

Approximation Methods

Physics 130B, UCSD Fall 2009

Joel Broida

November 15, 2009

Contents

1	The Variation Method	1
1.1	The Variation Theorem	1
1.2	Excited States	7
1.3	Linear Variation Functions	8
1.3.1	Proof that the Roots of the Secular Equation are Real	17
2	Time-Independent Perturbation Theory	21
2.1	Perturbation Theory for a Nondegenerate Energy Level	21
2.2	Perturbation Theory for a Degenerate Energy Level	26
2.3	Perturbation Treatment of the First Excited States of Helium	37
2.4	Spin-Orbit Coupling and the Hydrogen Atom Fine Structure	46
2.4.1	Supplement: Miscellaneous Proofs	53
2.5	The Zeeman Effect	58
2.5.1	Strong External Field	59
2.5.2	Weak External Field	61
2.5.3	Intermediate-Field Case	62
2.5.4	Supplement: The Electromagnetic Hamiltonian	66
3	Time-Dependent Perturbation Theory	71
3.1	Transitions Between Two Discrete States	71
3.2	Transitions to a Continuum of States	80

1 The Variation Method

1.1 The Variation Theorem

The variation method is one approach to approximating the ground state energy of a system without actually solving the Schrödinger equation. It is based on the following theorem, sometimes called the **variation theorem**.

Theorem 1.1. *Let a system be described by a time-independent Hamiltonian H , and let φ be any normalized well-behaved function that satisfies the boundary conditions of the problem. If E_0 is the true ground state energy of the system, then*

$$\langle \varphi | H \varphi \rangle \geq E_0. \quad (1.1)$$

Proof. Consider the integral $I = \langle \varphi | (H - E_0) \varphi \rangle$. Then

$$I = \langle \varphi | H \varphi \rangle - E_0 \langle \varphi | \varphi \rangle = \langle \varphi | H \varphi \rangle - E_0.$$

We must show that $I \geq 0$. Let $\{\psi_n\}$ be the true (stationary state) solutions to the Schrödinger equation, so that $H\psi_n = E_n\psi_n$. By assumption, the ψ_n form a complete, orthonormal set, so we can write

$$\varphi = \sum_n a_n \psi_n$$

where $\langle \psi_n | \psi_m \rangle = \delta_{nm}$. Then

$$\begin{aligned} I &= \sum_n a_n^* \langle \psi_n | (H - E_0) \sum_m a_m \psi_m \rangle \\ &= \sum_{n,m} a_n^* a_m (\langle \psi_n | H \psi_m \rangle - E_0 \delta_{nm}) \\ &= \sum_{n,m} a_n^* a_m (E_m - E_0) \delta_{nm} \\ &= \sum_n |a_n|^2 (E_n - E_0). \end{aligned}$$

But $|a_n| \geq 0$ and $E_n > E_0$ for all $n > 0$ because E_0 is the ground state of the system. Therefore $I \geq 0$ as claimed. ■

Suppose we have a trial function φ that is not normalized. Then multiplying by a normalization constant N , equation (1.1) becomes $|N|^2 \langle \varphi | H \varphi \rangle \geq E_0$. But by definition we know that $1 = \langle N\varphi | N\varphi \rangle = |N|^2 \langle \varphi | \varphi \rangle$ so that $|N|^2 = 1/\langle \varphi | \varphi \rangle$, and hence our variation theorem becomes

$$\frac{\langle \varphi | H \varphi \rangle}{\langle \varphi | \varphi \rangle} \geq E_0. \quad (1.2)$$

The integral in (1.1) (or the ratio of integrals in (1.2)) is called the **variational integral**.

So the idea is to try a number of different trial functions, and see how low we can get the variational integral to go. Fortunately, the variational integral approaches E_0 a lot faster than φ approaches ψ_0 , so it is possible to get a good approximation to E_0 even with a poor φ . However, a common approach is to introduce arbitrary parameters and minimize the energy with respect to them.

Before continuing with an example, there are two points I need to make. First, I state without proof that *the bound stationary states of a one-dimensional system are characterized by having no nodes interior to the boundary points in the ground state (i.e., the wavefunction is never zero), and the number of nodes increases by one for each successive excited state*. While the proof of this statement is not particularly difficult (it's really a statement about Sturm-Liouville type differential equations), it would take us too far astray at the moment. If you are interested, a proof may be found in Messiah, *Quantum Mechanics*, Chapter III, Sections 8-12.

A related issue is the following: *In one dimension, the bound states are nondegenerate*. To prove this, suppose we have two degenerate states ψ_1 and ψ_2 , both with the same energy E . Multiply the Schrödinger equation for ψ_1 by ψ_2 :

$$-\frac{\hbar^2}{2m}\psi_2\frac{d^2\psi_1}{dx^2} + V\psi_1\psi_2 = E\psi_1\psi_2$$

and multiply the Schrödinger equation for ψ_2 by ψ_1 :

$$-\frac{\hbar^2}{2m}\psi_1\frac{d^2\psi_2}{dx^2} + V\psi_1\psi_2 = E\psi_1\psi_2.$$

Subtracting, we obtain

$$\psi_2\frac{d^2\psi_1}{dx^2} - \psi_1\frac{d^2\psi_2}{dx^2} = 0.$$

But then

$$\frac{d}{dx}\left[\psi_2\frac{d\psi_1}{dx} - \psi_1\frac{d\psi_2}{dx}\right] = \psi_2\frac{d^2\psi_1}{dx^2} - \psi_1\frac{d^2\psi_2}{dx^2} = 0$$

so that

$$\psi_2\frac{d\psi_1}{dx} - \psi_1\frac{d\psi_2}{dx} = \text{const}.$$

However, we know that $\psi \rightarrow 0$ as $x \rightarrow \pm\infty$, and hence the constant must equal zero. Rewriting this result we now have $d\ln\psi_1 = d\ln\psi_2$ or $\ln\psi_1 = \ln\psi_2 + \ln k$ where $\ln k$ is an integration constant. This is equivalent to $\psi_1 = k\psi_2$ so that ψ_1 and ψ_2 are linearly dependent and hence degenerate as claimed.

The second topic I need to address is the notion of classification by symmetry. So, let us consider the time-independent Schrödinger equation $H\psi = E\psi$, and suppose that the potential energy function $V(\mathbf{x})$ is symmetric, i.e.,

$$V(-\mathbf{x}) = V(\mathbf{x}).$$

Under these conditions, the total Hamiltonian is also symmetric:

$$H(-\mathbf{x}) = H(\mathbf{x}).$$

To understand the consequences of this, let us introduce an operator Π called the **parity operator**, defined by

$$\Pi f(\mathbf{x}) = f(-\mathbf{x})$$

where $f(\mathbf{x})$ is an arbitrary function. It is easy to see that Π is Hermitian because

$$\begin{aligned} \langle f | \Pi g \rangle &= \int_{-\infty}^{\infty} f(\mathbf{x})^* \Pi g(\mathbf{x}) d\mathbf{x} = \int_{-\infty}^{\infty} f(\mathbf{x})^* g(-\mathbf{x}) d\mathbf{x} \\ &= \int_{-\infty}^{\infty} f(-\mathbf{x})^* g(\mathbf{x}) d\mathbf{x} = \int_{-\infty}^{\infty} [\Pi f(\mathbf{x})]^* g(\mathbf{x}) d\mathbf{x} \\ &= \langle \Pi f | g \rangle \end{aligned}$$

where in going from the first line to the second we simply changed variables $\mathbf{x} \rightarrow -\mathbf{x}$. (I will use the symbol $d\mathbf{x}$ to denote the volume element in whatever n -dimensional space is under consideration.)

Now what can we say about the eigenvalues of Π ? Well, if $\Pi f = \lambda f$, then

$$\Pi^2 f = \Pi(\Pi f) = \lambda \Pi f = \lambda^2 f.$$

On the other hand, it is clear that

$$\Pi^2 f(\mathbf{x}) = \Pi(\Pi f(\mathbf{x})) = \Pi f(-\mathbf{x}) = f(\mathbf{x})$$

and hence we must have $\lambda^2 = 1$, so the eigenvalues of Π are ± 1 . Let us denote the corresponding eigenfunctions by f_{\pm} :

$$\Pi f_+ = f_+ \quad \text{and} \quad \Pi f_- = -f_-.$$

In other words,

$$f_+(-\mathbf{x}) = f_+(\mathbf{x}) \quad \text{and} \quad f_-(-\mathbf{x}) = -f_-(\mathbf{x}).$$

Thus f_+ is any even function, and f_- is any odd function. Note that what have shown is the existence of a Hermitian operator with only two eigenvalues, each of which is infinitely degenerate. (I leave it as an easy exercise for you to show that f_+ and f_- are orthogonal as they should be.)

Next, note that any $f(\mathbf{x})$ can always be written in the form

$$f(\mathbf{x}) = f_+(\mathbf{x}) + f_-(\mathbf{x})$$

where

$$f_+(\mathbf{x}) = \frac{f(\mathbf{x}) + f(-\mathbf{x})}{2} \quad \text{and} \quad f_-(\mathbf{x}) = \frac{f(\mathbf{x}) - f(-\mathbf{x})}{2}$$

are obviously symmetric and antisymmetric, respectively. Thus the eigenfunctions of the parity operator are complete, i.e., any function can be written as the sum of a symmetric function and an antisymmetric function.

It will be extremely convenient to now introduce the operators Π_{\pm} defined by

$$\Pi_{\pm} = \frac{1 \pm \Pi}{2}.$$

In terms of these operators, we can write

$$\Pi_{\pm} f = f_{\pm}.$$

It is easy to see that the operators Π_{\pm} satisfy the three properties

$$\begin{aligned}\Pi_{\pm}^2 &= \Pi_{\pm} \\ \Pi_+ \Pi_- &= \Pi_- \Pi_+ = 0 \\ \Pi_+ + \Pi_- &= 1.\end{aligned}$$

The operators Π_{\pm} are called **projection operators**.

Returning to our symmetric Hamiltonian, we observe that

$$\Pi(H(\mathbf{x})\psi(\mathbf{x})) = H(-\mathbf{x})\psi(-\mathbf{x}) = H(\mathbf{x})\psi(-\mathbf{x}) = H(\mathbf{x})\Pi\psi(\mathbf{x})$$

and thus the Hamiltonian commutes with the parity operator. But if $[H, \Pi] = 0$, then it is trivial to see that $[H, \Pi_{\pm}] = 0$ also, and therefore acting on $H\psi_E = E\psi_E$ with Π_{\pm} we see that

$$H\psi_{E+} = E\psi_{E+}$$

and

$$H\psi_{E-} = E\psi_{E-}.$$

Thus the stationary states in a symmetric potential can always be classified according to their parity, i.e., they can always be chosen to have a definite symmetry. Moreover, since, as we saw above, the bound states in one dimension are nondegenerate, it follows that *each bound state in a one-dimensional symmetric potential must be either even or odd.*

Example 1.1. Let us find a trial function for a particle in a one-dimensional box of length l . Since the true wavefunction vanishes at the ends $x = 0$ and $x = l$, our trial function must also have this property. A simple (un-normalized) function that obeys these boundary conditions is

$$\varphi = x(l - x) \quad \text{for } 0 \leq x \leq l$$

and $\varphi = 0$ outside the box.

The integrals in equation (1.2) are

$$\begin{aligned}\langle \varphi | H \varphi \rangle &= -\frac{\hbar^2}{2m} \int_0^l x(l-x) \frac{d^2}{dx^2} x(l-x) dx \\ &= \frac{\hbar^2}{m} \int_0^l x(l-x) dx = \frac{\hbar^2 l^3}{6m}\end{aligned}$$

and

$$\langle \varphi | \varphi \rangle = \int_0^l x^2(l-x)^2 dx = \frac{l^5}{30}.$$

Therefore

$$E_0 \leq \frac{\langle \varphi | H \varphi \rangle}{\langle \varphi | \varphi \rangle} = 5 \frac{\hbar^2}{ml^2}.$$

For comparison, the exact solution has energy levels

$$E_n = \frac{n^2 \hbar^2 \pi^2}{2ml^2} \quad n = 1, 2, \dots$$

so the ground state ($n = 1$) has energy

$$\frac{\pi^2}{2} \frac{\hbar^2}{ml^2} = 4.9348 \frac{\hbar^2}{ml^2}$$

for an error of 1.3%. The figure below is a plot of the exact normalized ground state solution to the particle in a box together with the normalized trial function. You can see how closely the trial function is to the exact solution.

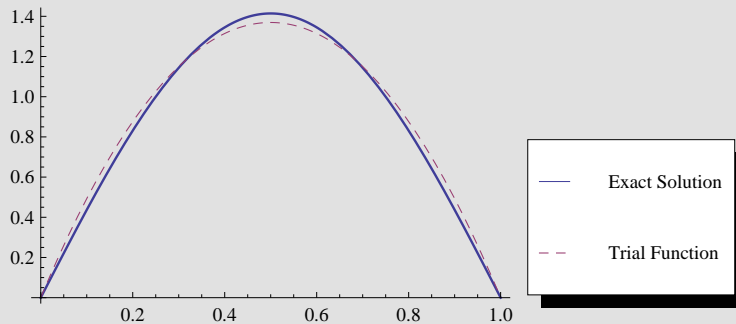


Figure 1: Plot of $\sqrt{2} \sin \pi x$ and $\sqrt{30}x(1-x)$.

Example 1.2. Let us construct a variation function with parameter for the one-dimensional harmonic oscillator, and find the optimal value for that parameter.

What do we know in general? First, the wavefunction must vanish as $x \rightarrow \pm\infty$. The most obvious function that satisfies this is e^{-x^2} . However, x has units of length, and we can only take the exponential of a dimensionless quantity (think of the power series expansion for e^{-x^2}). However, if we include a constant α with dimensions of length⁻², then $e^{-\alpha x^2}$ is satisfactory from a dimensional standpoint. In addition, since the potential $V = \frac{1}{2}kx^2$ is symmetric, we know that the eigenstates will have a definite parity. And since the ground state has no nodes, it must be an even function (since an odd function has a node at the origin). Thus the trial function $\varphi = e^{-\alpha x^2}$ has all of our desired properties.

Since φ is unnormalized, we use equation (1.2). The Hamiltonian is

$$-\frac{\hbar^2}{2m} \frac{d^2}{dx^2} + \frac{1}{2}m\omega^2 x^2$$

and hence

$$\begin{aligned} \langle \varphi | H \varphi \rangle &= -\frac{\hbar^2}{2m} \int_{-\infty}^{\infty} e^{-\alpha x^2} \frac{d^2 e^{-\alpha x^2}}{dx^2} dx + \frac{1}{2}m\omega^2 \int_{-\infty}^{\infty} x^2 e^{-2\alpha x^2} dx \\ &= -\frac{\hbar^2}{2m} \int_{-\infty}^{\infty} [4\alpha^2 x^2 e^{-2\alpha x^2} - 2\alpha e^{-2\alpha x^2}] dx + \frac{1}{2}m\omega^2 \int_{-\infty}^{\infty} x^2 e^{-2\alpha x^2} dx \\ &= \left[\frac{-2\hbar^2 \alpha^2}{m} + \frac{1}{2}m\omega^2 \right] \int_{-\infty}^{\infty} x^2 e^{-2\alpha x^2} dx + \frac{\hbar^2 \alpha}{m} \int_{-\infty}^{\infty} e^{-2\alpha x^2} dx. \end{aligned}$$

The second integral is easy (and you should already know the answer):

$$\int_{-\infty}^{\infty} e^{-2\alpha x^2} dx = \sqrt{\frac{\pi}{2\alpha}}.$$

Using this, the first integral is also easy. Letting $\beta = 2\alpha$ we have

$$\begin{aligned} \int_{-\infty}^{\infty} x^2 e^{-2\alpha x^2} dx &= \int_{-\infty}^{\infty} x^2 e^{-\beta x^2} dx = -\frac{\partial}{\partial \beta} \int_{-\infty}^{\infty} e^{-\beta x^2} dx \\ &= -\frac{\partial}{\partial \beta} \sqrt{\frac{\pi}{\beta}} = \frac{1}{2} \frac{\pi^{1/2}}{\beta^{3/2}} \\ &= \frac{1}{2} \frac{\pi^{1/2}}{(2\alpha)^{3/2}}. \end{aligned}$$

After a little algebra, we now arrive at

$$\langle \varphi | H \varphi \rangle = \frac{\hbar^2 \pi^{1/2} \alpha^{1/2}}{2^{3/2} m} + \frac{m\omega^2 \pi^{1/2} \alpha^{-3/2}}{2^{7/2}}.$$

And the denominator in equation (1.2) is just

$$\langle \varphi | \varphi \rangle = \int_{-\infty}^{\infty} e^{-2\alpha x^2} dx = \sqrt{\frac{\pi}{2\alpha}}.$$

Thus our variational integral becomes

$$W := \frac{\langle \varphi | H \varphi \rangle}{\langle \varphi | \varphi \rangle} = \frac{\hbar^2 \alpha}{2m} + \frac{m\omega^2}{8\alpha}.$$

To minimize this with respect to α we set $dW/d\alpha = 0$ and solve for α :

$$\frac{\hbar^2}{2m} - \frac{m\omega^2}{8\alpha^2} = 0$$

or

$$\alpha = \pm \frac{m\omega}{2\hbar}.$$

The negative root must be rejected because otherwise $\varphi = e^{-\alpha x^2}$ would be divergent. Substituting the positive root for α into our expression for W yields

$$W = \frac{1}{2}\hbar\omega$$

which is the exact ground state harmonic oscillator energy. This isn't surprising, because up to normalization, our φ with $\alpha = m\omega/2\hbar$ is just the exact ground state harmonic oscillator wave function.

1.2 Excited States

So far all we have discussed is how to approximate the ground-state energy of a system. Now we want to take a look at how to go about approximating the energy of an excited state. Let us assume that the stationary states of our system are numbered so that

$$E_0 \leq E_1 \leq E_2 \leq \dots$$

If $\{\psi_n\}$ is a complete set of orthonormal eigenstates of H , then our normalized trial function can be written $\varphi = \sum_n a_n \psi_n$ where $a_n = \langle \psi_n | \varphi \rangle$. Then as we have seen

$$\langle \varphi | H \varphi \rangle = \sum_{n,m} a_n^* a_m E_m \langle \psi_n | \psi_m \rangle = \sum_{n,m} a_n^* a_m E_m \delta_{nm} = \sum_{n=0}^{\infty} |a_n|^2 E_n$$

and

$$\langle \varphi | \varphi \rangle = \sum_{n=0}^{\infty} |a_n|^2 = 1.$$

Suppose we restrict ourselves to trial functions that are orthogonal to the true ground-state wavefunction ψ_0 . Then $a_0 = \langle \psi_0 | \varphi \rangle = 0$ and we are left with

$$\langle \varphi | H \varphi \rangle = \sum_{n=1}^{\infty} |a_n|^2 E_n \quad \text{and} \quad \langle \varphi | \varphi \rangle = \sum_{n=1}^{\infty} |a_n|^2 = 1.$$

For $n \geq 1$ we have $E_n \geq E_1$ so that $|a_n|^2 E_n \geq |a_n|^2 E_1$ and hence

$$\sum_{n=1}^{\infty} |a_n|^2 E_n \geq \sum_{n=1}^{\infty} |a_n|^2 E_1 = E_1 \sum_{n=1}^{\infty} |a_n|^2 = E_1.$$

This gives us our desired result

$$\langle \varphi | H \varphi \rangle \geq E_1 \quad \text{if } \langle \psi_0 | \varphi \rangle = 0 \text{ and } \langle \varphi | \varphi \rangle = 1. \quad (1.3)$$

While equation (1.3) gives an upper bound on the energy E_1 of the first excited state, it depends on the restriction $\langle \psi_0 | \varphi \rangle = 0$ which can be problematic. However, for some systems this is not a difficult requirement to achieve even though we don't know the exact ground-state wavefunction. For example, a one-dimensional problem with a symmetric potential has a ground-state wavefunction that is always even, while the first excited state is always odd. This means that any (normalized) trial function φ that is an odd function will automatically satisfy $\langle \psi_0 | \varphi \rangle = 0$.

It is also possible to extend this approach to approximating the energy levels of higher excited states. In particular, if we somehow choose the trial function φ so that

$$\langle \psi_0 | \varphi \rangle = \langle \psi_1 | \varphi \rangle = \cdots = \langle \psi_n | \varphi \rangle = 0,$$

then, following exactly the same argument as above, it is easy to see that if $\langle \varphi | \varphi \rangle = 1$ we have

$$\langle \varphi | H \varphi \rangle \geq E_{n+1}.$$

For example, consider any particle moving under a central potential $V(r)$ (e.g., the hydrogen atom). Then the Schrödinger equation factors into a radial equation that depends on $V(r)$ times an angular equation (that is independent of V) with solutions that are just the spherical harmonics $Y_l^m(\theta, \phi)$. It may very well be that we can't solve the radial equation with this potential, but we know that spherical harmonics with different values of l are orthogonal. Thus, we can get an upper bound to the energy of the lowest state with a particular angular momentum l by choosing a trial function that contains the factor Y_l^m .

1.3 Linear Variation Functions

The approach that we are now going to describe is probably the most common method of finding approximate molecular wave functions. A **linear variation function** φ is a linear combination of n linearly independent functions f_i :

$$\varphi = \sum_{i=1}^n c_i f_i.$$

The functions f_i are called **basis functions**, and they must obey the boundary conditions of the problem. The coefficients c_i are to be determined by minimizing the variational integral.

We shall restrict ourselves to a real φ , so the functions f_i and coefficients c_i are taken to be real. Later we will remove this requirement. Furthermore, note that the basis functions are not generally orthogonal since they are not necessarily the eigenfunctions of any operator. Let us define the **overlap integrals** S_{ij} by

$$S_{ij} := \langle f_i | f_j \rangle = \int f_i^* f_j \, d\mathbf{x}$$

(where the asterisk on f_i isn't necessary because we are assuming that our basis functions are real). Then (remember that the c_i are real)

$$\langle \varphi | \varphi \rangle = \sum_{i,j=1}^n c_i c_j \langle f_i | f_j \rangle = \sum_{i,j=1}^n c_i c_j S_{ij}.$$

Next, we define the integrals

$$H_{ij} := \langle f_i | H f_j \rangle = \int f_i^* H f_j \, d\mathbf{x}$$

so that

$$\langle \varphi | H \varphi \rangle = \sum_{i,j=1}^n c_i c_j \langle f_i | H f_j \rangle = \sum_{i,j=1}^n c_i c_j H_{ij}.$$

Then the variation theorem (1.2) becomes

$$W = \frac{\langle \varphi | H \varphi \rangle}{\langle \varphi | \varphi \rangle} = \frac{\sum_{i,j=1}^n c_i c_j H_{ij}}{\sum_{i,j=1}^n c_i c_j S_{ij}}$$

or

$$W \sum_{i,j=1}^n c_i c_j S_{ij} = \sum_{i,j=1}^n c_i c_j H_{ij}. \quad (1.4)$$

Now W is a function of the n c_i 's, and we know that $W \geq E_0$. In order to minimize W with respect to all of the c_k 's, we must require that at the minimum we have

$$\frac{\partial W}{\partial c_k} = 0; \quad k = 1, \dots, n.$$

Taking the derivative of (1.4) with respect to c_k and using

$$\frac{\partial c_i}{\partial c_k} = \delta_{ik}$$

we have

$$\frac{\partial W}{\partial c_k} \sum_{i,j=1}^n c_i c_j S_{ij} + W \sum_{i,j=1}^n (\delta_{ik} c_j + c_i \delta_{jk}) S_{ij} = \sum_{i,j=1}^n (\delta_{ik} c_j + c_i \delta_{jk}) H_{ij}$$

or (since $\partial W/\partial c_k = 0$)

$$W \sum_{j=1}^n c_j S_{kj} + W \sum_{i=1}^n c_i S_{ik} = \sum_{j=1}^n c_j H_{kj} + \sum_{i=1}^n c_i H_{ik}.$$

However, the basis functions f_i are real so we have

$$S_{ik} = \int f_i f_k d\mathbf{x} = S_{ki}$$

and since H is Hermitian (and $H(\mathbf{x})$ is real) we also have

$$H_{ik} = \langle f_i | H f_k \rangle = \langle H f_i | f_k \rangle = \langle f_k | H f_i \rangle^* = \langle f_k | H f_i \rangle = H_{ki}.$$

Therefore, because the summation indices are dummy indices, we see that the two terms on each side of the last equation are identical, and we are left with

$$W \sum_{j=1}^n c_j S_{kj} = \sum_{j=1}^n c_j H_{kj}$$

or

$$\sum_{j=1}^n (H_{kj} - W S_{kj}) c_j = 0; \quad k = 1, \dots, n. \quad (1.5)$$

This is just a system of n homogeneous linear equations in n unknowns (the n coefficients c_j), and hence for a nontrivial solution to exist (we don't want all of the c_j 's to be zero) we must have the **secular equation**

$$\det(H_{kj} - W S_{kj}) = 0. \quad (1.6)$$

(You can think of this as a system of the form $\sum_j a_{kj} x_j = 0$ where the matrix $A = (a_{jk})$ must be singular or else A^{-1} would exist and then the equation $Ax = 0$ would imply that $x = 0$. The requirement that A be singular is equivalent to the requirement that $\det A = 0$.) Written out, equation (1.6) looks like

$$\begin{vmatrix} H_{11} - W S_{11} & H_{12} - W S_{12} & \cdots & H_{1n} - W S_{1n} \\ H_{21} - W S_{21} & H_{22} - W S_{22} & \cdots & H_{2n} - W S_{2n} \\ \vdots & \vdots & & \vdots \\ H_{n1} - W S_{n1} & H_{n2} - W S_{n2} & \cdots & H_{nn} - W S_{nn} \end{vmatrix} = 0.$$

The determinant in (1.6) is a polynomial in W of degree n , and it can be proved that all n roots of this equation are real. (The proof is given at the end of this section for those who are interested.) Let us arrange the roots in order of increasing value as

$$W_0 \leq W_1 \leq \cdots \leq W_{n-1}.$$

Similarly, we number the bound states of the system so that the corresponding true energies of these bound states are also arranged in increasing order:

$$E_0 \leq E_1 \leq \cdots \leq E_{n-1} \leq E_n \leq \cdots .$$

From the variation theorem we know that $E_0 \leq W_0$. Furthermore, it can also be proved (see the homework) that

$$E_i \leq W_i \quad \text{for each } i = 0, \dots, n-1 .$$

In other words, the linear variation method provides upper bounds for the energies of the lowest n bound states of the system. It can be shown that increasing the number of basis functions used (and hence increasing the number of states whose energies are approximated), the better the accuracy of the previously calculated energies.

Once we have found the n roots W_i , we can substitute them one-at-a-time back into equation (1.5) and solve for the coefficients $c_j^{(i)}$, where the superscript denotes that fact that this particular set of coefficients applies to the root W_i . (Again, this is just like finding the eigenvector corresponding to a given eigenvalue.) Note also that all we can really find is the ratios of the coefficients, say relative to c_1 , and then fix c_1 by normalization.

There are some tricks that can simplify the solution of equation (1.6). For example, if we choose the basis functions to be orthonormal, then $S_{kj} = \delta_{kj}$. If the originally chosen set of basis functions isn't orthonormal, we can always use the Gram-Schmidt process to construct an orthonormal set. Also, we can make some of the off-diagonal H_{kj} 's vanish if we choose our basis functions to be eigenfunctions of some other Hermitian operator A that commutes with H . This because of the following theorem:

Theorem 1.2. *Let f_i and f_j be eigenfunctions of a Hermitian operator A corresponding to the eigenvalues $a_i \neq a_j$. If H is an operator that commutes with A , then*

$$H_{ji} = \langle f_j | H f_i \rangle = 0 .$$

Proof. Let us first assume that the eigenvalue a_i is nondegenerate. Then $A f_i = a_i f_i$ and

$$A(H f_i) = H A f_i = a_i (H f_i) .$$

Thus $H f_i$ is in the eigenspace V_{a_i} of A corresponding to the eigenvalue a_i . But a_i is nondegenerate so that the eigenspace is one-dimensional and spanned by f_i . Hence we must have $H f_i = b_i f_i$ for some scalar b_i . Recalling that eigenfunctions belonging to distinct eigenvalues of a Hermitian operator are orthogonal, we have

$$\langle f_j | H f_i \rangle = b_i \langle f_j | f_i \rangle = 0 .$$

Now assume that the eigenvalue a_i is degenerate. This means that the eigenspace V_{a_i} has dimension greater than one, say $\dim V_{a_i} = n$. Then V_{a_i} has a basis g_1, \dots, g_n consisting of eigenvectors of A corresponding to the eigenvalue a_i , i.e., $Ag_k = a_i g_k$ for each $k = 1, \dots, n$. So if Hf_i is in V_{a_i} , then $Hf_i = \sum_{k=1}^n c_k g_k$ for some expansion coefficients c_k . But then we again have

$$\langle f_j | Hf_i \rangle = \sum_{k=1}^n c_k \langle f_j | g_k \rangle = 0$$

because the eigenfunctions f_j and g_k belong to the distinct eigenvalues a_j and a_i respectively. ■

Another (possibly easier) way to prove Theorem 1.2 is this. Let $Af_i = a_i f_i$ and $Af_j = a_j f_j$ where $a_i \neq a_j$. (In other words, f_i and f_j belong to different eigenspaces of A .) Then on the one hand we have

$$\langle f_j | H A f_i \rangle = a_i \langle f_j | H f_i \rangle$$

while on the other hand, we can use the fact that H and A commute along with the fact that A is Hermitian and hence has real eigenvalues, to write

$$\langle f_j | H A f_i \rangle = \langle f_j | A H f_i \rangle = \langle A f_j | H f_i \rangle = a_j \langle f_j | H f_i \rangle.$$

Equating these results shows that $(a_i - a_j) \langle f_j | H f_i \rangle = 0$. Therefore, if $a_i \neq a_j$, we must have $\langle f_j | H f_i \rangle = 0$.

Finally, it is left as a homework problem to show that equations (1.5) and (1.6) also hold if the variation function is in fact allowed to be complex.

Example 1.3. In Example 1.1 we constructed the trial function $\varphi = x(l - x)$ for the ground state of the one-dimensional particle in a box. Let us now construct a linear variation function $\varphi = \sum_i c_i f_i$ to approximate the energies of the first four states. This means that we need at least four independent functions f_i that obey the boundary conditions of vanishing at the ends of the box. While there are an infinite number of possibilities, we want to limit ourselves to integrals that are easy to evaluate.

We begin by taking

$$f_1 = x(l - x),$$

and another simple function that obeys the proper boundary conditions is

$$f_2 = x^2(l - x)^2.$$

If the origin were chosen to be at the center of the box, we know that the exact solutions would have a definite parity, alternating between even and odd functions, starting with the even ground state. To see that both f_1 and f_2 are even functions,

we shift the origin to the center of the box by changing variables to $x' = x - l/2$. Then $x = x' + l/2$ and we find

$$f_1 = (x' + l/2)(l/2 - x') \quad \text{and} \quad f_2 = (x' + l/2)^2(l/2 - x')^2$$

which shows that f_1 and f_2 are both clearly even functions of x' .

Since both f_1 and f_2 are even functions, if we took $\varphi = c_1 f_1 + c_2 f_2$ we would end up with an upper bound for the two lowest energy even states (the $n = 1$ and $n = 3$ states). In order to also approximate the odd $n = 2$ and $n = 4$ states, we must add in two odd functions. Thus we need two functions that vanish at $x = 0, x = l$ and $x = l/2$. Two functions that satisfy these requirements are

$$f_3 = x(l - x)(l/2 - x)$$

and

$$f_4 = x^2(l - x)^2(l/2 - x).$$

By again changing variables as we did for f_1 and f_2 , you can easily show that f_3 and f_4 are indeed odd functions. Note also that the four functions we have chosen are linearly independent as they must be.

One of the advantages in choosing our functions to have a definite parity is that many of the integrals that occur in equation (1.6) will vanish. In particular, since any integral of an odd function over an even interval is identically zero, and since the product of an even function with an odd function is odd, it should be clear that

$$\begin{aligned} S_{13} = S_{31} = 0 & & S_{14} = S_{41} = 0 \\ S_{23} = S_{32} = 0 & & S_{24} = S_{42} = 0. \end{aligned}$$

Furthermore, since the functions have a definite parity, they are eigenfunctions of the parity operator Π with $\Pi f_{1,2} = +f_{1,2}$ and $\Pi f_{3,4} = -f_{3,4}$. And since the potential is symmetric, we have $[\Pi, H] = 0$ so that by Theorem 1.2 we know that $H_{ij} = 0$ if one index refers to an even function and the other refers to an odd function:

$$\begin{aligned} H_{13} = H_{31} = 0 & & H_{14} = H_{41} = 0 \\ H_{23} = H_{32} = 0 & & H_{24} = H_{42} = 0. \end{aligned}$$

With these simplifications, (1.6) becomes

$$\begin{vmatrix} H_{11} - WS_{11} & H_{12} - WS_{12} & 0 & 0 \\ H_{21} - WS_{21} & H_{22} - WS_{22} & 0 & 0 \\ 0 & 0 & H_{33} - WS_{33} & H_{34} - WS_{34} \\ 0 & 0 & H_{43} - WS_{43} & H_{44} - WS_{44} \end{vmatrix} = 0.$$

Since the determinant of a block diagonal matrix is the product of the determinants of the blocks, we can find all four roots by finding the two roots of each of the

following equations:

$$\begin{vmatrix} H_{11} - WS_{11} & H_{12} - WS_{12} \\ H_{21} - WS_{21} & H_{22} - WS_{22} \end{vmatrix} = 0 \quad (1.7a)$$

$$\begin{vmatrix} H_{33} - WS_{33} & H_{34} - WS_{34} \\ H_{43} - WS_{43} & H_{44} - WS_{44} \end{vmatrix} = 0. \quad (1.7b)$$

Let the roots of (1.7a) be denoted W_1, W_3 . These are the approximations to the energies of the $n = 1$ and $n = 3$ even states. Similarly, the roots W_2, W_4 of (1.7b) are the approximations to the odd energy states $n = 2$ and $n = 4$. Once we have the roots W_i , we substitute them one-at-a-time back into equation (1.5) to determine the set of coefficients $c_j^{(i)}$ corresponding to that particular root. In the particular case of W_1 , this yields the set of equations

$$(H_{11} - W_1 S_{11})c_1^{(1)} + (H_{12} - W_1 S_{12})c_2^{(1)} = 0 \quad (1.8a)$$

$$(H_{21} - W_1 S_{21})c_1^{(1)} + (H_{22} - W_1 S_{22})c_2^{(1)} = 0$$

$$(H_{33} - W_1 S_{33})c_3^{(1)} + (H_{34} - W_1 S_{34})c_4^{(1)} = 0 \quad (1.8b)$$

$$(H_{43} - W_1 S_{43})c_3^{(1)} + (H_{44} - W_1 S_{44})c_4^{(1)} = 0.$$

Now, W_1 was a root of (1.7a), so the determinant of the coefficients in (1.8a) must vanish, and we have a nontrivial solution for $c_1^{(1)}$ and $c_2^{(1)}$. However, W_1 was not a root of (1.7b), so the determinant of the coefficients in (1.8b) does not vanish, and hence there is only the trivial solution $c_3^{(1)} = c_4^{(1)} = 0$. Thus the trial function for W_1 is $\varphi_1 = c_1^{(1)} f_1 + c_2^{(1)} f_2$. Exactly the same reasoning applies to the other three roots, and we have the trial functions

$$\begin{aligned} \varphi_1 &= c_1^{(1)} f_1 + c_2^{(1)} f_2 & \varphi_3 &= c_1^{(3)} f_1 + c_2^{(3)} f_2 \\ \varphi_2 &= c_3^{(2)} f_3 + c_4^{(2)} f_4 & \varphi_4 &= c_3^{(4)} f_3 + c_4^{(4)} f_4. \end{aligned}$$

So we see that the even states ψ_1 and ψ_3 are approximated by the trial functions φ_1 and φ_3 consisting of linear combinations of the even functions f_1 and f_2 . Similarly, the odd states ψ_2 and ψ_4 are approximated by the trial functions φ_2 and φ_4 that are linear combinations of the odd functions f_3 and f_4 .

To proceed any further, we need to evaluate the non-zero integrals H_{ij} and S_{ij} . From Example 1.1 we can immediately write down H_{11} and S_{11} . The rest of the integrals are also straight-forward to evaluate, and the result is

$$\begin{aligned} H_{11} &= \hbar^2 l^3 / 6m & H_{12} = H_{21} &= \hbar^2 l^5 / 30m & H_{22} &= \hbar^2 l^7 / 105m \\ H_{33} &= \hbar^2 l^5 / 40m & H_{44} &= \hbar^2 l^9 / 1260m & H_{34} = H_{43} &= \hbar^2 l^7 / 280m \\ S_{11} &= l^5 / 30 & S_{12} = S_{21} &= l^7 / 140 & S_{22} &= l^9 / 630 \\ S_{33} &= l^7 / 840 & S_{44} &= l^{11} / 27720 & S_{34} = S_{43} &= l^9 / 5040. \end{aligned}$$

Substituting these results into equation (1.7a) to determine W_1 and W_3 we have

$$\begin{vmatrix} \frac{\hbar^2 l^3}{6m} - \frac{l^5}{30}W & \frac{\hbar^2 l^5}{30m} - \frac{l^7}{140}W \\ \frac{\hbar^2 l^5}{30m} - \frac{l^7}{140}W & \frac{\hbar^2 l^7}{105m} - \frac{l^9}{630}W \end{vmatrix} = 0.$$

To evaluate this, it is easiest to recall that multiplying any single row of a determinant by some scalar is the same as multiplying the original determinant by that same scalar. (This is an obvious consequence of the definition

$$\det A = \sum_{i_1, \dots, i_n=1}^n \varepsilon^{i_1 \dots i_n} a_{1i_1} \dots a_{ni_n}.)$$

Since the right hand side of this equation is zero, we don't change anything by multiplying any row in this determinant by some constant. Multiplying the first row by $420m/l^3$ and the second row by $1260m/l^5$ we obtain

$$\begin{vmatrix} 70\hbar^2 - 14ml^2W & 14\hbar^2 l^2 - 3ml^4W \\ 42\hbar^2 - 9ml^2W & 12\hbar^2 l^2 - 2ml^4W \end{vmatrix} = 0 \quad (1.9)$$

or

$$ml^4W^2 - 56ml^2\hbar^2W + 252\hbar^4 = 0.$$

The roots of this quadratic are

$$W_{1,3} = (\hbar^2/ml^2)(28 \pm \sqrt{532}) = 4.93487\hbar^2/ml^2, 51.0651\hbar^2/ml^2.$$

Similarly, substituting the values for H_{ij} and S_{ij} into (1.7b) results in

$$W_{2,4} = (\hbar^2/ml^2)(60 \pm \sqrt{1620}) = 19.7508\hbar^2/ml^2, 100.249\hbar^2/ml^2.$$

For comparison, the first four exact solutions $E_n = n^2\hbar^2\pi^2/2ml^2$ are

$$E_n = 4.9348\hbar^2/ml^2, 19.7392\hbar^2/ml^2, 44.4132\hbar^2/ml^2, 78.9568\hbar^2/ml^2$$

so the errors are (in the order of increasing energy levels) 0.0014%, 0.059%, 15.0% and 27.0%. As expected, we did great for $n = 1$ and $n = 2$, but not so great for $n = 3$ and $n = 4$.

We still have to find the approximate wave functions that correspond to each of the W_i 's. We want to substitute $W_1 = 4.93487\hbar^2/ml^2$ into equations (1.8a) and use the integrals we have already evaluated. However, it is somewhat easier to note that the coefficients of $c_{1,2}^{(1)}$ in equations (1.8a) are equivalent to the entries in equation (1.9). Furthermore, as we have already noted, all we can find is the ratio of the c_i 's, so the two equations in (1.9) are equivalent, and we only need to use either one of them. (That the equations are equivalent is a consequence of the fact that the determinant (1.9) is zero, so the rows must be linearly dependent. Hence we get no new information by using both rows.)

So choosing the first row we have

$$70\hbar^2 - 14ml^2W_1 = 70\hbar^2 - 14ml^2(4.93487\hbar^2/ml^2) = 0.91182\hbar^2$$

$$14\hbar^2l^2 - 3ml^4W_1 = 14\hbar^2l^2 - 3ml^4(4.93487\hbar^2/ml^2) = -0.80461\hbar^2l^2$$

so that

$$c_2^{(1)} = \frac{0.91182\hbar^2}{0.80461\hbar^2l^2}c_1^{(1)} = 1.133c_1^{(1)}/l^2.$$

To fix the value of $c_1^{(1)}$ we use the normalization condition:

$$\begin{aligned} 1 &= \langle \varphi_1 | \varphi_1 \rangle = \langle c_1^{(1)} f_1 + c_2^{(1)} f_2 | c_1^{(1)} f_1 + c_2^{(1)} f_2 \rangle \\ &= [c_1^{(1)}]^2 S_{11} + 2c_1^{(1)} c_2^{(1)} S_{12} + [c_2^{(1)}]^2 S_{22} \\ &= [c_1^{(1)}]^2 \left[S_{11} + 2 \cdot \frac{1.133}{l^2} S_{12} + \frac{(1.133)^2}{l^4} S_{22} \right] \\ &= [c_1^{(1)}]^2 \left[\frac{l^5}{30} + 2 \cdot \frac{1.133}{l^2} \frac{l^7}{140} + \frac{(1.133)^2}{l^4} \frac{l^9}{630} \right] \\ &= 0.05156 [c_1^{(1)}]^2 l^5 \end{aligned}$$

and hence $c_1^{(1)} = 4.404l^{-5/2}$.

Putting this all together we finally obtain

$$\begin{aligned} \varphi_1 &= 4.404l^{-5/2}f_1 + 4.990l^{-9/2}f_2 \\ &= 4.404l^{-5/2}x(l-x) + 4.990l^{-9/2}x^2(l-x)^2 \\ &= l^{-1/2}[4.404(x/l)(1-x/l) + 4.990(x/l)^2(1-x/l)^2]. \end{aligned}$$

As you can see from the plot below, the function φ_1 is almost identical to the exact solution $\psi_1 = \sqrt{2} \sin \pi x/l$:

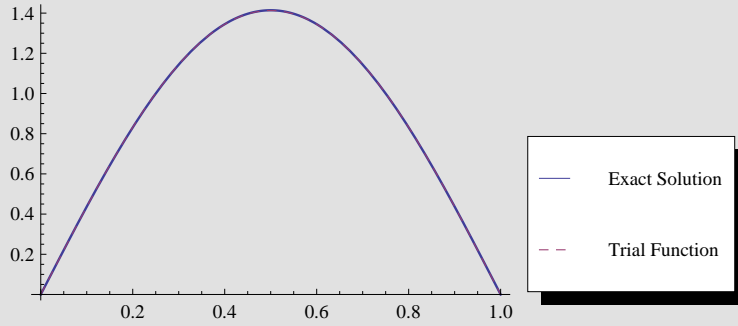


Figure 2: Plot of ψ_1 and φ_1 vs x/l .

Repeating all of this with the other roots W_2, W_3 and W_4 we eventually arrive at

$$\varphi_2 = l^{-1/2}[16.78(x/l)(1-x/l)(1/2-x/l) + 71.85(x/l)^2(1-x/l)^2(1/2-x/l)]$$

$$\varphi_3 = l^{-1/2}[28.65(x/l)(1-x/l) - 132.7(x/l)^2(1-x/l)^2]$$

$$\varphi_4 = l^{-1/2}[98.99(x/l)(1-x/l)(1/2-x/l) - 572.3(x/l)^2(1-x/l)^2(1/2-x/l)]$$

1.3.1 Proof that the Roots of the Secular Equation are Real

In this section we will prove that the roots of the polynomial in W defined by equation (1.6) are in fact real. In order to show this, we must first review some basic linear algebra.

Let V be a vector space over \mathbb{C} . By an **inner product** on V (sometimes called the **Hermitian inner product**), we mean a mapping $\langle \cdot, \cdot \rangle : V \times V \rightarrow \mathbb{C}$ such that for all $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$ and $a, b \in \mathbb{C}$ we have

- (IP1) $\langle a\mathbf{u} + b\mathbf{v}, \mathbf{w} \rangle = a^*\langle \mathbf{u}, \mathbf{w} \rangle + b^*\langle \mathbf{v}, \mathbf{w} \rangle$;
- (IP2) $\langle \mathbf{u}, \mathbf{v} \rangle = \langle \mathbf{v}, \mathbf{u} \rangle^*$;
- (IP3) $\langle \mathbf{u}, \mathbf{u} \rangle \geq 0$ and $\langle \mathbf{u}, \mathbf{u} \rangle = 0$ if and only if $\mathbf{u} = \mathbf{0}$.

If $\{\mathbf{e}_i\}$ is a basis for V , then in terms of components we have

$$\langle \mathbf{u}, \mathbf{v} \rangle = \sum_{i,j} u_i^* v_j \langle \mathbf{e}_i, \mathbf{e}_j \rangle := \sum_{i,j} u_i^* v_j g_{ij}$$

where we have defined the (square) matrix $G = (g_{ij}) = \langle \mathbf{e}_i, \mathbf{e}_j \rangle$. As a matrix product, we may write

$$\langle \mathbf{u}, \mathbf{v} \rangle = \mathbf{u}^{*T} G \mathbf{v} .$$

I emphasize that this is the most general inner product on V , and any inner product can be written in this form. (For example, if V is a real space and $g_{ij} = \langle \mathbf{e}_i, \mathbf{e}_j \rangle = \delta_{ij}$, then we obtain the usual Euclidean inner product on V .) Notice that

$$g_{ij} = \langle \mathbf{e}_i, \mathbf{e}_j \rangle = \langle \mathbf{e}_j, \mathbf{e}_i \rangle^* = g_{ji}^*$$

and hence $G = G^\dagger$ so that G is in fact a Hermitian matrix. (Some of you may realize that in the case where V is a real vector space, the matrix G is just the usual **metric** on V .)

Now, given an inner product, we may define a **norm** on V by $\|\mathbf{u}\| = \langle \mathbf{u}, \mathbf{u} \rangle^{1/2}$. Note that because of condition (IP3), we have $\|\mathbf{u}\| \geq 0$ and $\|\mathbf{u}\| = 0$ if and only if $\mathbf{u} = \mathbf{0}$. This imposes a condition on G because

$$\|\mathbf{u}\|^2 = \langle \mathbf{u}, \mathbf{u} \rangle = \mathbf{u}^{*T} G \mathbf{u} = \sum_{i,j} u_i^* u_j g_{ij} \geq 0$$

and equality holds if and only if $\mathbf{u} = \mathbf{0}$. A Hermitian matrix G with the property that $\mathbf{u}^*T G \mathbf{u} > 0$ for all $\mathbf{u} \neq \mathbf{0}$ is said to be **positive definite**.

It is important to realize that conversely, given a positive definite Hermitian matrix G , we can define an inner product by $\langle \mathbf{u}, \mathbf{v} \rangle = \mathbf{u}^*T G \mathbf{v}$. That this is true follows easily by reversing the above steps.

Another fundamental concept is that of the kernel of a linear transformation (or matrix). If T is a linear transformation, we define the **kernel** of T to be the set

$$\text{Ker } T = \{ \mathbf{u} \in V : T \mathbf{u} = \mathbf{0} \}.$$

A linear transformation whose kernel is zero is said to be **nonsingular**.

The reason the kernel is so useful is that it allows us to determine whether or not a linear transformation is an isomorphism (i.e., one-to-one). A linear transformation T on V is said to be **one-to-one** if $\mathbf{u} \neq \mathbf{v}$ implies $T \mathbf{u} \neq T \mathbf{v}$. An equivalent way to say this is that $T \mathbf{u} = T \mathbf{v}$ implies $\mathbf{u} = \mathbf{v}$ (this is the contrapositive statement). Thus, if $T \mathbf{u} = T \mathbf{v}$, then using the linearity of T we see that $\mathbf{0} = T \mathbf{u} - T \mathbf{v} = T(\mathbf{u} - \mathbf{v})$ and hence $\mathbf{u} - \mathbf{v} \in \text{Ker } T$. But if $\text{Ker } T = \{ \mathbf{0} \}$, then we in fact have $\mathbf{u} = \mathbf{v}$ so that T is an isomorphism. Conversely, if T is an isomorphism, then we must have $\text{Ker } T = \{ \mathbf{0} \}$. This is because T is one-to-one, and any linear transformation has the property that $T \mathbf{0} = \mathbf{0}$. (Because $T \mathbf{u} = T(\mathbf{u} + \mathbf{0}) = T \mathbf{u} + T \mathbf{0}$ so that $T \mathbf{0} = \mathbf{0}$.)

Now suppose that T is a nonsingular surjective (i.e., onto) linear transformation on V . Such a T is said to be a **bijection**. You should already know that the matrix representation $A = (a_{ij})$ of T with respect to the basis $\{ \mathbf{e}_i \}$ for V is defined by

$$T \mathbf{e}_i = \sum_j \mathbf{e}_j a_{ji}.$$

This is frequently written as $A = [T]_{\mathbf{e}}$. Then the fact that T is a bijection simply means that the matrix A is invertible (i.e., that A^{-1} exists).

(Actually, if $T : U \rightarrow V$ is a nonsingular (one-to-one) linear transformation between two finite-dimensional vector spaces of equal dimensions, then it is automatically surjective. This is a consequence of the well-known **rank theorem** which says

$$\text{rank } T + \dim \text{Ker } T = \dim U$$

where $\text{rank } T$ is another term for the dimension of the image of T . Therefore, if $\text{Ker } T = \{ \mathbf{0} \}$ we have $\dim \text{Ker } T = 0$ so that $\text{rank } T = \dim U = \dim V$. The proof of the rank theorem is also not hard: Let $\dim U = n$, and let $\{ w_1, \dots, w_k \}$ be a basis for $\text{Ker } T$. Extend this to a basis $\{ w_1, \dots, w_n \}$ for U . Then $\text{Im } T$ is spanned by $\{ T w_{k+1}, \dots, T w_n \}$, and it is easy to see that these are linearly independent. Thus $\dim U = n = k + (n - k) = \dim \text{Ker } T + \dim \text{Im } T$.)

Note that if G is positive definite, then we must have $\text{Ker } G = \{ \mathbf{0} \}$. This is because if $\mathbf{u} \neq \mathbf{0}$ and $G \mathbf{u} = \mathbf{0}$, we would have $\langle \mathbf{u}, \mathbf{u} \rangle = \mathbf{u}^*T G \mathbf{u} = 0$ in contradiction to the assumed positive definiteness of G . Thus a positive definite matrix is necessarily nonsingular.

Let us take a more careful look at $S_{ij} = \langle f_i | f_j \rangle$. I claim that the matrix $S = (S_{ij})$ is positive definite. To show this, I will prove a general result. Suppose

I have n linearly independent (complex) vectors $\mathbf{v}_1, \dots, \mathbf{v}_n$, and I construct the nonsingular matrix M whose columns are just the vectors \mathbf{v}_i . Letting v_{ij} denote the j th component of the vector \mathbf{v}_i , we have

$$M = \begin{bmatrix} v_{11} & v_{21} & \cdots & v_{n1} \\ v_{12} & v_{22} & \cdots & v_{n2} \\ \vdots & \vdots & & \vdots \\ v_{1n} & v_{2n} & \cdots & v_{nn} \end{bmatrix}.$$

From this we see that

$$M^\dagger = \begin{bmatrix} v_{11}^* & v_{12}^* & \cdots & v_{1n}^* \\ v_{21}^* & v_{22}^* & \cdots & v_{2n}^* \\ \vdots & \vdots & & \vdots \\ v_{n1}^* & v_{n2}^* & \cdots & v_{nn}^* \end{bmatrix}$$

and therefore

$$M^\dagger M = \begin{bmatrix} \langle \mathbf{v}_1 | \mathbf{v}_1 \rangle & \langle \mathbf{v}_1 | \mathbf{v}_2 \rangle & \cdots & \langle \mathbf{v}_1 | \mathbf{v}_n \rangle \\ \langle \mathbf{v}_2 | \mathbf{v}_1 \rangle & \langle \mathbf{v}_2 | \mathbf{v}_2 \rangle & \cdots & \langle \mathbf{v}_2 | \mathbf{v}_n \rangle \\ \vdots & \vdots & & \vdots \\ \langle \mathbf{v}_n | \mathbf{v}_1 \rangle & \langle \mathbf{v}_n | \mathbf{v}_2 \rangle & \cdots & \langle \mathbf{v}_n | \mathbf{v}_n \rangle \end{bmatrix}. \quad (1.10)$$

A matrix of this form is called a **Gram matrix**.

If I denote the Hermitian matrix $M^\dagger M$ by S , then for any vector $\mathbf{c} \neq \mathbf{0}$ we have

$$\langle \mathbf{c} | S \mathbf{c} \rangle = \langle \mathbf{c} | M^\dagger M \mathbf{c} \rangle = \langle M \mathbf{c} | M \mathbf{c} \rangle = \|M \mathbf{c}\|^2 > 0$$

so that S is positive definite. That this is strictly greater than zero (and not greater than or equal to zero) follows from the fact that M is nonsingular so its kernel is $\{\mathbf{0}\}$, together with the assumption that $\mathbf{c} \neq \mathbf{0}$. In other words, *any matrix of the form (1.10) is positive definite*.

But this is exactly what we had when we defined $S_{ij} = \langle f_i | f_j \rangle = \langle i | j \rangle$, where the linearly independent functions f_i define a basis for a vector space. In other words, what we really have is $f_i = \mathbf{v}_i$ so that the matrix $M^\dagger M$ defined above is exactly the matrix S defined by $S_{ij} = \langle i | j \rangle$.

With all of this formalism out of the way, it is now easy to show that the roots of the secular equation are real. Let us write equation (1.5) in matrix form as

$$H \mathbf{c} = W S \mathbf{c}$$

so that

$$\langle \mathbf{c} | H \mathbf{c} \rangle = W \langle \mathbf{c} | S \mathbf{c} \rangle.$$

On the other hand, using the fact that H is Hermitian and S is real and symmetric, we can write

$$\langle \mathbf{c} | H \mathbf{c} \rangle = \langle H \mathbf{c} | \mathbf{c} \rangle = \langle W S \mathbf{c} | \mathbf{c} \rangle = W^* \langle S \mathbf{c} | \mathbf{c} \rangle = W^* \langle \mathbf{c} | S \mathbf{c} \rangle.$$

Thus we have

$$(W - W^*) \langle \mathbf{c} | S \mathbf{c} \rangle = 0$$

which implies $W = W^*$ because $\mathbf{c} \neq \mathbf{0}$ so that $\langle \mathbf{c} | S \mathbf{c} \rangle > 0$.

Note that this proof is also valid in the case where φ is complex because (1.5) still holds, and $S = M^\dagger M$ is Hermitian so that $\langle S \mathbf{c} | \mathbf{c} \rangle = \langle \mathbf{c} | S \mathbf{c} \rangle$.

2 Time-Independent Perturbation Theory

2.1 Perturbation Theory for a Nondegenerate Energy Level

Suppose that we want to solve the time-independent Schrödinger equation $H\psi_n = E_n\psi_n$, but the Hamiltonian is too complicated for us to find an exact solution. However, let us suppose that the Hamiltonian can be written in the form

$$H = H^0 + \lambda H'$$

where we know the exact solutions to $H^0\psi_n^{(0)} = E_n^{(0)}\psi_n^{(0)}$. (We will use a superscript 0 to denote the energies and eigenstates of the unperturbed Hamiltonian H^0 .) The additional term H' is called a **perturbation**, and it must in some sense be considered small relative to H^0 . The dimensionless parameter λ is redundant, but is introduced for mathematical convenience; it will not remain a part of our final solution. For example, the unperturbed Hamiltonian H^0 could be the (free) hydrogen atom, and the perturbation H' could represent the interaction energy $e\mathbf{E} \cdot \mathbf{r}$ of the electron with an electric field \mathbf{E} . (This leads to an energy level shift called the **Stark effect**.)

The full (i.e., interacting or perturbed) Schrödinger equation is written

$$H\psi_n = (H^0 + \lambda H')\psi_n = E_n\psi_n \quad (2.1)$$

and the unperturbed equation is

$$H^0\psi_n^{(0)} = E_n^{(0)}\psi_n^{(0)}. \quad (2.2)$$

We think of the parameter λ as varying from 0 to 1, and taking the system smoothly from the unperturbed system described by H^0 to the fully interacting system described by H . And *as long as we are discussing nondegenerate states*, we can think of each unperturbed state $\psi_n^{(0)}$ as undergoing a smooth transition to the exact state ψ_n . In other words,

$$\lim_{\lambda \rightarrow 0} \psi_n = \psi_n^{(0)} \quad \text{and} \quad \lim_{\lambda \rightarrow 0} E_n = E_n^{(0)}.$$

Since the states $\psi_n = \psi_n(\lambda, \mathbf{x})$ and energies $E_n = E_n(\lambda)$ depend on λ , let us expand both in a Taylor series about $\lambda = 0$:

$$\begin{aligned} \psi_n &= \psi_n^{(0)} + \lambda \left(\frac{\partial \psi_n}{\partial \lambda} \right)_{\lambda=0} + \frac{\lambda^2}{2!} \left(\frac{\partial^2 \psi_n}{\partial \lambda^2} \right)_{\lambda=0} + \dots \\ E_n &= E_n^{(0)} + \lambda \left(\frac{dE_n}{d\lambda} \right)_{\lambda=0} + \frac{\lambda^2}{2!} \left(\frac{d^2 E_n}{d\lambda^2} \right)_{\lambda=0} + \dots \end{aligned}$$

Now introduce the notation

$$\psi_n^{(k)} = \frac{1}{k!} \frac{\partial^k \psi_n}{\partial \lambda^k} \Big|_{\lambda=0} \quad E_n^{(k)} = \frac{1}{k!} \frac{d^k E_n}{d\lambda^k} \Big|_{\lambda=0}$$

so we can write

$$\psi_n = \psi_n^{(0)} + \lambda\psi_n^{(1)} + \lambda^2\psi_n^{(2)} + \dots \quad (2.3a)$$

$$E_n = E_n^{(0)} + \lambda E_n^{(1)} + \lambda^2 E_n^{(2)} + \dots \quad (2.3b)$$

For each $k = 1, 2, \dots$ we call $\psi_n^{(k)}$ and $E_n^{(k)}$ the **kth-order correction** to the wavefunction and energy. We assume that the series converges for $\lambda = 1$, and that the first few terms give a good approximation to the exact solutions.

It will be convenient to simplify some of our notation, so integrals such as $\langle \psi_n^{(j)} | \psi_n^{(k)} \rangle$ will simply be written $\langle n^{(j)} | n^{(k)} \rangle$. We assume that the unperturbed states are orthonormal so that

$$\langle m^{(0)} | n^{(0)} \rangle = \delta_{mn}$$

and we also choose our normalization so that

$$\langle n^{(0)} | n \rangle = 1. \quad (2.4)$$

If this last condition on ψ_n isn't satisfied, then multiplying ψ_n by $\langle n^{(0)} | n \rangle^{-1}$ will ensure that it is. Since multiplying the Schrödinger equation $H\psi_n = E_n\psi_n$ by a constant doesn't change E_n , this has no effect on the energy levels. If so desired, at the end of the calculation we can always re-normalize ψ_n in the usual way.

Substituting (2.3a) into (2.4) yields

$$1 = \langle n^{(0)} | n^{(0)} \rangle + \lambda \langle n^{(0)} | n^{(1)} \rangle + \lambda^2 \langle n^{(0)} | n^{(2)} \rangle + \dots$$

Now, it is a general result that if you have a power series equation of the form $\sum_{n=0}^{\infty} a_n x^n = 0$ for all x , then $a_n = 0$ for all n . That $a_0 = 0$ follows by letting $x = 0$. Now take the derivative with respect to x and let $x = 0$ to obtain $a_1 = 0$. Taking the derivative again and letting $x = 0$ yields $a_2 = 0$. Clearly we can continue this procedure to arrive at $a_n = 0$ for all n . Applying this result to the above power series in λ and using the fact that $\langle n^{(0)} | n^{(0)} \rangle = 1$ we conclude that

$$\langle n^{(0)} | n^{(k)} \rangle = 0 \quad \text{for all } k = 1, 2, \dots \quad (2.5)$$

We now substitute equations (2.3) into the Schrödinger equation (2.1):

$$\begin{aligned} (H^0 + \lambda H')(\psi_n^{(0)} + \lambda\psi_n^{(1)} + \lambda^2\psi_n^{(2)} + \dots) \\ = (E_n^{(0)} + \lambda E_n^{(1)} + \lambda^2 E_n^{(2)} + \dots)(\psi_n^{(0)} + \lambda\psi_n^{(1)} + \lambda^2\psi_n^{(2)} + \dots) \end{aligned}$$

or, grouping powers of λ ,

$$\begin{aligned} H^0\psi_n^{(0)} + \lambda(H^0\psi_n^{(1)} + H'\psi_n^{(0)}) + \lambda^2(H^0\psi_n^{(2)} + H'\psi_n^{(1)}) + \dots \\ = E_n^{(0)}\psi_n^{(0)} + \lambda(E_n^{(0)}\psi_n^{(1)} + E_n^{(1)}\psi_n^{(0)}) \\ + \lambda^2(E_n^{(0)}\psi_n^{(2)} + E_n^{(1)}\psi_n^{(1)} + E_n^{(2)}\psi_n^{(0)}) + \dots \end{aligned}$$

Again ignoring questions of convergence, we can equate powers of λ on both sides of this equation. For λ^0 we simply have

$$H^0 \psi_n^{(0)} = E_n^{(0)} \psi_n^{(0)} \quad (2.6a)$$

which doesn't tell us anything new. For λ^1 we have

$$H^0 \psi_n^{(1)} + H' \psi_n^{(0)} = E_n^{(0)} \psi_n^{(1)} + E_n^{(1)} \psi_n^{(0)}$$

or

$$(H^0 - E_n^{(0)}) \psi_n^{(1)} = (E_n^{(1)} - H') \psi_n^{(0)}. \quad (2.6b)$$

For λ^2 we have

$$H^{(0)} \psi_n^{(2)} + H' \psi_n^{(1)} = E_n^{(0)} \psi_n^{(2)} + E_n^{(1)} \psi_n^{(1)} + E_n^{(2)} \psi_n^{(0)}$$

or

$$(H^0 - E_n^{(0)}) \psi_n^{(2)} = (E_n^{(1)} - H') \psi_n^{(1)} + E_n^{(2)} \psi_n^{(0)}. \quad (2.6c)$$

And in general we have for $k \geq 1$

$$(H^0 - E_n^{(0)}) \psi_n^{(k)} = (E_n^{(1)} - H') \psi_n^{(k-1)} + E_n^{(2)} \psi_n^{(k-2)} + \dots + E_n^{(k)} \psi_n^{(0)}. \quad (2.6d)$$

Notice that at each step along the way, $\psi_n^{(k)}$ is determined by $\psi_n^{(k-1)}$, $\psi_n^{(k-2)}$, \dots , $\psi_n^{(0)}$. We can also add an arbitrary multiple of $\psi_n^{(0)}$ to each $\psi_n^{(k)}$ without affecting the left side of these equations. Hence we can choose this multiple so that $\langle n^{(0)} | n^{(k)} \rangle = 0$ for $k \geq 1$, which is the same result as we had in (2.5).

Now using the hermiticity of H^0 and the fact that $E_n^{(0)}$ is real, we have

$$\langle n^{(0)} | H^0 n^{(k)} \rangle = \langle H^0 n^{(0)} | n^{(k)} \rangle = E_n^{(0)} \langle n^{(0)} | n^{(k)} \rangle = 0 \quad \text{for } k \geq 1.$$

Then multiplying (2.6d) from the left by $\psi_n^{(0)*}$ and integrating, we see that the left-hand side vanishes, and we are left with (since $\langle n^{(0)} | n^{(0)} \rangle = 1$)

$$0 = -\langle n^{(0)} | H' n^{(k-1)} \rangle + E_n^{(k)}$$

or

$$E_n^{(k)} = \langle n^{(0)} | H' n^{(k-1)} \rangle \quad \text{for } k \geq 1. \quad (2.7)$$

In particular, we have the extremely important result for the first order energy correction to the n th state

$$\boxed{E_n^{(1)} = \langle n^{(0)} | H' n^{(0)} \rangle = \int \psi_n^{(0)*} H' \psi_n^{(0)} d\mathbf{x}.} \quad (2.8)$$

Letting $\lambda = 1$ in (2.3b), we see that to first order, the energy of the n th state is given by

$$E_n \approx E_n^{(0)} + E_n^{(1)} = E_n^{(0)} + \int \psi_n^{(0)*} H' \psi_n^{(0)} d\mathbf{x}.$$

Example 2.1. Let the unperturbed system be the free harmonic oscillator, with ground-state wavefunction

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

and energy levels

$$E_n^{(0)} = \left(n + \frac{1}{2}\right)\hbar\omega.$$

Now consider the anharmonic oscillator with Hamiltonian

$$H = H^0 + H' := H^0 + ax^3 + bx^4.$$

The first-order energy correction to the ground state is given by

$$E_0^{(1)} = \langle n^{(0)} | H' | n^{(0)} \rangle = \left(\frac{m\omega}{\pi\hbar}\right)^{1/2} \int_{-\infty}^{\infty} e^{-m\omega x^2/\hbar} (ax^3 + bx^4) dx.$$

However, the integral over x^3 vanishes by symmetry (the integral of an odd function over an even interval), and we are left with

$$\begin{aligned} E_0^{(1)} &= b \left(\frac{m\omega}{\pi\hbar}\right)^{1/2} \int_{-\infty}^{\infty} x^4 e^{-m\omega x^2/\hbar} dx = b \left(\frac{\alpha}{\pi}\right)^{1/2} \int_{-\infty}^{\infty} x^4 e^{-\alpha x^2} dx \\ &= b \left(\frac{\alpha}{\pi}\right)^{1/2} \frac{\partial^2}{\partial \alpha^2} \int_{-\infty}^{\infty} e^{-\alpha x^2} dx = b \left(\frac{\alpha}{\pi}\right)^{1/2} \frac{\partial^2}{\partial \alpha^2} \left(\frac{\pi}{\alpha}\right)^{1/2} \\ &= \frac{3b}{4\alpha^2} = \frac{3b}{4} \frac{\hbar^2}{m^2\omega^2}. \end{aligned}$$

Thus, to first order, the ground state energy of the anharmonic oscillator is given by

$$E_0 \approx E_0^{(0)} + E_0^{(1)} = \frac{1}{2}\hbar\omega + \frac{3b}{4} \frac{\hbar^2}{m^2\omega^2}.$$

Now let's find the first-order correction to the wavefunction. Since the unperturbed states $\psi_n^{(0)}$ form a complete orthonormal set, we may expand $\psi_n^{(1)}$ in terms of them as

$$\psi_n^{(1)} = \sum_m a_m \psi_m^{(0)}$$

where

$$a_m = \langle m^{(0)} | n^{(1)} \rangle.$$

(It would be way too cluttered to try and label these expansion coefficients to denote the fact that they also refer to the first-order correction of the n th state.) Then for

$m \neq n$, we multiply (2.6b) from the left by $\psi_m^{(0)}$ and integrate:

$$\langle m^{(0)} | (H^0 - E_n^{(0)}) n^{(1)} \rangle = E_n^{(1)} \langle m^{(0)} | n^{(0)} \rangle - \langle m^{(0)} | H' n^{(0)} \rangle$$

or (since $H^0 \psi_m^{(0)} = E_m^{(0)} \psi_m^{(0)}$ and $\langle m^{(0)} | n^{(0)} \rangle = 0$ for $m \neq n$)

$$(E_m^{(0)} - E_n^{(0)}) \langle m^{(0)} | n^{(1)} \rangle = -\langle m^{(0)} | H' n^{(0)} \rangle.$$

Therefore

$$a_m = \langle m^{(0)} | n^{(1)} \rangle = \frac{\langle m^{(0)} | H' n^{(0)} \rangle}{E_n^{(0)} - E_m^{(0)}} \quad \text{for } m \neq n. \quad (2.9)$$

You should realize that this last step was where the assumed nondegeneracy of the states came in. In order for us to divide by $E_n^{(0)} - E_m^{(0)}$, we must assume that it is nonzero. This is true as long as $m \neq n$ implies that $E_n^{(0)} \neq E_m^{(0)}$. Since $a_n = \langle n^{(0)} | n^{(1)} \rangle = 0$ (this is equation (2.5)), we finally obtain

$$\psi_n^{(1)} = \sum_{m \neq n} \frac{\langle m^{(0)} | H' n^{(0)} \rangle}{E_n^{(0)} - E_m^{(0)}} \psi_m^{(0)}. \quad (2.10)$$

Now that we have the first-order correction to the wavefunction, it is easy to get the second-order correction to the energy. Using (2.10) in (2.7) with $k = 2$ we immediately have

$$E_n^{(2)} = \sum_{m \neq n} \frac{\langle m^{(0)} | H' n^{(0)} \rangle \langle n^{(0)} | H' m^{(0)} \rangle}{E_n^{(0)} - E_m^{(0)}} = \sum_{m \neq n} \frac{|\langle n^{(0)} | H' m^{(0)} \rangle|^2}{E_n^{(0)} - E_m^{(0)}}. \quad (2.11)$$

The last term we will compute is the second-order correction to the wavefunction. We again expand in terms of the $\psi_n^{(0)}$ as

$$\psi_n^{(2)} = \sum_m b_m \psi_m^{(0)}$$

where $b_m = \langle m^{(0)} | n^{(2)} \rangle$. Multiplying (2.6c) from the left by $\psi_m^{(0)*}$ and integrating we have (assuming $m \neq n$)

$$(E_m^{(0)} - E_n^{(0)}) \langle m^{(0)} | n^{(2)} \rangle = E_n^{(1)} \langle m^{(0)} | n^{(1)} \rangle - \langle m^{(0)} | H' n^{(1)} \rangle$$

or

$$b_m = \langle m^{(0)} | n^{(2)} \rangle = \frac{E_n^{(1)}}{E_m^{(0)} - E_n^{(0)}} \langle m^{(0)} | n^{(1)} \rangle - \frac{\langle m^{(0)} | H' n^{(1)} \rangle}{E_m^{(0)} - E_n^{(0)}}.$$

Now use (2.9) in the first term on the right-hand side and use (2.10) in the second term to write

$$b_m = -\frac{E_n^{(1)} \langle m^{(0)} | H' n^{(0)} \rangle}{(E_n^{(0)} - E_m^{(0)})^2} - \sum_{k \neq n} \frac{\langle m^{(0)} | H' k^{(0)} \rangle \langle k^{(0)} | H' n^{(0)} \rangle}{(E_m^{(0)} - E_n^{(0)})(E_n^{(0)} - E_k^{(0)})}.$$

Using (2.8) we finally obtain

$$\begin{aligned} \psi_n^{(2)} = & \sum_{m \neq n} \sum_{k \neq n} \frac{\langle m^{(0)} | H' k^{(0)} \rangle \langle k^{(0)} | H' n^{(0)} \rangle}{(E_n^{(0)} - E_m^{(0)})(E_n^{(0)} - E_k^{(0)})} \psi_m^{(0)} \\ & - \sum_{m \neq n} \frac{\langle m^{(0)} | H' n^{(0)} \rangle \langle n^{(0)} | H' n^{(0)} \rangle}{(E_n^{(0)} - E_m^{(0)})^2} \psi_m^{(0)}. \end{aligned} \quad (2.12)$$

Let me make several points. First, recall that because of equation (2.4), our states are not normalized. Second, be sure to realize that the sums in equations (2.10), (2.11) and (2.12) are over *states*, and not energy levels. If some of the energy levels other than the n th are degenerate, then we must include a term in each of these sums for each linearly independent wavefunction corresponding to the degenerate energy level. The reason for this is that the expansions of $\psi_n^{(1)}$ and $\psi_n^{(2)}$ were in terms of a *complete* set of functions, and hence we must be sure to include all linearly independent states in the sums. Furthermore, if there happens to be a continuum of states in the unperturbed system, then we must also include an integral over these so that we have included *all* linearly independent states in our expansion.

2.2 Perturbation Theory for a Degenerate Energy Level

We now turn to the perturbation treatment of a degenerate energy level, meaning that there are multiple unperturbed states that all have the same energy. If we let d be the degree of degeneracy, then we have states $\psi_1^{(0)}, \dots, \psi_d^{(0)}$ satisfying the unperturbed Schrödinger equation

$$H^0 \psi_n^{(0)} = E_n^{(0)} \psi_n^{(0)} \quad (2.13a)$$

with

$$E_1^{(0)} = E_2^{(0)} = \dots = E_d^{(0)}. \quad (2.13b)$$

You must be careful with the notation here, because we don't want to clutter it up with too many indices. Even though we write $E_1^{(0)}, \dots, E_d^{(0)}$, this does *not* mean that these are necessarily the d lowest-lying states that satisfy the unperturbed Schrödinger equation. We are referring here to a single degenerate energy level.

The interacting (or perturbed) Schrödinger equation is

$$H \psi_n = (H^0 + \lambda H') \psi_n = E_n \psi_n.$$

In our treatment of a nondegenerate energy level, we assumed that $\lim_{\lambda \rightarrow 0} E_n = E_n^{(0)}$ and $\lim_{\lambda \rightarrow 0} \psi_n = \psi_n^{(0)}$ where the state $\psi_n^{(0)}$ was unique. However, in the case of degeneracy, the second of these does not hold. While it is true that as λ goes to zero we still have

$$\lim_{\lambda \rightarrow 0} E_n = E_n^{(0)}$$

the presence of the perturbation generally splits the degenerate energy level into multiple distinct states. However, there are varying degrees of splitting, and while the perturbation may completely remove the degeneracy, it may also only partially remove it or have no effect at all. This is illustrated in the figure below.

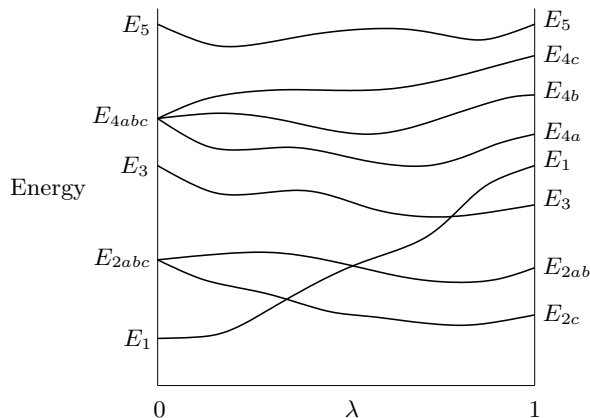


Figure 3: Splitting of energy levels due to a perturbation.

The important point to realize here is that in the limit $\lambda \rightarrow 0$, the state ψ_n does not necessarily go to a unique $\psi_n^{(0)}$, but rather only to some linear combination of the normalized degenerate states $\psi_1^{(0)}, \dots, \psi_d^{(0)}$. This is because any such linear combination

$$c_1\psi_1^{(0)} + c_2\psi_2^{(0)} + \dots + c_d\psi_d^{(0)}$$

will satisfy (2.13a) with the same eigenvalue $E_n^{(0)}$. Thus there are an infinite number of such linear combinations made up of these d linearly independent normalized eigenfunctions, and any of them will work as the unperturbed state.

For example, recall that the hydrogen atom states are labeled ψ_{nlm} where the energy only depends on n and l , and the factor $e^{im\phi}$ makes the wave function complex for $m \neq 0$. The $2p$ states correspond to $n = 2$ and $l = 1$, and these are broken into the wave functions $2p_1$ and $2p_{-1}$. However, instead of these complex wave functions, we can take the *real* linear combinations defined by

$$\psi_{2p_x} = \frac{1}{\sqrt{2}}(\psi_{2p_1} + \psi_{2p_{-1}})$$

and

$$\psi_{2p_y} = \frac{1}{i\sqrt{2}}(\psi_{2p_1} - \psi_{2p_{-1}})$$

which have the same energies. For most purposes in chemistry, these real wave functions are much more convenient to work with. And while the $2p_0$, $2p_1$ and $2p_{-1}$ states are degenerate, the presence of an electric or magnetic field will split the

degeneracy because the interaction term in the Hamiltonian depends on the spin of the electron (i.e., the m value).

Returning to our problem, all we can say is that

$$\lim_{\lambda \rightarrow 0} \psi_n = \sum_{i=1}^d c_i \psi_i^{(0)}, \quad 1 \leq n \leq d.$$

Hence the first thing we must do is determine the correct zeroth-order wave functions, which we denote by $\phi_n^{(0)}$. In other words,

$$\phi_n^{(0)} := \lim_{\lambda \rightarrow 0} \psi_n = \sum_{i=1}^d c_i \psi_i^{(0)}, \quad 1 \leq n \leq d \quad (2.14)$$

where each $\phi_n^{(0)}$ has a different set of coefficients c_i . (These should be labeled $c_i^{(n)}$, but I'm trying to keep it simple.) Note that since $H^0 \psi_i^{(0)} = E_d^{(0)} \psi_i^{(0)}$ for each $i = 1, \dots, d$ it follows that

$$H^0 \phi_n^{(0)} = E_d^{(0)} \phi_n^{(0)}. \quad (2.15)$$

For the d -fold degenerate case, we proceed as in the nondegenerate case, except that now we use $\phi_n^{(0)}$ instead of $\psi_n^{(0)}$ for the zeroth-order wave function. Then equations (2.3) become

$$\psi_n = \phi_n^{(0)} + \lambda \psi_n^{(1)} + \lambda^2 \psi_n^{(2)} + \dots \quad (2.16a)$$

$$E_n = E_d^{(0)} + \lambda E_n^{(1)} + \lambda^2 E_n^{(2)} + \dots \quad (2.16b)$$

where we have used (2.13b). Equations (2.16) apply for each $n = 1, \dots, d$. As in the nondegenerate case, we substitute these into the Schrödinger equation $H\psi_n = E_n\psi_n$ and equate powers of λ . This is exactly the same as we had before, except that now we have $\phi_n^{(0)}$ instead of $\psi_n^{(0)}$, so we can immediately write down the results from equations (2.6).

Equating the coefficients of λ^0 we have $H^0 \phi_n^{(0)} = E_d^{(0)} \phi_n^{(0)}$. Since for each $n = 1, \dots, d$ the linear combination $\phi_n^{(0)}$ is an eigenstate of H^0 with eigenvalue $E_d^{(0)}$ (this is just the statement of equation (2.15)), this doesn't give us any new information.

From the coefficients of λ^1 we have (for each $n = 1, \dots, d$)

$$(H^0 - E_d^{(0)})\psi_n^{(1)} = (E_n^{(1)} - H')\phi_n^{(0)}. \quad (2.17)$$

Multiplying this from the left by $\phi_n^{(0)*}$ and integrating we have (here I'm not using $n^{(0)}$ as a shorthand for $\psi_n^{(0)}$ to make sure there is no confusion with $\phi_n^{(0)}$)

$$\langle \phi_n^{(0)} | H^0 \psi_n^{(0)} \rangle - E_d^{(0)} \langle \phi_n^{(0)} | \psi_n^{(0)} \rangle = E_n^{(1)} \langle \phi_n^{(0)} | \phi_n^{(0)} \rangle - \langle \phi_n^{(0)} | H' \phi_n^{(0)} \rangle.$$

Using (2.15) we see that the left-hand side of this equation vanishes, so assuming that the correct zeroth-order wave functions are normalized, we arrive at the first order correction to the energy

$$E_n^{(1)} = \langle \phi_n^{(0)} | H' \phi_n^{(0)} \rangle. \quad (2.18)$$

This is similar to the nondegenerate result (2.8) except that now we use the correct zeroth-order wave functions. Of course, in order to evaluate these integrals, we must know the functions $\phi_n^{(0)}$ which, so far, we don't.

So, for any $1 \leq m \leq d$, we multiply (2.17) from the left by one of the d -fold degenerate unperturbed wave functions $\psi_m^{(0)}$ and integrate to obtain

$$\langle \psi_m^{(0)} | H^0 \psi_n^{(1)} \rangle - E_d^{(0)} \langle \psi_m^{(0)} | \psi_n^{(1)} \rangle = E_n^{(1)} \langle \psi_m^{(0)} | \phi_n^{(0)} \rangle - \langle \psi_m^{(0)} | H' \phi_n^{(0)} \rangle.$$

Since $H^0 \psi_m^{(0)} = E_d^{(0)} \psi_m^{(0)}$, we see that the left-hand side of this equation vanishes, and we are left with

$$\langle \psi_m^{(0)} | H' \phi_n^{(0)} \rangle - E_n^{(1)} \langle \psi_m^{(0)} | \phi_n^{(0)} \rangle = 0, \quad m = 1, \dots, d.$$

There is no loss of generality in assuming that the zeroth-order wave functions $\psi_i^{(0)}$ of the degenerate level are orthonormal, so we take

$$\langle \psi_m^{(0)} | \psi_i^{(0)} \rangle = \delta_{mi} \quad \text{for } m, i = 1, \dots, d. \quad (2.19)$$

(If the zeroth-order wave functions $\psi_i^{(0)}$ aren't orthonormal, then apply the Gram-Schmidt process to construct an orthonormal set. Since the new orthonormal functions are just linear combinations of the original set, and the correct zeroth-order functions $\phi_n^{(0)}$ are linear combinations of the $\psi_i^{(0)}$, the $\phi_n^{(0)}$ will just be different linear combinations of the new orthonormal functions.) Then substituting the definition (2.14) for $\phi_n^{(0)}$ we have

$$\sum_{i=1}^d c_i \langle \psi_m^{(0)} | H' \psi_i^{(0)} \rangle - E_n^{(1)} \sum_{i=1}^d c_i \langle \psi_m^{(0)} | \psi_i^{(0)} \rangle = 0$$

or

$$\sum_{i=1}^d (H'_{mi} - E_n^{(1)} \delta_{mi}) c_i = 0, \quad m = 1, \dots, d \quad (2.20a)$$

where

$$H'_{mi} = \langle \psi_m^{(0)} | H' \psi_i^{(0)} \rangle.$$

This is just another homogeneous system of d equations in the d unknowns c_i . In fact, if we let \mathbf{c} be the vector with components c_i , then we can write (2.20a) in matrix form as

$$H' \mathbf{c} = E_n^{(1)} \mathbf{c} \quad (2.20b)$$

which shows that this is nothing more than an eigenvalue equation for the matrix H' acting on the d -dimensional eigenspace of degenerate wave functions.

As usual, if (2.20a) is to have a nontrivial solution, we must have the secular equation

$$\boxed{\det(H'_{mi} - E_n^{(1)} \delta_{mi}) = 0.} \quad (2.21)$$

Written out, this looks like

$$\begin{vmatrix} H'_{11} - E_n^{(1)} & H'_{12} & \cdots & H'_{1d} \\ H'_{21} & H'_{22} - E_n^{(1)} & \cdots & H'_{2d} \\ \vdots & \vdots & \ddots & \vdots \\ H'_{d1} & H'_{d2} & \cdots & H'_{dd} - E_n^{(1)} \end{vmatrix} = 0.$$

This is a polynomial of degree d in $E_n^{(1)}$, and the d roots $E_1^{(1)}, E_2^{(1)}, \dots, E_d^{(1)}$ are the first-order corrections to the energy of the d -fold degenerate unperturbed state. So, we solve (2.21) for the eigenvalues $E_n^{(1)}$, and use these in (2.20b) to solve for the eigenvectors \mathbf{c} . These then define the correct zeroth-order wave functions according to (2.14).

Again, note that all we are doing is finding the eigenvalues and eigenvectors of the matrix H'_{mi} . And since H' is Hermitian, eigenvectors belonging to distinct eigenvalues are orthogonal. But each eigenvector \mathbf{c} has components that are just the expansion coefficients in (2.14), and therefore (reverting to a more complete notation)

$$\begin{aligned} \langle \phi_m^{(0)} | H' \phi_n^{(0)} \rangle &= \sum_{i,j=1}^d c_i^{(m)*} \langle \psi_i^{(0)} | H' \psi_j^{(0)} \rangle c_j^{(n)} = \sum_{i,j=1}^d c_i^{(m)*} H'_{ij} c_j^{(n)} \\ &= \mathbf{c}^{(m)\dagger} H' \mathbf{c}^{(n)} = E_n^{(1)} \mathbf{c}^{(m)\dagger} \mathbf{c}^{(n)} \\ &= E_n^{(1)} \langle \mathbf{c}^{(m)} | \mathbf{c}^{(n)} \rangle \end{aligned}$$

or

$$\langle \phi_m^{(0)} | H' \phi_n^{(0)} \rangle = E_n^{(1)} \delta_{mn} \quad (2.22)$$

where we assume that the eigenvectors are normalized.

In the case where $m = n$, we arrive back at (2.18). What about the case $m \neq n$? Recall that in our treatment of nondegenerate perturbation theory, the reason we had to assume the nondegeneracy was because equations (2.10) and (2.11) would blow up if there were another state $\psi_m^{(0)}$ with the same energy as $\psi_n^{(0)}$. However, in that case, we would be saved if the numerator also went to zero, and that is precisely what happens if we use the correct zeroth-order wave functions. Essentially then, the degenerate case proceeds just like the nondegenerate case, except that we must use the correct zeroth-order wave functions.

Returning to (2.21), if all d roots are distinct, then we have completely split the degeneracy into d distinct levels

$$E_d^{(0)} + E_1^{(1)}, \quad E_d^{(0)} + E_2^{(1)}, \quad \dots, \quad E_d^{(0)} + E_d^{(1)}.$$

If not all of the roots are distinct, then we have only partly removed the degeneracy (at least to first order). We will assume that all d roots are distinct, and hence that the degeneracy has been completely lifted in first order.

Now that we have the d roots $E_n^{(1)}$, we can take them one-at-a-time and plug back into the system of equations (2.20a) and solve for c_2, \dots, c_d in terms of c_1 . (Recall that because the determinant of the coefficient matrix of the system (2.20a) is zero, the d equations in (2.20a) are linearly dependent, and hence we can only find $d-1$ of the unknowns in terms of one of them.) Finally, we fix c_1 by normalization, using equations (2.14) and (2.19):

$$1 = \langle \phi_n^{(0)} | \phi_n^{(0)} \rangle = \sum_{i,j=1}^d c_i^* c_j \langle \psi_i^{(0)} | \psi_j^{(0)} \rangle = \sum_{i,j=1}^d c_i^* c_j \delta_{ij} = \sum_{i=1}^d |c_i|^2. \quad (2.23)$$

Also be sure to realize that we obtain a separate set of coefficients c_i for each root $E_n^{(1)}$. This is how we get the d independent zeroth-order wave functions.

Obviously, finding the roots of (2.21) is a difficult problem in general. However, under some special conditions, the problem may be much more tractable. The best situation would be if all off-diagonal elements H_{mi} , $m \neq i$ vanished. Then the determinant is just the product of the diagonal elements, and the d roots are simply $E_n^{(1)} = H'_{mm}$ for $m = 1, \dots, d$ or

$$E_1^{(1)} = H'_{11}, \quad E_2^{(1)} = H'_{22}, \quad \dots, \quad E_d^{(1)} = H'_{dd}.$$

Let us assume that all d roots are distinct. Taking the root $E_n^{(1)} = E_1^{(1)} = H'_{11}$ as a specific example, (2.20a) becomes the set of $d-1$ equations

$$\begin{aligned} (H'_{22} - E_1^{(1)})c_2 &= 0 \\ (H'_{33} - E_1^{(1)})c_3 &= 0 \\ &\vdots \\ (H'_{dd} - E_1^{(1)})c_d &= 0. \end{aligned}$$

Since $E_1^{(1)} = H'_{11} \neq H'_{mm}$ for $m = 2, 3, \dots, d$, it follows that $c_2 = c_3 = \dots = c_d = 0$. Normalization then implies that $c_1 = 1$, and the corresponding zeroth-order wave function defined by (2.14) is $\phi_1^{(0)} = \psi_1^{(0)}$. Clearly this applies to any of the d roots, so we have

$$\phi_i^{(0)} = \psi_i^{(0)}, \quad i = 1, \dots, d.$$

Thus we have shown that when the secular equation is diagonal and the d matrix elements H'_{mm} are all distinct, then the initial wave functions $\psi_i^{(0)}$ are the correct zeroth-order wave functions $\phi_i^{(0)}$.

Another situation that lends itself to a relatively simple solution is when the secular determinant is block diagonal. For example, in the case where $d = 4$ we

would have

$$\begin{vmatrix} H'_{11} - E_n^{(1)} & H'_{12} & 0 & 0 \\ H'_{21} & H'_{22} - E_n^{(1)} & 0 & 0 \\ 0 & 0 & H'_{33} - E_n^{(1)} & H'_{34} \\ 0 & 0 & H'_{43} & H'_{44} - E_n^{(1)} \end{vmatrix} = 0.$$

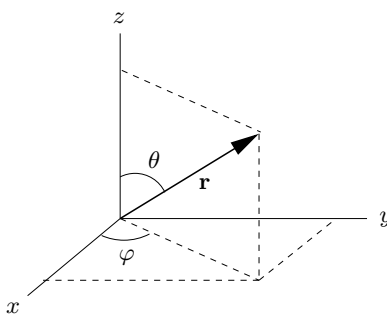
This is of the same form as we had in Example 1.3 (except with $S_{ij} = \delta_{ij}$). Exactly the same reasoning we used to show that two of the variation functions were linear combinations of f_1 and f_2 and two of the variation functions were linear combinations of f_3 and f_4 now shows that the correct zeroth-order wave functions are of the form

$$\begin{aligned} \phi_1^{(0)} &= c_1^{(1)} \psi_1^{(0)} + c_2^{(1)} \psi_2^{(0)} & \phi_2^{(0)} &= c_1^{(2)} \psi_1^{(0)} + c_2^{(2)} \psi_2^{(0)} \\ \phi_3^{(0)} &= c_3^{(3)} \psi_3^{(0)} + c_4^{(3)} \psi_4^{(0)} & \phi_4^{(0)} &= c_3^{(4)} \psi_3^{(0)} + c_4^{(4)} \psi_4^{(0)} \end{aligned}$$

Is there any way we can choose our initial wave functions $\psi_i^{(0)}$ to make things easier? Well, referring back to Theorem 1.2, suppose we have a Hermitian operator A that commutes with both H^0 and H' . If we choose our initial wave functions to be eigenfunctions of both A and H^0 , then the off-diagonal matrix elements $H'_{ij} = \langle \psi_i^{(0)} | H' \psi_j^{(0)} \rangle$ will vanish if $\psi_i^{(0)}$ and $\psi_j^{(0)}$ belong to different eigenspaces of A . Therefore, if the functions $\psi_i^{(0)}$ all have different eigenvalues of A , the secular determinant will be diagonal so that the $\phi_i^{(0)} = \psi_i^{(0)}$.

If more than one $\psi_i^{(0)}$ belongs to a given eigenvalue a_k of A (in other words, $\dim V_{a_k} > 1$), then this subcollection will form a block in the secular determinant. So in general, we will have a secular determinant that is block diagonal where each block has size $\dim V_{a_k}$. In this case, each correct zeroth-order wave function will be a linear combination of those $\psi_i^{(0)}$ that belong to the same eigenvalue of A .

Before proceeding with an example, let me prove a very important and useful property of the spherical harmonics. The parity operation is $\mathbf{r} \rightarrow -\mathbf{r}$, and in spherical coordinates, this is equivalent to $\theta \rightarrow \pi - \theta$ and $\varphi \rightarrow \varphi + \pi$.



Indeed, we know that (for the unit sphere) $z = \cos \theta$, and from the figure we see that $-z$ would be at $\pi - \theta$. Similarly, a point on the x -axis at $\varphi = 0$ goes to the point $-x$ at $\varphi = \pi$. Alternatively, letting $\theta \rightarrow \pi - \theta$ in $x = \sin \theta \cos \varphi$ doesn't change x , so in order to have $x \rightarrow -x$ we need $\cos \varphi \rightarrow -\cos \varphi$ which is accomplished by letting $\varphi \rightarrow \varphi + \pi$.

Now observe that under parity, $\mathbf{r} \rightarrow -\mathbf{r}$ and $\mathbf{p} \rightarrow -\mathbf{p}$, so that $\mathbf{L} = \mathbf{r} \times \mathbf{p}$ is unchanged. Thus angular momentum is a pseudo-vector, as you probably already knew. But this means that the parity operation Π commutes with the quantum mechanical operator \mathbf{L} , so that the three operators \mathbf{L}^2 , L_z and Π are mutually commuting, and the eigenfunctions $Y_l^m(\theta, \varphi)$ of angular momentum can be chosen to have a definite parity. Note also that since Π and \mathbf{L} commute, it follows that Π and L_{\pm} commute, so acting on any Y_l^m with L_{\pm} won't change its parity.

Look at the explicit form of the state Y_l^l :

$$Y_l^l(\theta, \varphi) = (-1)^l \left[\frac{(2l+1)!}{4\pi} \right]^{1/2} \frac{1}{2^l l!} (\sin \theta)^l e^{il\varphi}.$$

Letting $\theta \rightarrow \pi - \theta$ we have $(\sin \theta)^l \rightarrow (\sin \theta)^l$, but under $\varphi \rightarrow \varphi + \pi$ we have $e^{il\varphi} \rightarrow e^{i(l\pi + \varphi)} = (-1)^l e^{il\varphi}$. Therefore, under parity we see that $Y_l^l \rightarrow (-1)^l Y_l^l$. But we can get to any Y_l^m by repeatedly applying L_- to Y_l^l , and since this doesn't change the parity of Y_l^m we have the extremely useful result

$$\boxed{\Pi Y_l^m(\theta, \varphi) = (-1)^l Y_l^m(\theta, \varphi).} \quad (2.24)$$

Example 2.2 (Stark Effect). In this example we will take a look at the effect of a *uniform* electric field $\mathbf{E} = \mathcal{E} \hat{\mathbf{z}}$ on a hydrogen atom, where the unperturbed Hamiltonian is given by

$$H^0 = \frac{p^2}{2m} - \frac{e^2}{r}.$$

and $\mathbf{r} = \mathbf{r}_1 - \mathbf{r}_2$ is the relative position vector from the proton to the electron. We first need to find the perturbing potential energy.

The force on a particle of charge q in an electric field $\mathbf{E} = -\nabla\phi$ is $\mathbf{F} = q\mathbf{E} = -q\nabla\phi$ where $\phi(\mathbf{r})$ is the electric potential. On the other hand, the force is also given in terms of the potential energy $V(\mathbf{r})$ by $\mathbf{F} = -\nabla V$, and hence $\nabla V = q\nabla\phi$ so that

$$\int_0^{\mathbf{r}} \nabla V \cdot d\mathbf{r} = q \int_0^{\mathbf{r}} \nabla\phi \cdot d\mathbf{r}$$

or

$$V(\mathbf{r}) - V(\mathbf{0}) = q[\phi(\mathbf{r}) - \phi(\mathbf{0})].$$

If we take $V(\mathbf{0}) = \phi(\mathbf{0}) = 0$, then we have

$$V(\mathbf{r}) = q\phi(\mathbf{r}).$$

Thus the interaction Hamiltonian H' consists of both the energy $e\phi(\mathbf{r}_2)$ of the proton and the energy $-e\phi(\mathbf{r}_1)$ of the electron, and therefore

$$H' = e[\phi(\mathbf{r}_2) - \phi(\mathbf{r}_1)].$$

But the electric field is constant so that

$$\int_{\mathbf{r}_1}^{\mathbf{r}_2} \mathbf{E} \cdot d\mathbf{r} = \mathbf{E} \cdot (\mathbf{r}_2 - \mathbf{r}_1) = -\mathbf{E} \cdot (\mathbf{r}_1 - \mathbf{r}_2) = -\mathbf{E} \cdot \mathbf{r} = -\mathcal{E}z$$

while we also have

$$\int_{\mathbf{r}_1}^{\mathbf{r}_2} \mathbf{E} \cdot d\mathbf{r} = - \int_{\mathbf{r}_1}^{\mathbf{r}_2} \nabla\phi \cdot d\mathbf{r} = -[\phi(\mathbf{r}_2) - \phi(\mathbf{r}_1)].$$

Hence the final form of our perturbation is $H' = e\mathbf{E} \cdot \mathbf{r}$ or

$$H' = e\mathcal{E}z.$$

Note also that if we define the **electric dipole moment** $\boldsymbol{\mu}_e = e(\mathbf{r}_2 - \mathbf{r}_1) = -e\mathbf{r}$, then H' can be called a **dipole interaction** because

$$H' = -\boldsymbol{\mu}_e \cdot \mathbf{E}.$$

Let us first consider the ground state ψ_{100} of the hydrogen atom. This state is nondegenerate, so the first-order energy correction to the ground state is, from equation (2.8),

$$E_{100}^{(1)} = \langle \psi_{100} | e\mathcal{E}z | \psi_{100} \rangle = e\mathcal{E} \langle \psi_{100} | z | \psi_{100} \rangle.$$

But H^0 is parity invariant, so the states ψ_{nlm} all have a definite parity $(-1)^l$. Then $E_{100}^{(1)}$ is the integral of an odd function over an even interval, and hence it vanishes:

$$E_{100}^{(1)} = 0.$$

In fact, this shows that any nondegenerate state of the hydrogen atom has no first-order Stark effect.

Now consider the $n = 2$ levels of hydrogen. This is a four-fold degenerate state consisting of the wave functions ψ_{200} , ψ_{210} , ψ_{211} and ψ_{21-1} . Since the parity of the states is given by $(-1)^l$, we see that the $l = 0$ state has even parity while the $l = 1$ states are odd.

However, it is not hard to see that $[H', L_z] = 0$. This is either a consequence of the fact that H' is a function of $z = \cos\theta$ while $L_z = -i\hbar\partial/\partial\varphi$, or you can note that $[L_i, r_j] = i\sum_k \varepsilon_{ijk} r_k$ so that $[L_z, z] = 0$. Either way, we have

$$\begin{aligned} 0 &= \langle \psi_{n'l'm'} | [H', L_z] | \psi_{nlm} \rangle = \langle \psi_{n'l'm'} | H' L_z - L_z H' | \psi_{nlm} \rangle \\ &= \hbar(m - m') \langle \psi_{n'l'm'} | H' | \psi_{nlm} \rangle \end{aligned}$$

and hence we have the **selection rule**

$$\langle \psi_{nl'm'} | H' | \psi_{nlm} \rangle = 0 \quad \text{if } m \neq m'.$$

(This is an example of Theorem 1.2.) This shows that H' can only connect states with the same m values. And since H' has odd parity, it can only connect states with opposite parities, i.e., in the present case it can only connect an $l = 0$ state with an $l = 1$ state.

Suppressing the index $n = 2$, we *order* our basis states ψ_{lm} as $\{\psi_{00}, \psi_{10}, \psi_{11}, \psi_{1-1}\}$. (In other words, the rows and columns are labeled by these functions in this order.) Then the secular equation (2.21) becomes (also writing E instead of $E^{(1)}$ for simplicity)

$$\begin{vmatrix} -E & \langle \psi_{00} | H' | \psi_{10} \rangle & 0 & 0 \\ \langle \psi_{10} | H' | \psi_{00} \rangle & -E & 0 & 0 \\ 0 & 0 & -E & 0 \\ 0 & 0 & 0 & -E \end{vmatrix} = 0$$

or (since it's block diagonal)

$$[E^2 - (H'_{12})^2]E^2 = 0$$

where

$$H'_{12} = \langle \psi_{00} | H' | \psi_{10} \rangle = \langle \psi_{10} | H' | \psi_{00} \rangle = H'_{21}$$

because both H' and the wave functions are real. Therefore the roots of the secular equation are

$$E_n^{(1)} = \pm H'_{12}, 0, 0.$$

For our wave functions we have $\psi_{nlm} = R_{nl} Y_l^m$ or

$$\begin{aligned} \psi_{200} &= \left(\frac{1}{2a_0^3} \right)^{1/2} \left(1 - \frac{r}{2a_0} \right) e^{-r/2a_0} Y_0^0 \\ \psi_{210} &= \left(\frac{1}{24a_0^3} \right)^{1/2} \frac{r}{a_0} e^{-r/2a_0} Y_1^0 \end{aligned}$$

where a_0 is the **Bohr radius** defined by $a_0 = \hbar^2/m_e e^2$, and hence

$$\begin{aligned} H'_{12} &= \langle \psi_{200} | e^{\mathcal{E}} z | \psi_{210} \rangle \\ &= e^{\mathcal{E}} \int (2a_0)^{-3} \frac{2}{\sqrt{3}} e^{-r/a_0} \frac{r}{a_0} \left(1 - \frac{r}{2a_0} \right) z Y_0^{0*} Y_1^0 r^2 dr d\Omega. \end{aligned}$$

But

$$Y_0^{0*} = \frac{1}{\sqrt{4\pi}} \quad \text{and} \quad z = r \cos \theta = r \sqrt{\frac{4\pi}{3}} Y_1^{0*}$$

so that using

$$\int d\Omega Y_l^{m*} Y_l^{m'} = \delta_{ll'} \delta_{mm'}$$

we have

$$\begin{aligned} H'_{12} &= e\mathcal{E}(2a_0)^{-3} \frac{2}{3a_0} \int e^{-r/a_0} r^4 \left(1 - \frac{r}{2a_0}\right) Y_1^{0*} Y_1^0 dr d\Omega \\ &= e\mathcal{E}(2a_0)^{-3} \frac{2}{3a_0} \int_0^\infty \left(r^4 - \frac{r^5}{2a_0}\right) e^{-r/a_0} dr. \end{aligned}$$

Using the general result

$$\begin{aligned} \int_0^\infty r^n e^{-\alpha r} dr &= (-1)^n \frac{\partial^n}{\partial \alpha^n} \int_0^\infty e^{-\alpha r} dr \\ &= (-1)^n \frac{\partial^n}{\partial \alpha^n} \alpha^{-1} \\ &= \frac{n!}{\alpha^{n+1}} \end{aligned}$$

we finally arrive at

$$H'_{12} = -3e\mathcal{E}a_0.$$

Now we need to find the corresponding eigenvectors \mathbf{c} that will specify the correct zeroth-order wave functions. These are the solutions to the system of equations $H'\mathbf{c} = E_n^{(1)}\mathbf{c}$ for each value of $E_n^{(1)}$ (see equation (2.20b)). Let $E_1^{(1)} = H'_{12}$. Then the eigenvector $\mathbf{c}^{(1)}$ satisfies

$$\begin{bmatrix} -H'_{12} & H'_{12} & 0 & 0 \\ H'_{12} & -H'_{12} & 0 & 0 \\ 0 & 0 & -H'_{12} & 0 \\ 0 & 0 & 0 & -H'_{12} \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \\ c_4 \end{bmatrix} = 0.$$

This implies that $c_1 = c_2$ and $c_3 = c_4 = 0$. Normalizing we have $c_1 = c_2 = 1/\sqrt{2}$ so that

$$\varphi_1^{(0)} = \frac{1}{\sqrt{2}}(\psi_{200} + \psi_{210}).$$

Next we let $E_2^{(1)} = -H'_{12}$. Now the eigenvector $\mathbf{c}^{(2)}$ satisfies

$$\begin{bmatrix} H'_{12} & H'_{12} & 0 & 0 \\ H'_{12} & H'_{12} & 0 & 0 \\ 0 & 0 & H'_{12} & 0 \\ 0 & 0 & 0 & H'_{12} \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \\ c_4 \end{bmatrix} = 0$$

so that $c_1 = -c_2$ and $c_3 = c_4 = 0$. Again, normalization yields $c_1 = -c_2 = 1/\sqrt{2}$ and hence

$$\varphi_2^{(0)} = \frac{1}{\sqrt{2}}(\psi_{200} - \psi_{210}).$$

Finally, for the two degenerate roots $E_3^{(1)} = E_4^{(1)} = 0$ we have

$$\begin{bmatrix} 0 & H'_{12} & 0 & 0 \\ H'_{12} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \\ c_4 \end{bmatrix} = 0$$

so that $c_1 = c_2 = 0$ while c_3 and c_4 are completely arbitrary. Thus we can simply choose

$$\varphi_3^{(0)} = \psi_{211} \quad \text{and} \quad \varphi_4^{(0)} = \psi_{21-1}.$$

In summary, the correct zeroth-order wave functions for treating the Stark effect are $\varphi_1^{(0)}$ which gets a first-order energy shift of $-3e\mathcal{E}a_0$, the wave function $\varphi_2^{(0)}$ which gets a first-order energy shift of $+3e\mathcal{E}a_0$, and the original degenerate states $\varphi_3^{(0)} = \psi_{211}$ and $\varphi_4^{(0)} = \psi_{21-1}$ which remain degenerate to this order.

2.3 Perturbation Treatment of the First Excited States of Helium

The helium atom consists of a nucleus with two protons and two neutrons, and two orbiting electrons. If we take the nuclear charge to be $+Ze$ instead of $+2e$, then our discussion will apply equally well to helium-like ions such as H^- , Li^+ or Be^{2+} . Neglecting terms such as spin-orbit coupling, the Hamiltonian is

$$H = -\frac{\hbar^2}{2m_e}\nabla_1^2 - \frac{\hbar^2}{2m_e}\nabla_2^2 - \frac{Ze^2}{r_1} - \frac{Ze^2}{r_2} + \frac{e^2}{r_{12}} \quad (2.25)$$

where r_i is the distance to electron i , r_{12} is the distance from electron 1 to electron 2, and ∇_i^2 is the Laplacian with respect to the coordinates of electron i . The Schrödinger equation is thus a function of six variables, the three coordinates for each of the two electrons. (Technically, the electron mass m_e is the reduced mass $m = m_e M / (m_e + M)$ where M is the mass of the nucleus. But $M \gg m_e$ so that $m \approx m_e$. If this isn't familiar to you, we will treat two-body problems such as this in detail when we discuss identical particles.)

Because of the term e^2/r_{12} the Schrödinger equation isn't separable, and we must resort to approximation methods. We write

$$H = H^0 + H'$$

where

$$H^0 = H_1^0 + H_2^0 = -\frac{\hbar^2}{2m_e}\nabla_1^2 - \frac{Ze^2}{r_1} - \frac{\hbar^2}{2m_e}\nabla_2^2 - \frac{Ze^2}{r_2} \quad (2.26)$$

is the sum of two independent hydrogen atom Hamiltonians, and

$$H' = \frac{e^2}{r_{12}}. \quad (2.27)$$

We can now use separation of variables to write the unperturbed wave function $\Psi(\mathbf{r}_1, \mathbf{r}_2)$ as a product

$$\Psi(\mathbf{r}_1, \mathbf{r}_2) = \psi_1(\mathbf{r}_1)\psi_2(\mathbf{r}_2).$$

In this case we have the time-independent equation

$$H^0\Psi = (H_1^0 + H_2^0)\psi_1\psi_2 = \psi_2H_1^0\psi_1 + \psi_1H_2^0\psi_2 = E\psi_1\psi_2$$

so that dividing by $\psi_1\psi_2$ yields

$$\frac{H_1^0\psi_1}{\psi_1} = E - \frac{H_2^0\psi_2}{\psi_2}.$$

Since the left side of this equation is a function of \mathbf{r}_1 only, and the right side is a function of \mathbf{r}_2 only, each side must in fact be equal to a constant, and we can write

$$E = E_1 + E_2$$

where each E_i is the energy of a hydrogenlike wave function:

$$E_1 = -\frac{Z^2}{n_1^2} \frac{e^2}{2a_0} \quad E_2 = -\frac{Z^2}{n_2^2} \frac{e^2}{2a_0}$$

and a_0 is the Bohr radius

$$a_0 = \frac{\hbar^2}{m_e e^2} = 0.529 \text{ \AA}.$$

In other words, we have the unperturbed zeroth-order energies

$$E^{(0)} = -Z^2 \left(\frac{1}{n_1^2} + \frac{1}{n_2^2} \right) \frac{e^2}{2a_0}, \quad n_1 = 1, 2, \dots, \quad n_2 = 1, 2, \dots \quad (2.28)$$

Correspondingly, the zeroth-order wave functions are products of the usual hydrogenlike wave functions.

The lowest excited states of helium have $n_1 = 1, n_2 = 2$ or $n_1 = 2, n_2 = 1$. Then from (2.28) we have (for $Z = 2$)

$$E^{(0)} = -2^2 \left(\frac{1}{1^2} + \frac{1}{2^2} \right) \frac{e^2}{2a_0} = -5(13.606 \text{ eV}) = -68.03 \text{ eV}.$$

For $n = 2$, the possible values of l are $l = 0, 1$, and since there are $2l + 1$ values of m_l , we see that the $n = 2$ level of a hydrogenlike atom is fourfold degenerate. (This just says that the $2s$ and $2p$ states have the same energy.) Thus the first excited unperturbed state of He is eightfold degenerate, and the eight unperturbed wave functions are

$$\begin{aligned}\psi_1^{(0)} &= 1s(1)2s(2) & \psi_2^{(0)} &= 2s(1)1s(2) & \psi_3^{(0)} &= 1s(1)2p_x(2) & \psi_4^{(0)} &= 2p_x(1)1s(2) \\ \psi_5^{(0)} &= 1s(1)2p_y(2) & \psi_6^{(0)} &= 2p_y(1)1s(2) & \psi_7^{(0)} &= 1s(1)2p_z(2) & \psi_8^{(0)} &= 2p_z(1)1s(2)\end{aligned}$$

Here the notation $1s(1)2s(2)$ means, for example, that electron 1 is in the $1s$ state and electron 2 is in the $2s$ state. I have also chosen to use the real hydrogenlike wave functions $2p_x$, $2p_y$ and $2p_z$ which are defined as linear combinations of the complex wave functions $2p_0$, $2p_1$ and $2p_{-1}$:

$$\begin{aligned}2p_x &:= \frac{1}{\sqrt{2}}(2p_1 + 2p_{-1}) = \frac{1}{4\sqrt{2\pi}} \left(\frac{Z}{a_0}\right)^{5/2} r e^{-Zr/2a_0} \sin \theta \cos \phi \\ &= \frac{1}{4\sqrt{2\pi}} \left(\frac{Z}{a_0}\right)^{5/2} x e^{-Zr/2a_0}\end{aligned}\tag{2.29a}$$

$$\begin{aligned}2p_y &:= \frac{1}{i\sqrt{2}}(2p_1 - 2p_{-1}) = \frac{1}{4\sqrt{2\pi}} \left(\frac{Z}{a_0}\right)^{5/2} r e^{-Zr/2a_0} \sin \theta \sin \phi \\ &= \frac{1}{4\sqrt{2\pi}} \left(\frac{Z}{a_0}\right)^{5/2} y e^{-Zr/2a_0}\end{aligned}\tag{2.29b}$$

$$\begin{aligned}2p_z &:= 2p_0 = \frac{1}{4\sqrt{2\pi}} \left(\frac{Z}{a_0}\right)^{5/2} r e^{-Zr/2a_0} \cos \theta \\ &= \frac{1}{4\sqrt{2\pi}} \left(\frac{Z}{a_0}\right)^{5/2} z e^{-Zr/2a_0}\end{aligned}\tag{2.29c}$$

This is perfectly valid since any linear combination of solutions with a given energy is also a solution with that energy. (However, the $2p_x$ and $2p_y$ functions are *not* eigenfunctions of L_z since they are linear combinations of eigenfunctions with different values of m_l .) These real hydrogenlike wave functions are more convenient for many purposes in constructing chemical bonds and molecular wave functions. In fact, you have probably seen these wave functions in more elementary chemistry courses. For example, a contour plot in the plane (i.e., a cross section) of a real $2p$ wave function is shown in Figure 4 below. (Let $\phi = \pi/2$ in any of equations (2.29).) The three-dimensional orbital is obtained by rotating this plot about the horizontal axis, so we see that the actual shape of a real $2p$ orbital (i.e., a one-electron wave function) is two separated, distorted ellipsoids.

It is not hard to verify that the real $2p$ wave functions are orthonormal, and hence the eight degenerate wave functions $\psi_i^{(0)}$ are also orthonormal as required by equation (2.19). The secular determinant contains $8^2 = 64$ elements. However, H'

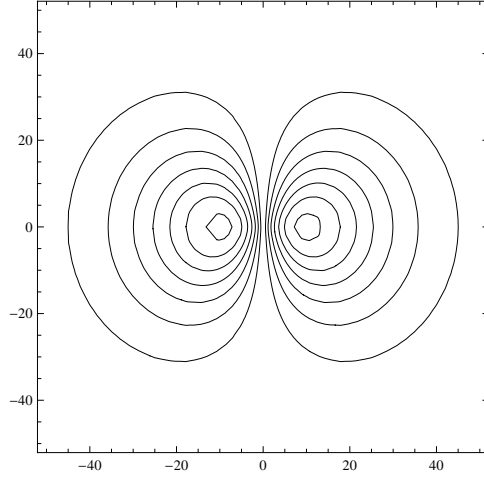


Figure 4: Contour plot in the plane of a real $2p$ wave function.

is real, as are the $\psi_i^{(0)}$, so that $H'_{ij} = H'_{ji}$ and the determinant is symmetric about the main diagonal. This cuts the number of integrals almost in half.

Even better, by using parity we can easily show that most of the H'_{ij} are zero. Indeed, the perturbing Hamiltonian $H' = e^2/r_{12}$ is an even function of \mathbf{r} since

$$r_{12} = [(x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2]^{1/2}$$

and this is unchanged if $\mathbf{r}_1 \rightarrow -\mathbf{r}_1$ and $\mathbf{r}_2 \rightarrow -\mathbf{r}_2$. Also, the hydrogenlike s -wave functions depend only on $r = |\mathbf{r}|$ and hence are invariant under $\mathbf{r} \rightarrow -\mathbf{r}$. Furthermore, you can see from the above forms that the $2p$ wave functions are *odd* under parity since they depend on r and either x , y or z . Hence, since we are integrating over all space, any integral with only a single factor of $2p$ must vanish:

$$H'_{13} = H'_{14} = H'_{15} = H'_{16} = H'_{17} = H'_{18} = 0$$

and

$$H'_{23} = H'_{24} = H'_{25} = H'_{26} = H'_{27} = H'_{28} = 0.$$

Now consider an integral such as

$$H'_{35} = \int_{-\infty}^{\infty} 1s(1)2p_x(2) \frac{e^2}{r_{12}} 1s(1)2p_y(2) d\mathbf{r}_1 d\mathbf{r}_2.$$

If we let $x_1 \rightarrow -x_1$ and $x_2 \rightarrow -x_2$, then r_{12} is unchanged as are $1s(1)$ and $2p_y(2)$. However, $2p_x(2)$ changes sign, and the net result is that the integrand is an odd function under this transformation. Hence it is not hard to see that the integral vanishes. This lets us conclude that

$$H'_{35} = H'_{36} = H'_{37} = H'_{38} = 0$$

and

$$H'_{45} = H'_{46} = H'_{47} = H'_{48} = 0.$$

Similarly, by considering the transformation $y_1 \rightarrow -y_1$ and $y_2 \rightarrow -y_2$, it follows that

$$H'_{57} = H'_{58} = H'_{67} = H'_{68} = 0.$$

With these simplifications, the secular equation becomes

$$\begin{vmatrix} b_{11} & H'_{12} & 0 & 0 & 0 & 0 & 0 & 0 \\ H'_{12} & b_{22} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & b_{33} & H'_{34} & 0 & 0 & 0 & 0 \\ 0 & 0 & H'_{34} & b_{44} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & b_{55} & H'_{56} & 0 & 0 \\ 0 & 0 & 0 & 0 & H'_{56} & b_{66} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & b_{77} & H'_{78} \\ 0 & 0 & 0 & 0 & 0 & 0 & H'_{78} & b_{88} \end{vmatrix} = 0 \quad (2.30)$$

where

$$b_{ii} = H'_{ii} - E^{(1)}, \quad i = 1, 2, \dots, 8.$$

Since the secular determinant is in block-diagonal form with 2×2 blocks on the diagonal, the same logic that we used in Example 1.3 would seem to tell us that the correct zeroth-order wave functions have the form

$$\begin{aligned} \phi_1^{(0)} &= c_1 \psi_1^{(0)} + c_2 \psi_2^{(0)} & \phi_2^{(0)} &= \bar{c}_1 \psi_1^{(0)} + \bar{c}_2 \psi_2^{(0)} \\ \phi_3^{(0)} &= c_3 \psi_3^{(0)} + c_4 \psi_4^{(0)} & \phi_4^{(0)} &= \bar{c}_3 \psi_3^{(0)} + \bar{c}_4 \psi_4^{(0)} \\ \phi_5^{(0)} &= c_5 \psi_5^{(0)} + c_6 \psi_6^{(0)} & \phi_6^{(0)} &= \bar{c}_5 \psi_5^{(0)} + \bar{c}_6 \psi_6^{(0)} \\ \phi_7^{(0)} &= c_7 \psi_7^{(0)} + c_8 \psi_8^{(0)} & \phi_8^{(0)} &= \bar{c}_7 \psi_7^{(0)} + \bar{c}_8 \psi_8^{(0)} \end{aligned}$$

where the barred and unbarred coefficients distinguish between the two roots of each second-order determinant. However, while that argument applies to the upper 2×2 determinant (i.e., the first two equations of the system), it doesn't apply to the whole determinant in this case. This is because it turns out (as we will see below) that the lower three 2×2 determinants are identical. Therefore, their pairs of roots are the same, and all we can say is that there are two six-dimensional eigenspaces. In other words, all we can say is that for each of the two roots and for each $n = 3, 4, \dots, 8$, the function $\phi_n^{(0)}$ will be a linear combination of $\psi_3^{(0)}, \dots, \psi_8^{(0)}$. However, we can choose any basis we wish for this six-dimensional space, so we choose the three two-dimensional orthonormal $\phi_n^{(0)}$'s as shown above.

The first determinant is

$$\begin{vmatrix} H'_{11} - E^{(1)} & H'_{12} \\ H'_{12} & H'_{22} - E^{(1)} \end{vmatrix} = 0 \quad (2.31)$$

where

$$H'_{11} = \int 1s(1)2s(2) \frac{e^2}{r_{12}} 1s(1)2s(2) d\mathbf{r}_1 d\mathbf{r}_2 = \int [1s(1)]^2 [2s(2)]^2 \frac{e^2}{r_{12}} d\mathbf{r}_1 d\mathbf{r}_2$$

$$H'_{22} = \int [1s(2)]^2 [2s(1)]^2 \frac{e^2}{r_{12}} d\mathbf{r}_1 d\mathbf{r}_2 .$$

Since the integration variables are just dummy variables, it is pretty obvious that letting $\mathbf{r}_1 \leftrightarrow \mathbf{r}_2$ shows that

$$H'_{11} = H'_{22} .$$

Similarly, it is easy to see that

$$H'_{33} = H'_{44} \quad H'_{55} = H'_{66} \quad H'_{77} = H'_{88} .$$

The integral H'_{11} is sometimes denoted by J_{1s2s} and called a **Coulomb integral**:

$$H'_{11} = J_{1s2s} = \int [1s(1)]^2 [2s(2)]^2 \frac{e^2}{r_{12}} d\mathbf{r}_1 d\mathbf{r}_2 .$$

The reason for the name is that this represents the electrostatic energy of repulsion between an electron with the probability density function $[1s]^2$ and an electron with probability density function $[2s]^2$. The integral H'_{12} is denoted by K_{1s2s} and called an **exchange integral**:

$$H'_{12} = K_{1s2s} = \int 1s(1)2s(2) \frac{e^2}{r_{12}} 2s(1)1s(2) d\mathbf{r}_1 d\mathbf{r}_2 .$$

Here the functions to the left and right of H' differ from each other by the exchange of electrons 1 and 2. The general definitions of the Coulomb and exchange integrals are

$$J_{ij} = \langle f_i(1)f_j(2) | e^2/r_{12} | f_i(1)f_j(2) \rangle$$

$$K_{ij} = \langle f_i(1)f_j(2) | e^2/r_{12} | f_j(1)f_i(2) \rangle$$

where the range of integration is over the full range of spatial coordinates of particles 1 and 2, and the functions f_i, f_j are spatial orbitals.

Substituting these integrals into (2.31) we have

$$\begin{vmatrix} J_{1s2s} - E^{(1)} & K_{1s2s} \\ K_{1s2s} & J_{1s2s} - E^{(1)} \end{vmatrix} = 0 \quad (2.32)$$

or

$$J_{1s2s} - E^{(1)} = \pm K_{1s2s}$$

and hence the two roots are

$$E_1^{(1)} = J_{1s2s} - K_{1s2s} \quad \text{and} \quad E_2^{(1)} = J_{1s2s} + K_{1s2s} .$$

Just as in Example 1.3, we substitute $E_1^{(1)}$ back into (2.20a) to write

$$K_{1s2s}c_1 + K_{1s2s}c_2 = 0$$

$$K_{1s2s}c_1 + K_{1s2s}c_2 = 0$$

and hence $c_2 = -c_1$. Normalizing $\phi_1^{(0)}$ we have (using the orthonormality of the $\psi_i^{(0)}$)

$$\langle \phi_1^{(0)} | \phi_1^{(0)} \rangle = \langle c_1\psi_1^{(0)} - c_1\psi_2^{(0)} | c_1\psi_1^{(0)} - c_1\psi_2^{(0)} \rangle = |c_1|^2 + |c_2|^2 = 1$$

so that $c_1 = 1/\sqrt{2}$. Thus the zeroth-order wave function corresponding to $E_1^{(1)}$ is

$$\phi_1^{(0)} = 2^{-1/2}[\psi_1^{(0)} - \psi_2^{(0)}] = 2^{-1/2}[1s(1)2s(2) - 2s(1)1s(2)].$$

Similarly, the wave function corresponding to $E_2^{(1)}$ is easily found to be

$$\phi_2^{(0)} = 2^{-1/2}[\psi_1^{(0)} + \psi_2^{(0)}] = 2^{-1/2}[1s(1)2s(2) + 2s(1)1s(2)].$$

This takes care of the first determinant in (2.30), but we still have the remaining three to handle.

First look at the integrals H'_{33} and H'_{55} :

$$H'_{33} = \int 1s(1)2p_x(2) \frac{e^2}{r_{12}} 1s(1)2p_x(2) d\mathbf{r}_1 d\mathbf{r}_2$$

$$H'_{55} = \int 1s(1)2p_y(2) \frac{e^2}{r_{12}} 1s(1)2p_y(2) d\mathbf{r}_1 d\mathbf{r}_2.$$

The only difference between these is the $2p(2)$ orbital, and the only difference between the $2p_x$ and $2p_y$ orbitals is their spatial orientation. Since the $1s$ orbitals are spherically symmetric, it should be clear that these integrals are the same. Formally, in H'_{33} we can change variables by letting $x_1 \rightarrow y_1$, $y_1 \rightarrow x_1$, $x_2 \rightarrow y_2$ and $y_2 \rightarrow x_2$. This leaves r_{12} unchanged, and transforms H'_{33} into H'_{55} . The same argument shows that $H'_{77} = H'_{33}$ also. Hence we have

$$H'_{33} = H'_{55} = H'_{77} = \int 1s(1)2p_z(2) \frac{e^2}{r_{12}} 1s(1)2p_z(2) d\mathbf{r}_1 d\mathbf{r}_2 := J_{1s2p}.$$

A similar argument shows that we also have equal exchange integrals:

$$H'_{34} = H'_{56} = H'_{78} = \int 1s(1)2p_z \frac{e^2}{r_{12}} 2p_z(1)1s(2) d\mathbf{r}_1 d\mathbf{r}_2 := K_{1s2p}.$$

Thus the remaining three determinants in (2.30) are the same and have the form

$$\begin{vmatrix} J_{1s2p} - E^{(1)} & K_{1s2p} \\ K_{1s2p} & J_{1s2p} - E^{(1)} \end{vmatrix} = 0.$$

But this is the same as (2.32) if we replace $2s$ by $2p$, and hence we can immediately write down the solutions:

$$\begin{aligned} E_3^{(1)} &= E_5^{(1)} = E_7^{(1)} = J_{1s2p} - K_{1s2p} \\ E_4^{(1)} &= E_6^{(1)} = E_8^{(1)} = J_{1s2p} + K_{1s2p} \end{aligned}$$

and

$$\begin{aligned} \phi_3^{(0)} &= 2^{-1/2}[1s(1)2p_x(2) - 1s(2)2p_x(1)] \\ \phi_4^{(0)} &= 2^{-1/2}[1s(1)2p_x(2) + 1s(2)2p_x(1)] \\ \phi_5^{(0)} &= 2^{-1/2}[1s(1)2p_y(2) - 1s(2)2p_y(1)] \\ \phi_6^{(0)} &= 2^{-1/2}[1s(1)2p_y(2) + 1s(2)2p_y(1)] \\ \phi_7^{(0)} &= 2^{-1/2}[1s(1)2p_z(2) - 1s(2)2p_z(1)] \\ \phi_8^{(0)} &= 2^{-1/2}[1s(1)2p_z(2) + 1s(2)2p_z(1)] \end{aligned}$$

So what has happened? Starting from the eight degenerate (unperturbed) states $\psi_i^{(0)}$ that would exist in the absence of electron-electron repulsion, we find that including this repulsion term splits the degenerate states into two nondegenerate levels associated with the configuration $1s2s$, and two triply degenerate levels associated with the configuration $1s2p$. Interestingly, going to higher-order energy corrections will not completely remove the degeneracy, and in fact it takes the application of an external magnetic field to do so.

In order to evaluate the Coulomb and exchange integrals in the expressions for $E^{(1)}$ we need to use the expansion

$$\frac{1}{r_{12}} = \sum_{l=0}^{\infty} \sum_{m=-l}^l \frac{4\pi}{2l+1} \frac{r_{<}^l}{r_{>}^{l+1}} [Y_l^m(\theta_1, \varphi_1)]^* Y_l^m(\theta_2, \varphi_2) \quad (2.33)$$

where $r_{<}$ means the smaller of r_1 and r_2 and $r_{>}$ is the larger of these. The details of this type of integral are left to the homework, and the results are

$$\begin{aligned} J_{1s2s} &= \frac{17}{81} \frac{Ze^2}{a_0} = 11.42 \text{ eV} & J_{1s2p} &= \frac{59}{243} \frac{Ze^2}{a_0} = 13.21 \text{ eV} \\ K_{1s2s} &= \frac{16}{729} \frac{Ze^2}{a_0} = 1.19 \text{ eV} & K_{1s2p} &= \frac{112}{6561} \frac{Ze^2}{a_0} = 0.93 \text{ eV} \end{aligned}$$

where we used $Z = 2$ and $e^2/2a_0 = 13.606 \text{ eV}$. Recalling that $E^{(0)} = -68.03 \text{ eV}$ we obtain

$$\begin{aligned} E^{(0)} + E_1^{(1)} &= E^{(0)} + J_{1s2s} - K_{1s2s} = -57.8 \text{ eV} \\ E^{(0)} + E_2^{(1)} &= E^{(0)} + J_{1s2s} + K_{1s2s} = -55.4 \text{ eV} \\ E^{(0)} + E_3^{(1)} &= E^{(0)} + J_{1s2p} - K_{1s2p} = -55.7 \text{ eV} \\ E^{(0)} + E_4^{(1)} &= E^{(0)} + J_{1s2p} + K_{1s2p} = -53.9 \text{ eV} . \end{aligned}$$

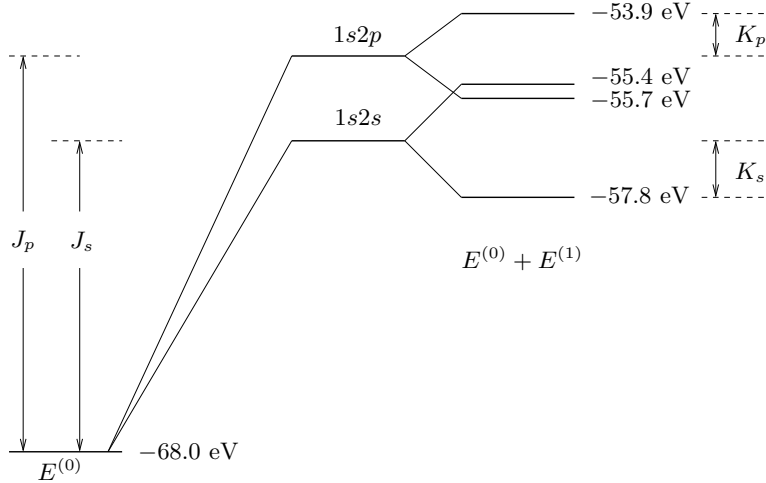


Figure 5: The first excited levels of the helium atom.

(See Figure 5 below.) The first-order energy corrections place the lower $1s2p$ level below the upper $1s2s$ level, which disagrees with the actual helium spectrum. This is due to the neglect of higher-order corrections. Since the electron-electron repulsion is not a small quantity, this is not surprising.

Finally, let us look at the sources of the degeneracy of the original eight zeroth-order wave functions and the reason for the partial lifting of this degeneracy. There are three types of degeneracy to consider: (1) The degeneracy between states with the same n but different values of l . The $2s$ and $2p$ functions have the same energy. (2) The degeneracy between wave functions with the same n and l but different values of m_l . The $2p_x$, $2p_y$ and $2p_z$ functions have the same energy. (This could just as well have been the $2p_0$, $2p_1$ and $2p_{-1}$ complex functions.) (3) There is an **exchange degeneracy** between functions that differ only in the exchange of electrons between the orbitals. For example, $\psi_1^{(0)} = 1s(1)2s(2)$ and $\psi_2^{(0)} = 1s(2)2s(1)$ have the same energy.

By introducing the electron-electron perturbation $H' = e^2/r_{12}$ we removed the degeneracy associated with l and the exchange degeneracy, but not the degeneracy due to m_l . To understand the reason for the lifting of the l degeneracy, realize that a $2s$ electron has a greater probability than a $2p$ electron of being closer to the nucleus than a $1s$ electron, and hence a $2s$ electron is not as effectively shielded from the nucleus by the $1s$ electrons as a $2p$ electron is. Since the energy levels are given by

$$E = -\frac{Z^2}{n^2} \frac{e^2}{2a_0}$$

we see that a larger nuclear charge means a lower energy, and hence the $2s$ electron has a lower energy than the $2p$ electron. This is also evident from the Coulomb integrals, where we see that J_{1s2s} is less than J_{1s2p} . These integrals represent the

electrostatic repulsion of their respective charge distributions: when the $2s$ electron penetrates the $1s$ charge distribution it only feels a repulsion due to the unpenetrated portion of the $1s$ distribution. Therefore the $1s$ - $2s$ electrostatic repulsion is less than the $1s$ - $2p$ repulsion, and the $1s2s$ levels lies below the $1s2p$ levels. So we see that the interelectronic repulsion in many-electron atoms lifts the l degeneracy, and the orbital energies for the same value of n increase with increasing l .

To understand the removal of the exchange degeneracy, note that the original zeroth-order wave functions specified which electron went into which orbital. Since the secular determinant wasn't diagonal, these couldn't have been the correct zeroth-order wave functions. In fact, the correct zeroth-order wave functions do not assign a specific electron to a specific orbital, as is evident from the form of each $\phi_i^{(0)}$. This is a consequence of the indistinguishability of identical particles, and will be discussed at length a little later in this course. Since, for example, $\phi_1^{(0)}$ and $\phi_2^{(0)}$ have different energies, the exchange degeneracy is removed by using the correct zeroth-order wave functions.

2.4 Spin–Orbit Coupling and the Hydrogen Atom Fine Structure

The Hamiltonian

$$H^0 = -\frac{\hbar^2}{2m} \left(\frac{\partial^2}{\partial r^2} + \frac{2}{r} \frac{\partial}{\partial r} \right) + \frac{1}{2mr^2} \mathbf{L}^2 - \frac{e^2}{r} \quad (2.34)$$

used to derive the hydrogen atom wave functions ψ_{nlm} that we have worked with so far consists of the kinetic energy of the electron plus the potential energy of the Coulomb force binding the electron and proton together. (Recall that in this equation, m is really the reduced mass $m = m_e M_p / (m_e + M_p) \approx m_e$.) While this works very well, the actual Hamiltonian is somewhat more complicated than this. In this section we derive an additional term in the Hamiltonian that is due to a coupling between the orbital angular momentum \mathbf{L} and the spin angular momentum \mathbf{S} .

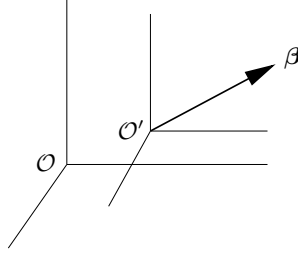
The discussion that follows is a somewhat heuristic approach to deriving an interaction term that agrees with experiment. You shouldn't take the physical picture too seriously. However, the basic idea is simple enough. From the point of view of the electron, the moving nucleus (i.e., a proton) generates a current that is the source of a magnetic field \mathbf{B} . This current is proportional to the electron's angular momentum \mathbf{L} . The interaction energy of a magnetic moment $\boldsymbol{\mu}$ with this magnetic field is $-\boldsymbol{\mu} \cdot \mathbf{B}$. Since the magnetic moment of an electron is proportional to its spin \mathbf{S} , we see that the interaction energy will be proportional to $\mathbf{L} \cdot \mathbf{S}$.

With the above disclaimer, the interaction term we are looking for is due to the fact that from the point of view of the electron, the moving hydrogen nucleus (the proton) forms a current, and thus generates a magnetic field. From special relativity, we know that the electric and magnetic fields are related by a Lorentz transformation so that

$$\mathbf{B}_\perp = \gamma(\mathbf{B}'_\perp + \boldsymbol{\beta} \times \mathbf{E}') \quad \mathbf{B}_\parallel = \mathbf{B}'_\parallel$$

$$\mathbf{E}_{\perp} = \gamma(\mathbf{E}'_{\perp} - \boldsymbol{\beta} \times \mathbf{B}') \qquad \mathbf{E}_{\parallel} = \mathbf{E}'_{\parallel}$$

where $\boldsymbol{\beta} = \mathbf{v}/c$ is the velocity of the primed frame with respect to the unprimed frame, $\gamma = (1 - \beta^2)^{-1/2}$ and \perp, \parallel refer to the components perpendicular or parallel to $\boldsymbol{\beta}$.



We let the primed frame be the proton rest frame, and note that there is no \mathbf{B}' field in the proton's frame due to the proton itself. Also, if $\beta \ll 1$, then $\gamma \approx 1$ and we then have

$$\mathbf{B} = \boldsymbol{\beta} \times \mathbf{E}' \qquad \text{and} \qquad \mathbf{E} = \mathbf{E}' .$$

If \mathbf{v} is the electron's velocity with respect to the lab (or the proton), then $\boldsymbol{\beta} = -\mathbf{v}/c$ so the field felt by the electron is

$$\mathbf{B} = -\frac{\mathbf{v}}{c} \times \mathbf{E}' . \qquad (2.35)$$

The electric field \mathbf{E}' due to the proton is

$$\mathbf{E}' = \frac{e}{r^2} \hat{\mathbf{r}} = \frac{e}{r^3} \mathbf{r} \qquad (2.36)$$

where $e > 0$ and \mathbf{r} is the position vector from the proton to the electron.

From basic electrodynamics, we know that the energy of a particle with magnetic moment $\boldsymbol{\mu}$ in a magnetic field \mathbf{B} is given by (see the end of this section)

$$W = -\boldsymbol{\mu} \cdot \mathbf{B} \qquad (2.37)$$

so we need to know $\boldsymbol{\mu}$. Consider a particle of charge q moving in a circular orbit. It forms an effective current

$$I = \frac{\Delta q}{\Delta t} = \frac{q}{2\pi r/v} = \frac{qv}{2\pi r} .$$

By definition, the magnetic moment has magnitude

$$\mu = \frac{I}{c} \times \text{area} = \frac{qv}{2\pi r c} \cdot \pi r^2 = \frac{qvr}{2c} .$$

But the angular momentum of the particle is $L = mvr$ so we conclude that the magnetic moment due to orbital motion is

$$\boldsymbol{\mu}_l = \frac{q}{2mc} \mathbf{L} . \qquad (2.38)$$

The ratio of $\boldsymbol{\mu}$ to \mathbf{L} is called the **gyromagnetic ratio**.

While the above derivation of (2.38) was purely classical, we know that the electron also possesses an intrinsic spin angular momentum. Let us hypothesize that the electron magnetic moment associated with this spin is of the form

$$\boldsymbol{\mu}_s = g \frac{-e}{2mc} \mathbf{S}.$$

The constant g is found by experiment to be very close to 2. (However, the relativistic Dirac equation predicts that g is exactly 2. Higher order corrections in quantum electrodynamics predict a slightly different value, and the measurement of $g - 2$ is one of the most accurate experimental result in all of physics.)

So we now have the electron magnetic moment given by

$$\boldsymbol{\mu}_s = \frac{-e}{mc} \mathbf{S} \quad (2.39)$$

and hence the interaction energy of the electron with the magnetic field of the proton is (using equations (2.35) and (2.36))

$$W = -\boldsymbol{\mu}_s \cdot \mathbf{B} = +\frac{e}{mc} \mathbf{S} \cdot \mathbf{B} = -\frac{e}{mc} \mathbf{S} \cdot \left(\frac{e}{r^3 c} \mathbf{v} \times \mathbf{r} \right) = \frac{e^2}{m^2 c^2 r^3} \mathbf{S} \cdot (\mathbf{r} \times \mathbf{p})$$

or

$$W = \frac{e^2}{m^2 c^2 r^3} \mathbf{S} \cdot \mathbf{L}. \quad (2.40)$$

Alternatively, we can write W in another form as follows. If we assume that the electron moves in a spherically symmetric potential field, then the force $-e\mathbf{E}$ on the electron may be written as the negative gradient of this potential energy:

$$-e\mathbf{E} = -\nabla V(r) = -\frac{dV}{dr} \hat{\mathbf{r}} = -\frac{\mathbf{r}}{r} \frac{dV}{dr}.$$

Using this in (2.35) we have

$$\mathbf{B} = -\frac{\mathbf{v}}{c} \times \mathbf{r} \frac{1}{er} \frac{dV}{dr} = \frac{1}{mc} \mathbf{r} \times \mathbf{p} \frac{1}{er} \frac{dV}{dr}$$

and hence

$$W = -\boldsymbol{\mu}_s \cdot \mathbf{B} = \frac{e}{m^2 c^2} \mathbf{S} \cdot (\mathbf{r} \times \mathbf{p}) \frac{1}{er} \frac{dV}{dr}$$

or

$$W = \frac{1}{m^2 c^2} \frac{1}{r} \frac{dV}{dr} \mathbf{S} \cdot \mathbf{L}. \quad (2.41)$$

However, we have made one major mistake. The classical equation that leads to (2.37) is

$$\frac{d\mathbf{L}}{dt} = \mathbf{N} = \boldsymbol{\mu} \times \mathbf{B} \quad (2.42)$$

where \mathbf{L} is the angular momentum of the particle in its rest frame, \mathbf{N} is the applied torque, and \mathbf{B} is the magnetic field in that frame. But this only applies if the

electron's rest frame isn't rotating. If it is, then the left side of this equation isn't valid (i.e., it isn't equal to only the applied torque), and we must use the correct (operator) expression from classical mechanics:

$$\left(\frac{d}{dt}\right)_{\text{lab}} = \left(\frac{d}{dt}\right)_{\text{rot}} + \boldsymbol{\omega} \times . \quad (2.43)$$

(If you don't know this result, I will derive it at the end of this section so you can see what is going on and why.)

For the electron, (2.42) gives $d\mathbf{S}/dt$ in the lab frame, so in the electron's frame we must use

$$\left(\frac{d\mathbf{S}}{dt}\right)_{\text{rot}} = \left(\frac{d\mathbf{S}}{dt}\right)_{\text{lab}} - \boldsymbol{\omega}_T \times \mathbf{S} \quad (2.44)$$

where $\boldsymbol{\omega}_T$ is called the **Thomas precessional frequency**. Thus we see that the change in the spin angular momentum of the electron, $(d\mathbf{S}/dt)_{\text{rot}}$, is given by the change due to the applied torque $\boldsymbol{\mu} \times \mathbf{B}$ minus an effect due to the rotation of the coordinate system:

$$\left(\frac{d\mathbf{S}}{dt}\right)_{\text{rot}} = \boldsymbol{\mu} \times \mathbf{B} - \boldsymbol{\omega}_T \times \mathbf{S} = -\frac{e}{mc}\mathbf{S} \times \mathbf{B} + \mathbf{S} \times \boldsymbol{\omega}_T$$

or

$$\left(\frac{d\mathbf{S}}{dt}\right)_{\text{rot}} = \mathbf{S} \times \left(-\frac{e\mathbf{B}}{mc} + \boldsymbol{\omega}_T\right). \quad (2.45)$$

This is the analogue of (2.42), so the analogue of (2.37) is

$$W = -\mathbf{S} \cdot \left(-\frac{e\mathbf{B}}{mc} + \boldsymbol{\omega}_T\right) = \frac{e}{mc}\mathbf{S} \cdot \mathbf{B} - \mathbf{S} \cdot \boldsymbol{\omega}_T. \quad (2.46)$$

Note that the first term is what we already calculated in equation (2.40). What we need to know is the Thomas factor $\mathbf{S} \cdot \boldsymbol{\omega}_T$. This is not a particularly easy calculation to do exactly, so we will give a very simplified derivation. (See Jackson, *Classical Electrodynamics*, Chapter 11 if you want a careful derivation.)

Basically, Thomas precession can be attributed to time dilation, i.e., observers on the electron and proton disagree on the time required for one particle to make a revolution about the other. Let T be the time required for a revolution according to the electron, and let it be T' according to the proton. Then $T' = \gamma T$ where $\gamma = (1 - \beta^2)^{-1/2}$. (Note that a circular orbit means an acceleration, so even this isn't really correct.) Then the electron and proton each measure orbital angular velocities of $2\pi/T$ and $2\pi/T'$ respectively.

To the electron, its spin \mathbf{S} maintains its direction in space, but to the proton, it appears to precess at a rate equal to the difference in angular velocities, or

$$\begin{aligned} \omega_T &= \frac{2\pi}{T} - \frac{2\pi}{T'} = 2\pi \left(\frac{1}{T} - \frac{1}{T'}\right) = 2\pi \left(\frac{\gamma}{T'} - \frac{1}{T'}\right) \\ &= \frac{2\pi}{T'} \left[(1 - \beta^2)^{-1/2} - 1\right] \approx \frac{2\pi}{T'} \frac{\beta^2}{2}. \end{aligned}$$

But in general we know that $\omega = v/r$ and hence

$$\frac{2\pi}{T'} = \frac{v}{r} = \frac{mvr}{mr^2} = \frac{L}{mr^2}$$

and therefore

$$\omega_T = \frac{L}{mr^2} \frac{\beta^2}{2} = \frac{L}{mr^2} \frac{v^2}{2c^2} = \frac{1}{2} \frac{L}{m^2 c^2} \frac{1}{r} \frac{mv^2}{r}.$$

We also know that $F = ma$, where for circular motion we have an inward directed acceleration $a = v^2/r$. Since $\mathbf{F} = -\nabla V$, we have

$$\mathbf{F} = -\frac{mv^2}{r} \hat{\mathbf{r}} = -\frac{dV}{dr} \hat{\mathbf{r}}$$

and we can write

$$\boldsymbol{\omega}_T = \frac{1}{2} \frac{1}{m^2 c^2} \frac{1}{r} \frac{dV}{dr} \mathbf{L}. \quad (2.47)$$

From this we see that $\mathbf{S} \cdot \boldsymbol{\omega}_T$ is just one-half the energy given by equation (2.41), and equation (2.46) shows that it is subtracted off. Therefore the correct spin-orbit energy is given by

$$W = \frac{1}{2m^2 c^2} \frac{1}{r} \frac{dV}{dr} \mathbf{L} \cdot \mathbf{S} \quad (2.48a)$$

or, from (2.40) with a slight change of notation,

$$H_{so} = \frac{e^2}{2m^2 c^2 r^3} \mathbf{L} \cdot \mathbf{S}. \quad (2.48b)$$

Calculating the spin-orbit interaction energy E_{so} by finding the eigenfunctions and eigenvalues of the Hamiltonian $H = H^0 + H_{so}$ is a difficult problem. Since the effect of H_{so} is small compared to H^0 (at least for the lighter atoms), we will estimate the value of E_{so} by using first-order perturbation theory. Then first-order energy shifts for the hydrogen atom will be the integrals

$$E_{so}^{(1)} \approx \langle \Psi | H_{so} | \Psi \rangle$$

where the hydrogen atom wave functions including spin are of the form

$$\Psi = R_{nl}(r) Y_l^m(\theta, \varphi) \chi(s).$$

From $\mathbf{J} = \mathbf{L} + \mathbf{S}$, we have $J^2 = L^2 + S^2 + 2\mathbf{L} \cdot \mathbf{S}$ so that

$$\mathbf{L} \cdot \mathbf{S} = \frac{1}{2}(J^2 - L^2 - S^2). \quad (2.49)$$

Note that neither \mathbf{L} nor \mathbf{S} separately commutes with $\mathbf{L} \cdot \mathbf{S}$, but you can easily show that $\mathbf{J} = \mathbf{L} + \mathbf{S}$ does in fact commute with $\mathbf{L} \cdot \mathbf{S}$. Because of this, we can choose our states to be simultaneous eigenfunctions of J^2 , J_z , L^2 and S^2 , all of which commute with H .

Since Y_l^m is an eigenfunction of L_z and χ is an eigenfunction of S_z , the wave function $Y_l^m\chi$ is an eigenfunction of $J_z = L_z + S_z$ but *not* of J^2 . However, by the usual addition of angular momentum problem, in this case \mathbf{L} and \mathbf{S} , we can construct simultaneous eigenfunctions ψ of J^2 , J_z , L^2 and S^2 . In this case we have $s = 1/2$, so we know that the resulting possible j values are $j = l - 1/2, l + 1/2$. The reason we want to do this is because there are $2(2l + 1)$ degenerate levels for a given n and l , where the additional factor of 2 comes from the two possible spin orientations.

Let us assume that we have constructed these eigenfunctions, and we now denote the hydrogen atom wave functions by

$$\Psi = R_{nl}(r)\psi(\theta, \varphi, s)$$

where, by (2.49)

$$\begin{aligned} \mathbf{L} \cdot \mathbf{S}\psi &= \frac{\hbar^2}{2}[j(j+1) - l(l+1) - s(s+1)]\psi \\ &= \frac{\hbar^2}{2}\left[j(j+1) - l(l+1) - \frac{3}{4}\right]\psi. \end{aligned}$$

Using this, our first-order energy estimate becomes

$$\begin{aligned} E_{\text{so}}^{(1)} &\approx \left\langle R_{nl}\psi \left| \frac{e^2}{2m^2c^2} \frac{1}{r^3} \mathbf{L} \cdot \mathbf{S} \right| R_{nl}\psi \right\rangle \\ &= \frac{e^2\hbar^2}{4m^2c^2} \left[j(j+1) - l(l+1) - \frac{3}{4} \right] \left\langle R_{nl} \left| \frac{1}{r^3} \right| R_{nl} \right\rangle \end{aligned} \quad (2.50)$$

where

$$\left\langle R_{nl}\psi \left| \frac{1}{r^3} \right| R_{nl}\psi \right\rangle = \left\langle R_{nl} \left| \frac{1}{r^3} \right| R_{nl} \right\rangle$$

because $\langle \psi | \psi \rangle = 1$. The integral in (2.50) is not at all hard to do if you use some clever tricks. I will show how to do it at the end of this section, and the answer is

$$\left\langle R_{nl} \left| \frac{1}{r^3} \right| R_{nl} \right\rangle = \frac{1}{a_0^3 n^3 l(l+1/2)(l+1)} \quad (2.51)$$

where the Bohr radius is

$$a_0 = \frac{\hbar^2}{me^2} = \frac{\hbar}{m\alpha} \quad (2.52)$$

and the **fine structure constant** is

$$\alpha = \frac{e^2}{\hbar c} \approx \frac{1}{137}. \quad (2.53)$$

Note that for $l = 0$ we also have $\mathbf{L} \cdot \mathbf{S} = 0$ anyway, so there is no spin-orbit energy.

Recall that the energy corresponding to H^0 is

$$E_n^{(0)} = -\frac{me^4}{2\hbar^2 n^2} = -\frac{mc^2 \alpha^2}{2n^2}. \quad (2.54a)$$

or

$$E_n^{(0)} = \frac{E_1^{(0)}}{n^2} = \frac{-13.6 \text{ eV}}{n^2}. \quad (2.54b)$$

Combining (2.50) and (2.51) we have

$$\begin{aligned} E_{\text{so}}^{(1)} &= \frac{e^2 \hbar^2}{4m^2 c^2 a_0^3 n^3} \left\{ \frac{[j(j+1) - l(l+1) - 3/4]}{l(l+1/2)(l+1)} \right\} \\ &= \frac{|E_n^{(0)}| \alpha^2}{2n} \left\{ \frac{[j(j+1) - l(l+1) - 3/4]}{l(l+1/2)(l+1)} \right\}. \end{aligned} \quad (2.55)$$

Since $j = l \pm 1/2$, this gives us the two corrections to the energy

$$E_{\text{so}}^{(1)} = \frac{|E_n^{(0)}| \alpha^2}{n} \left[\frac{1}{(2l+1)(l+1)} \right] \quad \text{for } j = l + 1/2 \text{ and } l \neq 0 \quad (2.56a)$$

$$E_{\text{so}}^{(1)} = -\frac{|E_n^{(0)}| \alpha^2}{n} \left[\frac{1}{l(2l+1)} \right] \quad \text{for } j = l - 1/2 \text{ and } l \neq 0. \quad (2.56b)$$

There is yet another correction to the hydrogen atom energy levels due to the relativistic contribution to the kinetic energy of the electron. The kinetic energy is really the difference between the total relativistic energy $E = (p^2 c^2 + m^2 c^4)^{1/2}$ and the rest energy mc^2 . To order p^4 this is

$$T = (p^2 c^2 + m^2 c^4)^{1/2} - mc^2 \approx \frac{p^2}{2m} - \frac{p^4}{8m^3 c^2}.$$

Since the Hamiltonian is the sum of kinetic and potential energies, we see from this that the term

$$H_{\text{rel}} = -\frac{p^4}{8m^3 c^2} \quad (2.57)$$

may be treated as a perturbation to the states ψ_{nlm} .

While the states ψ_{nlm} are in general degenerate, in this case we don't have to worry about it. The reason is that H_{rel} is rotationally invariant, so it's already diagonal in the ψ_{nlm} basis, and that is precisely what the zeroth-order wavefunctions $\varphi_n^{(0)}$ accomplish (see equation (2.22)). Therefore we can use simple first-order perturbation theory so that

$$E_{\text{rel}}^{(1)} = -\frac{1}{8m^3 c^2} \langle \psi_{nlm} | p^4 | \psi_{nlm} \rangle.$$

Using $H^0 = p^2/2m - e^2/r$ we can write

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

and therefore

$$E_{\text{rel}}^{(1)} = -\frac{1}{2mc^2} \left[(E_n^{(0)})^2 + 2E_n^{(0)} e^2 \left\langle \frac{1}{r} \right\rangle + e^4 \left\langle \frac{1}{r^2} \right\rangle \right]$$

where $\langle \cdot \rangle$ is shorthand for $\langle \psi_{nlm} | \cdot | \psi_{nlm} \rangle$. These integrals are not hard to evaluate (see the end of this section), and the result (in different forms) is

$$\begin{aligned} E_{\text{rel}}^{(1)} &= -\frac{(E_n^{(0)})^2}{2mc^2} \left[-3 + \frac{4n}{l+1/2} \right] \\ &= -\frac{|E_n^{(0)}|^2 \alpha^2}{n^2} \left[-\frac{3}{4} + \frac{n}{l+1/2} \right] \\ &= -\frac{1}{2} mc^2 \alpha^4 \left[-\frac{3}{4n^4} + \frac{1}{n^3(l+1/2)} \right]. \end{aligned} \tag{2.58}$$

Adding equations (2.56) and (2.58) we obtain the **fine structure energy shift**

$$\begin{aligned} E_{\text{fs}}^{(1)} &= -\frac{mc^2 \alpha^4}{2n^3} \left[-\frac{3}{4n} + \frac{1}{j+1/2} \right] \\ &= -\frac{|E_n^{(0)}|^2 \alpha^2}{n^2} \left[-\frac{3}{4} + \frac{n}{j+1/2} \right] \end{aligned} \tag{2.59}$$

which is valid for both $j = l \pm 1/2$. This is the first-order energy correction due to the “fine structure Hamiltonian”

$$H_{\text{fs}} = H_{\text{so}} + H_{\text{rel}}. \tag{2.60}$$

2.4.1 Supplement: Miscellaneous Proofs

Now let’s go back and prove several miscellaneous results stated in this section. The first thing we want to show is that the energy of a magnetic moment in a uniform magnetic field is given by $-\boldsymbol{\mu} \cdot \mathbf{B}$ where $\boldsymbol{\mu}$ for a loop of area A carrying current I is defined to have magnitude IA and pointing perpendicular to the loop in the direction of your thumb if the fingers of your right hand are along the direction of the current. To see this, we simply calculate the work required to rotate a current loop from its equilibrium position to the desired orientation.

Consider Figure 6 below, where the current flows counterclockwise out of the page at the bottom and into the page at the top. Let the loop have length a on the sides and b across the top and bottom, so its area is ab . The magnetic force on a current-carrying wire is

$$\mathbf{F}_B = \int I d\mathbf{l} \times \mathbf{B}$$

and hence the forces on the opposite “ a sides” of the loop cancel, and the force on the top and bottom “ b sides” is $F_B = IbB$. The equilibrium position of the loop is

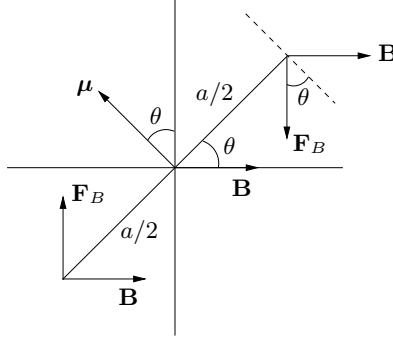


Figure 6: A current loop in a uniform magnetic field

horizontal, so the potential energy of the loop is the work required to rotate it from $\theta = 0$ to some value θ . This work is given by $W = \int \mathbf{F} \cdot d\mathbf{r}$ where \mathbf{F} is the force that I must apply against the magnetic field to rotate the loop.

Since the loop is rotating, the force I must apply at the top of the loop is in the direction of $\boldsymbol{\mu}$ and perpendicular to the loop, and hence has magnitude $F_B \cos \theta$. Then the work I do is (the factor of 2 takes into account both the top and bottom sides)

$$W = \int \mathbf{F} \cdot d\mathbf{r} = 2 \int F_B \cos \theta (a/2) d\theta = IabB \int_0^\theta \cos \theta d\theta = \mu B \sin \theta.$$

But note that $\boldsymbol{\mu} \cdot \mathbf{B} = \mu B \cos(90 + \theta) = -\mu B \sin \theta$, and therefore

$$W = -\boldsymbol{\mu} \cdot \mathbf{B}. \quad (2.61)$$

In this derivation, I never explicitly mentioned the torque on the loop due to \mathbf{B} . However, we see that

$$\begin{aligned} \|\mathbf{N}\| &= \|\mathbf{r} \times \mathbf{F}_B\| = 2(a/2)F_B \sin(90 + \theta) = IabB \sin(90 + \theta) \\ &= \mu B \sin(90 + \theta) = \|\boldsymbol{\mu} \times \mathbf{B}\| \end{aligned}$$

and therefore

$$\mathbf{N} = \boldsymbol{\mu} \times \mathbf{B}. \quad (2.62)$$

Note that $W = \int \|\mathbf{N}\| d\theta$.

Next I will prove equation (2.43). Let \mathbf{A} be a vector as seen in both the rotating and lab frames, and let $\{\mathbf{e}_i\}$ be a fixed basis in the rotating frame. Then (using the summation convention) $\mathbf{A} = A^i \mathbf{e}_i$ so that

$$\frac{d\mathbf{A}}{dt} = \frac{d}{dt}(A^i \mathbf{e}_i) = \frac{dA^i}{dt} \mathbf{e}_i + A^i \frac{d\mathbf{e}_i}{dt}.$$

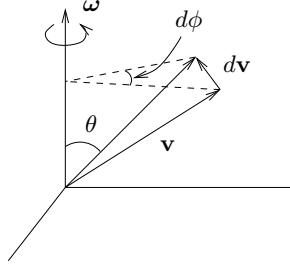
Now $(dA^i/dt)\mathbf{e}_i$ is the rate of change of \mathbf{A} with respect to the rotating frame, so we have

$$\frac{dA^i}{dt}\mathbf{e}_i = \left(\frac{d\mathbf{A}}{dt}\right)_{\text{rot}}.$$

And \mathbf{e}_i is a fixed basis vector in the frame that is rotating with respect to the lab frame. Then, just like any vector rotating in the lab with angular velocity $\boldsymbol{\omega}$, we have

$$\frac{d\mathbf{e}_i}{dt} = \boldsymbol{\omega} \times \mathbf{e}_i.$$

(See the figure below. Here $\omega = d\phi/dt$, and $dv = v \sin \theta d\phi$ so $dv/dt = v \sin \theta \omega$ or $d\mathbf{v}/dt = \boldsymbol{\omega} \times \mathbf{v}$.)



Then

$$A^i \frac{d\mathbf{e}_i}{dt} = A^i \boldsymbol{\omega} \times \mathbf{e}_i = \boldsymbol{\omega} \times A^i \mathbf{e}_i = \boldsymbol{\omega} \times \mathbf{A}.$$

Putting this all together we have

$$\left(\frac{d\mathbf{A}}{dt}\right)_{\text{lab}} = \left(\frac{d\mathbf{A}}{dt}\right)_{\text{rot}} + \boldsymbol{\omega} \times \mathbf{A}.$$

Equation (2.43) is just the ‘operator’ version of this result.

Finally, let me show how to evaluate the integrals $\langle 1/r \rangle$, $\langle 1/r^2 \rangle$ and $\langle 1/r^3 \rangle$ where the expectation values are taken with respect to the hydrogen atom wave functions ψ_{nlm} .

First, instead of $\langle 1/r \rangle$, consider $\langle \lambda/r \rangle$. This can be interpreted as the first-order correction to the energy due to the perturbation λ/r . But $H^0 = T + V = T - e^2/r$, so $H = H^0 + H' = H^0 + \lambda/r = T - (e^2 - \lambda)/r$, and this is just our original problem if we replace e^2 by $e^2 - \lambda$ everywhere. In particular, the *exact* energy solution is then

$$E_n(\lambda) = -\frac{m(e^2 - \lambda)^2}{2\hbar^2 n^2} = -\frac{me^4}{2\hbar^2 n^2} + \lambda \frac{me^2}{\hbar^2 n^2} - \lambda^2 \frac{m}{2\hbar^2 n^2}.$$

But another way of looking at this is as the expansion of $E_n(\lambda)$ given in (2.3b):

$$\begin{aligned} E_n &= E_n^{(0)} + \lambda \left(\frac{dE_n}{d\lambda}\right)_{\lambda=0} + \frac{\lambda^2}{2!} \left(\frac{d^2 E_n}{d\lambda^2}\right)_{\lambda=0} + \dots \\ &= E_n^{(0)} + \lambda E_n^{(1)} + \lambda^2 E_n^{(2)} + \dots \end{aligned}$$

where the first-order correction $E_n^{(1)} = \langle H' \rangle$ is just the term linear in λ . Therefore, letting $\lambda \rightarrow 1$, we have $\langle 1/r \rangle = \langle H' \rangle = me^2/\hbar^2 n^2$ or

$$\left\langle \frac{1}{r} \right\rangle = \frac{1}{a_0 n^2}. \quad (2.63)$$

Note that if you have the exact solution $E_n(\lambda)$, you can obtain $E_n^{(1)}$ by simply evaluating $\lambda(dE_n/d\lambda)_{\lambda=0}$.

Before continuing, let me rewrite the hydrogen atom Hamiltonian as follows:

$$\begin{aligned} H^0 &= -\frac{\hbar^2}{2m} \left(\frac{\partial^2}{\partial r^2} + \frac{2}{r} \frac{\partial}{\partial r} \right) + \frac{1}{2mr^2} \mathbf{L}^2 - \frac{e^2}{r} \\ &= \frac{p_r^2}{2m} + \frac{\mathbf{L}^2}{2mr^2} - \frac{e^2}{r} \end{aligned} \quad (2.64)$$

where I have defined the “radial momentum” p_r by

$$p_r = -i\hbar \left(\frac{\partial}{\partial r} + \frac{1}{r} \right).$$

Now consider $\langle \lambda/r^2 \rangle$. Again, letting $H = H^0 + H' = H^0 + \lambda/r^2$, we can still solve the problem exactly because all we are doing is modifying the centrifugal term

$$\frac{\mathbf{L}^2}{2mr^2} \rightarrow \frac{\mathbf{L}^2 + 2m\lambda}{2mr^2} \rightarrow \frac{\hbar^2 l(l+1) + 2m\lambda}{2mr^2} = \frac{\hbar^2 l'(l'+1)}{2mr^2}$$

where $l' = l'(\lambda)$ is a function of λ . (Just write $\hbar^2 l'(l'+1) = \hbar^2 l(l+1) + 2m\lambda$ and use the quadratic formula to find l' as a function of λ .)

Recall that the exact energies were defined by

$$E_n = -\frac{me^4}{2\hbar^2 n^2} = -\frac{me^4}{2\hbar^2 (k+l+1)^2}$$

where $k = 0, 1, 2, \dots$ was the integer that terminated the power series solution of the radial equation. Now what we have is

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

where (note $\lambda = 0$ implies $l' = l$)

$$E^{(1)} = \left. \frac{dE}{d\lambda} \right|_{\lambda=0} = \left. \frac{dl'}{d\lambda} \right|_{l'=l} \left. \frac{dE}{dl'} \right|_{l'=l}.$$

Then from the explicit form of $E(l')$ and the definition of n we have

$$\left. \frac{dE}{dl'} \right|_{l'=l} = \frac{me^4}{\hbar^2 (k+l+1)^3} = \frac{me^4}{\hbar^2 n^3}$$

and taking the derivative of $\hbar^2 l'(l+1) = \hbar^2 l(l+1) + 2m\lambda$ with respect to λ yields

$$\left. \frac{dl'}{d\lambda} \right|_{l'=l} = \frac{2m}{\hbar^2} \frac{1}{2l+1} = \frac{m}{\hbar^2} \frac{1}{(l+1/2)}.$$

Therefore

$$E^{(1)} = \frac{(me^2/\hbar^2)^2}{(l+1/2)n^3}$$

and $\langle \lambda/r^2 \rangle = \lambda E^{(1)}$ so that

$$\left\langle \frac{1}{r^2} \right\rangle = \frac{1}{a_0^2(l+1/2)n^3}. \quad (2.65)$$

The last integral to evaluate is $\langle 1/r^3 \rangle$. Since there is no term in H^0 that goes like $1/r^3$, we have to try something else. Note that $H^0\psi_{nlm} = E_n\psi_{nlm}$ so that

$$\langle [H^0, p_r] \rangle = \langle \psi_{nlm} | H^0 p_r - p_r H^0 | \psi_{nlm} \rangle = E_n \langle p_r \rangle - \langle p_r \rangle E_n = 0.$$

Using

$$\left[\frac{1}{r}, \frac{\partial}{\partial r} \right] = \frac{1}{r^2} \quad \text{and} \quad \left[\frac{1}{r^2}, \frac{\partial}{\partial r} \right] = \frac{2}{r^3}$$

(recall $[ab, c] = a[b, c] + [a, c]b$), it is easy to use (2.64) and show that

$$[H^0, p_r] = -\frac{i\hbar \mathbf{L}^2}{m r^3} + \frac{i\hbar e^2}{r^2}.$$

But now

$$\begin{aligned} 0 = \langle [H^0, p_r] \rangle &= -\frac{i\hbar}{m} \left\langle \frac{\mathbf{L}^2}{r^3} \right\rangle + i\hbar e^2 \left\langle \frac{1}{r^2} \right\rangle \\ &= -\frac{i\hbar^3 l(l+1)}{m} \left\langle \frac{1}{r^3} \right\rangle + i\hbar e^2 \left\langle \frac{1}{r^2} \right\rangle \end{aligned}$$

and therefore

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

or

$$\left\langle \frac{1}{r^3} \right\rangle = \frac{1}{a_0 l(l+1)} \left\langle \frac{1}{r^2} \right\rangle. \quad (2.66)$$

Combining this with (2.65) we have

$$\left\langle \frac{1}{r^3} \right\rangle = \frac{1}{a_0^3 l(l+1)(l+1/2)n^3} \quad (2.67)$$

2.5 The Zeeman Effect

In the previous section we studied the effect of an atomic electron's magnetic moment interacting with the magnetic field generated by the nucleus (a proton). In this section, I want to investigate what happens when a hydrogen atom is placed in a *uniform external* magnetic field \mathbf{B} . These types of interactions are generally referred to as the **Zeeman effect**, and they were instrumental in the discovery of spin. (Pieter Zeeman and H.A. Lorentz shared the second Nobel prize in physics in 1902. For a very interesting summary of the history of spin, read Chapter 10 in the text *Quantum Mechanics* by Hendrik Hamerka.)

The hydrogen atom Hamiltonian, including fine structure, is given by

$$H = H^0 + H_{\text{fs}} = H^0 + H_{\text{so}} + H_{\text{rel}}$$

where

$$H^0 = -\frac{\hbar^2}{2m} \left(\frac{\partial^2}{\partial r^2} + \frac{2}{r} \frac{\partial}{\partial r} \right) + \frac{1}{2mr^2} \mathbf{L}^2 - \frac{e^2}{r} \quad (\text{equation (2.34)})$$

$$H_{\text{so}} = \frac{e^2}{2m^2 c^2 r^3} \mathbf{L} \cdot \mathbf{S} \quad (\text{equation (2.48b)})$$

$$H_{\text{rel}} = -\frac{p^4}{8m^3 c^2} \quad (\text{equation (2.57)}).$$

(And where I'm approximating the reduced mass by the electron mass m_e .) The easy way to include the presence of an external field \mathbf{B} is to simply add an interaction energy

$$H_{\text{mag}} = -\boldsymbol{\mu}_{\text{tot}} \cdot \mathbf{B}$$

where, from equations (2.38) and (2.39), we know that the total magnetic moment for a hydrogenic electron is

$$\boldsymbol{\mu}_{\text{tot}} = \boldsymbol{\mu}_l + \boldsymbol{\mu}_s = -\frac{e}{2m_e c} (\mathbf{L} + 2\mathbf{S}) = -\frac{e}{2m_e c} (\mathbf{J} + \mathbf{S}). \quad (2.68)$$

However, the correct way to arrive at this is to rewrite the Hamiltonian taking into account the presence of an electromagnetic field. For those who are interested, I work through this approach at the end of this section.

In any case, the Hamiltonian for a hydrogen atom in an external uniform magnetic field is then

$$H = H^0 + H_{\text{so}} + H_{\text{rel}} + H_{\text{mag}}.$$

There are really three cases to consider. (I'll ignore H_{rel} for now because it's a correction to the kinetic energy and irrelevant to this discussion.) The first is when \mathbf{B} is strong enough that H_{mag} is large relative to H_{so} . In this case we can treat H_{so} as a perturbation on the states defined by $H^0 + H_{\text{mag}}$, where these states are simultaneous eigenfunctions of L^2 , S^2 , L_z and S_z (rather than J^2 and J_z). The reason that \mathbf{J} is not a good quantum number is that the external field exerts a

torque $\boldsymbol{\mu}_{\text{tot}} \times \mathbf{B}$ on the total magnetic moment, and this is equivalent to a changing total angular momentum $d\mathbf{J}/dt$. Thus \mathbf{J} is not conserved, and in fact precesses about \mathbf{B} . In addition, if there is a spin-orbit interaction, then this *internal* field causes \mathbf{L} and \mathbf{S} to precess about \mathbf{J} .

The second case is when \mathbf{B} is weak and H_{so} dominates H_{mag} . In this situation, H_{mag} is treated as a perturbation on the states defined by $H^0 + H_{\text{so}}$. As we saw in our discussion of H_{so} , in this case we must choose our states to be eigenfunctions of L^2 , S^2 , J^2 and J_z because \mathbf{L} and \mathbf{S} are not conserved separately, even though $\mathbf{J} = \mathbf{L} + \mathbf{S}$ is conserved. (Neither \mathbf{L} nor \mathbf{S} alone commutes with $\mathbf{L} \cdot \mathbf{S}$, but $[J_i, \mathbf{L} \cdot \mathbf{S}] = 0$ and hence J^2 commutes with H .)

And the third and most difficult case is when both H_{so} and H_{mag} are roughly equivalent. Under this “intermediate-field” situation, we must take them together and use degenerate perturbation theory to break the degeneracies of the basis states.

2.5.1 Strong External Field

Let us first consider the case where the external magnetic field is much stronger than the internal field felt by the electron and due to its orbital motion. Taking $\mathbf{B} = B\hat{\mathbf{z}}$ we have

$$H_{\text{mag}} = \frac{eB}{2m_e c} (L_z + 2S_z). \quad (2.69)$$

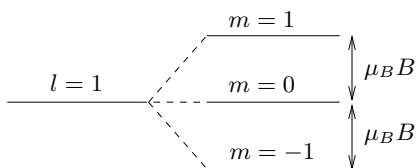
If we first ignore spin, then the first-order correction to the hydrogen atom energy levels is

$$E_{nlm}^{(1)} = \left\langle \psi_{nlm} \left| \frac{eB}{2m_e c} L_z \right| \psi_{nlm} \right\rangle = \frac{e\hbar}{2m_e c} Bm := \mu_B Bm$$

where

$$\mu_B = \frac{e\hbar}{2m_e c} = 5.79 \times 10^{-9} \text{ eV/gauss} = 9.29 \times 10^{-21} \text{ erg/gauss}$$

is called the (electron) **Bohr magneton**. Thus we see that for a given l , the $(2l+1)$ -fold degeneracy is lifted. For example, the 3-fold degenerate $l = 1$ state is split into three states, with an energy difference of $\mu_B B$ between states:

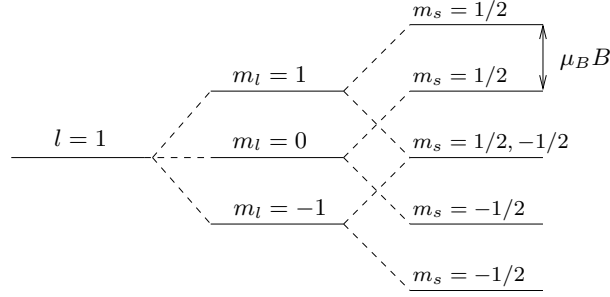


This strong field case is sometimes called the **Paschen-Back effect**.

If we now include spin, then

$$E_{nlm_l m_s}^{(1)} = \mu_B B(m_l + 2m_s) \quad (2.70)$$

where $m_s = \pm 1/2$. This yields the further splitting (or lifting of degeneracies) sometimes called the **anomalous Zeeman effect**:



This gives us the energy levels $E_n^{(0)} + E_{nlm_l m_s}^{(1)}$ where $E_n^{(0)}$ is given by (2.54a).

However, since the basis states we used here are just the usual hydrogen atom wave functions, it is easy to include further corrections due to both H_{so} and the relativistic correction H_{rel} discussed in Section 2.4. We simply apply first-order perturbation theory using these as the perturbing potentials. For H_{rel} , we can simply use the result (2.58). However, we can't just use equations (2.56) for H_{so} because they were derived using the eigenfunctions of J^2 which don't apply when there is a strong external magnetic field.

To get around this problem, we simply calculate $\langle \psi_{nlm_l m_s} | \mathbf{L} \cdot \mathbf{S} | \psi_{nlm_l m_s} \rangle$. We have

$$\mathbf{L} \cdot \mathbf{S} = L_x S_x + L_y S_y + L_z S_z$$

where $L_x = (L_+ + L_-)/2$ and $L_y = (L_+ - L_-)/2i$ with similar results for S_x and S_y . Using these, it is quite easy to see that the orthogonality of the eigenfunctions yields

$$\langle \psi | L_x S_x | \psi \rangle = \langle \psi | L_y S_y | \psi \rangle = 0$$

while

$$\langle \psi_{nlm_l m_s} | L_z S_z | \psi_{nlm_l m_s} \rangle = \hbar^2 m_l m_s. \quad (2.71)$$

Combining the results for H_{rel} and H_{so} we obtain the following corrections to the “unperturbed” energies $E_n^{(0)} + E_{nlm_l m_s}^{(1)}$:

$$E_{\text{rel}}^{(1)} + E_{\text{so}}^{(1)} = \frac{mc^2 \alpha^4}{2n^3} \left[\frac{3}{4n} - \frac{1}{l+1/2} \right] + \frac{e^2}{2m^2 c^2} \hbar^2 m_l m_s \frac{1}{a_0^3 n^3 l(l+1)(l+1/2)}$$

where we used equations (2.58), (2.48b), (2.67) and (2.71). After a little algebra, which I leave to you, we arrive at

$$\begin{aligned} E_{\text{rel}}^{(1)} + E_{\text{so}}^{(1)} &= \frac{me^4 \alpha^2}{2\hbar^2 n^3} \left\{ \frac{3}{4n} - \left[\frac{l(l+1) - m_l m_s}{l(l+1)(l+1/2)} \right] \right\} \\ &= -E_1^{(0)} \frac{\alpha^2}{n^3} \left\{ \frac{3}{4n} - \left[\frac{l(l+1) - m_l m_s}{l(l+1)(l+1/2)} \right] \right\}. \end{aligned} \quad (2.72)$$

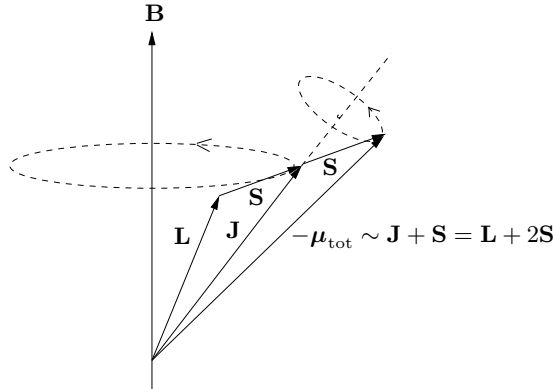
2.5.2 Weak External Field

Now we turn to the second case where the external field is weak relative to the spin-orbit term. As we discussed above, now we must take our basis states to be eigenfunctions of L^2 , S^2 , J^2 and J_z .

For a many-electron atom, there are basically two ways to calculate the total \mathbf{J} . The first way is to calculate $\mathbf{L} = \sum \mathbf{L}_i$ and $\mathbf{S} = \sum \mathbf{S}_i$ and then evaluate $\mathbf{J} = \mathbf{L} + \mathbf{S}$. This is called **L - S or Russel-Saunders coupling**. It is applicable to the lighter elements where interelectronic repulsion energies are significantly greater than the spin-orbit interaction energies. This is because if the spin-orbit coupling is weak, then \mathbf{L} and \mathbf{S} “almost” commute with $H^0 + H_{so}$.

The second way is to first calculate $\mathbf{J}_i = \mathbf{L}_i + \mathbf{S}_i$ so that $\mathbf{J} = \sum \mathbf{J}_i$. This is called **j - j coupling**. It is used for heavier elements where the electrons are moving very rapidly, and hence there is a strong spin-orbit interaction. Because of this, \mathbf{L} and \mathbf{S} no longer commute with H , even though \mathbf{J} does so. This type of coupling is also more difficult to use, so we will deal only with the L - S scheme.

Here is the physical situation:



Since \mathbf{J} commutes with $H^0 + H_{so}$, it is conserved (and hence is fixed in space), even though \mathbf{L} and \mathbf{S} are not. This means that \mathbf{L} and \mathbf{S} both precess about \mathbf{J} . If the applied external \mathbf{B} field is much weaker than the internal field, then \mathbf{J} will precess much more slowly about \mathbf{B} than \mathbf{L} and \mathbf{S} precess about \mathbf{J} . We need to evaluate the correction (2.69) in first-order perturbation theory.

Since our basis states are eigenfunctions of J^2 and J_z but not L_z and S_z , we can't directly evaluate the expectation value of $L_z + 2S_z = J_z + S_z$. The correct way to handle this is to use the Wigner-Eckart theorem, which is rather beyond the scope of this course. Instead, we will use a physical argument that gets us to the same answer.

We note that since \mathbf{L} and \mathbf{S} (and hence μ_{tot}) precess rapidly about \mathbf{J} , the time average of the Hamiltonian $H_{\text{mag}}^{\text{av}} = -\langle \mu_{\text{tot}} \cdot \mathbf{B} \rangle$ will be the same as $-\langle \mu_{\text{tot}} \rangle \cdot \mathbf{B}$. But

the average of $\boldsymbol{\mu}_{\text{tot}}$ is just its component along \mathbf{J} , which is

$$\langle \boldsymbol{\mu}_{\text{tot}} \rangle = (\boldsymbol{\mu}_{\text{tot}} \cdot \hat{\mathbf{J}}) \hat{\mathbf{J}} = \frac{\boldsymbol{\mu}_{\text{tot}} \cdot \mathbf{J}}{J^2} \mathbf{J}.$$

Using $\mathbf{L} = \mathbf{J} - \mathbf{S}$ so that $L^2 = J^2 + S^2 - 2\mathbf{S} \cdot \mathbf{J}$ we have

$$(\mathbf{J} + \mathbf{S}) \cdot \mathbf{J} = J^2 + \mathbf{S} \cdot \mathbf{J} = J^2 + \frac{1}{2}(J^2 + S^2 - L^2).$$

Then since $\mathbf{B} = B\hat{\mathbf{z}}$, we now have

$$\begin{aligned} H_{\text{mag}}^{\text{av}} &= -B \langle \boldsymbol{\mu}_{\text{tot}} \rangle \cdot \hat{\mathbf{z}} = \frac{eB}{2m_e c} \frac{(\mathbf{J} + \mathbf{S}) \cdot \mathbf{J}}{J^2} J_z \\ &= \frac{eB J_z}{2m_e c} \left[1 + \frac{J^2 + S^2 - L^2}{2J^2} \right]. \end{aligned}$$

Our basis states are simultaneous eigenstates of L^2 , S^2 , J^2 and J_z , so the average energy $E_{\text{mag}}^{\text{av}}$ is given by the first-order correction

$$\begin{aligned} E_{\text{mag}}^{\text{av}} &= \frac{e\hbar B m_j}{2m_e c} \left[1 + \frac{j(j+1) + s(s+1) - l(l+1)}{2j(j+1)} \right] \\ &= \frac{e\hbar B m_j}{2m_e c} \left[\frac{3}{2} + \frac{3/4 - l(l+1)}{2j(j+1)} \right] \tag{2.73} \\ &:= \mu_B B m_j g_J \end{aligned}$$

where the **Landé g-factor** g_J is defined by

$$g_J = 1 + \frac{j(j+1) + s(s+1) - l(l+1)}{2j(j+1)}.$$

The total energy of a hydrogen atom in a uniform magnetic field is now given by the sum of the ground state energy $E_n^{(0)}$ (equation (2.54a)), the fine-structure correction $E_{fs}^{(1)}$ (equation (2.59)) and $E_{\text{mag}}^{\text{av}}$ (equation (2.73)).

2.5.3 Intermediate-Field Case

Finally, we consider the intermediate-field case where the internal and external magnetic fields are approximately the same. In this situation, we must apply degenerate perturbation theory to the degenerate “unperturbed” states $\psi_{nlm_l m_s}$ by treating $H' = H_{\text{fs}} + H_{\text{mag}}$ as a perturbation. It is easiest to simply work out an example.

As we saw in our discussion of spin-orbit coupling, it is best to work in the basis in which our states are simultaneous eigenstates of L^2 , S^2 , J^2 and J_z . (The choice of basis has no effect on the eigenvalues of $H_{\text{fs}} + H_{\text{mag}}$, and the eigenvalues are just

what we are looking for when we solve (2.21).) Let us consider the hydrogen atom state with $n = 2$, so that $l = 0, 1$. Since $s = 1/2$, the possible j values are

$$0 \otimes \frac{1}{2} + 1 \otimes \frac{1}{2} = \frac{1}{2} + \frac{3}{2} \oplus \frac{1}{2}$$

or $j = 1/2, 3/2, 1/2$. Our basis states $|l s j m_j\rangle$ are given in terms of the states $|l s m_l m_s\rangle$ using the appropriate Clebsch-Gordan coefficients (which you can look up or calculate for yourself).

For $l = 0$ we have $j = 1/2$ so $m_j = \pm 1/2$ and we have the two states

$$\begin{aligned}\psi_1 &:= |0 \frac{1}{2} \frac{1}{2} \frac{1}{2}\rangle &= |0 \frac{1}{2} 0 \frac{1}{2}\rangle \\ \psi_2 &:= |0 \frac{1}{2} \frac{1}{2} -\frac{1}{2}\rangle &= |0 \frac{1}{2} 0 -\frac{1}{2}\rangle\end{aligned}$$

where the first state in each line is the state $|l s j m_j\rangle$, and the second state in each line is the linear combination of states $|l s m_l m_s\rangle$ with Clebsch-Gordan coefficients. (For $l = 0$ the C-G coefficients are just 1.)

For $l = 1$ we have the four states with $j = 3/2$ and the two states with $j = 1/2$ (which we order with a little hindsight so the determinant (2.21) turns out block diagonal):

$$\begin{aligned}\psi_3 &:= |1 \frac{1}{2} \frac{3}{2} \frac{3}{2}\rangle &= |1 \frac{1}{2} 1 \frac{1}{2}\rangle \\ \psi_4 &:= |1 \frac{1}{2} \frac{3}{2} -\frac{3}{2}\rangle &= |1 \frac{1}{2} -1 -\frac{1}{2}\rangle \\ \psi_5 &:= |1 \frac{1}{2} \frac{3}{2} \frac{1}{2}\rangle &= \sqrt{\frac{2}{3}} |1 \frac{1}{2} 0 \frac{1}{2}\rangle + \sqrt{\frac{1}{3}} |1 \frac{1}{2} 1 -\frac{1}{2}\rangle \\ \psi_6 &:= |1 \frac{1}{2} \frac{3}{2} \frac{1}{2}\rangle &= -\sqrt{\frac{1}{3}} |1 \frac{1}{2} 0 \frac{1}{2}\rangle + \sqrt{\frac{2}{3}} |1 \frac{1}{2} 1 -\frac{1}{2}\rangle \\ \psi_7 &:= |1 \frac{1}{2} \frac{3}{2} -\frac{1}{2}\rangle &= \sqrt{\frac{1}{3}} |1 \frac{1}{2} -1 \frac{1}{2}\rangle + \sqrt{\frac{2}{3}} |1 \frac{1}{2} 0 -\frac{1}{2}\rangle \\ \psi_8 &:= |1 \frac{1}{2} \frac{1}{2} -\frac{1}{2}\rangle &= -\sqrt{\frac{2}{3}} |1 \frac{1}{2} -1 \frac{1}{2}\rangle + \sqrt{\frac{1}{3}} |1 \frac{1}{2} 0 -\frac{1}{2}\rangle.\end{aligned}$$

Now we need to evaluate the matrices of $H_{\text{fs}} = H_{\text{so}} + H_{\text{rel}}$ and H_{mag} in the $|j m_j\rangle$ basis $\{\psi_i\}$. Since $H_{\text{rel}} \sim p^4$, it's already diagonal in the $|j m_j\rangle$ basis. And since $H_{\text{so}} \sim \mathbf{S} \cdot \mathbf{L} = (1/2)(J^2 - L^2 - S^2)$, it's also diagonal in the $|j m_j\rangle$ basis. Therefore H_{fs} is diagonal and its contribution is given by (2.59):

$$\begin{aligned}\langle j m_j | H_{\text{fs}} | j m_j \rangle &= -\frac{|E_n^{(0)}| \alpha^2}{n^2} \left[-\frac{3}{4} + \frac{n}{j + 1/2} \right] \\ &= -\frac{|E_1^{(0)}| \alpha^2}{16} \left[\frac{2}{j + 1/2} - \frac{3}{4} \right]\end{aligned}$$

where I used $E_n^{(0)} = E_1^{(0)}/n^2$ and let $n = 2$. For states with $j = 1/2$, this gives a contribution

$$\langle \psi_i | H_{\text{fs}} | \psi_i \rangle = -\frac{5|E_1^{(0)}| \alpha^2}{64} := -5\xi \quad \text{for } i = 1, 2, 6, 8 \quad (2.74a)$$

and for states with $j = 3/2$ this is

$$\langle \psi_i | H_{\text{fs}} | \psi_i \rangle = -\frac{|E_1^{(0)}| \alpha^2}{64} := -\xi \quad \text{for } i = 3, 4, 5, 7. \quad (2.74b)$$

Next, we easily see that the first four states ψ_1 – ψ_4 are eigenstates of $H_{\text{mag}} \sim L_z + 2S_z$ (since they each contain only a single factor $|l s m_l m_s\rangle$). Hence H_{mag} is already diagonal in this 4×4 block, and so contributes the diagonal terms

$$\langle \psi_i | H_{\text{mag}} | \psi_i \rangle = \mu_B B (m_l + 2m_s) := \beta (m_l + 2m_s) \quad \text{for } i = 1, 2, 3, 4.$$

For the remaining four states ψ_5 – ψ_8 we must explicitly evaluate the matrix elements. For example,

$$\begin{aligned} H_{\text{mag}} | \psi_5 \rangle &= \frac{\mu_B B}{\hbar} (L_z + 2S_z) \left\{ \sqrt{\frac{2}{3}} \left| 1 \frac{1}{2} 0 \frac{1}{2} \right\rangle + \sqrt{\frac{1}{3}} \left| 1 \frac{1}{2} 1 -\frac{1}{2} \right\rangle \right\} \\ &= \mu_B B \left\{ 1 \cdot \sqrt{\frac{2}{3}} \left| 1 \frac{1}{2} 0 \frac{1}{2} \right\rangle + 0 \cdot \sqrt{\frac{1}{3}} \left| 1 \frac{1}{2} 1 -\frac{1}{2} \right\rangle \right\} \\ &= \mu_B B \sqrt{\frac{2}{3}} \left| 1 \frac{1}{2} 0 \frac{1}{2} \right\rangle \end{aligned}$$

and therefore (using the orthonormality of the states $|l s m_l m_s\rangle$)

$$\langle \psi_5 | H_{\text{mag}} | \psi_5 \rangle = \frac{2}{3} \mu_B B := \frac{2}{3} \beta$$

and

$$\langle \psi_6 | H_{\text{mag}} | \psi_5 \rangle = \langle \psi_5 | H_{\text{mag}} | \psi_6 \rangle = -\frac{\sqrt{2}}{3} \mu_B B := -\frac{\sqrt{2}}{3} \beta.$$

Also,

$$\langle \psi_6 | H_{\text{mag}} | \psi_6 \rangle = \left\langle \psi_6 \left| -\sqrt{\frac{1}{3}} \mu_B B \left| 1 \frac{1}{2} 0 \frac{1}{2} \right\rangle \right. \right\rangle = \frac{1}{3} \mu_B B := \frac{1}{3} \beta.$$

Since all other matrix elements with ψ_5 and ψ_6 vanish, there is a 2×2 block corresponding to the subspace spanned by ψ_5 and ψ_6 . Similarly, there is a 2×2 block corresponding to the subspace spanned by ψ_7 and ψ_8 with

$$\langle \psi_7 | H_{\text{mag}} | \psi_7 \rangle = -\frac{2}{3} \beta$$

$$\langle \psi_8 | H_{\text{mag}} | \psi_7 \rangle = \langle \psi_7 | H_{\text{mag}} | \psi_8 \rangle = -\frac{\sqrt{2}}{3} \beta$$

$$\langle \psi_8 | H_{\text{mag}} | \psi_8 \rangle = -\frac{1}{3} \beta.$$

$$\begin{aligned}
E_6^{(1)} &= E_2^{(0)} - 3\xi + \frac{\beta}{2} - \sqrt{4\xi^2 + \frac{2}{3}\xi\beta + \frac{1}{4}\beta^2} \\
E_7^{(1)} &= E_2^{(0)} - 3\xi - \frac{\beta}{2} + \sqrt{4\xi^2 - \frac{2}{3}\xi\beta + \frac{1}{4}\beta^2} \\
E_8^{(1)} &= E_2^{(0)} - 3\xi - \frac{\beta}{2} - \sqrt{4\xi^2 - \frac{2}{3}\xi\beta + \frac{1}{4}\beta^2}
\end{aligned}$$

For $i = 1, 2, 3, 4$ the energy $E_i^{(1)}$ corresponds to ψ_i . But for $i = 5, 6$ the energy $E_i^{(1)}$ corresponds to some linear combination of ψ_5 and ψ_6 , and similarly for $i = 7, 8$ the energy E_i corresponds to a linear combination of ψ_7 and ψ_8 . (This is the essential content of Section 2.2.)

It is easy to see that for $\beta = 0$ (i.e., $\mathbf{B} = 0$), these energies reduce to E_{fs} given by (2.74), and for very large β , we obtain the Paschen-Back energies given by (2.70). Thus our results have the correct limiting behavior. See Figure 7 below.

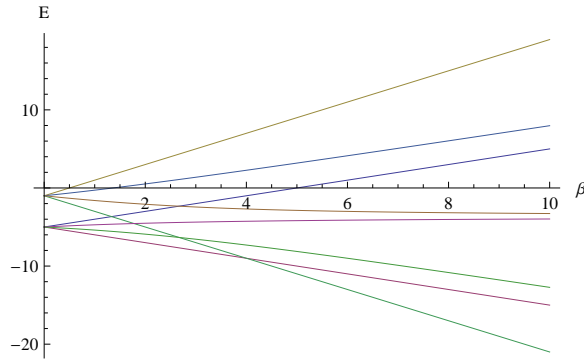


Figure 7: Intermediate-field energy corrections as a function of \mathbf{B} for $n = 2$.

2.5.4 Supplement: The Electromagnetic Hamiltonian

In a proper derivation of the Lagrange equations of motion, one starts from d'Alembert's principle of virtual work, and derives Lagrange's equations

$$\frac{d}{dt} \frac{\partial T}{\partial \dot{q}_i} - \frac{\partial T}{\partial q_i} = Q_i \tag{2.75}$$

where the q_i are generalized coordinates, $T = T(q_i, \dot{q}_i)$ is the kinetic energy and $Q_i = \sum_j F_j (\partial x_j / \partial q_i)$ is a generalized force. In the particular case that Q_i is derivable from a conservative force $F_j = -\partial V / \partial x_j$, then we have $Q_i = -\partial V / \partial q_i$. Since the potential energy V is assumed to be independent of \dot{q}_i , we can replace $\partial T / \partial \dot{q}_i$ by $\partial(T - V) / \partial \dot{q}_i$ and we arrive at the usual Lagrange's equations

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{q}_i} - \frac{\partial L}{\partial q_i} = 0 \tag{2.76}$$

where $L = T - V$. However, even if there is no potential function V , we can still arrive at this result if there exists a function $U = U(q_i, \dot{q}_i)$ such that the generalized forces may be written as

$$Q_i = -\frac{\partial U}{\partial q_i} + \frac{d}{dt} \frac{\partial U}{\partial \dot{q}_i}$$

because defining $L = T - U$ we again arrive at equation (2.76). The function U is called a **generalized potential** or a **velocity dependent potential**. We now seek such a function to describe the force on a charged particle in an electromagnetic field.

Recall from electromagnetism that the Lorentz force law is given by

$$\mathbf{F} = q \left(\mathbf{E} + \frac{\mathbf{v}}{c} \times \mathbf{B} \right)$$

or

$$\mathbf{F} = q \left(-\nabla\phi - \frac{1}{c} \frac{\partial \mathbf{A}}{\partial t} + \frac{\mathbf{v}}{c} \times (\nabla \times \mathbf{A}) \right)$$

where $\mathbf{E} = -\nabla\phi - (1/c)\partial\mathbf{A}/\partial t$ and $\mathbf{B} = \nabla \times \mathbf{A}$. Our goal is to write this in the form

$$F_i = -\frac{\partial U}{\partial x_i} + \frac{d}{dt} \frac{\partial U}{\partial \dot{x}_i}$$

for a suitable U . All it takes is some vector algebra. We have

$$\begin{aligned} [\mathbf{v} \times (\nabla \times \mathbf{A})]_i &= \varepsilon_{ijk} \varepsilon^{klm} v^j \partial_l A_m = (\delta_i^l \delta_j^m - \delta_i^m \delta_j^l) v^j \partial_l A_m \\ &= v^j \partial_i A_j - v^j \partial_j A_i = v^j \partial_i A_j - (\mathbf{v} \cdot \nabla) A_i. \end{aligned}$$

But x^i and \dot{x}^j are independent variables (in other words, \dot{x}^j has no explicit dependence on x^i) so that

$$v^j \partial_i A_j = \dot{x}^j \frac{\partial A_j}{\partial x^i} = \frac{\partial}{\partial x^i} (\dot{x}^j A_j) = \frac{\partial}{\partial x^i} (\mathbf{v} \cdot \mathbf{A})$$

and we have

$$[\mathbf{v} \times (\nabla \times \mathbf{A})]_i = \frac{\partial}{\partial x^i} (\mathbf{v} \cdot \mathbf{A}) - (\mathbf{v} \cdot \nabla) A_i.$$

But we also have

$$\frac{dA_i}{dt} = \frac{\partial A_i}{\partial x^j} \frac{dx^j}{dt} + \frac{\partial A_i}{\partial t} = v^j \frac{\partial A_i}{\partial x^j} + \frac{\partial A_i}{\partial t} = (\mathbf{v} \cdot \nabla) A_i + \frac{\partial A_i}{\partial t}$$

so that

$$(\mathbf{v} \cdot \nabla) A_i = \frac{dA_i}{dt} - \frac{\partial A_i}{\partial t}$$

and therefore

$$[\mathbf{v} \times (\nabla \times \mathbf{A})]_i = \frac{\partial}{\partial x^i} (\mathbf{v} \cdot \mathbf{A}) - \frac{dA_i}{dt} + \frac{\partial A_i}{\partial t}.$$

But we can write $A_i = \partial(v^j A_j)/\partial v^i = \partial(\mathbf{v} \cdot \mathbf{A})/\partial v^i$ which gives us

$$[\mathbf{v} \times (\nabla \times \mathbf{A})]_i = \frac{\partial}{\partial x^i}(\mathbf{v} \cdot \mathbf{A}) - \frac{d}{dt} \frac{\partial}{\partial v^i}(\mathbf{v} \cdot \mathbf{A}) + \frac{\partial A_i}{\partial t}.$$

The Lorentz force law can now be written in the form

$$\begin{aligned} F_i &= q \left(-\frac{\partial \phi}{\partial x^i} - \frac{1}{c} \frac{\partial A_i}{\partial t} + \frac{1}{c} [\mathbf{v} \times (\nabla \times \mathbf{A})]_i \right) \\ &= q \left(-\frac{\partial \phi}{\partial x^i} - \frac{1}{c} \frac{\partial A_i}{\partial t} + \frac{1}{c} \frac{\partial}{\partial x^i}(\mathbf{v} \cdot \mathbf{A}) - \frac{1}{c} \frac{d}{dt} \frac{\partial}{\partial v^i}(\mathbf{v} \cdot \mathbf{A}) + \frac{1}{c} \frac{\partial A_i}{\partial t} \right) \\ &= q \left[-\frac{\partial}{\partial x^i} \left(\phi - \frac{\mathbf{v}}{c} \cdot \mathbf{A} \right) - \frac{d}{dt} \frac{\partial}{\partial v^i} \left(\frac{\mathbf{v}}{c} \cdot \mathbf{A} \right) \right]. \end{aligned}$$

Since ϕ is independent of \mathbf{v} we can write

$$-\frac{d}{dt} \frac{\partial}{\partial v^i} \left(\frac{\mathbf{v}}{c} \cdot \mathbf{A} \right) = \frac{d}{dt} \frac{\partial}{\partial v^i} \left(\phi - \frac{\mathbf{v}}{c} \cdot \mathbf{A} \right)$$

so that

$$F_i = q \left[-\frac{\partial}{\partial x^i} \left(\phi - \frac{\mathbf{v}}{c} \cdot \mathbf{A} \right) + \frac{d}{dt} \frac{\partial}{\partial v^i} \left(\phi - \frac{\mathbf{v}}{c} \cdot \mathbf{A} \right) \right]$$

or

$$F_i = -\frac{\partial U}{\partial x_i} + \frac{d}{dt} \frac{\partial U}{\partial \dot{x}_i}$$

where $U = q(\phi - \mathbf{v}/c \cdot \mathbf{A})$. This shows that U is a generalized potential and that the Lagrangian for a particle of charge q in an electromagnetic field is

$$L = T - q\phi + \frac{q}{c} \mathbf{v} \cdot \mathbf{A} \quad (2.77a)$$

or

$$L = \frac{1}{2} m v^2 - q\phi + \frac{q}{c} \mathbf{v} \cdot \mathbf{A}. \quad (2.77b)$$

From this, the canonical momentum is defined by $p_i = \partial L / \partial \dot{x}_i = \partial L / \partial v_i$ so that

$$\mathbf{p} = m\mathbf{v} + \frac{q}{c} \mathbf{A}.$$

Using this, the Hamiltonian is then given by

$$\begin{aligned} H &= \sum p_i \dot{x}_i - L = \mathbf{p} \cdot \mathbf{v} - L \\ &= m v^2 + \frac{q}{c} \mathbf{A} \cdot \mathbf{v} - \frac{1}{2} m v^2 + q\phi - \frac{q}{c} \mathbf{A} \cdot \mathbf{v} \\ &= \frac{1}{2} m v^2 + q\phi \\ &= \frac{1}{2m} \left(\mathbf{p} - \frac{q}{c} \mathbf{A} \right)^2 + q\phi. \end{aligned}$$

This is the basis for the oft heard statement that to include electromagnetic forces, you need to make the replacement $\mathbf{p} \rightarrow \mathbf{p} - (q/c)\mathbf{A}$. Including any other additional potential energy terms, the Hamiltonian becomes

$$H = \frac{1}{2m} \left(\mathbf{p} - \frac{q}{c}\mathbf{A} \right)^2 + q\phi + V(\mathbf{r}). \quad (2.78)$$

Let's evaluate (2.78) for the case of a *uniform* magnetic field. Since $\mathbf{B} = \nabla \times \mathbf{A}$, it is not hard to verify that

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

will work (I'll work it out, but you could also just plug into a vector identity if you take the time to look it up):

$$\begin{aligned} [\nabla \times (\mathbf{r} \times \mathbf{B})]_i &= \varepsilon_{ijk}\varepsilon_{klm}\partial_j(x_l B_m) \\ &= (\delta_{il}\delta_{jm} - \delta_{im}\delta_{jl})[\delta_{jl}B_m + x_l\partial_j B_m] \\ &= B_i - 3B_i = -2B_i \end{aligned}$$

where I used $\partial_j x_l = \delta_{jl}$, $\delta_{jl}\delta_{lj} = \delta_{jj} = 3$ and $\partial_j B_m = 0$ since \mathbf{B} is uniform. This shows that $\mathbf{B} = (-1/2)[\nabla \times (\mathbf{r} \times \mathbf{B})] = \nabla \times \mathbf{A}$ as claimed. Note also that for this \mathbf{B} we have

$$-2\nabla \cdot \mathbf{A} = \nabla \cdot (\mathbf{r} \times \mathbf{B}) = \varepsilon_{ijk}\partial_i(x_j B_k) = \varepsilon_{ijk}\delta_{ij}B_k = 0$$

because $\varepsilon_{ijk}\delta_{ij} = \varepsilon_{iik} = 0$. Hence $\nabla \cdot \mathbf{A} = 0$.

Before writing out (2.78), let me use this last result to show that

$$(\mathbf{p} \cdot \mathbf{A})\psi = -i\hbar\nabla \cdot (\mathbf{A}\psi) = -i\hbar(\nabla \cdot \mathbf{A})\psi + i\hbar\mathbf{A} \cdot \nabla\psi = (\mathbf{A} \cdot \mathbf{p})\psi$$

and hence $\mathbf{p} \cdot \mathbf{A} = \mathbf{A} \cdot \mathbf{p}$. (Note this shows that $\mathbf{p} \cdot \mathbf{A} = \mathbf{A} \cdot \mathbf{p}$ even if \mathbf{B} is not uniform if we are using the Coulomb gauge $\nabla \cdot \mathbf{A} = 0$.) Now using this, we have

$$\begin{aligned} \frac{1}{2m} \left(\mathbf{p} - \frac{q}{c}\mathbf{A} \right)^2 &= \frac{1}{2m} \left[p^2 - \frac{q}{c}(\mathbf{p} \cdot \mathbf{A} + \mathbf{A} \cdot \mathbf{p}) + \frac{q^2}{c^2}A^2 \right] \\ &= \frac{p^2}{2m} - \frac{q}{mc}\mathbf{A} \cdot \mathbf{p} + \frac{q^2}{2mc^2}A^2. \end{aligned}$$

But (thinking of the scalar triple product as a determinant and switching two rows)

$$\begin{aligned} \frac{q}{mc}\mathbf{A} \cdot \mathbf{p} &= -\frac{q}{2mc}(\mathbf{r} \times \mathbf{B}) \cdot \mathbf{p} = +\frac{q}{2mc}\mathbf{B} \cdot (\mathbf{r} \times \mathbf{p}) \\ &= \frac{q}{2mc}\mathbf{B} \cdot \mathbf{L}. \end{aligned}$$

And using (I'll leave the proof to you)

$$A^2 = \frac{1}{4}(\mathbf{r} \times \mathbf{B}) \cdot (\mathbf{r} \times \mathbf{B}) = \frac{1}{4}[r^2 B^2 - (\mathbf{r} \cdot \mathbf{B})^2]$$

we obtain

$$\frac{1}{2m} \left(\mathbf{p} - \frac{q}{c} \mathbf{A} \right)^2 = \frac{p^2}{2m} - \frac{q}{2mc} \mathbf{B} \cdot \mathbf{L} + \frac{q^2}{8mc^2} [r^2 B^2 - (\mathbf{r} \cdot \mathbf{B})^2].$$

Let's compare the relative magnitudes of the $\mathbf{B} \cdot \mathbf{L}$ term and the quadratic (last) term for an electron. Taking $r^2 \approx a_0^2$ and $L \sim \hbar$, we have

$$\begin{aligned} \frac{(e^2/8mc^2)r^2 B^2}{(e/2mc)\mathbf{B} \cdot \mathbf{L}} &= \frac{(e^2/8mc^2)a_0^2 B^2}{(e/2mc)\hbar B} = \frac{1}{4} \frac{e^2}{\hbar c} \frac{B}{e/a_0^2} \\ &= \frac{1}{4} \frac{1}{137} \frac{B}{(4.8 \times 10^{-10} \text{ esu}) / (0.5 \times 10^{-8} \text{ cm})^2} \\ &= \frac{B}{9 \times 10^9 \text{ gauss}}. \end{aligned}$$

Since magnetic fields in the lab are of order 10^4 gauss or less, we see that the quadratic term is negligible in comparison.

Referring back to (2.38), we see that

$$\frac{q}{2mc} \mathbf{L} = \boldsymbol{\mu}_l$$

where, for an electron, we have $q = -e$. And as we have also seen, for spin we must postulate a magnetic moment of the form

$$\boldsymbol{\mu}_s = g \frac{q}{2mc} \mathbf{S}$$

where $g = 2$ for an electron (and $g = 5.59$ for a proton). Therefore, an electron has a total magnetic moment

$$\boldsymbol{\mu}_{\text{tot}} = -\frac{e}{2m_e c} (\mathbf{L} + 2\mathbf{S})$$

as we stated in (2.68).

Combining our results, the Hamiltonian for a hydrogen atom in a uniform external magnetic field is then given by

$$H = \frac{p^2}{2m_e} - \frac{e^2}{r} - \boldsymbol{\mu}_{\text{tot}} \cdot \mathbf{B} = H^0 - \boldsymbol{\mu}_{\text{tot}} \cdot \mathbf{B} = H^0 + H'$$

where we are taking $q\phi + V(\mathbf{r}) = 0 - e^2/r$, and m_e in this equation is really the reduced mass, which is approximately the same as the electron mass.

3 Time-Dependent Perturbation Theory

3.1 Transitions Between Two Discrete States

We now turn our attention to the situation where the perturbation depends on time. In this situation, we assume that the system is originally in some definite state, and that applying a time-dependent external force then induces a transition to another state. For example, shining electromagnetic radiation on an atom in its ground state will (may) cause it to undergo a transition to a higher energy state. We assume that the external force is weak enough that perturbation theory applies.

There are several ways to deal with this problem, and everyone seems to have their own approach. We shall follow a method that is closely related to the time-independent method that we employed.

To begin, suppose

$$H = H^0 + H'(t)$$

and that we have the orthonormal solutions

$$H^0 \varphi_n = E_n \varphi_n$$

with

$$\varphi_n(t) = \varphi_n e^{-iE_n t/\hbar}.$$

Note that we no longer need to add a superscript 0 to the energies, because with a time-dependent Hamiltonian there is no energy conservation and hence we are not looking for energy corrections.

We would like to solve the time-dependent Schrödinger equation

$$H\psi(t) = [H^0 + H'(t)]\psi(t) = i\hbar \frac{\partial \psi(t)}{\partial t}. \quad (3.1)$$

In this case, the solutions φ_n still form a complete set (they describe every possible state available to the system), the difference being that now the state $\psi(t)$ that results from the perturbation will depend on time. So let us write

$$\psi(t) = \sum_k c_k(t) e^{-iE_k t/\hbar} \varphi_k. \quad (3.2)$$

The reason for this form is that we want the time-dependent coefficients $c_n(t)$ to reduce to constants if $H'(t) = 0$. In other words, so $H'(t) \rightarrow 0$ implies $\psi(t) \rightarrow \varphi(t)$. Our goal is to find the probability that if the system is in an eigenstate $\varphi_i = \psi(0)$ at time $t = 0$, it will be found in the eigenstate φ_f at a later time t . This probability is given by

$$\mathcal{P}_{if}(t) = |\langle \varphi_f | \psi(t) \rangle|^2 = |c_f(t)|^2 \quad (3.3)$$

where $\langle \psi(t) | \psi(t) \rangle = 1$ implies

$$\sum_k |c_k(t)|^2 = 1.$$

Using (3.2) in (3.1) we obtain

$$\sum_k c_k(t) e^{-iE_k t/\hbar} [E_k + H'(t)] \varphi_k = \sum_k i\hbar \left[\dot{c}_k(t) - \frac{iE_k}{\hbar} c_k(t) \right] e^{-iE_k t/\hbar} \varphi_k$$

or

$$i\hbar \sum_k \dot{c}_k(t) e^{-iE_k t/\hbar} \varphi_k = \sum_k H'(t) c_k(t) e^{-iE_k t/\hbar} \varphi_k. \quad (3.4)$$

But $\langle \varphi_n | \varphi_k \rangle = \delta_{nk}$ so that

$$i\hbar \dot{c}_n(t) e^{-iE_n t/\hbar} = \sum_k \langle \varphi_n | H'(t) | \varphi_k \rangle c_k(t) e^{-iE_k t/\hbar}.$$

Defining the **Bohr angular frequency**

$$\omega_{nk} = \frac{E_n - E_k}{\hbar} \quad (3.5)$$

we can write

$$\dot{c}_n(t) = \frac{1}{i\hbar} \sum_k \langle \varphi_n | H'(t) | \varphi_k \rangle c_k(t) e^{i\omega_{nk} t}. \quad (3.6a)$$

This set of equations for $c_n(t)$ is exact and completely equivalent to the original Schrödinger equation (3.1). Defining

$$H'_{nk}(t) = \langle \varphi_n | H'(t) | \varphi_k \rangle$$

we may write out (3.6a) in matrix form as (for a finite number of terms)

$$i\hbar \begin{bmatrix} \dot{c}_1(t) \\ \dot{c}_2(t) \\ \vdots \\ \dot{c}_n(t) \end{bmatrix} = \begin{bmatrix} H'_{11} & H'_{12} e^{i\omega_{12} t} & \dots & H'_{1n} e^{i\omega_{1n} t} \\ H'_{21} e^{i\omega_{21} t} & H'_{22} & \dots & H'_{2n} e^{i\omega_{2n} t} \\ \vdots & \vdots & \ddots & \vdots \\ H'_{n1} e^{i\omega_{n1} t} & H'_{n2} e^{i\omega_{n2} t} & \dots & H'_{nn} \end{bmatrix} \begin{bmatrix} c_1(t) \\ c_2(t) \\ \vdots \\ c_n(t) \end{bmatrix}. \quad (3.6b)$$

As we did in the time-independent case, we now let $H'(t) \rightarrow \lambda H'(t)$, and expand $c_k(t)$ in a power series in λ :

$$c_k(t) = c_k^{(0)}(t) + \lambda c_k^{(1)}(t) + \dots. \quad (3.7)$$

Inserting this into (3.6a) yields

$$\begin{aligned} \dot{c}_n^{(0)}(t) + \lambda \dot{c}_n^{(1)}(t) + \lambda^2 \dot{c}_n^{(2)}(t) + \dots \\ = \frac{1}{i\hbar} \sum_k H'_{nk}(t) [\lambda c_k^{(0)}(t) + \lambda^2 c_k^{(1)}(t) + \lambda^3 c_k^{(2)}(t) + \dots] e^{i\omega_{nk} t}. \end{aligned}$$

Equating powers of λ , for λ^0 we have

$$\dot{c}_n^{(0)}(t) = 0 \quad (3.8a)$$

and for λ^{s+1} with $s \geq 0$ we have

$$\dot{c}_n^{(s+1)}(t) = \frac{1}{i\hbar} \sum_k H'_{nk}(t) c_k^{(s)}(t) e^{i\omega_{nk}t}. \quad (3.8b)$$

In principle, these may be solved successively. Solving (3.8a) gives $c_k^{(0)}(t)$, and using this in (3.8b) then gives $c_n^{(1)}(t)$. Then putting these back into (3.8b) again yields $c_n^{(2)}(t)$, and in principle this can be continued to any desired order.

Let us assume that the system is initially in the state φ_i , so that

$$c_n(0) = \delta_{ni}. \quad (3.9a)$$

Since this must be true for all λ , we have

$$c_n^{(0)}(0) = \delta_{ni} \quad (3.9b)$$

and

$$c_n^{(s)}(0) = 0 \quad \text{for } s \geq 1. \quad (3.9c)$$

From (3.8a) we see that the zeroth-order coefficients are constant in time, so we have

$$c_n^{(0)}(t) = \delta_{ni} \quad (3.9d)$$

and the zeroth-order solutions are completely determined.

Using (3.9b) in (3.8b) we obtain, to first order,

$$\dot{c}_n^{(1)}(t) = \frac{1}{i\hbar} \sum_k H'_{nk}(t) \delta_{ki} e^{i\omega_{nk}t} = \frac{1}{i\hbar} H'_{ni}(t) e^{i\omega_{ni}t}$$

so that

$$c_n^{(1)}(t) = \frac{1}{i\hbar} \int_0^t H'_{ni}(t') e^{i\omega_{ni}t'} dt' \quad (3.10)$$

where the constant of integration is zero by (3.9c). Using (3.9d) and (3.10) in (3.2) yields $\psi(t)$ to first order:

$$\psi(t) = \varphi_i e^{-iE_i t/\hbar} + \lambda \sum_k \left(\frac{1}{i\hbar} \int_0^t H'_{ki}(t') e^{i\omega_{ki}t'/\hbar} dt' \right) e^{-iE_k t} \varphi_k.$$

From (3.3) we know that the transition probability to the state φ_f is given by

$$\mathcal{P}_{if}(t) = |\langle \varphi_f | \psi(t) \rangle|^2 = |c_f(t)|^2$$

where $c_f(t) = c_f^{(0)}(t) + \lambda c_f^{(1)}(t) + \dots$. We will only consider transitions to states φ_f that are distinct from the initial state φ_i , and hence $c_f^{(0)}(t) = 0$. Then the first-order transition probability is

$$\mathcal{P}_{if}(t) = \lambda^2 |c_f^{(1)}(t)|^2$$

or, from (3.10) and letting $\lambda \rightarrow 1$,

$$\mathcal{P}_{if}(t) = \frac{1}{\hbar^2} \left| \int_0^t H'_{fi}(t') e^{i\omega_{fi}t'} dt' \right|^2. \quad (3.11)$$

A minor point is that our initial conditions could equally well be defined at $t \rightarrow -\infty$. In this case, the lower limit on the above integrals would obviously be $-\infty$ rather than 0.

Example 3.1. Consider a one-dimensional harmonic oscillator of a particle of charge q with characteristic frequency ω . Let this oscillator be placed in an electric field that is turned on and off so that its potential energy is given by

$$H'(t) = q\mathcal{E}xe^{-t^2/\tau^2}$$

where τ is a constant. If the particle starts out in its ground state, let us find the probability that it will be in its first excited state after a time $t \gg \tau$.

Since $t \gg \tau$, we may as well take $t \rightarrow \pm\infty$ as limits. From (3.11), we see that we must evaluate the integral

$$I = \int_{-\infty}^{\infty} H'_{10}(t') e^{i\omega_{10}t'} dt'$$

where

$$H'_{10}(t) = q\mathcal{E}e^{-t^2/\tau^2} \langle \psi_1 | x | \psi_0 \rangle$$

and $E_n = \hbar\omega(n + 1/2)$ so that $\omega_{10} = (E_1 - E_0)/\hbar\omega = 1$. Then (keeping ω_{10} for generality at this point)

$$\begin{aligned} I &= q\mathcal{E} \langle \psi_1 | x | \psi_0 \rangle \int_{-\infty}^{\infty} e^{-t^2/\tau^2} e^{i\omega_{10}t} dt \\ &= q\mathcal{E} \langle \psi_1 | x | \psi_0 \rangle \int_{-\infty}^{\infty} e^{-(1/\tau^2)(t^2 - i\omega_{10}\tau^2 t)} dt \\ &= q\mathcal{E} \langle \psi_1 | x | \psi_0 \rangle e^{-\omega_{10}^2\tau^2/4} \int_{-\infty}^{\infty} e^{-(1/\tau^2)(t - i\omega_{10}\tau^2/2)^2} dt \\ &= q\mathcal{E} \langle \psi_1 | x | \psi_0 \rangle e^{-\omega_{10}^2\tau^2/4} \int_{-\infty}^{\infty} e^{-(1/\tau^2)u^2} du \\ &= q\mathcal{E} \langle \psi_1 | x | \psi_0 \rangle e^{-\omega_{10}^2\tau^2/4} \sqrt{\pi\tau^2}. \end{aligned}$$

The easy way to do the spatial integral is to use the harmonic oscillator ladder operators. From

$$x = \sqrt{\frac{\hbar}{2m\omega}} (a + a^\dagger)$$

where

$$a\psi_n = \sqrt{n}\psi_{n-1} \quad \text{and} \quad a^\dagger\psi_n = \sqrt{n+1}\psi_{n+1}$$

we have

$$\langle \psi_1 | x | \psi_0 \rangle = \sqrt{\frac{\hbar}{2m\omega}} \langle \psi_1 | a^\dagger \psi_0 \rangle = \sqrt{\frac{\hbar}{2m\omega}} \langle \psi_1 | \psi_1 \rangle = \sqrt{\frac{\hbar}{2m\omega}}.$$

Therefore

$$I = q\mathcal{E}\tau \sqrt{\frac{\pi\hbar}{2m\omega}} e^{-\omega_{10}^2\tau^2/4}$$

so that

$$\mathcal{P}_{01}(t \rightarrow \infty) = \frac{\pi q^2 \mathcal{E}^2 \tau^2}{2m\hbar\omega} e^{-\omega_{10}^2\tau^2/2} = \frac{\pi q^2 \mathcal{E}^2 \tau^2}{2m\hbar\omega} e^{-\tau^2/2}.$$

Note that as $\tau \rightarrow \infty$ (i.e., the electric field is turned on very slowly), we have $\mathcal{P}_{01} \rightarrow 0$. This shows that the system adjusts “adiabatically” to the field and is not shocked into a transition.

Example 3.2. Let us consider a **harmonic perturbation** of the form

$$H'(t) = V_0(\mathbf{r}) \cos \omega t, \quad t \geq 0.$$

Note that letting $\omega = 0$ we obtain the constant perturbation $H'(t) = V_0(\mathbf{r})$ as a special case. It just isn't much harder to treat the more general situation, which represents the interaction of the system with an electromagnetic wave of frequency ω .

If we define

$$V_{fi} = \langle \varphi_f | V_0(\mathbf{r}) | \varphi_i \rangle,$$

then

$$H'_{fi} = \langle \varphi_f | V_0(\mathbf{r}) \cos \omega t | \varphi_i \rangle = \langle \varphi_f | V_0(\mathbf{r}) | \varphi_i \rangle \cos \omega t = V_{fi} \cos \omega t.$$

Using $\cos \omega t = (e^{i\omega t} + e^{-i\omega t})/2i$, we then have

$$\begin{aligned} \int_0^t H'_{fi}(t') e^{i\omega_{fi}t'} dt' &= \frac{V_{fi}}{2i} \int_0^t (e^{i\omega t'} + e^{-i\omega t'}) e^{i\omega_{fi}t'} dt' \\ &= \frac{V_{fi}}{2i} \int_0^t (e^{i(\omega_{fi}+\omega)t'} + e^{i(\omega_{fi}-\omega)t'}) dt' \\ &= \frac{V_{fi}}{2i} \left[\frac{e^{i(\omega_{fi}+\omega)t} - 1}{i(\omega_{fi} + \omega)} + \frac{e^{i(\omega_{fi}-\omega)t} - 1}{i(\omega_{fi} - \omega)} \right]. \end{aligned}$$

Inserting this into (3.11), we can write

$$\mathcal{P}_{if}(t; \omega) = \frac{|V_{fi}|^2}{4\hbar^2} \left| \frac{1 - e^{i(\omega_{fi}+\omega)t}}{\omega_{fi} + \omega} + \frac{1 - e^{i(\omega_{fi}-\omega)t}}{\omega_{fi} - \omega} \right|^2 \quad (3.12)$$

where I'm specifically including ω as an argument of \mathcal{P}_{if} because the transition probability depends on ω .

Let us consider the special case of a constant (i.e., time-independent) perturbation, $\omega = 0$. In this case, (3.12) reduces to

$$\mathcal{P}_{if}(t; 0) = \frac{|V_{fi}|^2}{\hbar^2 \omega_{fi}^2} |1 - e^{i\omega_{fi}t}|^2 = \frac{|V_{fi}|^2}{\hbar^2 \omega_{fi}^2} 2(1 - \cos \omega_{fi}t).$$

Using the elementary identity

$$\cos A = \cos(A/2 + A/2) = \cos^2 A/2 - \sin^2 A/2 = 1 - 2 \sin^2 A/2$$

we can write the transition probability as

$$\mathcal{P}_{if}(t; 0) = \frac{|V_{fi}|^2}{\hbar^2} \left[\frac{\sin \omega_{fi}t/2}{\omega_{fi}/2} \right]^2 := \frac{|V_{fi}|^2}{\hbar^2} F(t; \omega_{fi}). \quad (3.13)$$

The function

$$F(t; \omega_{fi}) = \left[\frac{\sin \omega_{fi}t/2}{\omega_{fi}/2} \right]^2 = t^2 \left[\frac{\sin \omega_{fi}t/2}{\omega_{fi}t/2} \right]^2$$

has amplitude equal to t^2 , and zeros at $\omega_{fi} = 2\pi n/t$. See Figure 8 below.

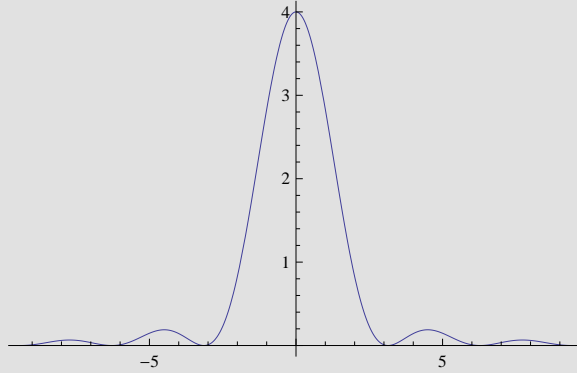


Figure 8: Plot of $F(t; \omega_{fi})$ vs ω_{fi} for $t = 2$.

The main peak lies between zeros at $\pm 2\pi/t$, so its width goes like $1/t$ while its height goes like t^2 , and hence its area grows like t .

It is also interesting to see how the transition probability depends on time.

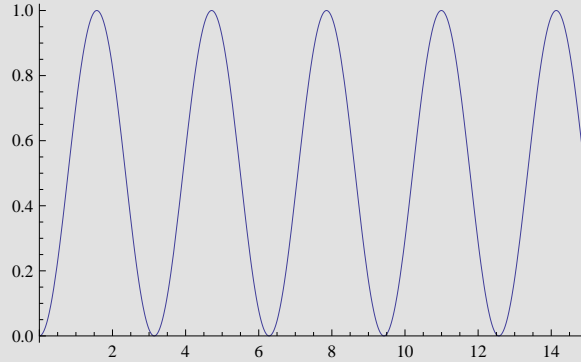


Figure 9: Plot of $F(t; \omega_{fi})$ vs t for $\omega_{fi} = 2$.

Here we see clearly that for times $t = 2\pi n/\omega_{fi}$ the transition probability is zero, and the system is certain to be in its initial state. Because of this oscillatory behavior, the greatest probability for a transition is to allow the perturbation to act only for a short time π/ω_{fi} .

For future reference, let me make a (very un-rigorous but useful) mathematical observation. From Figure 8, we see that as $t \rightarrow \infty$, the function $F(t, \omega) = t^2[(\sin \omega t/2)/(\omega t/2)]^2$ has an amplitude t^2 that also goes to infinity, and a width $4\pi/t$ centered at $\omega = 0$ that goes to zero. Then if we include $F(t, \omega)$ inside the integral of a smooth function $f(\omega)$, the only contribution to the integral will come where $\omega = 0$. Using the well-known result

$$\int_{-\infty}^{\infty} \frac{\sin^2 x}{x^2} dx = \pi$$

we have (with $x = \omega t/2$ so $dx = (t/2)d\omega$)

$$\lim_{t \rightarrow \infty} \int_{-\infty}^{\infty} f(\omega) t^2 \left[\frac{\sin \omega t/2}{\omega t/2} \right]^2 d\omega = 2t f(0) \int_{-\infty}^{\infty} \frac{\sin^2 x}{x^2} dx = 2\pi t f(0)$$

and hence we conclude that

$$F(t; \omega) = \left[\frac{\sin \omega t/2}{\omega/2} \right]^2 = t^2 \left[\frac{\sin \omega t/2}{\omega t/2} \right]^2 \xrightarrow{t \rightarrow \infty} 2\pi t \delta(\omega). \quad (3.14)$$

Example 3.3. Let us take a look at equation (3.12) when $\omega \approx \omega_{fi}$. This is called a **resonance phenomenon**. We will assume that $\omega \geq 0$ by definition, and we will consider the case where $\omega_{fi} > 0$. The alternative case where $\omega_{fi} < 0$ can be treated in an analogous manner.

We begin by rewriting the two complex terms in (3.12). For the first we have

$$\begin{aligned} A_+ &:= \frac{1 - e^{i(\omega_{fi} + \omega)t}}{\omega_{fi} + \omega} = e^{i(\omega_{fi} + \omega)t/2} \left[\frac{e^{-i(\omega_{fi} + \omega)t/2} - e^{i(\omega_{fi} + \omega)t/2}}{\omega_{fi} + \omega} \right] \\ &= -i e^{i(\omega_{fi} + \omega)t/2} \left[\frac{\sin(\omega_{fi} + \omega)t/2}{(\omega_{fi} + \omega)/2} \right] \end{aligned}$$

and similarly for the second

$$A_- := \frac{1 - e^{i(\omega_{fi} - \omega)t}}{\omega_{fi} - \omega} = -i e^{i(\omega_{fi} - \omega)t/2} \left[\frac{\sin(\omega_{fi} - \omega)t/2}{(\omega_{fi} - \omega)/2} \right]$$

If $\omega \approx \omega_{fi}$, then A_- dominates and is called the **resonant term**, while the term A_+ is called the **anti-resonant term**. (These terms would be switched if we were considering the case $\omega_{fi} < 0$.)

We are considering the case where $|\omega - \omega_{fi}| \ll |\omega_{fi}|$, so A_+ can be neglected in comparison to A_- . Under these conditions, (3.12) becomes

$$\begin{aligned} \mathcal{P}_{if}(t; \omega) &= \frac{|V_{fi}|^2}{4\hbar^2} |A_-|^2 = \frac{|V_{fi}|^2}{4\hbar^2} \left[\frac{\sin(\omega_{fi} - \omega)t/2}{(\omega_{fi} - \omega)/2} \right]^2 \\ &:= \frac{|V_{fi}|^2}{4\hbar^2} F(t; \omega_{fi} - \omega). \end{aligned} \quad (3.15)$$

A plot of $F(t; \omega_{fi} - \omega)$ as a function of ω would be identical to Figure 8 except that the peak would be centered over the point $\omega = \omega_{fi}$. In particular, $F(t; \omega_{fi} - \omega)$ has a maximum value of t^2 , and a width between its first two zeros of

$$\Delta\omega = \frac{4\pi}{t}. \quad (3.16)$$

Here is another way to view Example 3.3. Let us consider a time-dependent potential of the form

$$H'(t) = V_0(\mathbf{r}) e^{\pm i\omega t}. \quad (3.17)$$

Then

$$\begin{aligned} \int_0^t H'_{fi}(t') e^{i\omega_{fi}t'} dt' &= V_{fi} \int_0^t e^{i(\omega_{fi} \pm \omega)t'} dt' = V_{fi} \frac{e^{i(\omega_{fi} \pm \omega)t} - 1}{i(\omega_{fi} \pm \omega)} \\ &= V_{fi} e^{i(\omega_{fi} \pm \omega)t/2} \frac{\sin(\omega_{fi} \pm \omega)t/2}{(\omega_{fi} \pm \omega)/2} \end{aligned}$$

and (3.11) becomes

$$\mathcal{P}_{if}(t) = \frac{|V_{fi}|^2}{\hbar^2} \left[\frac{\sin(\omega_{fi} \pm \omega)t/2}{(\omega_{fi} \pm \omega)/2} \right]^2. \quad (3.18)$$

As $t \rightarrow \infty$, we can use (3.14) to write

$$\lim_{t \rightarrow \infty} \mathcal{P}_{if}(t) = \frac{2\pi}{\hbar} |V_{fi}|^2 \delta(E_f - E_i \pm \hbar\omega)t$$

where we used the general result $\delta(ax) = (1/|a|)\delta(x)$ so that $\delta(\omega) = \delta(E/\hbar) = \hbar\delta(E)$. Note that the transition probability grows linearly with time. We can write this as

$$\mathcal{P}_{if}(t \rightarrow \infty) = \Gamma_{i \rightarrow f} t \quad (3.19a)$$

where the **transition rate** (i.e., the transition probability per unit time) is defined by

$$\Gamma_{i \rightarrow f} = \frac{2\pi}{\hbar} |V_{fi}|^2 \delta(E_f - E_i \pm \hbar\omega). \quad (3.19b)$$

(The result (3.19b) differs from (3.15) by a factor of 4 in the denominator. This is because in Example 3.2 we used $\cos \omega t$ which contains the terms $(1/2)e^{\pm i\omega t}$.)

Because of the delta function, we only get transitions in those cases where $|E_f - E_i| = \hbar\omega$, which is simply a statement of energy conservation. Assuming that $E_f > E_i$, in the case of a potential of the form $V_0 e^{+i\omega t}$, we have $E_f = E_i + \hbar\omega$ so the system has *emitted* a quantum of energy. And in the case where we have a potential of the form $V_0 e^{-i\omega t}$, we have $E_f = E_i - \hbar\omega$ so the system has *absorbed* a quantum of energy.

In Example 3.3, we saw that resonance occurs when $\omega = \omega_{fi}$. Since we are considering the case where $\omega_{fi} = (E_f - E_i)/\hbar \geq 0$, this means that resonance is at the point where $E_f = E_i + \hbar\omega$. In other words, a system with energy E_i undergoes a resonant *absorption* of a quantum of energy $\hbar\omega$ to transition to a state with energy E_f . Had we started with the case where $\omega_{fi} < 0$, we would have found that the system underwent a resonant induced *emission* of the same quantum of energy $\hbar\omega$, so that $E_f = E_i - \hbar\omega$.

Also recall that in Example 3.3, we neglected A_+ relative to A_- . Noting that $|A_+(\omega)|^2 = |A_-(-\omega)|^2$, it is easy to see that a plot of $|A_+|^2$ is exactly the same as a plot of $|A_-|^2$ reflected about the vertical axis $\omega = 0$. See Figure 10 below. Note that both of these curves have a width $\Delta\omega = 4\pi/t$ that narrows as time increases.

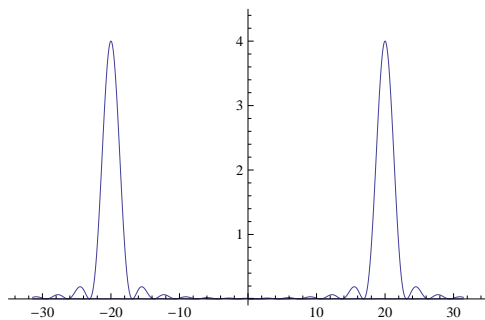


Figure 10: Plot of $|A_+|^2$ and $|A_-|^2$ vs ω for $t = 2$ and $\omega_{fi} = 20$.

In addition, we see that A_+ will be negligible relative to A_- as long as they are well-separated, in other words, as long as

$$2|\omega_{fi}| \gg \Delta\omega.$$

Since $\Delta\omega = 4\pi/t$, this is equivalent to requiring

$$t \gg \frac{1}{|\omega_{fi}|} \approx \frac{1}{\omega}.$$

Physically, this means that the perturbation must act over a long enough time interval t for the system to oscillate enough that it indeed appears sinusoidal.

On the other hand, in both Examples 3.2 and 3.3, the transition probability $\mathcal{P}_{if}(t; \omega)$ has a maximum value proportional to t^2 . Since this approaches infinity as $t \rightarrow \infty$, and since a probability always has to be less than or equal to 1, there is clearly something wrong. One answer is that the first order approximation we are using has a limited time range. In Example 3.3, resonance occurs when $\omega = \omega_{fi}$, in which case

$$\mathcal{P}_{if}(t; \omega = \omega_{fi}) = \frac{|V_{fi}|^2}{4\hbar^2} t^2.$$

So in order for our first-order approximation to be valid, we must have

$$t \ll \frac{\hbar}{|V_{fi}|}.$$

Combining this with the previous paragraph, we conclude that

$$\frac{1}{|\omega_{fi}|} \ll \frac{\hbar}{|V_{fi}|}.$$

This is the same as

$$\hbar|\omega_{fi}| = |E_f - E_i| \gg |V_{fi}| = \langle \varphi_f | V_0 | \varphi_i \rangle$$

and hence the energy difference between the initial and final states must be much larger than the matrix element V_{fi} between these states.

3.2 Transitions to a Continuum of States

In the previous section we considered the transition probability $\mathcal{P}_{if}(t)$ from an initial state φ_i to a final state φ_f . But in the real experimental world, detectors generally observe transitions over a (at least) small range of energies and over a finite range of incident angles. Thus, we should treat not a single final state φ_f , but rather a group (or continuum) of closely spaced states centered about some φ_f . Since the area under the curve in Figure 8 grows like t , we expect that the transition probability to a set of states with approximately the same energy as φ_f to grow linearly with time. (We saw this for a transition to a single state in equation (3.19a).)

Let us now generalize (3.19b) to a more physically realistic detector. After all, no physical transition rate can go like a delta function. To get a good idea of what to expect, we first consider the perturbation (3.17) and the resulting transition probability (3.18).

For a physically realistic detector, instead of a transition to a single final state we must consider all transitions to a group of final states centered about E_f :

$$\begin{aligned}\mathcal{P}(t) &= \sum_{E_f \in \Delta E_f} \frac{|V_{fi}|^2}{\hbar^2} \left[\frac{\sin(\omega_{fi} \pm \omega)t/2}{(\omega_{fi} \pm \omega)/2} \right]^2 \\ &= \sum_{E_f \in \Delta E_f} |V_{fi}|^2 \left[\frac{\sin(E_f - E_i \pm \hbar\omega)t/2\hbar}{(E_f - E_i \pm \hbar\omega)/2} \right]^2\end{aligned}$$

where the sum is over all states with energies in the range ΔE_f . We assume that the final states are very closely spaced, and hence may be treated as a continuum of states. In that case, the sum may be converted to an integral over the interval ΔE_f by writing the number of states with energy between E_f and $E_f + dE_f$ as $\rho(E_f) dE_f$, where $\rho(E_f)$ is called the **density of final states**. It is just the number of states per unit energy. Then

$$\mathcal{P}(t) = \int_{E_f - \Delta E_f/2}^{E_f + \Delta E_f/2} \rho(E_f) dE_f |V_{fi}|^2 \left[\frac{\sin(E_f - E_i \pm \hbar\omega)t/2\hbar}{(E_f - E_i \pm \hbar\omega)/2} \right]^2. \quad (3.20)$$

As t becomes very large, we have seen that the term in brackets becomes sharply peaked about $E_f = E_i \mp \hbar\omega$, and hence we may assume that $\rho(E_f)$ and $|V_{fi}|$ are essentially constant over the region of integration, which we may also let go to $\pm\infty$. Changing variables to $x = (E_f - E_i \pm \hbar\omega)t/2\hbar$ we then have

$$\mathcal{P}(t) = \rho(E_f) |V_{fi}|^2 \frac{2t}{\hbar} \int_{-\infty}^{\infty} \frac{\sin^2 x}{x^2} dx = \frac{2\pi}{\hbar} \rho(E_f) |V_{fi}|^2 t.$$

Defining the **transition rate** $\Gamma = d\mathcal{P}/dt$ we finally arrive at

$$\Gamma = \frac{2\pi}{\hbar} \rho(E_f) |V_{fi}|^2 \Big]_{E_f = E_i \mp \hbar\omega} \quad (3.21)$$

which is called **Fermi's golden rule**.

A completely equivalent way to write this is to take equations (3.19) and write

$$\mathcal{P}(t) = \sum_{\text{final states}} \mathcal{P}_{if}(t) = \sum_{\text{final states}} \Gamma_{i \rightarrow f} t = \Gamma t$$

where

$$\Gamma_{i \rightarrow f} = \frac{2\pi}{\hbar} |V_{fi}|^2 \delta(E_f - E_i \pm \hbar\omega)$$

and

$$\Gamma = \sum_{\text{final states}} \Gamma_{i \rightarrow f}.$$

If you wish, you can then replace the sum over states by an integral over energies if you include a density of states factor $\rho(E)$. This has the same effect as simply using (3.14) in (3.20) to write

$$\begin{aligned}\mathcal{P}(t) &= \int_{E_f - \Delta E_f/2}^{E_f + \Delta E_f/2} \rho(E_f) dE_f |V_{fi}|^2 \frac{2\pi}{\hbar} t \delta(E_f - E_i \pm \hbar\omega) \\ &= \left\{ \frac{2\pi}{\hbar} |V_{fi}|^2 \int_{E_f - \Delta E_f/2}^{E_f + \Delta E_f/2} \rho(E_f) \delta(E_f - E_i \pm \hbar\omega) dE_f \right\} t \\ &= \Gamma t.\end{aligned}$$

Example 3.4. Let us consider a simple, one-dimensional model of photo-ionization, in which a particle of charge e in its ground state ψ_0 in a potential $U(x)$ is irradiated by light of frequency ω , and hence is ejected into the continuum.

To keep things simple, we first assume that the wavelength of the incident light is much longer than atomic dimensions. Under these conditions, the electric field of the light may be considered uniform in space, but harmonic in time. (The magnetic field of the light exerts a force that is of order v/c less than the electric force, and may be neglected.) Since we are treating the absorption of energy, we write the electric field as $\mathbf{E} = \mathcal{E}e^{-i\omega t}\hat{\mathbf{x}}$. Using $\mathbf{E} = -\nabla\varphi$ we have

$$\int \mathbf{E} \cdot d\mathbf{x} = \mathcal{E}e^{-i\omega t} \int dx = \mathcal{E}e^{-i\omega t}x = - \int \nabla\varphi \cdot d\mathbf{x} = -\varphi(x)$$

so that $\varphi(x) = -\mathcal{E}e^{-i\omega t}x$. From Example 2.2 we know that the interaction energy of the particle in the electric field is given by $e\varphi(x)$, and hence the perturbation is

$$H'(x, t) = -e\mathcal{E}xe^{-i\omega t} = V_0(x)e^{-i\omega t}.$$

The second assumption we shall make is that the frequency ω is large enough that the final state energy E_f is very large compared to $U(x)$, and therefore we may treat the final state of the ejected particle as a plane wave (i.e., a free particle of definite energy and momentum).

We need to find the density of final states and the normalization of these states. The standard trick to accomplishing this is to consider our system to be in a box of length L , and then letting $L \rightarrow \infty$. By a proper choice of boundary conditions, this will give us a discrete set of normalizable states. However, we can't treat this like a "particle in a box," because such states must vanish at the walls, and a state of definite momentum can't vanish. Therefore, we employ the mathematical (but non-physical) trick of assuming **periodic boundary conditions**, whereby the walls are taken to lie at x_0 and $x_0 + L$ together with $\psi(x_0 + L) = \psi(x_0)$.

The free particle plane waves are of the form $e^{ipx/\hbar}$, so our periodic boundary conditions become

$$e^{ip(x_0+L)/\hbar} = e^{ipx_0}$$

so that $e^{ipL/\hbar} = 1$ and hence

$$p = \sqrt{2mE} = \frac{2\pi n\hbar}{L}; \quad n = 0, \pm 1, \pm 2, \dots$$

This shows that the momentum (and hence energy) of the particle takes on discrete values. Note that as L gets larger and larger, the spacing of the states becomes closer and closer, and in the limit $L \rightarrow \infty$ they become the usual free particle continuum states of definite momentum. This is the justification for using periodic boundary conditions. Finally, the normalization condition $\int_{x_0}^{x_0+L} |\psi|^2 dx = 1$ implies that the normalized wave functions are then

$$\psi_E = \frac{1}{\sqrt{L}} e^{i\sqrt{2mE}x/\hbar}.$$

The next thing we need to do is find the density of states $\rho(E)$, which is defined as the number of states with an energy between E and $E + dE$, i.e., $\rho(E) = dN/dE$. Consider a state with energy E defined by

$$\sqrt{2mE} = \frac{2\pi N\hbar}{L}$$

so that

$$N = \frac{L}{2\pi\hbar} \sqrt{2mE}.$$

From $n = 0, \pm 1, \pm 2, \dots, \pm N$, we see that there are $2N + 1$ states with energy less than or equal to E . Calling this number $N(E)$, we have

$$N(E) = 2N + 1 = \frac{L}{\pi\hbar} \sqrt{2mE} + 1.$$

But then

$$\begin{aligned} N(E + dE) &= \frac{L}{\pi\hbar} \sqrt{2m(E + dE)} + 1 = \frac{L}{\pi\hbar} \sqrt{2mE} \sqrt{1 + dE/E} + 1 \\ &\approx \frac{L}{\pi\hbar} \sqrt{2mE} (1 + dE/2E) + 1 = N(E) + \frac{L}{2\pi\hbar} \sqrt{\frac{2m}{E}} dE \end{aligned}$$

and hence

$$dN = N(E + dE) - N(E) = \frac{L}{2\pi\hbar} \sqrt{\frac{2m}{E}} dE.$$

Directly from the definition of $\rho(E)$ we then have

$$\rho(E) = \frac{L}{2\pi\hbar} \sqrt{\frac{2m}{E}}. \quad (3.22)$$

Now we turn to the matrix element V_{fi} . The initial state is the normalized wave function ψ_0 with energy $E_0 = -\epsilon$ where ϵ is the binding energy. The final state is the normalized free particle state ψ_{E_f} with energy $E_f = E_0 + \hbar\omega = \hbar\omega - \epsilon$. Then

$$V_{fi} = -\mathcal{E} \langle \psi_{E_f} | ex | \psi_0 \rangle = -\mathcal{E} \int \frac{1}{\sqrt{L}} e^{-i\sqrt{2mE_f}x/\hbar} ex \psi_0 dx.$$

Note that this is the quantum mechanical average of the energy of an electric dipole in a uniform electric field \mathcal{E} .

Putting all of this together in (3.21), we have the transition probability

$$\begin{aligned}\Gamma &= \frac{2\pi}{\hbar} \frac{L}{2\pi\hbar} \sqrt{\frac{2m}{E_f}} e^2 \mathcal{E}^2 \frac{1}{L} \left| \int e^{-i\sqrt{2mE_f}x/\hbar} \psi_0 dx \right|^2 \\ &= \frac{e^2 \mathcal{E}^2}{\hbar^2} \sqrt{\frac{2m}{E_f}} \left| \int e^{-i\sqrt{2mE_f}x/\hbar} \psi_0 dx \right|^2.\end{aligned}\quad (3.23)$$

Note that the box size L has canceled out of the final result, as it must.

Let's actually evaluate the integral in (3.23) for the specific example of a particle in a square well potential. Recall that the solutions to this problem consist of sines and cosines inside the well, and exponentially decaying solutions outside. To simplify the calculation, we assume first that the well is so narrow that the ground state is the only bound state (a cosine wave function), and second, that this state is only very slightly bound, so that its wave function extends far beyond the edges of the well. By making the well so narrow, we can simply replace the cosine wave function inside the well by extending the exponential wave functions back to the origin.

With these additional simplifications, the normalized ground state wave function is

$$\psi_0 = \left(\frac{2m\epsilon}{\hbar^2} \right)^{1/4} e^{-\sqrt{2m\epsilon}|x|/\hbar}$$

where ϵ is the binding energy. Then the integral in (3.23) becomes

$$\begin{aligned}\int_{-\infty}^{\infty} e^{-i\sqrt{2mE_f}x/\hbar} \psi_0 dx &= \left(\frac{2m\epsilon}{\hbar^2} \right)^{1/4} \int_{-\infty}^{\infty} e^{-\sqrt{2m}(\sqrt{\epsilon}|x|+i\sqrt{E_f}x)/\hbar} dx \\ &= \left(\frac{2m\epsilon}{\hbar^2} \right)^{1/4} \left\{ \int_{-\infty}^0 e^{\sqrt{2m}(\sqrt{\epsilon}-i\sqrt{E_f})x/\hbar} dx \right. \\ &\quad \left. + \int_0^{\infty} e^{-\sqrt{2m}(\sqrt{\epsilon}+i\sqrt{E_f})x/\hbar} dx \right\}.\end{aligned}$$

Using

$$\int_{-\infty}^0 e^{ax} dx = \frac{\partial}{\partial a} \int_{-\infty}^0 e^{ax} dx = \frac{\partial}{\partial a} \frac{1}{a} = -\frac{1}{a^2}$$

and

$$\int_0^{\infty} e^{-bx} dx = -\frac{\partial}{\partial b} \int_0^{\infty} e^{-bx} dx = -\frac{\partial}{\partial b} \frac{1}{b} = \frac{1}{b^2}$$

we have

$$\int_{-\infty}^{\infty} e^{-i\sqrt{2mE_f}x/\hbar} \psi_0 dx = \left(\frac{2m\epsilon}{\hbar^2} \right)^{1/4} \frac{\hbar^2}{2m} \left\{ \frac{1}{(\sqrt{\epsilon}+i\sqrt{E_f})^2} - \frac{1}{(\sqrt{\epsilon}-i\sqrt{E_f})^2} \right\}$$

$$= \left(\frac{2m\epsilon}{\hbar^2} \right)^{1/4} \frac{\hbar^2}{2m} \frac{(-4i)\sqrt{\epsilon E_f}}{(\epsilon + E_f)^2}.$$

Hence equation (3.23) becomes

$$\Gamma = \frac{8\hbar e^2 \mathcal{E}^2}{m} \frac{\epsilon^{3/2} E_f^{1/2}}{(\epsilon + E_f)^4}$$

where $E_f = \hbar\omega - \epsilon$, or $\epsilon + E_f = \hbar\omega$. Since our second initial assumption was essentially that $\hbar\omega \gg \epsilon$, we can replace E_f in the numerator by $\hbar\omega$, leaving us with the final result

$$\Gamma = \frac{8e^2 \mathcal{E}^2}{m\hbar^5/2} \frac{\epsilon^{3/2}}{\omega^{7/2}}.$$

What this means is that if we have a collection of \mathcal{N} particles of charge e and mass m in their ground state in a potential well with binding energy ϵ , and they are placed in an electromagnetic wave of frequency ω and electric vector \mathcal{E} , then the number of photoelectrons with energy $\hbar\omega - \epsilon$ produced per second is $\mathcal{N}\Gamma$.

Now that we have an idea of what the density of states means and how to use the golden rule, let us consider a somewhat more general three-dimensional problem. We will consider an atomic decay $\varphi_i \rightarrow \varphi_f$, with the emission of a particle (photon, electron etc.), whose detection is far from the atom, and hence may be described by a plane wave

$$\psi(\mathbf{r}, t) = \frac{1}{\sqrt{V}} e^{i(\mathbf{p}\cdot\mathbf{r} - \omega_{\mathbf{p}}t)}.$$

(At the end of our derivation, we will generalize to multiple particles in the final state.) Here V is the volume of a box that contains the entire system, and the factor $1/\sqrt{V}$ is necessary to normalize the wave function. If we take the box to be very large, its shape doesn't matter, so we take it to be a cube of side L . In order to determine the allowed momenta, we impose periodic boundary conditions:

$$\psi(x + L, y, z) = \psi(x, y, z)$$

and similarly for y and z . Then $e^{ip_x L/\hbar} = e^{ip_y L/\hbar} = e^{ip_z L/\hbar} = 1$ so that we must have

$$p_x = \frac{2\pi\hbar}{L} n_x; \quad p_y = \frac{2\pi\hbar}{L} n_y; \quad p_z = \frac{2\pi\hbar}{L} n_z$$

where each $n_i = 0, \pm 1, \pm 2, \dots$.

Our real detector will measure all incoming momenta in a range \mathbf{p} to $\mathbf{p} + \delta\mathbf{p}$, and hence we want to calculate the transition rate to all final states in this range. Thus we want

$$\Gamma = \sum_{\delta\mathbf{p}} \Gamma_{i \rightarrow f}(\mathbf{p})$$

where $\Gamma_{i \rightarrow f}(\mathbf{p})$ is given by (3.19b). Since each momentum state is described by the triple of integers (n_x, n_y, n_z) , this is equivalent to the sum

$$\Gamma = \sum_{\delta n_x, \delta n_y, \delta n_z} \Gamma_{i \rightarrow f}(\mathbf{n}) \rightarrow \int d^3 n \Gamma_{i \rightarrow f}(\mathbf{n})$$

where we have gone over to an integral in the limit of a very large box, so that compared to L , each δn_i becomes an infinitesimal dn_i . Noting that

$$d^3 n = dn_x dn_y dn_z = \left(\frac{L}{2\pi\hbar} \right)^3 dp_x dp_y dp_z = \frac{V}{(2\pi\hbar)^3} d^3 p \quad (3.24)$$

we then have (from (3.19b))

$$\Gamma = \frac{2\pi}{\hbar} \int \frac{V}{(2\pi\hbar)^3} d^3 p |M_{fi}|^2 \delta(E_f - E_i + E) \quad (3.25)$$

where we have assumed that the emitted particle has energy E (which is essentially the integration variable), and we changed notation slightly to $|M_{fi}|^2 = |\langle \varphi_f | H'(t) | \varphi_i \rangle|^2$ where $H'(t) = V_0(\mathbf{r})e^{+i\omega t}$ as in (3.17).

If we let $d\Omega_{\mathbf{p}} = d \cos \theta_{\mathbf{p}} d\phi_{\mathbf{p}}$ be the element of solid angle about the direction defined by \mathbf{p} , then

$$\begin{aligned} \Gamma &= \frac{2\pi}{\hbar} \int d\Omega_{\mathbf{p}} \int \frac{V}{(2\pi\hbar)^3} p^2 dp |M_{fi}|^2 \delta(E_f - E_i + E) \\ &= \frac{2\pi}{\hbar} \int d\Omega_{\mathbf{p}} \int \frac{V}{(2\pi\hbar)^3} dE \left(\frac{p^2 dp}{dE} \right) |M_{fi}|^2 \delta(E_f - E_i + E) \\ &= \frac{2\pi}{\hbar} \int d\Omega_{\mathbf{p}} \frac{V}{(2\pi\hbar)^3} \left[\left(\frac{p^2 dp}{dE} \right) |M_{fi}|^2 \right]_{E=E_i-E_f}. \end{aligned} \quad (3.26)$$

Here the integral is over $\Omega_{\mathbf{p}}$, and is to cover whatever solid angle range we wish to include. This could be just a small detector angle, or as large as 4π to include all emitted particles. The quantity in brackets is evaluated at $E = E_i - E_f$ as required by the energy conserving delta function. And the factor of V in the numerator will be canceled by the normalization factor $(1/\sqrt{V})^2$ coming from $|M_{fi}|^2$ and due to the outgoing plane wave particle.

From (3.24) we see that

$$d\Omega_{\mathbf{p}} \frac{V}{(2\pi\hbar)^3} \left(\frac{p^2 dp}{dE} \right) = \frac{V}{(2\pi\hbar)^3} \frac{d^3 p}{dE} = \frac{d^3 n}{dE} := \rho(E) \quad (3.27)$$

where the **density of states** $\rho(E)$ is defined as the number of states per unit of energy. Note that in the case of a photon (i.e., a massless particle) we have $E = pc$ so that

$$\frac{p^2 dp}{dE} = \frac{p^2}{c} = \frac{E^2}{c^3} = \frac{\hbar^2}{c^3} \omega^2$$

where we used the alternative relation $E = \hbar\omega$. And in the case of a massive particle, we have $E = p^2/2m$ and

$$\frac{p^2 dp}{dE} = p^2 \frac{m}{p} = mp = m\sqrt{2mE}.$$

You should compare (3.27) using these results to (3.22). In all cases, the density of states goes like $1/E$ as it should.

In terms of the density of states, (3.26) may be written

$$\Gamma = \frac{2\pi}{\hbar} \rho(E) |M_{fi}|^2 \Big]_{E=E_i-E_f}. \quad (3.28)$$

This is the golden rule for the emission of a particle of energy E . If the final state contains several particles labeled by k , then (3.25) becomes

$$\Gamma = \frac{2\pi}{\hbar} \int_{\text{indep } p_k} \prod_k \frac{V d^3 p_k}{(2\pi\hbar)^3} |M_{fi}|^2 \delta\left(E_f - E_i + \sum_k E_k\right)$$

where the integral is over all *independent* momenta, since the energy conserving delta function is a condition on the total momenta of the emitted particles, and hence eliminates a degree of freedom. However, the product of **phase space** factors $V d^3 p_k / (2\pi\hbar)^3$ is over *all* particles in the final state. Alternatively, we may leave the integral over *all* momenta if we include an energy conserving delta function in addition:

$$\Gamma = \frac{2\pi}{\hbar} \int \prod_k \frac{V d^3 p_k}{(2\pi\hbar)^3} |M_{fi}|^2 \delta\left(E_f + \sum_k E_k - E_i\right) \delta\left(\mathbf{p}_f + \sum_k \mathbf{p}_k - \mathbf{p}_i\right).$$