.

This perturbing potential corresponds to monochromatic(single wavelength) electromagnetic radiation.

For an initial state $|\psi(0)\rangle = |i\rangle$ where $\hat{H}_0 |i\rangle = \varepsilon_i |i\rangle$ the probability, at any time $t \ge T$, that the atom will have a final energy ε_f is

$$P_{i \to f}(T) = \left| a_f^{(1)}(T) \right|^2$$

to first order, where we have from our earlier derivation

$$a_f^{(1)}(T) = \frac{\langle f|\left(-q\vec{r}\cdot\vec{E}_0\right)|i\rangle}{\hbar}\left[\frac{1-e^{i(\omega_{fi}-\omega)T}}{\omega_{fi}-\omega}+\frac{1-e^{i(\omega_{fi}+\omega)T}}{\omega_{fi}+\omega}\right]$$

If $\varepsilon_f > \varepsilon_i$ this gives the probability amplitude for absorbing radiation and if $\varepsilon_f < \varepsilon_i$ this gives the probability amplitude for emitting radiation.

In the limit $T \to \infty$, as we saw earlier, $P_{i\to f}(t) = \Gamma_{i\to f}\, t$ which gives

$$\Gamma_{i\to f} = \frac{2\pi q^2}{\hbar}\frac{\left|\langle f|\vec{r}\cdot\vec{E}_0|i\rangle\right|^2}{4}\left[\delta(\varepsilon_i-\varepsilon_f-\hbar\omega)+\delta(\varepsilon_i-\varepsilon_f+\hbar\omega)\right]$$

In this expression

$$\delta(\varepsilon_i-\varepsilon_f-\hbar\omega) \to \text{ absorption}$$
$$\delta(\varepsilon_i-\varepsilon_f+\hbar\omega) \to \text{ emission}$$

We then have the transition rates

$$\Gamma^a_{i\to f} = \Gamma^e_{i\to f} = \frac{2\pi q^2}{\hbar}\frac{\left|\langle f|\vec{r}\cdot\vec{E}_0|i\rangle\right|^2}{4}$$

These expressions are *zero/nonzero* or transitions are *not-allowed/allowed* depending on the matrix element of the perturbation, i.e.,

These relationships between the quantum numbers of the initial and final states that tell us whether or not a transition is allowed are called *selection rules*.

To determine the selection rules for one-electron atoms we only need to consider matrix elements of the form

$$\langle n'\ell' m_{\ell'} m_{s'} | \vec{r} \cdot \vec{E}_0 | n\ell m_\ell m_s \rangle$$

or we need to look at three matrix elements, namely,

$$\langle n'\ell' m_{\ell'} m_{s'} | \hat{x} | n\ell m_\ell m_s \rangle, \langle n'\ell' m_{\ell'} m_{s'} | \hat{y} | n\ell m_\ell m_s \rangle \text{ and } \langle n'\ell' m_{\ell'} m_{s'} | \hat{z} | n\ell m_\ell m_s \rangle$$

Now

$$z = r\cos\theta \propto rY_{10} = x_3$$
$$x = r\sin\theta\cos\varphi \propto r(Y_{11} + Y_{1,-1}) = x_1$$
$$y = r\sin\theta\sin\varphi \propto r(Y_{11} - Y_{1,-1}) = x_2$$

A typical matrix element, therefore, will have a term like the following:

$$\langle sm_{s'} \mid sm_s \rangle \int R_{n'\ell'}(r) R_{n\ell}(r) r^3 dr \int Y^*_{\ell' m\ell'} Y_{\ell m_\ell} Y_{1m} d\Omega$$

where $m = \pm 1, 0$.

The radial integral would equal zero only by accident implying that it is not part of the general selection rules, which must come from the other terms.

The term $\langle sm_{s'} \mid sm_s \rangle = \delta_{ms'\ ms}$ gives us a simple selection rule (this simple rule arises here because the interaction does not depend on spin).

We have

$$m_{s'} = m_s \rightarrow \text{ the } \Delta m_s = 0 \qquad \textbf{SELECTION RULE}$$

The rest of the selection rules come from the angular integration terms

$$\int Y^*_{\ell' m_{\ell'}} Y_{\ell m_\ell} Y_{1m} d\Omega$$

This integral equals 0 unless $\ell + \ell' + 1 =$ an even integer.

This rule follows from parity considerations.

For any angular integration over all angles to be nonzero, the integrand must be even under the parity operation.

Now

$$Y_{\ell m} \to (-1)^{\ell} Y_{\ell m}$$

under the parity operation.

Therefore

$$Y^*_{\ell' m \ell'} Y_{\ell m_\ell} Y_{1m} \to (-1)^{\ell+\ell'+1} Y^*_{\ell' m \ell'} Y_{\ell m_\ell} Y_{1m}$$

which gives the stated rule.

The θ–integration says that we must have

$$|\ell' - \ell| \leq 1 \leq \ell' + \ell$$

This corresponds to thinking of the integrand as made up of two states, namely,

$$\langle \ell' m_{\ell'} | \left(|\ell m_\ell\rangle \otimes |1m\rangle \right)$$

Our angular momentum addition rules say that

$$|\ell m_\ell\rangle \otimes |1m\rangle = |\ell+1\rangle \oplus |\ell\rangle \oplus |\ell-1\rangle$$

and the selection rule then follows from the orthogonality condition.

The two ℓ –rules when combined imply the selection rule

$$\Delta m_\ell = m_{\ell'} - m_\ell = \pm 1, 0 \qquad \textbf{SELECTION RULE}$$

Thus, for transitions within the electric dipole approximation, as defined above, we have the SELECTION RULES

$$\Delta m_\ell = m_{\ell'} - m_\ell = \pm 1\,,\, 0$$
$$\Delta \ell = \ell' - \ell = \pm 1$$
$$\Delta m_s = 0 = \Delta s$$

The derivation is more complex for multi-electron atoms due the complexity of the wave function, but it can be shown that, in general, the SELECTION RULES in the electric dipole approximation are

1. parity changes
2. $\Delta(\sum \ell_i) = \pm 1$
3. $\Delta S = 0$
4. $\Delta M_S = 0$
5. $\Delta L = \pm 1, 0$ (possibility of 0 which is not allowed for 1-electron atoms)
6. $\Delta M_L = \pm 1, 0$
7. $\Delta J = \pm 1, 0$
8. $\Delta M_J = \pm 1, 0$
9. $J = 0 \rightarrow J = 0$ is strictly forbidden

A Real Physical Process - Ionization

We now carry out the calculations of the transition rate for the ionization(transition to the continuum) of hydrogen by electromagnetic radiation.

The initial state of the electron is the ground state of hydrogen

$$|i\rangle = |100\rangle \text{ with energy } \varepsilon_i = \varepsilon_{100} = -\frac{e^2}{2a_0}$$

The final state of the electron in the ionization process is a free particle state(ionized electron)

$$|f\rangle = \left|\vec{k}\right\rangle \text{ with energy } \varepsilon_f = \varepsilon_k = \frac{\hbar^2 k^2}{2m}$$

We then have

$$\vec{p}_{op}\left|\vec{k}\right\rangle = \hbar\vec{k}\left|\vec{k}\right\rangle$$

which says this is a momentum eigenstate also (for free particles $[\hat{H}, \vec{p}_{op}] = 0$ and momentum and energy have the same eigenstates).

The momentum is given by $\vec{p} = \hbar\vec{k}$.

Since $\varepsilon_f > \varepsilon_i$ this is an absorption process and thus the transition rate is given by

$$\Gamma_{0\to\vec{k}} = \text{transitionrate} = \lim_{t\to\infty} \frac{P_{0\to\vec{k}}(t)}{t}$$

$$= \frac{2\pi}{\hbar}\left|\left\langle\vec{k}\right|\hat{V}\left|100\right\rangle\right|^2 \delta(\varepsilon_k - \varepsilon_{100} - \hbar\omega)$$

where ω = frequency of the electromagnetic radiation and

$\hat{V} = -e\vec{r}\cdot\vec{\mathcal{E}}$ (electric dipole approximation)

$\vec{\mathcal{E}}$ = the electric field vector

We define

$$d\Gamma = \text{ rate of transition into a small solid angle } d\Omega$$
$$= \sum_{\vec{k} \text{ in } d\Omega} \Gamma_{0\to\vec{k}}$$

For convenience we use a common trick and assume that the universe is a large box (side = L, volume = L^3).

This allows us to normalize the plane wave states associated with the free electron.

We have

$$\langle \vec{r} \mid \vec{k} \rangle = Ae^{i\vec{k}\cdot\vec{r}}$$
$$\langle \vec{k} \mid \vec{k} \rangle = 1 = \int d^3\vec{r} \langle \vec{k} \mid \vec{r} \rangle \langle \vec{r} \mid \vec{k} \rangle$$
$$= A^2 \int d^3\vec{r} e^{-i\vec{k}\cdot\vec{r}} e^{i\vec{k}\cdot\vec{r}} = A^2 \int d^3\vec{r} = A^2 L^3$$

or

$$A = \frac{1}{L^{3/2}}$$

Now there are

$$L^3 \frac{d^3\vec{k}}{(2\pi)^3} = \frac{L^3}{(2\pi)^3} d\Omega k^2 dk = \frac{L^3 mk}{(2\pi)^3 \hbar^2} d\Omega d\varepsilon_k$$

states in the volume $d^3\vec{k}$ of phase space.

This implies that there are

$$\frac{L^3 mk}{(2\pi)^3 \hbar^2}$$

states per unit energy per unit solid angle.

Therefore

$$d\Gamma = \sum_{\vec{k} \text{in} d\Omega} \Gamma_{0\to\vec{k}} \to d\Omega \int_0^\infty \frac{L^3 mk}{(2\pi)^3 \hbar^2} d\varepsilon_k \frac{\pi e^2}{2\hbar} \left|\left\langle \vec{k} \right| \vec{r} \cdot \vec{\mathcal{E}} \left| 100 \right\rangle\right|^2 \delta(\varepsilon_k - \varepsilon_{100} - \hbar\omega)$$

Doing the integration (using the delta function) we get

$$d\Gamma = d\Omega \frac{L^3 mke^2}{16\pi^2 \hbar^3} \left|\left\langle \vec{k} \right| \vec{r} \cdot \vec{\mathcal{E}} \left| 100 \right\rangle\right|^2$$

where

$$\varepsilon_k = \varepsilon_{100} + \hbar\omega = \frac{\hbar^2 k^2}{2m}$$

$$k = \left(\frac{2m\omega}{\hbar} - \frac{1}{a_0^2} \right)^{1/2}$$

Evaluation of the Matrix Element

We have

$$\begin{aligned}
\left\langle \vec{k} \right| \vec{r} \cdot \vec{\mathcal{E}} \left| 100 \right\rangle &= \int d^3\vec{r} \left\langle \vec{k} \mid \vec{r} \right\rangle \left\langle \vec{r} \right| \vec{r} \cdot \vec{\mathcal{E}} \left| 100 \right\rangle \\
&= \frac{1}{L^{3/2}} \int d^3\vec{r} e^{-i\vec{k}\cdot\vec{r}} \vec{r} \cdot \vec{\mathcal{E}} \left\langle \vec{r} \mid 100 \right\rangle \\
&= \frac{1}{L^{3/2}} \int d^3\vec{r} e^{-i\vec{k}\cdot\vec{r}} \vec{r} \cdot \vec{\mathcal{E}} R_{10}(r) Y_{00} \\
&= \frac{2}{a_0^{3/2} L^{3/2}} \int d^3\vec{r} e^{-i\vec{k}\cdot\vec{r}} e^{-\frac{r}{a_0}} \vec{r} \cdot \vec{\mathcal{E}}
\end{aligned}$$

We can arbitrarily choose $\vec{k} = k\hat{z}$ which gives

$$\left\langle \vec{k} \right| \vec{r} \cdot \vec{\mathcal{E}} \left| 100 \right\rangle = \frac{2}{a_0^{3/2} L^{3/2}} \int d\Omega dr r^2 e^{-ikr\cos\theta - \frac{r}{a_0}} \vec{r} \cdot \vec{\mathcal{E}}$$

using $\vec{k} \cdot \vec{R} = kz = kr\cos\sigma$.

Now in spherical-polar coordinates we have

$$\vec{\mathcal{E}} = \mathcal{E}\left(\sin\theta_\varepsilon \cos\varphi_\varepsilon \hat{e}_x + \sin\theta_\varepsilon \sin\varphi_\varepsilon \hat{e}_z + \cos\theta_\varepsilon \hat{e}_z\right)$$
$$\vec{r} = r\left(\sin\theta\cos\varphi\hat{e}_x + \sin\theta\sin\varphi\hat{e}_y + \cos\theta\hat{e}_z\right)$$

so that

$$\vec{\mathcal{E}}\cdot\vec{r} = \mathcal{E}r\left(\cos\theta\cos\theta_\varepsilon + \sin\theta\sin\theta_\varepsilon\cos(\varphi-\varphi_\varepsilon)\right)$$

We then have

$$\left\langle\vec{k}\middle|\vec{r}\cdot\vec{\mathcal{E}}\middle|100\right\rangle$$
$$= \frac{2\varepsilon}{a_0^{3/2}L^{3/2}}\int_0^{2\pi} d\varphi \int_0^{\pi}\sin\theta d\theta \int dr r^3 e^{-ikr\cos\theta-\frac{r}{a_0}}\left[\cos\theta\cos\theta_\varepsilon + \sin\theta\sin\theta_\varepsilon\cos(\varphi-\varphi_\varepsilon)\right]$$

Since

$$\int_0^{2\pi} d\varphi\cos\varphi = \int_0^{2\pi} d\varphi\sin\varphi = 0$$

the $\phi-$ integration wipes out the $\cos(\phi-\phi_\varepsilon)$ term.

Letting $x = \cos\theta$ we then have

$$\left\langle\vec{k}\middle|\vec{r}\cdot\vec{\mathcal{E}}\middle|100\right\rangle = \frac{4\pi\mathcal{E}\cos\theta_\varepsilon}{a_0^{3/2}L^{3/2}}\int_{-1}^{1}\left[\int_0^{\infty} dr r^3 e^{-ikrx-\frac{r}{a_0}}\right]x dx$$

Now

$$\int_0^{\infty} dr r^3 e^{-ikrx-\frac{r}{a_0}} = \int_0^{\infty} dr r^3 e^{-\alpha r} = \frac{3!}{\alpha^4} = \frac{6}{\left(ikx+\frac{1}{a_0}\right)^4}$$

Therefore, we have

$$\langle \vec{k} | \vec{r} \cdot \vec{\mathcal{E}} | 100 \rangle = \frac{4\pi \mathcal{E} \cos\theta_\varepsilon}{a_0^{3/2} L^{3/2}} \int_{-1}^{1} \frac{6x}{\left(ikx + \frac{1}{a_0}\right)^4} dx$$

$$= \frac{4\pi \mathcal{E} \cos\theta_\varepsilon}{a_0^{3/2} L^{3/2}} \frac{16 k a_0^5}{(1 + k^2 a_0^2)^3}$$

and

$$\left| \langle \vec{k} | \vec{r} \cdot \vec{\mathcal{E}} | 100 \rangle \right|^2 = \frac{4096 \pi^2 \mathcal{E}^2 \cos^2\theta_\varepsilon a_0^7 k^2}{L^3 (1 + k^2 a_0^2)^6}$$

where

$$k^2 = \frac{2m\omega}{\hbar} - \frac{1}{a_0^2}$$

Finally, we get

$$\Gamma_{ionization} = \int d\Gamma = \frac{L^3 m k e^2}{16\pi^2 \hbar^3} \int d\Omega_{\vec{k}} \left| \langle \vec{k} | \vec{r} \cdot \vec{\mathcal{E}} | 100 \rangle \right|^2$$

where

$d\Omega_{\vec{k}}$ = integration over the angles of $\vec{k}$

(varies direction of arbitrary z - axis)

However, varying the z-direction is the same as keeping $\vec{E}$ fixed and integrating over $d\Omega_{\vec{E}}$.

Therefore, we have

$$\Gamma_{ionization} = \frac{L^3 m k e^2}{16\pi^2 \hbar^3} \frac{4096 \pi^2 \mathcal{E}^2 a_0^7 k^2}{L^3 (1 + k^2 a_0^2)^6} \int d\Omega_{\vec{\mathcal{E}}} \cos^2\theta_\varepsilon$$

$$= \frac{64\pi m e^2 \mathcal{E}^2 a_0^7}{3\hbar^3} \frac{k^3}{(1 + k^2 a_0^2)^6}$$

Now

$$1 + k^2 a_0^2 = \frac{2m\omega a_0^2}{\hbar}$$

and letting

$$\omega_0 = \frac{\hbar}{2ma_0^2}$$

we find that

$$1 + k^2 a_0^2 = \frac{\omega}{\omega_0} \text{ and } k = \frac{1}{a_0}\left(\frac{\omega}{\omega_0} - 1\right)^{1/2}$$

and we get

$$\Gamma_{ionization} = \frac{64\pi e^2 \mathcal{E}^2 a_0^3}{3\hbar}\left(\frac{\omega_0}{\omega}\right)^6 \left(\frac{\omega}{\omega_0} - 1\right)^{3/2}$$

Thus, there exists a threshold energy for this process, i.e., it cannot occur unless the energy of the photon is greater than a minimum amount, which makes physical sense.

In particular, we must have $\omega \geq \omega_0$ so that

$$\hbar\omega_0 = |\varepsilon_{100}| = E_{ionization} = \text{ minimum} = \text{threshold}$$

We also get the correct 6^{th}–power term in the answer which agrees with experiment.

End of class.

5. Perturbing the 2-dimensional Infinite Square Well

Consider a particle in a 2-dimensional infinite square well given by

$$V(x,y) = \begin{cases} 0 & \text{for } 0 \le x \le a \text{ and } 0 \le y \le a \\ \infty & \text{otherwise} \end{cases}$$

(a) What are the energy eigenvalues and eigenkets for the three lowest levels?

For the unperturbed system we have

$$V(x,y) = \begin{cases} 0 & \text{for } 0 \le x \le a \text{ and } 0 \le y \le a \\ \infty & \text{otherwise} \end{cases}$$

$$\hat{H} = -\frac{\hbar^2}{2m}\nabla^2 + V(x) \quad , \quad E_{n_x n_y} = \frac{\pi^2\hbar^2}{2ma^2}\left(n_x^2 + n_y^2\right)$$
$$\psi_{n_x n_y}(x,y) = \frac{2}{a}\sin\frac{n_x \pi x}{a}\sin\frac{n_y \pi y}{a} \quad , \quad n_x\,,\, n_y = 1,2,..$$

Ground state:

$$n_x = n_y = 1 \quad , \quad E_{11}^{(0)} = \frac{\pi^2\hbar^2}{ma^2}$$
$$\psi_{11}^{(0)}(x,y) = \frac{2}{a}\sin\frac{\pi x}{a}\sin\frac{\pi y}{a} \quad , \quad \text{non - degenerate}$$

This is a non-degenerate level.

1^{st} **Excited state**:

$$n_x = 1,\, n_y = 2 \quad , \quad E_{12}^{(0)} = \frac{5\pi^2\hbar^2}{2ma^2} \quad , \quad \psi_{12}^{(0)}(x,y) = \frac{2}{a}\sin\frac{\pi x}{a}\sin\frac{2\pi y}{a}$$
$$n_x = 2,\, n_y = 1 \quad , \quad E_{21}^{(0)} = \frac{5\pi^2\hbar^2}{2ma^2} \quad , \quad \psi_{21}^{(0)}(x,y) = \frac{2}{a}\sin\frac{2\pi x}{a}\sin\frac{\pi y}{a}$$

This is a 2-fold degenerate level.

2^{nd} **Excited state**:

$$n_x = n_y = 2 \quad , \quad E_{22}^{(0)} = \frac{4\pi^2\hbar^2}{ma^2} \quad , \quad \psi_{11}^{(0)}(x,y) = \frac{2}{a}\sin\frac{2\pi x}{a}\sin\frac{2\pi y}{a}$$

This is a non-degenerate level.

(b) We now add a perturbation given by

$$V_1(x,y) = \begin{cases} \lambda xy & \text{for } 0 \le x \le a \text{ and } 0 \le y \le a \\ 0 & \text{otherwise} \end{cases}$$

Determine the first order energy shifts for the three lowest levels for $\lambda \ll 1$.

First-order corrections: the perturbation is

$$V_1(x,y) = \begin{cases} \lambda xy & \text{for } 0 \le x \le a \text{ and } 0 \le y \le a \\ 0 & \text{otherwise} \end{cases}$$

Non-degenerate states:

$$\langle 11|\, V_1\, |11\rangle = \int_0^a \int_0^a \tfrac{4}{a^2}\lambda xy \sin^2\tfrac{\pi x}{a}\sin^2\tfrac{\pi y}{a}dxdy = \tfrac{1}{4}\lambda a^2$$
$$\langle 22|\, V_1\, |22\rangle = \int_0^a \int_0^a \tfrac{4}{a^2}\lambda xy \sin^2\tfrac{2\pi x}{a}\sin^2\tfrac{2\pi y}{a}dxdy = \tfrac{1}{4}\lambda a^2$$

Therefore, to 1^{st} order:

$$E_{11} = \tfrac{\pi^2\hbar^2}{ma^2} + \tfrac{1}{4}\lambda a^2$$
$$E_{22} = \tfrac{4\pi^2\hbar^2}{ma^2} + \tfrac{1}{4}\lambda a^2$$

Degenerate states: we must diagonalize a 2×2 matrix. We have

$$\langle 12|\, V_1\, |12\rangle = \int_0^a \int_0^a \tfrac{4}{a^2}\lambda xy \sin^2\tfrac{\pi x}{a}\sin^2\tfrac{2\pi y}{a}dxdy = \tfrac{1}{4}\lambda a^2 = \langle 22|\, V_1\, |22\rangle$$
$$\langle 12|\, V_1\, |21\rangle = \int_0^a \int_0^a \tfrac{4}{a^2}\lambda xy \sin\tfrac{2\pi x}{a}\sin\tfrac{\pi y}{a}\sin\tfrac{\pi x}{a}\sin\tfrac{2\pi y}{a}dxdy = \tfrac{256}{81\pi^4}\lambda a^2 = \langle 21|\, V_1\, |12\rangle$$

Therefore, the degenerate submatirx (to be diagonalized) is

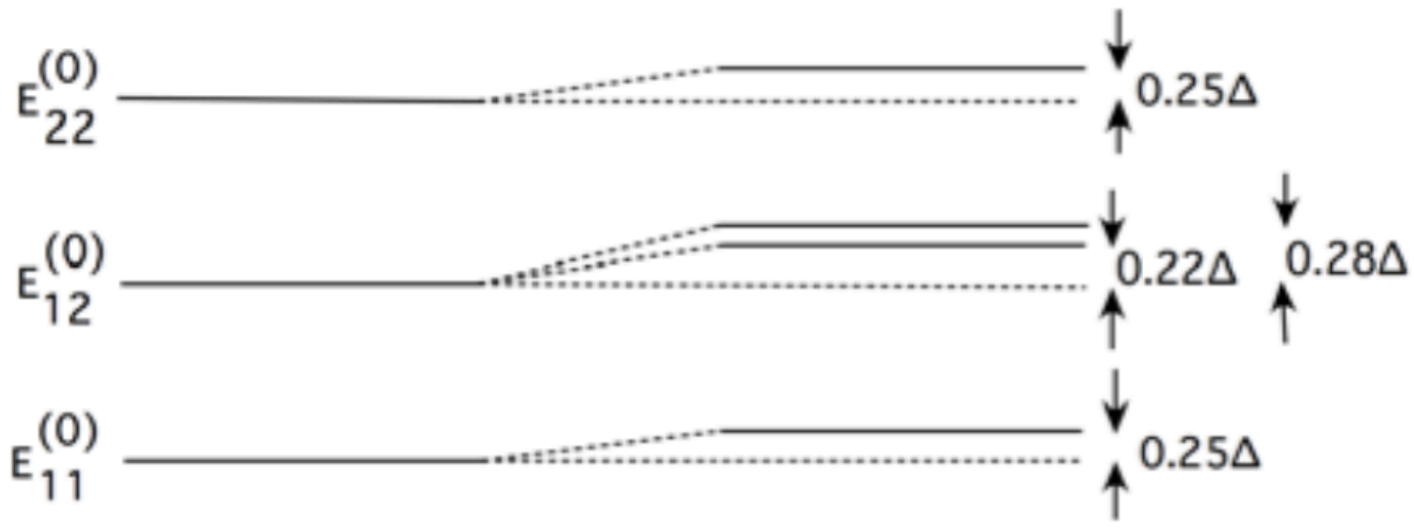
$$\frac{1}{4\pi^4}\lambda a^2 \begin{pmatrix} \pi^4 & \frac{1024}{81} \\ \frac{1024}{81} & \pi^4 \end{pmatrix}$$

This has eigenvalues

$$E_\pm = \frac{\lambda a^2}{4\pi^4}\left(\pi^4 \pm \frac{1024}{81}\right) = \begin{cases} 0.28\lambda a^2 \\ 0.22\lambda a^2 \end{cases}$$

(c) Draw an energy diagram with and without the perturbation for the three energy states, Make sure to specify which unperturbed state is connected to which perturbed state.

Therefore, the perturbed energy levels look like:



11. Finite Spatial Extent of the Nucleus

In most discussions of atoms, the nucleus is treated as a positively charged point particle. In fact, the nucleus does possess a finite size with a radius given approximately by the empirical formula

$$R \approx r_0 A^{1/3}$$

where $r_0 = 1.2 \times 10^{-13}\, cm$ (i.e., 1.2 Fermi) and A is the atomic weight or number (essentially the number of protons and neutrons in the nucleus). A reasonable assumption is to take the total nuclear charge $+Ze$ as being uniformly distributed over the entire nuclear volume (assumed to be a sphere).

(a) Derive the following expression for the electrostatic potential energy of an electron in the field of the *finite* nucleus:

$$V(r) = \begin{cases} -\frac{Ze^2}{r} & \text{for } r > R \\ \frac{Ze^2}{R}\left(\frac{r^2}{2R^2} - \frac{3}{2}\right) & \text{for } r < R \end{cases}$$

Draw a graph comparing this potential energy and the point nucleus potential energy.

For $r > R$, Gauss' law says that the sphere of charge acts like a point charge at the center of the sphere. Therefore,

$$V(r) = -\frac{Ze^2}{r} r > R$$

For $r < R$, Gauss's law says that only the charge inside contributes. Therefore,

$$\oint_S \vec{\varepsilon} \cdot d\vec{A} = \varepsilon_r 4\pi r^2 = 4\pi(\text{chargeinside}) = 4\pi Ze\frac{r^3}{R^3}$$
$$\varepsilon_r = Ze\frac{r}{R^3} \rightarrow \varphi = Ze\frac{r^2}{2R^3} + C = textpotential$$
$$\rightarrow V(r) = Ze^2\frac{r^2}{2R^3} + C = \text{ potential energy}$$

Now, continuity at $r - R$ gives

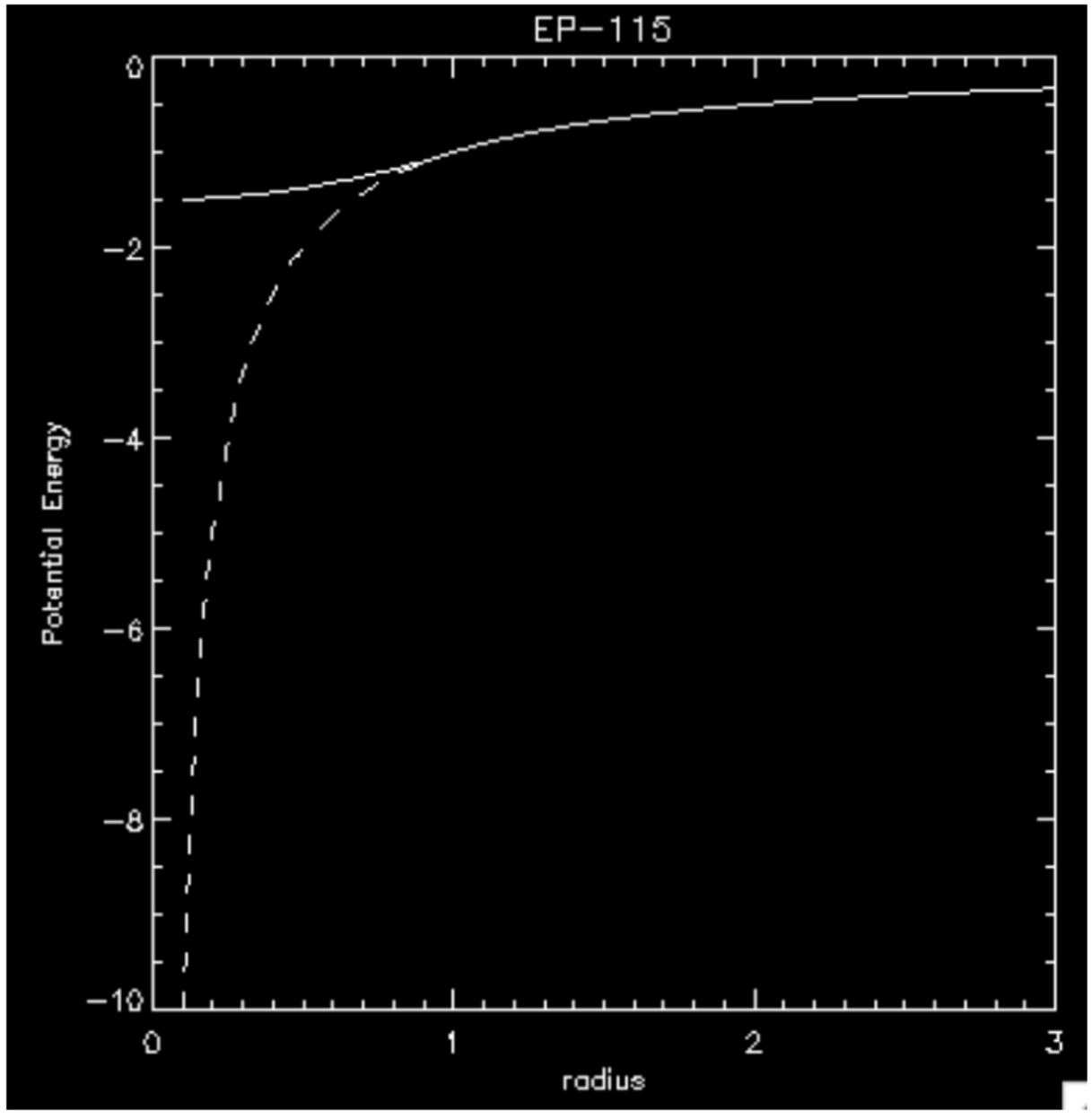
$$Ze^2 \frac{R^2}{2R^3} + C = -\frac{Ze^2}{R} \rightarrow C = -\frac{3Ze^2}{2R}$$

Therefore,

$$V(r) = \begin{cases} -\frac{Ze^2}{r} & r > R \\ \frac{Ze^2}{R}\left(\frac{r^2}{2R^2} - \frac{3}{2}\right) & r < R \end{cases}$$

The graph below chooses $Ze^2 = 1$ and $R = 1$ or

$$V(r) = \begin{cases} -\frac{1}{r} & r > 1 \\ \frac{r^2}{2} - \frac{3}{2} & r < 1 \end{cases}$$



(b) Since you know the solution of the point nucleus problem, choose this as the unperturbed Hamiltonian $\hat{H}_0$ and construct a perturbation Hamiltonian $\hat{H}_1$ such that the total Hamiltonian contains the $V(r)$ derived above. Write an expression for $\hat{H}_1$.

We then have

$$\hat{H} = \frac{\vec{p}^2}{2m} + V(r) = \frac{\vec{p}^2}{2m} - \frac{Ze^2}{r} + \left(V(r) + \frac{Ze^2}{r}\right)$$
$$\hat{H}_0 = \frac{\vec{p}^2}{2m} - \frac{Ze^2}{r} \rightarrow \text{ point charge at origin (hydrogen - like)}$$

$$\hat{H}_1 = V(r) + \frac{Ze^2}{r} = \begin{cases} 0 & r > R \\ \frac{Ze^2}{R}\left(\frac{r^2}{2R^2} - \frac{3}{2}\right) + \frac{Ze^2}{r} & r < R \end{cases}$$

(c) Calculate(remember that $R \ll a_0 =$ Bohr radius) the $1^{st}-$order perturbed energy for the $1s$ $(n\ell m) = (100)$ state obtaining an expression in terms of Z and fundamental constants. How big is this result compared to the ground state energy of hydrogen? How does it compare to hyperfine splitting?

The first-order correction to the ground state energy iis

$$E_0^{(1)} = \int d^3r \psi_0^{*(0)}(\vec{r}) H_1(r) \psi_0^{(0)}(\vec{r}) \quad ,$$
$$\psi_0^{(0)}(\vec{r}) = \sqrt{\frac{\alpha^3}{\pi}} e^{-\alpha r} \, , \quad (\ell = 0) \, , \quad \alpha = \frac{Z}{a_0}$$

Now

$$\frac{R}{a_0} = \frac{r_0}{a_0} A^{1/3} \approx 10^{-5} A^{1/3} << 1$$

Therefore, we can approximate

$$e^{-\alpha r/a_0} \approx 1 - \frac{\alpha r}{a_0} \text{for} r < R \text{ (only region perturbation is nonzero)}$$

so that we get

$$\begin{aligned}
E_0^{(1)} &= 4\alpha^3 \int_0^R r^2 dr \left(1 - \frac{2\alpha r}{a_0} + \left(\frac{2\alpha r}{a_0}\right)^2\right)\left(\frac{Ze^2}{R}\left(\frac{r^2}{2R^2} - \frac{3}{2}\right) + \frac{Ze^2}{r}\right) \\
&= 4\alpha^3 \left(\begin{array}{c} \left(\frac{Ze^2}{R}\left(\frac{R^5}{10R^2} - \frac{3R^3}{6}\right) + \frac{Ze^2R^2}{2}\right) - \frac{2\alpha}{a_0}\left(\frac{Ze^2}{R}\left(\frac{R^6}{12R^2} - \frac{3R^4}{8}\right)\right) \\ + \frac{Ze^2R^3}{3} + \left(\frac{2\alpha}{a_0}\right)^2 \left(\frac{Ze^2}{R}\left(\frac{R^7}{14R^2} - \frac{3R^5}{10}\right) + \frac{Ze^2R^4}{4}\right) \end{array} \right) \\
&= 4\alpha^3 \left(\frac{1}{10} Ze^2R^2 - \frac{\alpha}{12a_0} Ze^2R^3 + \frac{3}{140}\left(\frac{2\alpha}{a_0}\right)^2 Ze^2R^4 \right) \\
&\approx \frac{2}{5} Ze^2R^2\alpha^3 = \frac{4}{5}\frac{Z^4e^2}{2a_0}\left(\frac{R}{a_0}\right)^2
\end{aligned}$$

The effect compared to the ground states energy of hydrogen ($Z = 1$) is

$$\frac{\left|E_0^{(1)}\right|}{\left|E_0^{(0)}\right|} = \frac{\frac{4}{5}\frac{e^2}{2a_0}\left(\frac{R}{a_0}\right)^2}{\frac{e^2}{2a_0}} = \frac{4}{5}\left(\frac{R}{a_0}\right)^2 \approx 10^{-10}$$

This is the same size as the hyperfine splitting due to magnetic moment interactions.

17. Hydrogen atom + xy perturbation

An electron moves in a Coulomb field centered at the origin of coordinates. The first excited state ($n = 2$) is 4−fold degenerate. Consider what happens in the presence of a non-central perturbation

$$V_{pert} = f(r)xy$$

where $f(r)$ is some function only of r, which falls off rapidly as $r \to \infty$. To first order, this perturbation splits the 4−fold degenerate level into several distinct levels (some might still be degenerate).

(a) How many levels are there?

(b) What is the degeneracy of each?

(c) Given the energy shift, call it ΔE, for one of the levels, what are the values of the shifts for all the others?

The four degenerate unperturbed wave functions are

$$\begin{aligned}
\langle \vec{r} \mid 2,0,0 \rangle &= R_{20}(r) Y_{0,0}(\theta, \phi) \\
\langle \vec{r} \mid 2,1,1 \rangle &= R_{21}(r) Y_{1,1}(\theta, \phi) \\
\langle \vec{r} \mid 2,1,0 \rangle &= R_{21}(r) Y_{1,0}(\theta, \phi) \\
\langle \vec{r} \mid 2,1,-1 \rangle &= R_{21}(r) Y_{1,-1}(\theta, \phi)
\end{aligned}$$

They all correspond to the same unperturbed energy

$$E_2^{(0)} = -\frac{e^2}{8a_0} \rightarrow \text{ 4 - fold degenerate level}$$

We need to calculate the 4×4 degenerate submatrix for the perturbation

$$V = f(r)xy = f(r)r^2 \sin^2\theta \sin\phi \cos\phi$$

with matrix elements

$$\begin{aligned}
V_{\ell' m';\ell m} &= \langle \ell' m' | \, \hat{V} \, | \ell m \rangle \\
&= \int R_{2\ell'}(r) R_{2\ell}(r) r^4 f(r) dr \int Y^*_{\ell' m'}(\theta, \phi) Y_{\ell m}(\theta, \phi) \sin^3\theta \sin\phi \cos\phi d\theta d\phi
\end{aligned}$$

The necessary spherical harmonics are

$$\begin{aligned}
Y_{0,0} &= \tfrac{1}{\sqrt{4\pi}} \quad , \quad Y_{1,1} = \left(\tfrac{3}{8\pi}\right)^{1/2} \sin\theta e^{i\phi} \\
Y_{1,0} &= \left(\tfrac{3}{4\pi}\right)^{1/2} \cos\theta \quad , \quad Y_{1,-1} = \left(\tfrac{3}{8\pi}\right)^{1/2} \sin\theta e^{-i\phi}
\end{aligned}$$

Now the ϕ integrations take the form

$$\int_0^{2\pi} \sin\phi\cos\phi d\phi = 0$$

or

$$\int_0^{2\pi} e^{i(m-m')\phi} \sin\phi\cos\phi d\phi$$

which is zero unless

$$m = 1 = -m' \quad or \quad m = -1 = -m'$$

Therefore, the only nonzero matrix elements are

$$\begin{aligned} V_{1,-1;1,1} &= V^*_{1,1;1,-1} \\ &= \int R_{21}(r)R_{21}(r)r^4 f(r)dr \int Y^*_{1,-1}(\theta,\phi)Y_{11}(\theta,\phi)\sin^3\theta\sin\phi\cos\phi d\theta d\phi \\ &= \frac{3}{8\pi}\int_0^\infty R^2_{21}(r)r^4 f(r)dr \int_0^\pi \sin^5\theta d\theta \int_0^{2\pi} \sin\phi\cos\phi e^{-2i\phi} d\phi = iA \end{aligned}$$

where

$$A = \frac{1}{5}\int_0^\infty R^2_{21}(r)r^4 f(r)dr$$

Therefore the V matrix look like

$$\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & iA \\ 0 & 0 & -iA & 0 \end{pmatrix} \rightarrow eigenvalues = 0, 0, \pm A$$

Therefore the original 4-fold degenerate level splits into three new levels with the remaining degeneracy as indicated.

next