

Chapter 4. Some Important Tools of Theory

For all but the most elementary problems, many of which serve as fundamental approximations to the real behavior of molecules (e.g., the Hydrogenic atom, the harmonic oscillator, the rigid rotor, particles in boxes), the Schrödinger equation can not be solved exactly. It is therefore extremely useful to have tools that allow one to approach these insoluble problems by solving other Schrödinger equations that can be trusted to reasonably describe the solutions of the impossible problem. The approaches discussed in this Chapter are the most important tools of this type.

4.1. Perturbation Theory

In most practical applications of quantum mechanics to molecular problems, one is faced with the harsh reality that the Schrödinger equation pertinent to the problem at hand cannot be solved exactly. To illustrate how desperate this situation is, I note that neither of the following two Schrödinger equations has ever been solved exactly (meaning analytically):

1. The Schrödinger equation for the two electrons moving about the He nucleus:

$$[-\hbar^2/2m_e \nabla_1^2 - \hbar^2/2m_e \nabla_2^2 - 2e^2/r_1 - 2e^2/r_2 + e^2/r_{1,2}] \psi = E \psi,$$

2. The Schrödinger equation for the two electrons moving in an H₂ molecule even if the locations of the two nuclei (labeled A and B) are held clamped as in the Born-Oppenheimer approximation:

$$[-\hbar^2/2m_e \nabla_1^2 - \hbar^2/2m_e \nabla_2^2 - e^2/r_{1,A} - e^2/r_{2,A} - e^2/r_{1,B} - e^2/r_{2,B} + e^2/r_{1,2}] \psi = E \psi.$$

These two problems are examples of what is called the “three-body problem” meaning solving for the behavior of three bodies moving relative to one another. Motions of the sun, earth, and moon (even neglecting all the other planets and their moons) constitute

another three-body problem. None of these problems, even the classical Newton's equation for the sun, earth, and moon, have ever been solved exactly. So, what does one do when faced with trying to study real molecules using quantum mechanics?

There are two very powerful tools that one can use to “sneak up” on the solutions to the desired equations by first solving an easier model problem and then using the solutions to this problem to approximate the solutions to the real Schrödinger problem of interest. For example, to solve for the energies and wave functions of a boron atom, one could use hydrogenic 1s orbitals (but with $Z = 5$) and hydrogenic 2s and 2p orbitals (with $Z = 3$ to account for the screening of the full nuclear charge by the two 1s electrons) as a starting point. To solve for the vibrational energies of a diatomic molecule whose energy vs. bond length $E(R)$ is known, one could use the Morse oscillator wave functions and energies as starting points. But, once one has decided on a reasonable model to use, how does one connect this model to the real system of interest? Perturbation theory and the variational method are the two tools that are most commonly used for this purpose, and it is these two tools that are covered in this Chapter.

The perturbation theory approach provides a set of analytical expressions for generating a sequence of approximations to the true energy E and true wave function ψ . This set of equations is generated, for the most commonly employed perturbation method, Rayleigh-Schrödinger perturbation theory (RSPT), as follows. First, one decomposes the true Hamiltonian H into a so-called zeroth-order part H^0 (this is the Hamiltonian of the model problem used to represent the real system) and the difference $(H-H^0)$, which is called the perturbation and usually denoted V :

$$H = H^0 + V.$$

It is common to associate with the perturbation V a strength parameter λ , which could, for example, be associated with the strength of the electric field when the perturbation results from the interaction of the molecule of interest with an electric field. In such cases, it is usual to write the decomposition of H as

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

A fundamental assumption of perturbation theory is that the wave functions and energies for the full Hamiltonian H can be expanded in a Taylor series involving various powers of the perturbation parameter λ . Hence, one writes the energy E and the wave function ψ as zeroth-, first-, second-, etc, order pieces which form the unknowns in this method:

$$E = E^0 + E^1 + E^2 + E^3 + \dots$$

$$\psi = \psi^0 + \psi^1 + \psi^2 + \psi^3 + \dots$$

with E^n and ψ^n being proportional to λ^n . Next, one substitutes these expansions of E , H and ψ into $H\psi = E\psi$. This produces one equation whose right and left hand sides both contain terms of various “powers” in the perturbation λ . For example, terms of the form $E^1 \psi^2$ and $V \psi^2$ and $E^0 \psi^3$ are all of third power (also called third order). Next, one equates the terms on the left and right sides that are of the same order. This produces a set of equations, each containing all the terms of a given order. The zeroth, first, and second-order such equations are given below:

$$H^0 \psi^0 = E^0 \psi^0,$$

$$H^0 \psi^1 + V \psi^0 = E^0 \psi^1 + E^1 \psi^0$$

$$H^0 \psi^2 + V \psi^1 = E^0 \psi^2 + E^1 \psi^1 + E^2 \psi^0.$$

It is straightforward to see that the n^{th} order expression in this sequence of equations can be written as

$$H^0 \psi^n + V \psi^{n-1} = E^0 \psi^n + E^1 \psi^{n-1} + E^2 \psi^{n-2} + E^3 \psi^{n-3} + \dots + E^n \psi^0.$$

The zeroth-order equation simply instructs us to solve the model Schrödinger equation to obtain the zeroth-order wave function ψ^0 and its zeroth-order energy E^0 . Since H^0 is a Hermitian operator, it has a complete set of such eigenfunctions, which we label $\{\psi_k^0\}$ and $\{E_k^0\}$. One of these states will be the one we are interested in studying (e.g., we might be interested in the effect of an external electric field on the 2s state of the hydrogen atom), but, as will become clear soon, we actually have to find the full set of $\{\psi_k^0\}$ and $\{E_k^0\}$ (e.g., we need to also find the 1s, 2p, 3s, 3p, 3d, etc. states of the hydrogen atom when studying the electric field's effect on the 2s state).

In the first-order equation, the unknowns are ψ^1 and E^1 (recall that V is assumed to be known because it is the difference between the Hamiltonian one wants to solve and the model Hamiltonian H^0). To solve the first-order and higher-order equations, one expands each of the corrections to the wave function ψ of interest in terms of the complete set of wave functions of the zeroth-order problem $\{\psi_j^0\}$. As noted earlier, this means that one must solve $H^0 \psi_j^0 = E_j^0 \psi_j^0$ not just for the zeroth-order state one is interested in (denoted ψ^0 above) but for all of the other zeroth-order states $\{\psi_j^0\}$. For example, expanding ψ^1 in this manner gives:

$$\psi^1 = \sum_j C_j^1 \psi_j^0$$

Now, the unknowns in the first-order equation become E^1 and the C_j^1 expansion coefficients. To solve $H^0 \psi^1 + V \psi^0 = E^0 \psi^1 + E^1 \psi^0$, one proceeds as follows:

1. First, one multiplies this equation on the left by the complex conjugate of the zeroth-order function for the state of interest ψ^0 and integrates over the variables on which the wave functions depend. This gives

$$\langle \psi^0 | H^0 | \psi^1 \rangle + \langle \psi^0 | V | \psi^0 \rangle = E^0 \langle \psi^0 | \psi^1 \rangle + E^1 \langle \psi^0 | \psi^0 \rangle.$$

The first and third terms cancel one another because $H^0 \psi^0 = E^0 \psi^0$, and the fourth term reduces to E^1 because ψ^0 is assumed to be normalized. This allows the above equation to be written as

$$E^1 = \langle \psi^0 | V | \psi^0 \rangle$$

which is the RSPT expression for E^1 . It says the first-order correction to the energy E^0 of the unperturbed state can be evaluated by computing the average value of the perturbation with respect to the unperturbed wave function ψ^0 .

2. Returning to the first-order equation and multiplying on the left by the complex conjugate of one of the other zeroth-order functions ψ_j^0 gives

$$\langle \psi_j^0 | H^0 | \psi^1 \rangle + \langle \psi_j^0 | V | \psi^0 \rangle = E^0 \langle \psi_j^0 | \psi^1 \rangle + E^1 \langle \psi_j^0 | \psi^0 \rangle.$$

Using $H^0 \psi_j^0 = E_j^0 \psi_j^0$, the first term reduces to $E_j^0 \langle \psi_j^0 | \psi^1 \rangle$, and the fourth term vanishes because ψ_j^0 is orthogonal to ψ^0 because these two functions are different eigenfunctions of H^0 . This reduces the equation to

$$E_j^0 \langle \psi_j^0 | \psi^1 \rangle + \langle \psi_j^0 | V | \psi^0 \rangle = E^0 \langle \psi_j^0 | \psi^1 \rangle$$

The unknown in this expression is $\langle \psi_j^0 | \psi^1 \rangle$, which is the expansion coefficient C_j^1 for the expansion of ψ^1 in terms of the zeroth-order functions $\{\psi_j^0\}$. In RSPT, one assumes that the only contribution of ψ^0 to the full wave function ψ occurs in zeroth-order; this is referred to as assuming intermediate normalization of ψ . In other words, $\langle \psi^0 | \psi \rangle = 1$ because $\langle \psi^0 | \psi^0 \rangle = 1$ and $\langle \psi^0 | \psi^n \rangle = 0$ for $n = 1, 2, 3, \dots$. So, the coefficients $\langle \psi_j^0 | \psi^1 \rangle$ appearing in the above equation are all one needs to describe ψ^1 .

3. If the state of interest ψ^0 is non-degenerate in zeroth-order (i.e., none of the other E_j^0 is equal to E^0), this equation can be solved for the needed expansion coefficients

$$\langle \psi_j^0 | \psi^1 \rangle = \frac{\langle \psi_j^0 | V | \psi^0 \rangle}{E^0 - E_j^0}$$

which allow the first-order wave function to be written as

$$\psi^1 = \sum_j \psi_j^0 \frac{\langle \psi_j^0 | V | \psi^0 \rangle}{E^0 - E_j^0}$$

where the index J is restricted such that ψ_j^0 not equal the state ψ^0 you are interested in.

4. However, if one or more of the zeroth-order energies E_j^0 is equal to E^0 , an additional step needs to be taken before the above expression for ψ^1 can be used. If one were to try to solve $E_j^0 \langle \psi_j^0 | \psi^1 \rangle + \langle \psi_j^0 | V | \psi^0 \rangle = E^0 \langle \psi_j^0 | \psi^1 \rangle$ without taking this extra step, the $\langle \psi_j^0 | \psi^1 \rangle$ values for those states with $E_j^0 = E^0$ could not be determined because the first and third terms would cancel and the equation would read $\langle \psi_j^0 | V | \psi^0 \rangle = 0$. The way RSPT deals with this paradox is realize that, within a set of N degenerate states, any N orthogonal combinations of these states will also be degenerate. So RSPT assumes that one has already chosen the degenerate sets of zeroth-order states to make $\langle \psi_j^0 | V | \psi_k^0 \rangle = 0$ for $K \neq J$. This extra step is carried out in practice by forming the matrix representation of V in the original set of degenerate zeroth-order states and then finding the unitary transformation among these states that diagonalizes this matrix. These transformed states are then what one uses as ψ_j^0 and ψ^0 in the RSPT expressions. This means that the paradoxical result $\langle \psi_j^0 | V | \psi^0 \rangle = 0$ is indeed obeyed by this choice of states, so one does not need to determine the coefficients $\langle \psi_j^0 | \psi^1 \rangle$ for ψ_j^0 belonging to the degenerate zeroth-order states (i.e., these coefficients can be assumed to be zero). The bottom line is that the expression

$$\psi^1 = \sum_j \psi_j^0 \frac{\langle \psi_j^0 | V | \psi^0 \rangle}{E^0 - E_j^0}$$

remains valid, but the summation index J is now restricted to exclude any members of the zeroth-order states that are degenerate with ψ^0 .

To obtain the expression for the second-order correction to the energy of the state of interest, one returns to

$$H^0 \psi^2 + V \psi^1 = E^0 \psi^2 + E^1 \psi^1 + E^2 \psi^0$$

Multiplying on the left by the complex conjugate of ψ^0 and integrating yields

$$\langle \psi^0 | H^0 | \psi^2 \rangle + \langle \psi^0 | V | \psi^1 \rangle = E^0 \langle \psi^0 | \psi^2 \rangle + E^1 \langle \psi^0 | \psi^1 \rangle + E^2 \langle \psi^0 | \psi^0 \rangle.$$

The intermediate normalization condition causes the fourth term to vanish, and the first and third terms cancel one another. Recalling the fact that ψ^0 is normalized, the above equation reduces to

$$\langle \psi^0 | V | \psi^1 \rangle = E^2.$$

Substituting the expression obtained earlier for ψ^1 allows E^2 to be written as

$$E^2 = \sum_J \frac{|\langle \psi^0 | V | \psi_J^0 \rangle|^2}{E^0 - E_J^0}$$

where, as before, the sum over J is limited to states that are not degenerate with ψ^0 in zeroth-order.

These are the fundamental working equations of Rayleigh-Schrödinger perturbation theory. They instruct us to compute the average value of the perturbation taken over a probability distribution equal to $\psi^{0*} \psi^0$ to obtain the first-order correction to the energy E^1 . They also tell us how to compute the first-order correction to the wave function and the second-order energy in terms of integrals $\langle \psi^0 | V | \psi_J^0 \rangle$ coupling ψ^0 to other zeroth-order states and denominators involving energy differences $E^0 - E_J^0$.

An analogous approach is used to solve the second- and higher-order equations. For example, the equation for the n^{th} order energy and wave functions reads:

$$H^0 \psi^n + V \psi^{n-1} = E^0 \psi^n + E^1 \psi^{n-1} + E^2 \psi^{n-2} + E^3 \psi^{n-3} + \dots + E^n \psi^0$$

The n^{th} order energy is obtained by multiplying this equation on the left by ψ^{0*} and integrating over the relevant coordinates (and using the fact that ψ^0 is normalized and the intermediate normalization condition $\langle \psi^0 | \psi^m \rangle = 0$ for all $m > 0$):

$$\langle \psi^0 | V | \psi^{n-1} \rangle = E^n.$$

This allows one to recursively solve for higher and higher energy corrections once the various lower-order wave functions ψ^{n-1} are obtained. To obtain the expansion coefficients for the ψ^n expanded in terms of the zeroth-order states $\{\psi_j^0\}$, one multiplies the above n^{th} order equation on the left by ψ_j^0 (one of the zeroth-order states not equal to the state ψ^0 of interest) and obtains

$$E_j^0 \langle \psi_j^0 | \psi^n \rangle + \langle \psi_j^0 | V | \psi^{n-1} \rangle = E^0 \langle \psi_j^0 | \psi^n \rangle + E^1 \langle \psi_j^0 | \psi^{n-1} \rangle \\ + E^2 \langle \psi_j^0 | \psi^{n-2} \rangle + E^3 \langle \psi_j^0 | \psi^{n-3} \rangle + \dots + E^n \langle \psi_j^0 | \psi^0 \rangle.$$

The last term on the right-hand side vanishes because ψ_j^0 and ψ^0 are orthogonal. The terms containing the n^{th} order expansion coefficients $\langle \psi_j^0 | \psi^n \rangle$ can be brought to the left-hand side to produce the following equation for these unknowns:

$$E_j^0 \langle \psi_j^0 | \psi^n \rangle - E^0 \langle \psi_j^0 | \psi^n \rangle = - \langle \psi_j^0 | V | \psi^{n-1} \rangle + E^1 \langle \psi_j^0 | \psi^{n-1} \rangle \\ + E^2 \langle \psi_j^0 | \psi^{n-2} \rangle + E^3 \langle \psi_j^0 | \psi^{n-3} \rangle + \dots + E^n \langle \psi_j^0 | \psi^0 \rangle.$$

As long as the zeroth-order energy E_j^0 is not degenerate with E^0 (or, that the zeroth-order states have been chosen as discussed earlier to cause there to no contribution to ψ^n from such degenerate states), the above equation can be solved for the expansion coefficients $\langle \psi_j^0 | \psi^n \rangle$, which then define ψ^n .

The RSPT equations can be solved recursively to obtain even high-order energy and wave function corrections:

1. ψ^0 and E^0 and V are used to determine E^1 and ψ^1 as outlined above,
2. E^2 is determined from $\langle \psi^0 | V | \psi^{n-1} \rangle = E^n$ with $n = 2$, and the expansion coefficients of ψ^2 $\{ \langle \psi_j^0 | \psi^2 \rangle \}$ are determined from the above equation with $n = 2$,
3. E^3 (and higher E^n) are then determined from $\langle \psi^0 | V | \psi^{n-1} \rangle = E^n$ and the expansion coefficients of ψ^2 $\{ \langle \psi_j^0 | \psi^2 \rangle \}$ are determined from the above equation with $n = 2$.
4. This process can then be continued to higher and higher order.

Although modern quantum mechanics uses high-order perturbation theory in some cases, much of what the student needs to know is contained in the first- and second- order results to which I will therefore restrict our further attention. I recommend that students have in memory (their own brain, not a computer) the equations for E^1 , E^2 , and ψ^1 so they can make use of them even in qualitative applications of perturbation theory as we will discuss later in this Chapter. But, first, let's consider an example problem that illustrates how perturbation theory is used in a more quantitative manner.

4.1.1 An Example Problem

As we discussed earlier, an electron moving in a quasi-linear conjugated bond framework can be modeled as a particle in a box. An externally applied electric field of strength ϵ interacts with the electron in a fashion that can be described by adding the perturbation $V = e\epsilon\left(x - \frac{L}{2}\right)$ to the zeroth-order Hamiltonian. Here, x is the position of the electron in the box, e is the electron's charge, and L is the length of the box. The perturbation potential varies in a linear fashion across the box, so it acts to pull the electron to one side of the box.

First, we will compute the first-order correction to the energy of the $n=1$ state and the first-order wave function for the $n=1$ state. In the wave function calculation, we will only compute the contribution to ψ made by ψ_2^0 (this is just an approximation to keep things simple in this example). Let me now do all the steps needed to solve this part of the problem. Try to make sure you can do the algebra, but also make sure you understand how we are using the first-order perturbation equations.

The zeroth-order wave functions and energies are given by

$$\psi_n^0 = \left(\frac{2}{L}\right)^{\frac{1}{2}} \text{Sin}\left(\frac{n\pi x}{L}\right), \text{ and}$$

$$E_n^0 = \frac{\hbar^2 \pi^2 n^2}{2mL^2},$$

and the perturbation is

$$V = e\varepsilon\left(x - \frac{L}{2}\right).$$

The first-order correction to the energy for the state having $n = 1$ and denote ψ^0 is

$$\begin{aligned} E^1 &= \langle \psi^0 | V | \psi^0 \rangle = \left\langle \psi^0 \left| e\varepsilon\left(x - \frac{L}{2}\right) \right| \psi^0 \right\rangle \\ &= \left(\frac{2}{L}\right) \int_0^L \text{Sin}^2\left(\frac{\pi x}{L}\right) e\varepsilon\left(x - \frac{L}{2}\right) dx \\ &= \left(\frac{2e\varepsilon}{L}\right) \int_0^L \text{Sin}^2\left(\frac{\pi x}{L}\right) x dx - \left(\frac{2e\varepsilon}{L}\right) \frac{L}{2} \int_0^L \text{Sin}^2\left(\frac{\pi x}{L}\right) dx \end{aligned}$$

The first integral can be evaluated using the following identity with $a = \frac{\pi}{L}$:

$$\int_0^L \sin^2(ax) dx = \frac{x^2}{4} - \frac{x \sin(2ax)}{4a} - \frac{\cos(2ax)}{8a^2} \Big|_0^L = \frac{L^2}{4}$$

The second integral can be evaluated using the following identity with $\theta = \frac{\pi x}{L}$

and $d\theta = \frac{\pi}{L} dx$:

$$\int_0^L \sin^2\left(\frac{\pi x}{L}\right) dx = \frac{L}{\pi} \int_0^\pi \sin^2 \theta d\theta$$

$$\int_0^\pi \sin^2 \theta d\theta = -\frac{1}{4} \sin(2\theta) + \frac{\theta}{2} \Big|_0^\pi = \frac{\pi}{2}.$$

Making all of these appropriate substitutions we obtain:

$$E^1 = \left(\frac{2e\varepsilon}{L}\right) \left(\frac{L^2}{4} - \frac{L}{2} \frac{L}{\pi} \frac{\pi}{2}\right) = 0.$$

This result, that the first-order correction to the energy vanishes, could have been foreseen. In the expression for $E^1 = \langle \psi^0 | V | \psi^0 \rangle$, the product $\psi^{0*} \psi^0$ is an even function under reflection of x through the midpoint $x = L/2$; in fact, this is true for all of the particle-in-a-box wave functions. On the other hand, the perturbation $V = e\varepsilon \left(x - \frac{L}{2}\right)$ is an odd function under reflection through $x = L/2$. Thus, the integral $\langle \psi^0 | V | \psi^0 \rangle$ must vanish as its integrand is an odd function. This simple example illustrates how one can use symmetry to tell ahead of time whether the integrals $\langle \psi^0 | V | \psi^0 \rangle$ and $\langle \psi_j^0 | V | \psi^0 \rangle$ contributing to the first-order and higher-order energies and wave functions will vanish. The contribution to the first-order wave function made by the $n = 2$ state is given by

$$\psi^1 = \frac{\langle \psi^0 | e\epsilon \left(x - \frac{L}{2}\right) | \psi_2^0 \rangle \psi_2^0}{E^0 - E_2^0}$$

$$= \frac{\frac{2}{L} \langle \sin(\pi x/L) | (e\epsilon(x - L/2)) | \sin(2\pi x/L) \rangle \psi_2^0}{\frac{\hbar^2 \pi^2}{2mL^2} - \frac{\hbar^2 \pi^2 2^2}{2mL^2}}$$

The two integrals in the numerator involve

$$\int_0^L x \sin\left(\frac{2\pi x}{L}\right) \sin\left(\frac{\pi x}{L}\right) dx$$

and

$$\int_0^L \sin\left(\frac{2\pi x}{L}\right) \sin\left(\frac{\pi x}{L}\right) dx$$

Using the integral identities

$$\int x \cos(ax) dx = \frac{1}{a^2} \cos(ax) + \frac{x}{a} \sin(ax)$$

and

$$\int \cos(ax) dx = \frac{1}{a} \sin(ax),$$

we obtain the following:

$$\int_0^L \sin\left(\frac{2\pi x}{L}\right) \sin\left(\frac{\pi x}{L}\right) dx = \frac{1}{2} \left[\int_0^L \cos\left(\frac{\pi x}{L}\right) dx - \int_0^L \cos\left(\frac{3\pi x}{L}\right) dx \right]$$

$$= \frac{1}{2} \left[\frac{L}{\pi} \sin\left(\frac{\pi x}{L}\right) \Big|_0^L - \frac{L}{3\pi} \sin\left(\frac{3\pi x}{L}\right) \Big|_0^L \right] = 0$$

and

$$\begin{aligned} \int_0^L x \sin\left(\frac{2\pi x}{L}\right) \sin\left(\frac{\pi x}{L}\right) dx &= \frac{1}{2} \left[\int_0^L x \cos\left(\frac{\pi x}{L}\right) dx - \int_0^L x \cos\left(\frac{3\pi x}{L}\right) dx \right] \\ &= \frac{1}{2} \left[\left(\frac{L^2}{\pi^2} \cos\left(\frac{\pi x}{L}\right) + \frac{Lx}{\pi} \sin\left(\frac{\pi x}{L}\right) \right) \Big|_0^L - \left(\frac{L^2}{9\pi^2} \cos\left(\frac{3\pi x}{L}\right) + \frac{Lx}{3\pi} \sin\left(\frac{3\pi x}{L}\right) \right) \Big|_0^L \right] \\ &= \frac{-2L^2}{2\pi^2} - \frac{-2L^2}{18\pi^2} = \frac{L^2}{9\pi^2} - \frac{L^2}{\pi^2} = -\frac{8L^2}{9\pi^2}. \end{aligned}$$

Making all of these appropriate substitutions we obtain:

$$\psi^1 = \frac{32mL^3 e \varepsilon}{27\hbar^2 \pi^4} \sqrt{\frac{2}{L}} \sin(2\pi x / L)$$

for the first-order wave function (actually, only the $n = 2$ contribution). So, the wave function through first order (i.e., the sum of the zeroth- and first-order pieces) is

$$\psi^0 + \psi^1 = \sqrt{\frac{2}{L}} \sin(\pi x / L) + \frac{32mL^3 e \varepsilon}{27\hbar^2 \pi^4} \sqrt{\frac{2}{L}} \sin(2\pi x / L).$$

In Fig. 4.1 we show the $n = 1$ and $n = 2$ zeroth-order functions as well as the superposition function formed when the zeroth-order $n = 1$ and first-order $n = 1$ functions combine.

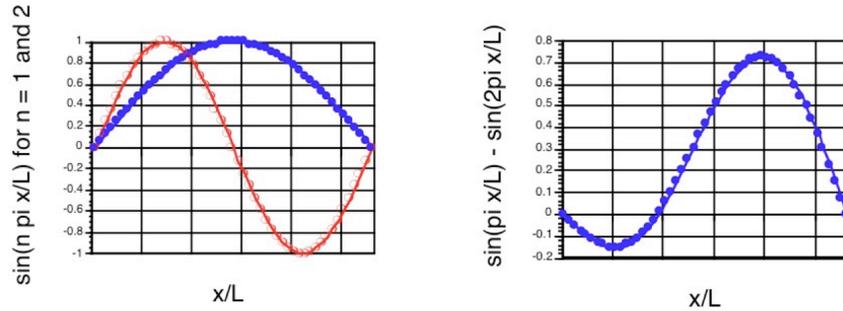


Figure 4.1 $n = 1$ (blue) and $n = 2$ (red) particle-in-a-box zeroth-order functions (left) and the $n = 1$ perturbed function through first order (right) arising from the electric field polarization.

Clearly, the external electric field acts to polarize the $n = 1$ wave function in a manner that moves its probability density toward the $x > L/2$ side of the box. The degree of polarization will depend on the strength of the applied electric field.

For such a polarized superposition wave function, there should be a net dipole moment induced in the system. We can evaluate this dipole moment by computing the expectation value of the dipole moment operator:

$$\mu_{\text{induced}} = -e \int \psi^* \left(x - \frac{L}{2} \right) \psi dx$$

with ψ being the sum of our zeroth- and first-order wave functions. In computing this integral, we neglect the term proportional to ϵ^2 because we are interested in only the term linear in ϵ because this is what gives the dipole moment. Again, allow me to do the algebra and see if you can follow.

$$\mu_{\text{induced}} = -e \int \psi^* \left(x - \frac{L}{2} \right) \psi dx,$$

where,

$$\psi = (\psi^0 + \psi^1).$$

$$\begin{aligned} \mu_{\text{induced}} &= -e \int_0^L (\psi^0 + \psi^1)^* \left(x - \frac{L}{2}\right) (\psi^0 + \psi^1) dx \\ &= -e \int_0^L \psi^{0*} \left(x - \frac{L}{2}\right) \psi^0 dx - e \int_0^L \psi^{1*} \left(x - \frac{L}{2}\right) \psi^0 dx \\ &\quad - e \int_0^L \psi^{0*} \left(x - \frac{L}{2}\right) \psi^1 dx - e \int_0^L \psi^{1*} \left(x - \frac{L}{2}\right) \psi^1 dx. \end{aligned}$$

The first integral is zero (we discussed this earlier when we used symmetry to explain why this vanishes). The fourth integral is neglected since it is proportional to ϵ^2 and we are interested in obtaining an expression for how the dipole varies linearly with ϵ . The second and third integrals are identical and can be combined to give:

$$\mu_{\text{induced}} = -2e \int_0^L \psi^{0*} \left(x - \frac{L}{2}\right) \psi^1 dx$$

Substituting our earlier expressions for

$$\psi^0 = \sqrt{\frac{2}{L}} \sin(\pi x / L)$$

and

$$\psi^1 = \frac{32mL^3 e \epsilon}{27\hbar^2 \pi^4} \sqrt{\frac{2}{L}} \sin(2\pi x / L)$$

we obtain:

$$\mu_{\text{induced}} = -2e \frac{32mL^3 e \epsilon}{27\hbar^2 \pi^4} \left(\frac{2}{L}\right) \int_0^L \sin\left(\frac{\pi x}{L}\right) \left(x - \frac{L}{2}\right) \sin\left(\frac{2\pi x}{L}\right) dx$$

These integrals are familiar from what we did to compute ψ^1 ; doing them we finally obtain:

$$\mu_{\text{induced}} = -2e \frac{32mL^3 e \epsilon \left(\frac{2}{L}\right) \left(\frac{-8L^2}{9\pi^2}\right)}{27\hbar^2 \pi^4} = \frac{mL^4 e^2 \epsilon 2^{10}}{\hbar^2 \pi^6 3^5}$$

Now. Let's compute the polarizability, α , of the electron in the $n=1$ state of the box, and try to understand physically why α should depend as it does upon the length of the box

L . To compute the polarizability, we need to know that $\alpha = \left. \frac{\partial \mu}{\partial \epsilon} \right|_{\epsilon=0}$. Using our

induced moment result above, we then find

$$\alpha = \left. \left(\frac{\partial \mu}{\partial \epsilon} \right) \right|_{\epsilon=0} = \frac{mL^4 e^2 2^{10}}{\hbar^2 \pi^6 3^5}$$

Notice that this finding suggests that the larger the box (i.e., the length of the conjugated molecule), the more polarizable the electron density. This result also suggests that the polarizability of conjugated polyenes should vary non-linearly with the length of the conjugated chain.

4.1.2 Other Examples

Let's consider a few more examples of how perturbation theory is used in chemistry, either quantitatively (i.e., to actually compute changes in energies and wave functions) or qualitatively (i.e., to interpret or anticipate how changes might alter energies or other properties).

1. The Stark effect

When a molecule is exposed to an electric field E , its electrons and nuclei experience a perturbation

$$V = E \cdot (e \sum_n Z_n R_n - e \sum_i r_i)$$

where Z_n is the charge of the n^{th} nucleus whose position is R_n , r_i is the position of the i^{th} electron, and e is the unit of charge. The effect of this perturbation on the energies is termed the Stark effect. The first-order change to the energy of this molecule is evaluated by calculating

$$E^1 = \langle \psi^* | V | \psi \rangle = \mathcal{E} \cdot \langle \psi | e \sum_n Z_n R_n - e \sum_i r_i | \psi \rangle$$

where ψ is the unperturbed wave function of the molecule (i.e., the wave function in the absence of the electric field). The quantity inside the integral is the electric dipole operator, so this integral is the dipole moment of the molecule in the absence of the field. For species that possess no dipole moment (e.g., non-degenerate states of atoms and centro-symmetric molecules), this first-order energy vanishes. It vanishes in the two specific cases mentioned because ψ is either even or odd under the inversion symmetry, but the product $\psi^* \psi$ is even, and the dipole operator is odd, so the integrand is odd and thus the integral vanishes.

If one is dealing with a degenerate state of a centro-symmetric system, things are different. For example, the 2s and 2p states of the hydrogen atom are degenerate, so, to apply perturbation theory one has to choose specific combinations that diagonalize the perturbation. This means one needs to first form the 2x2 matrix

$$\begin{pmatrix} \langle 2s | V | 2s \rangle & \langle 2s | V | 2p_z \rangle \\ \langle 2p_z | V | 2s \rangle & \langle 2p_z | V | 2p_z \rangle \end{pmatrix}$$

where z is taken to be the direction of the electric field. The diagonal elements of this matrix vanish due to parity symmetry, so the two eigenvalues are equal to

$$E_{\pm}^1 = \pm \langle 2s | V | 2p_z \rangle.$$

These are the two first-order (because they are linear in V and thus linear in \mathcal{E}) energies. So, in such degenerate cases, one can obtain linear Stark effects. The two corrected

zeroth-order wave functions corresponding to these two shifted energies are

$$\psi_{\pm}^0 = \frac{1}{\sqrt{2}}[2s \mp 2p_z]$$

and correspond to orbitals polarized into or away from the electric field.

The Stark effect example offers a good chance to explain a fundamental problem with applying perturbation theory. One of the basic assumptions of perturbation theory is that the unperturbed and perturbed Hamiltonians are both bounded from below (i.e., have a discrete lowest eigenvalues) and allow each eigenvalue of the unperturbed Hamiltonian to be connected to a unique eigenvalue of the perturbed Hamiltonian. Considering the example just discussed, we can see that these assumptions are not met for the Stark perturbation.

Consider the potential that an electron experiences within an atom or molecule close to a nucleus of charge Z . It is of the form (in atomic units where the energy is given in Hartrees ($1 \text{ H} = 27.21 \text{ eV}$) and distances in Bohr units ($1 \text{ Bohr} = 0.529 \text{ \AA}$))

$$V(r, \theta, \phi) = -\frac{Z}{r} - e\mathcal{E}r \cos \theta$$

where the first term is the Coulomb potential acting to attract the electron to the nucleus and the second is the electron-field potential assuming the field is directed along the z -direction. In Fig. 4.2 a we show this potential for a given value of the angle θ .

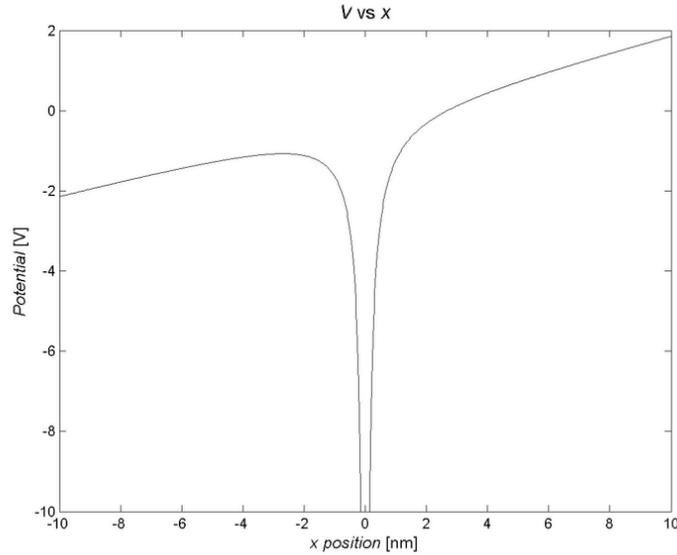


Figure 4.2 a Potential experienced by valence electron showing attraction to a nucleus located at the origin (the deep well) and the potential due to the applied electric field.

Along directions for which $\cos\theta$ is negative (to the right in Fig. 4.2 a), this potential becomes large and positive as the distance r of the electron from the nucleus increases; for bound states such as the 2s and 2p states discussed earlier, such regions are classically forbidden and the wave function exponentially decays in this direction. However, in directions along which $\cos\theta$ is positive, the potential is negative and strongly attractive for small- r (i.e., near the nucleus), then passes through a maximum (e.g., near $x = -2$ in Fig. 4.2 a) at

$$r_{\max} = \sqrt{\frac{Z}{e\mathcal{E}\cos\theta}}$$

where

$$V(r_{\max}) = -2\sqrt{e\mathcal{E}Z\cos\theta}$$

(ca. -1 eV in Fig. 4.2 a) and then decreases monotonically as r increases. In fact, this potential approaches $-\infty$ as r approaches ∞ as we see in the left portion of Fig. 4.2 a.

The bottom line is that the total potential with the electric field present violates the assumptions on which perturbation theory is based. However, it turns out that perturbation theory can be used in such cases under certain conditions. For example as applied to the Stark effect for the degenerate 2s and 2p levels of a hydrogenic atom (i.e., a one-electron system with nuclear charge Z), if the energy of the 2s and 2p states lies far below the maximum in the potential $V(r_{\max})$, perturbation theory can be used. We know the energies of hydrogenic ions vary with Z and with the principal quantum number n as

$$E_n(Z) = \frac{-13.6eV}{n^2 Z^2} = \frac{-1}{2n^2 Z^2} au.$$

So, as long as

$$\frac{-1}{2n^2 Z^2} \ll -2\sqrt{e\mathcal{E}Z \cos\theta}$$

the zeroth-order energy of the state will lie below the barrier on the potential surface. Because the wave function can penetrate this barrier, this state will no longer be a true bound state; it will be a metastable resonance state (recall, we studied such states in Chapter 1 where we learned about tunneling). However, if the zeroth-order energy lies far below the barrier, the extent of tunneling through the barrier will be small, so the state will have a long lifetime. In such cases, we can use perturbation theory to describe the effects of the applied electric field on the energies and wave functions of such metastable states, but we must realize that we are only allowed to do so in the limit of weak fields and for states that lie significantly below the barrier. In this case, perturbation theory describes the changes in the energy and wave function in regions of space where the zeroth-order wave function is bound, but does not describe at all the asymptotic part of the wave function where the electron is unbound.

Another example of Stark effects in degenerate cases arises when considering how polar diatomic molecules' rotational energies are altered by an electric field. The zeroth-order wave functions appropriate to such cases are given by

$$\psi = Y_{J,M}(\theta, \phi) \chi_v(R) \psi_e(r | R)$$

where the spherical harmonic $Y_{J,M}(\theta, \phi)$ is the rotational wave function, $\chi_v(R)$ is the vibrational function for level v , and $\psi_e(r | R)$ is the electronic wave function. The diagonal elements of the electric-dipole operator

$$\langle Y_{J,M}(\theta, \phi) \chi_v(R) \psi_e(r | R) | V | Y_{J,M}(\theta, \phi) \chi_v(R) \psi_e(r | R) \rangle$$

vanish because the vibrationally averaged dipole moment, which arises as

$$\langle \mu \rangle = \langle \chi_v(R) \psi_e(r | R) | e \sum_n Z_n R_n - e \sum_i r_i | \chi_v(R) \psi_e(r | R) \rangle$$

is a vector quantity whose component along the electric field \mathcal{E} is $\langle \mu \rangle \cos(\theta)$ (again taking the field to lie along the z -direction). Thinking of $\cos(\theta)$ as x , so $\sin(\theta) d\theta$ is dx , the integrals

$$\langle Y_{J,M}(\theta, \phi) | \cos \theta | Y_{J,M}(\theta, \phi) \rangle = \int Y_{J,M}^*(\theta, \phi) \cos \theta Y_{J,M}(\theta, \phi) \sin \theta d\theta d\phi = \int Y_{J,M}^*(\theta, \phi) x Y_{J,M}(\theta, \phi) dx d\phi = 0$$

because $|Y_{J,M}|^2$ is an even function of x (i.e., of $\cos(\theta)$). Because the angular dependence of the perturbation (i.e., $\cos \theta$) has no ϕ -dependence, matrix elements of the form

$$\int Y_{J,M}^*(\theta, \phi) \cos \theta Y_{J,M'}(\theta, \phi) \sin \theta d\theta d\phi = 0$$

also vanish. This means that if one were to form the $(2J+1) \times (2J+1)$ matrix representation of V for the $2J+1$ degenerate states $Y_{J,M}$ belonging to a given J , all of its elements would be zero. Thus the rotational energies of polar diatomic (or rigid linear polyatomic) molecules have no first-order Stark splittings.

There will, however, be second-order Stark splittings, in which case we need to

examine the terms that arise in the formula

$$E^2 = \sum_J \frac{|\langle \psi^0 | V | \psi_J^0 \rangle|^2}{E^0 - E_J^0}.$$

For a zeroth-order state $Y_{J,M}$, only certain other zeroth-order states will have non-vanishing coupling matrix elements $\langle \psi^0 | V | \psi_J^0 \rangle$. These non-zero integrals are governed by $\langle Y_{J,M} | \cos\theta | Y_{J',M'} \rangle$, which can be shown to be

$$\langle Y_{J,M} | \cos\theta | Y_{J',M'} \rangle = \left\{ \sqrt{\frac{(J+1)^2 - M^2}{(2J+1)(2J+3)}} \text{ for } J' = J+1; \sqrt{\frac{J^2 - M^2}{(2J-1)(2J+1)}} \text{ for } J' = J-1 \right\} \delta_{M,M'};$$

of course, if $J = 0$, the term $J' = J-1$ does not occur. The limitation that M must equal M' arises, as above, because the perturbation contains no terms dependent on the variable ϕ .

The limitation that $J' = J \pm 1$ comes from a combination of three conditions

- (i) angular momentum coupling, which you learned about in Chapter 2, tells us that $\cos\theta$, which happens to be proportional to $Y_{1,0}(\theta, \phi)$, can couple to $Y_{J,M}$ to generate terms having $J+1$, J , or $J-1$ for their J^2 quantum number but only M for their J_z quantum number,
- (ii) the $J+1$, J , and $J-1$ factors arising from the product $\cos\theta Y_{J,M}$ must match $Y_{J',M'}$ for the integral not to vanish because $\langle Y_{J,M} | Y_{J',M'} \rangle = \delta_{JJ'} \delta_{MM'}$,
- (iii) finally, the $J = J'$ terms will vanish because of the inversion symmetry ($\cos\theta$ is odd under inversion but $|Y_{J,M}|^2$ is even).

Using the fact that the perturbation is $\mathcal{E} \langle \mu \rangle \cos(\theta)$, these two non-zero matrix elements can be used to express the second-order energy for the J, M level as

$$E = \mathcal{E}^2 \langle \mu \rangle^2 \left\{ \frac{(J+1)^2 - M^2}{-2B(J+1)} + \frac{J^2 - M^2}{2BJ} \right\}$$

where h is Planck's constant and B is the rotational constant for the molecule

$$B = \frac{h}{8\pi^2 \mu r_e^2}$$

for a diatomic molecule of reduced mass μ and equilibrium bond length r_e .

Before moving on to another example, it is useful to point out some common threads that occur in many applications of perturbation theory and that will also be common to variational calculations that we discuss later in this Chapter. Once one has identified the specific zeroth-order state ψ^0 of interest, one proceeds as follows:

(i) The first-order energy $E^1 = \langle \psi^0 | V | \psi^0 \rangle$ is evaluated. In doing so, one should first make use of any symmetry (point group symmetry is treated later in this Chapter) such as inversion, angular momentum, spin, etc., to determine whether this expectation value will vanish by symmetry, in which case, we don't bother to consider this matrix element any more. We used this earlier when considering $\langle 2s | \cos\theta | 2s \rangle$, $\langle 2p_\sigma | \cos\theta | 2p_\sigma \rangle$, and $\langle Y_{J,M} | \cos\theta | Y_{J,M} \rangle$ to conclude that certain first-order energies are zero.

(ii). If E^1 vanishes (so the lowest-order effect is in second order) or if we want to examine higher-order corrections, we consider evaluating E^2 . Before doing so explicitly, we think about whether symmetry will limit the matrix elements $\langle \psi^0 | V | \psi_n^0 \rangle$ entering into the expression for E^2 . For example, in the case just studied, we saw that only other zeroth-order states having $J' = J + 1$ or $J' = J - 1$ gave non-vanishing matrix elements. In addition, because E^2 contains energy denominators ($E^0 - E_n^0$), we may choose to limit our calculation to those other zeroth-order states whose energies are close to our state of interest; this assumes that such states will contribute a dominant amount to the sum

$$\sum_n \frac{|\langle \psi_n^0 | V | \psi^0 \rangle|^2}{E^0 - E_n^0}$$

You will encounter many times when reading literature articles in which perturbation theory is employed situations in which researchers have focused attention on zeroth-order states that are close in energy to the state of interest and that have the correct symmetry to couple strongly (i.e., have substantial $\langle \psi^0 | V | \psi_n^0 \rangle$) to that state.

2. Electron-electron Coulomb repulsion

In one of the most elementary pictures of atomic electronic structure, one uses nuclear charge screening concepts to partially account for electron-electron interactions. For example, in $1s^2 2s^1$ Li, one might posit a zeroth-order wave function consisting of a product

$$\psi = \phi_{1s}(r_1)\alpha(1)\phi_{1s}(r_2)\beta(2)\phi_{2s}(r_3)\alpha(3)$$

in which two electrons occupy a 1s orbital and one electron occupies a 2s orbital. To find a reasonable form for the radial parts of these two orbitals, one could express each of them as a linear combination of (i) one orbital having hydrogenic 1s form with a nuclear charge of 3 and (ii) a second orbital of 2s form but with a nuclear charge of 1 (to account for the screening of the $Z = 3$ nucleus by the two inner-shell 1s electrons)

$$\phi_i(r) = C_i \chi_{1s,Z=1}(r) + D_i \chi_{2s,Z=3}(r)$$

where the index i labels the 1s and 2s orbitals to be determined. Next, one could determine the C_i and D_i expansion coefficients by requiring the ϕ_i to be approximate eigenfunctions of the Hamiltonian

$$h = -1/2\nabla^2 - \frac{3}{r}$$

that would be appropriate for an electron attracted to the Li nucleus but not experiencing any repulsions with other electrons. This would result in the following equation for the expansion coefficients:

$$\begin{pmatrix} \langle \chi_{1s,Z=1}(r) | -1/2\nabla^2 - \frac{3}{r} | \chi_{1s,Z=1}(r) \rangle & \langle \chi_{1s,Z=1}(r) | -1/2\nabla^2 - \frac{3}{r} | \chi_{2s,Z=3}(r) \rangle \\ \langle \chi_{1s,Z=1}(r) | -1/2\nabla^2 - \frac{3}{r} | \chi_{2s,Z=3}(r) \rangle & \langle \chi_{2s,Z=3}(r) | -1/2\nabla^2 - \frac{3}{r} | \chi_{2s,Z=3}(r) \rangle \end{pmatrix} \begin{pmatrix} C \\ D \end{pmatrix} =$$

$$\begin{pmatrix} \langle \chi_{1s,Z=1}(r) | \chi_{1s,Z=1}(r) \rangle & \langle \chi_{1s,Z=1}(r) | \chi_{2s,Z=3}(r) \rangle \\ \langle \chi_{1s,Z=1}(r) | \chi_{2s,Z=3}(r) \rangle & \langle \chi_{2s,Z=3}(r) | \chi_{2s,Z=3}(r) \rangle \end{pmatrix} \begin{pmatrix} C \\ D \end{pmatrix}.$$

This 2x2 matrix eigenvalue problem can be solved for the C_i and D_i coefficients and for the energies E_i of the 1s and 2s orbitals. The lower-energy solution will have $|C| > |D|$, and will be this model's description of the 1s orbital. The higher-energy solution will have $|D| > |C|$ and is the approximation to the 2s orbital.

Using these 1s and 2s orbitals and the 3-electron wave function they form

$$\psi = \phi_{1s}(r_1)\alpha(1)\phi_{1s}(r_2)\beta(2)\phi_{2s}(r_3)\alpha(3)$$

as a zeroth-order approximation, how do we then proceed to apply perturbation theory?

The full three-electron Hamiltonian

$$H = \sum_{i=1}^3 \left[-1/2\nabla_i^2 - \frac{3}{r_i} \right] + \sum_{i<j=1}^3 \frac{1}{r_{i,j}}$$

can be decomposed into a zeroth-order part

$$H^0 = \sum_{i=1}^3 \left[-1/2\nabla_i^2 - \frac{3}{r_i} \right]$$

and a perturbation

$$V = \sum_{i<j=1}^3 \frac{1}{r_{i,j}}.$$

The zeroth-order energy of the wave function

$$\psi = \phi_{1s}(r_1)\alpha(1)\phi_{1s}(r_2)\beta(2)\phi_{2s}(r_3)\alpha(3)$$

is

$$E^0 = 2E_{1s} + E_{2s}$$

where each of the E_{ns} are the energies obtained by solving the 2x2 matrix eigenvalue equation shown earlier. The first-order energy of this state can be written as

$$E^1 = \langle \phi_{1s}(r_1)\alpha(1)\phi_{1s}(r_2)\beta(2)\phi_{2s}(r_3)\alpha(3) | V | \phi_{1s}(r_1)\alpha(1)\phi_{1s}(r_2)\beta(2)\phi_{2s}(r_3)\alpha(3) \rangle = J_{1s,1s} + 2J_{1s,2s}$$

with the Coulomb interaction integrals being defined as

$$J_{a,b} = \int \phi_a^*(r)\phi_a(r) \frac{1}{|r-r'|} \phi_b^*(r)\phi_b(r) dr dr'$$

To carry out the 3-electron integral appearing in E^1 , one proceeds as follows. For the integral

$$\int [\phi_{1s}(r_1)\alpha(1)\phi_{1s}(r_2)\beta(2)\phi_{2s}(r_3)\alpha(3)]^* \frac{1}{r_{1,2}} \phi_{1s}(r_1)\alpha(1)\phi_{1s}(r_2)\beta(2)\phi_{2s}(r_3)\alpha(3) d1d2d3$$

one integrates over the 3 spin variables using $\langle \alpha | \alpha \rangle = 1$, $\langle \alpha | \beta \rangle = 0$ and $\langle \beta | \beta \rangle = 1$) and then integrates over the coordinate of the third electron using $\langle \phi_{2s} | \phi_{2s} \rangle = 1$ to obtain

$$\int [\phi_{1s}(r_1)\phi_{1s}(r_2)]^* \frac{1}{r_{1,2}} \phi_{1s}(r_1)\phi_{1s}(r_2) d1d2$$

which is $J_{1s,1s}$. The two $J_{1s,2s}$ integrals arise when carrying out similar integration for the terms arising from $(1/r_{1,3})$ and $(1/r_{2,3})$.

So, through first order, the energy of the Li atom at this level of treatment is given by

$$E^0 + E^1 = 2E_{1s} + E_{2s} + J_{1s,1s} + 2J_{1s,2s}.$$

The factor $2E_{1s} + E_{2s}$ contains the contributions from the kinetic energy and electron-nuclear Coulomb potential. The $J_{1s,1s} + 2J_{1s,2s}$ terms describe the Coulombic repulsions among the three electrons. Each of the Coulomb integrals $J_{i,j}$ can be interpreted as being equal to the Coulombic interaction between electrons (one at location \mathbf{r} ; the other at \mathbf{r}') averaged over the positions of these two electrons with their spatial probability distributions being given by $|\phi_i(\mathbf{r})|^2$ and $|\phi_j(\mathbf{r}')|^2$, respectively.

Although the example just considered is rather primitive, it introduces a point of view that characterizes one of the most commonly employed models for treating atomic and molecular electronic structure- the Hartree-Fock (HF) mean-field model, which we will discuss more in Chapter 6. In the HF model, one uses as a zeroth-order Hamiltonian

$$H^0 = \sum_{i=1}^3 \left[-1/2\nabla_i^2 - \frac{3}{r_i} + V_{HF}(r_i) \right]$$

consisting of a sum of one-electron terms containing the kinetic energy, the Coulomb attraction to the nucleus (I use the Li atom as an example here), and a potential $V_{HF}(\mathbf{r}_i)$. This potential, which is written in terms of Coulomb integrals similar to those we discussed earlier as well as so-called exchange integrals that we will discuss in Chapter 6, is designed to approximate the interaction of an electron at location \mathbf{r}_i with the other electrons in the atom or molecule. Because H^0 is one-electron additive, its eigenfunctions consist of products of eigenfunctions of the operator

$$h^0 = -1/2\nabla^2 - \frac{3}{r} + V_{HF}(r).$$

$V_{HF}(\mathbf{r}_i)$ offers an approximation to the true $1/r_{i,j}$ Coulomb interactions expressed in terms of a “smeared-out” electron distribution interacting with the electron at \mathbf{r}_i .

Perturbation theory is then used to treat the effect of the perturbation

$$V = \sum_{i < j = 1}^N \frac{1}{r_{i,j}} - \sum_{i=1}^N V_{HF}(r_i)$$

on the zeroth-order states. We say that the perturbation, often called the fluctuation potential, corrects for the difference between the instantaneous Coulomb interactions among the N electrons and the mean-field (average) interactions.

4.2. The Variational Method

Let us now turn to the other method that is used to solve Schrödinger equations approximately, the variational method. In this approach, one must again have some reasonable wave function ψ^0 that is used to approximate the true wave function. Within this approximate wave function, one imbeds one or more variables $\{\alpha_j\}$ that one subsequently varies to achieve a minimum in the energy of ψ^0 computed as an expectation value of the true Hamiltonian H:

$$E(\{\alpha_j\}) = \langle \psi^0 | H | \psi^0 \rangle / \langle \psi^0 | \psi^0 \rangle.$$

The optimal values of the α_j parameters are determined by making

$$dE/d\alpha_j = 0$$

To achieve the desired energy minimum (n.b., we also should verify that the second derivative matrix $(\partial^2 E / \partial \alpha_j \partial \alpha_l)$ has all positive eigenvalues, otherwise one may not have found the minimum).

The theoretical basis underlying the variational method can be understood through the following derivation. Suppose that someone knew the exact eigenstates (i.e., true ψ_K and true E_K) of the true Hamiltonian H. These states obey

$$H \psi_K = E_K \psi_K.$$

Because these true states form a complete set (it can be shown that the eigenfunctions of all the Hamiltonian operators we ever encounter have this property), our so-called “trial wave function” ψ^0 can, in principle, be expanded in terms of these ψ_K :

$$\psi^0 = \sum_K C_K \psi_K.$$

Before proceeding further, allow me to overcome one likely misconception. What I am going through now is only a derivation of the working formula of the variational method. The final formula will not require us to ever know the exact ψ_K or the exact E_K , but we are allowed to use them as tools in our derivation because we know they exist even if we never know them.

With the above expansion of our trial function in terms of the exact eigenfunctions, let us now substitute this into the quantity $\langle \psi^0 | H | \psi^0 \rangle / \langle \psi^0 | \psi^0 \rangle$ that the variational method instructs us to compute:

$$E = \langle \psi^0 | H | \psi^0 \rangle / \langle \psi^0 | \psi^0 \rangle = \langle \sum_K C_K \psi_K | H | \sum_L C_L \psi_L \rangle / \langle \sum_K C_K \psi_K | \sum_L C_L \psi_L \rangle.$$

Using the fact that the ψ_K obey $H\psi_K = E_K\psi_K$ and that the ψ_K are orthonormal (I hope you remember this property of solutions to all Schrödinger equations that we discussed earlier)

$$\langle \psi_K | \psi_L \rangle = \delta_{K,L}$$

the above expression reduces to

$$E = \sum_K \langle C_K \psi_K | H | C_K \psi_K \rangle / (\sum_K \langle C_K \psi_K | C_K \psi_K \rangle) = \sum_K |C_K|^2 E_K / \sum_K |C_K|^2.$$

One of the basic properties of the kind of Hamiltonians we encounter is that they have a lowest-energy state. Sometimes we say they are bounded from below, which means their energy states do not continue all the way to minus infinity. There are systems for which this is not the case (we saw one earlier when studying the Stark effect), but we will now assume that we are not dealing with such systems. This allows us to introduce the inequality $E_K \geq E_0$ which says that all of the energies are higher than or equal to the energy of the lowest state which we denote E_0 . Introducing this inequality into the above expression gives

$$E \geq \frac{\sum_K |C_K|^2 E_0}{\sum_K |C_K|^2} = E_0.$$

This means that the variational energy, computed as $\langle \psi^0 | H | \psi^0 \rangle / \langle \psi^0 | \psi^0 \rangle$ will lie above the true ground-state energy no matter what trial function ψ^0 we use.

The significance of the above result that $E \geq E_0$ is as follows. We are allowed to imbue into our trial wave function ψ^0 parameters that we can vary to make E , computed as $\langle \psi^0 | H | \psi^0 \rangle / \langle \psi^0 | \psi^0 \rangle$ as low as possible because we know that we can never make $\langle \psi^0 | H | \psi^0 \rangle / \langle \psi^0 | \psi^0 \rangle$ lower than the true ground-state energy. The philosophy then is to vary the parameters in ψ^0 to render E as low as possible, because the closer E is to E_0 the “better” is our variational wave function. Let me now demonstrate how the variational method is used in such a manner by solving an example problem.

4.2.1 An Example Problem

Suppose you are given a trial wave function of the form:

$$\phi = \frac{Z^3}{\pi a_0^3} \exp\left(\frac{-Z_e r_1}{a_0}\right) \exp\left(\frac{-Z_e r_2}{a_0}\right)$$

to represent a two-electron ion of nuclear charge Z and suppose that you are lucky enough that I have already evaluated the $\langle \psi^0 | H | \psi^0 \rangle / \langle \psi^0 | \psi^0 \rangle$ integral, which I’ll call W , for you and found

$$W = \left(Z_e^2 - 2ZZ_e + \frac{5}{8}Z_e \right) \frac{e^2}{a_0}.$$

Now, let's find the optimum value of the variational parameter Z_e for an arbitrary nuclear charge Z by setting $\frac{dW}{dZ_e} = 0$. After finding the optimal value of Z_e , we'll then find the optimal energy by plugging this Z_e into the above W expression. I'll do the algebra and see if you can follow.

$$W = \left(Z_e^2 - 2ZZ_e + \frac{5}{8}Z_e \right) \frac{e^2}{a_0}$$

$$\frac{dW}{dZ_e} = \left(2Z_e - 2Z + \frac{5}{8} \right) \frac{e^2}{a_0} = 0$$

$$2Z_e - 2Z + \frac{5}{8} = 0$$

$$2Z_e = 2Z - \frac{5}{8}$$

$$Z_e = Z - \frac{5}{16} = Z - 0.3125$$

(n.b., 0.3125 represents the shielding factor of one 1s electron to the other, reducing the optimal effective nuclear charge by this amount).

Now, using this optimal Z_e in our energy expression gives

$$W = Z_e \left(Z_e - 2Z + \frac{5}{8} \right) \frac{e^2}{a_0} ,$$

$$W = \left(Z - \frac{5}{16} \right) \left(\left(Z - \frac{5}{16} \right) - 2Z + \frac{5}{8} \right) \frac{e^2}{a_0}$$

$$W = \left(Z - \frac{5}{16} \right) \left(-Z + \frac{5}{16} \right) \frac{e^2}{a_0}$$

$$W = -\left(Z - \frac{5}{16}\right)\left(Z - \frac{5}{16}\right)\frac{e^2}{a_0} = -\left(Z - \frac{5}{16}\right)^2 \frac{e^2}{a_0}$$

$$= -(Z - 0.3125)^2(27.21) \text{ eV}$$

(n.b., since a_0 is the Bohr radius 0.529 \AA , $e^2/a_0 = 27.21 \text{ eV}$, or one atomic unit of energy).

Is this energy any good? The total energies of some two-electron atoms and ions have been experimentally determined to be as shown in the Table below.

Z = 1	H ⁻	-14.35 eV
Z = 2	He	-78.98 eV
Z = 3	Li ⁺	-198.02 eV
Z = 4	Be ⁺²	-371.5 eV
Z = 5	B ⁺³	-599.3 eV
Z = 6	C ⁺⁴	-881.6 eV
Z = 7	N ⁺⁵	-1218.3 eV
Z = 8	O ⁺⁶	-1609.5 eV

Using our optimized expression for W , let's now calculate the estimated total energies of each of these atoms and ions as well as the percent error in our estimate for each ion.

Z	Atom	Experimental	Calculated	% Error
Z = 1	H ⁻	-14.35 eV	-12.86 eV	10.38%
Z = 2	He	-78.98 eV	-77.46 eV	1.92%
Z = 3	Li ⁺	-198.02 eV	-196.46 eV	0.79%
Z = 4	Be ⁺²	-371.5 eV	-369.86 eV	0.44%
Z = 5	B ⁺³	-599.3 eV	-597.66 eV	0.27%
Z = 6	C ⁺⁴	-881.6 eV	-879.86 eV	0.19%
Z = 7	N ⁺⁵	-1218.3 eV	-1216.48 eV	0.15%
Z = 8	O ⁺⁶	-1609.5 eV	-1607.46 eV	0.13%

The energy errors are essentially constant over the range of Z , but produce a larger percentage error at small Z .

In 1928, when quantum mechanics was quite young, it was not known whether the isolated, gas-phase hydride ion, H^- , was stable with respect to loss of an electron to form a hydrogen atom. Let's compare our estimated total energy for H^- to the ground state energy of a hydrogen atom and an isolated electron (which is known to be -13.60 eV). When we use our expression for W and take $Z = 1$, we obtain $W = -12.86$ eV, which is greater than -13.6 eV ($H + e^-$), so this simple variational calculation erroneously predicts H^- to be unstable. More complicated variational treatments give a ground state energy of H^- of -14.35 eV, in agreement with experiment and agreeing that H^- is indeed stable with respect to electron detachment.

4.2.2 Another Example

A widely used example of is provided by the so-called linear variational method. Here one expresses the trial wave function a linear combination of so-called basis functions $\{\chi_j\}$

$$\psi = \sum_j C_j \chi_j.$$

Substituting this expansion into $\langle \psi | H | \psi \rangle$ and then making this quantity stationary with respect to variations in the C_i subject to the constraint that ψ remain normalized

$$1 = \langle \psi | \psi \rangle = \sum_i \sum_j C_i^* \langle \chi_i | \chi_j \rangle C_j$$

gives

$$\sum_j \langle \chi_i | H | \chi_j \rangle C_j = E \sum_j \langle \chi_i | \chi_j \rangle C_j.$$

This is a generalized matrix eigenvalue problem that we can write in matrix notation as

$$\mathbf{HC}=\mathbf{ESC}.$$

It is called a generalized eigenvalue problem because of the appearance of the overlap matrix \mathbf{S} on its right hand side.

This set of equations for the C_j coefficients can be made into a conventional eigenvalue problem as follows:

1. The eigenvectors \mathbf{v}_k and eigenvalues s_k of the overlap matrix are found by solving

$$\sum_j S_{i,j} v_{k,j} = s_k v_{k,i}$$

All of the eigenvalues s_k are positive because \mathbf{S} is a positive-definite matrix.

2. Next one forms the matrix $\mathbf{S}^{-1/2}$ whose elements are

$$S_{i,j}^{-1/2} = \sum_k v_{k,i} \frac{1}{\sqrt{s_k}} v_{k,j}$$

(another matrix $\mathbf{S}^{1/2}$ can be formed in a similar way replacing $\frac{1}{\sqrt{s_k}}$ with $\sqrt{s_k}$).

3. One then multiplies the generalized eigenvalue equation on the left by $\mathbf{S}^{-1/2}$ to obtain

$$\mathbf{S}^{-1/2}\mathbf{HC}=\mathbf{E} \mathbf{S}^{-1/2}\mathbf{SC}.$$

4. This equation is then rewritten, using $\mathbf{S}^{-1/2}\mathbf{S} = \mathbf{S}^{1/2}$ and $\mathbf{1}=\mathbf{S}^{-1/2}\mathbf{S}^{1/2}$ as

$$\mathbf{S}^{-1/2}\mathbf{H} \mathbf{S}^{-1/2} (\mathbf{S}^{1/2}\mathbf{C})=\mathbf{E} (\mathbf{S}^{1/2}\mathbf{C}).$$

This is a conventional eigenvalue problem in which the matrix is $\mathbf{S}^{-1/2}\mathbf{H} \mathbf{S}^{-1/2}$ and the eigenvectors are $(\mathbf{S}^{1/2}\mathbf{C})$.

The net result is that one can form $\mathbf{S}^{-1/2}\mathbf{H}\mathbf{S}^{-1/2}$ and then find its eigenvalues and eigenvectors. Its eigenvalues will be the same as those of the original generalized eigenvalue problem. Its eigenvectors ($\mathbf{S}^{1/2}\mathbf{C}$) can be used to determine the eigenvectors \mathbf{C} of the original problem by multiplying by $\mathbf{S}^{-1/2}$

$$\mathbf{C} = \mathbf{S}^{-1/2} (\mathbf{S}^{1/2}\mathbf{C}).$$

Although the derivation of the matrix eigenvalue equations resulting from the linear variational method was carried out as a means of minimizing $\langle \psi | \mathbf{H} | \psi \rangle$, it turns out that the solutions offer more than just an upper bound to the lowest true energy of the Hamiltonian. It can be shown that the n^{th} eigenvalue of the matrix $\mathbf{S}^{-1/2}\mathbf{H}\mathbf{S}^{-1/2}$ is an upper bound to the true energy of the n^{th} state of the Hamiltonian. A consequence of this is that, between any two eigenvalues of the matrix $\mathbf{S}^{-1/2}\mathbf{H}\mathbf{S}^{-1/2}$ there is at least one true energy of the Hamiltonian. This observation is often called the bracketing condition. The ability of linear variational methods to provide estimates to the ground- and excited-state energies from a single calculation is one of the main strengths of this approach.

4.3 Point Group Symmetry

It is assumed that the reader has previously learned, in undergraduate inorganic or physical chemistry classes, how symmetry arises in molecular shapes and structures and what symmetry elements are (e.g., planes, axes of rotation, centers of inversion, etc.). For the reader who feels, after reading this material, that additional background is needed, the texts by Eyring, Walter, and Kimball or by Atkins and Friedman can be consulted. We review and teach here only that material that is of direct application to symmetry analysis of molecular orbitals and vibrations and rotations of molecules. We use a specific example, the ammonia molecule, to introduce and illustrate the important aspects of point group symmetry because this example contains most of the complexities that arise in any application of group theory to molecular problems.

4.3.1 The C_{3v} Symmetry Group of Ammonia - An Example

The ammonia molecule NH_3 belongs, in its ground-state equilibrium geometry,

to the C_{3v} point group. Its symmetry operations consist of two C_3 rotations, C_3, C_3^2 (rotations by 120° and 240° , respectively about an axis passing through the nitrogen atom and lying perpendicular to the plane formed by the three hydrogen atoms), three vertical reflection operations, $\sigma_v, \sigma_v', \sigma_v''$, and the identity operation. Corresponding to these six operations are symmetry elements: the three-fold rotation axis, C_3 and the three symmetry planes σ_v, σ_v' and σ_v'' that contain the three NH bonds and the z-axis (see Fig. 4.3).

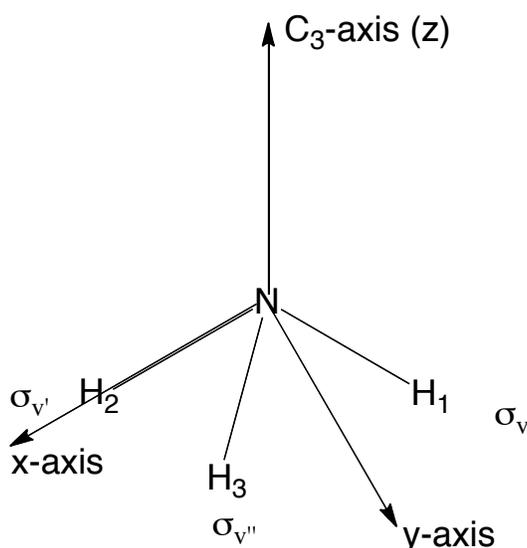


Figure 4.3 Ammonia Molecule and its Symmetry Elements

These six symmetry operations form a mathematical group. A group is defined as a set of objects satisfying four properties.

1. A combination rule is defined through which two group elements are combined to give a result that we call the product. The product of two elements in the group must also be a member of the group (i.e., the group is closed under the combination rule).
2. One special member of the group, when combined with any other member of the

group, must leave the group member unchanged (i.e., the group contains an identity element).

3. Every group member must have a reciprocal in the group. When any group member is combined with its reciprocal, the product is the identity element.
4. The associative law must hold when combining three group members (i.e., $(AB)C$ must equal $A(BC)$).

The members of symmetry groups are symmetry operations; the combination rule is successive operation. The identity element is the operation of doing nothing at all. The group properties can be demonstrated by forming a multiplication table. Let us label the rows of the table by the first operation and the columns by the second operation. Note that this order is important because most groups are not commutative. The C_{3v} group multiplication table is as follows:

	E	C_3	C_3^2	σ_v	σ_v'	σ_v''	Second operation
E	E	C_3	C_3^2	σ_v	σ_v'	σ_v''	
C_3	C_3	C_3^2	E	σ_v'	σ_v''	σ_v	
C_3^2	C_3^2	E	C_3	σ_v''	σ_v	σ_v'	
σ_v	σ_v	σ_v''	σ_v'	E	C_3^2	C_3	
σ_v'	σ_v'	σ_v	σ_v''	C_3	E	C_3^2	
σ_v''	σ_v''	σ_v'	σ_v	C_3^2	C_3	E	
First operation							

Note the reflection plane labels do not move. That is, although we start with H_1 in the σ_v plane, H_2 in σ_v' , and H_3 in σ_v'' , if H_1 moves due to the first symmetry operation, σ_v remains fixed and a different H atom lies in the σ_v plane.

4.3.2. Matrices as Group Representations

In using symmetry to help simplify molecular orbital (mo) or vibration/rotation energy-level identifications, the following strategy is followed:

1. A set of M objects belonging to the constituent atoms (or molecular fragments, in a more general case) is introduced. These objects are the orbitals of the individual atoms (or of the fragments) in the mo case; they are unit vectors along the Cartesian x , y , and z directions located on each of the atoms, and representing displacements along each of these directions, in the vibration/rotation case.
2. Symmetry tools are used to combine these M objects into M new objects each of which belongs to a specific symmetry of the point group. Because the Hamiltonian (electronic in the mo case and vibration/rotation in the latter case) commutes with the symmetry operations of the point group, the matrix representation of \mathbf{H} within the symmetry-adapted basis will be "block diagonal". That is, objects of different symmetry will not interact; only interactions among those of the same symmetry need be considered.

To illustrate such symmetry adaptation, consider symmetry adapting the $2s$ orbital of N and the three $1s$ orbitals of the three H atoms. We begin by determining how these orbitals transform under the symmetry operations of the C_{3v} point group. The act of each of the six symmetry operations on the four atomic orbitals can be denoted as follows:

$$\begin{array}{lcl} (S_N, S_1, S_2, S_3) & \xrightarrow{E} & (S_N, S_1, S_2, S_3) \\ & \xrightarrow{C_3} & (S_N, S_3, S_1, S_2) \\ & \xrightarrow{C_3^2} & (S_N, S_2, S_3, S_1) \\ & \xrightarrow{\sigma_v} & (S_N, S_1, S_3, S_2) \\ & \xrightarrow{\sigma_{v''}} & (S_N, S_3, S_2, S_1) \end{array}$$

$$\frac{\sigma_{v'}}{\rightarrow} \quad (S_N, S_2, S_1, S_3)$$

Here we are using the active view that a C_3 rotation rotates the molecule by 120° . The equivalent passive view is that the $1s$ basis functions are rotated -120° . In the C_3 rotation, S_3 ends up where S_1 began, S_1 , ends up where S_2 began and S_2 ends up where S_3 began.

These transformations can be thought of in terms of a matrix multiplying a vector with elements (S_N, S_1, S_2, S_3) . For example, if $D^{(4)}(C_3)$ is the representation matrix giving the C_3 transformation, then the above action of C_3 on the four basis orbitals can be expressed as:

$$D^{(4)}(C_3) \begin{pmatrix} S_N \\ S_1 \\ S_2 \\ S_3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} S_N \\ S_1 \\ S_2 \\ S_3 \end{pmatrix} = \begin{pmatrix} S_N \\ S_3 \\ S_1 \\ S_2 \end{pmatrix}$$

We can likewise write matrix representations for each of the symmetry operations of the C_{3v} point group:

$$D^{(4)}(C_3^2) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \end{pmatrix} \quad D^{(4)}(E) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

$$D^{(4)}(\sigma_v) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \quad D^{(4)}(\sigma_{v'}) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}$$

$$D^{(4)}(\sigma_v'') = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

It is easy to verify that a C_3 rotation followed by a σ_v reflection is equivalent to a σ_v' reflection alone. In other words

$$\sigma_v C_3 = \sigma_v', \quad \text{or} \quad \begin{array}{ccccccc} & S_1 & & S_3 & & S_3 & \\ & & \xrightarrow{C_3} & & \xrightarrow{\sigma_v} & & \\ & S_2 & & S_1 & & S_2 & S_1 \end{array}$$

Note that this same relationship is carried by the matrices:

$$D^{(4)}(\sigma_v) D^{(4)}(C_3) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} = D^{(4)}(\sigma_v')$$

Likewise we can verify that $C_3 \sigma_v = \sigma_v''$ directly and we can notice that the matrices also show the same identity:

$$D^{(4)}(C_3) D^{(4)}(\sigma_v) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = D^{(4)}(\sigma_v'')$$

In fact, one finds that the six matrices, $D^{(4)}(R)$, when multiplied together in all 36 possible ways, obey the same multiplication table as did the six symmetry operations.

We say the matrices form a representation of the group because the matrices have all the properties of the group.

4.3.3 Characters of Representations

One important property of a matrix is the sum of its diagonal elements which is called the trace of the matrix D and is denoted $\text{Tr}(D)$:

$$\text{Tr}(D) = \sum_i D_{i,i} = \chi.$$

So, χ is called the trace or character of the matrix. In the above example

$$\chi(E) = 4$$

$$\chi(C_3) = \chi(C_3^2) = 1$$

$$\chi(\sigma_v) = \chi(\sigma_v') = \chi(\sigma_v'') = 2.$$

The importance of the characters of the symmetry operations lies in the fact that they do not depend on the specific basis used to form the matrix. That is, they are invariant to a unitary or orthogonal transformation of the objects used to define the matrices. As a result, they contain information about the symmetry operation itself and about the space spanned by the set of objects. The significance of this observation for our symmetry adaptation process will become clear later.

Note that the characters of both rotations are the same as are the characters of all three reflections. Collections of operations having identical characters are called classes. Each operation in a class of operations has the same character as other members of the class. The character of a class depends on the space spanned by the basis of functions on which the symmetry operations act.

4.3.4. Another Basis and Another Representation

Above we used (S_N, S_1, S_2, S_3) as a basis. If, alternatively, we use the one-dimensional basis consisting of the $1s$ orbital on the N -atom, we obtain different characters, as we now demonstrate.

The act of the six symmetry operations on this S_N can be represented as follows:

$$\begin{array}{ccc}
S_N \xrightarrow{E} S_N & S_N \xrightarrow{C_3} S_N & S_N \xrightarrow{C_3^2} S_N; \\
S_N \xrightarrow{\sigma_v} S_N & S_N \xrightarrow{\sigma_{v'}} S_N & S_N \xrightarrow{\sigma_{v''}} S_N.
\end{array}$$

We can represent this group of operations in this basis by the one-dimensional set of matrices:

$$\begin{array}{ccc}
D^{(1)}(E) = 1; & D^{(1)}(C_3) = 1; & D^{(1)}(C_3^2) = 1, \\
D^{(1)}(\sigma_v) = 1; & D^{(1)}(\sigma_{v'}) = 1; & D^{(1)}(\sigma_{v''}) = 1.
\end{array}$$

Again we have

$$D^{(1)}(\sigma_v) D^{(1)}(C_3) = 1 \cdot 1 = D^{(1)}(\sigma_{v''}), \text{ and}$$

$$D^{(1)}(C_3) D^{(1)}(\sigma_v) = 1 \cdot 1 = D^{(1)}(\sigma_{v'}).$$

These six 1x1 matrices form another representation of the group. In this basis, each character is equal to unity. The representation formed by allowing the six symmetry operations to act on the 1s N-atom orbital is clearly not the same as that formed when the same six operations acted on the (S_N, S_1, S_2, S_3) basis. We now need to learn how to further analyze the information content of a specific representation of the group formed when the symmetry operations act on any specific set of objects.

4.3.5 Reducible and Irreducible Representations

1. Reducible Representations

Note that every matrix in the four dimensional group representation labeled $D^{(4)}$ has the so-called block diagonal form

1	0	0	0
---	---	---	---

0	A	B	C
0	D	E	F
0	G	H	I

This means that these $D^{(4)}$ matrices are really a combination of two separate group representations (mathematically, it is called a direct sum representation). We say that $D^{(4)}$ is reducible into a one-dimensional representation $D^{(1)}$ and a three-dimensional representation formed by the 3×3 submatrices that we will call $D^{(3)}$.

$$D^{(3)}(E) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad D^{(3)}(C_3) = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \quad D^{(3)}(C_3^2) = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}$$

$$D^{(3)}(\sigma_v) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \quad D^{(3)}(\sigma_v') = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} \quad D^{(3)}(\sigma_v'') = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

The characters of $D^{(3)}$ are $\chi(E) = 3$, $\chi(2C_3) = 0$, $\chi(3\sigma_v) = 1$. Note that we would have obtained this $D^{(3)}$ representation directly if we had originally chosen to examine the basis (S_1, S_2, S_3) alone; also note that these characters are equal to those of $D^{(4)}$ minus those of $D^{(1)}$.

2. A Change in Basis

Now let us convert to a new basis that is a linear combination of the original S_1, S_2, S_3 basis:

$$T_1 = S_1 + S_2 + S_3$$

$$T_2 = 2S_1 - S_2 - S_3$$

$$T_3 = S_2 - S_3$$

(Don't worry about how I constructed T_1 , T_2 , and T_3 yet. As will be demonstrated later, we form them by using symmetry projection operators defined below). We determine how the "T" basis functions behave under the group operations by allowing the operations to act on the S_j and interpreting the results in terms of the T_i . In particular,

$$(T_1, T_2, T_3) \xrightarrow{\sigma_v} (T_1, T_2, -T_3) \quad (T_1, T_2, T_3) \xrightarrow{E} (T_1, T_2, T_3);$$

$$(T_1, T_2, T_3) \xrightarrow{\sigma_{v'}} (S_3 + S_2 + S_1, 2S_3 - S_2 - S_1, S_2 - S_1) = (T_1, -1/2 T_2 - 3/2 T_3, -1/2 T_2 + 1/2 T_3);$$

$$(T_1, T_2, T_3) \xrightarrow{\sigma_{v''}} (S_2 + S_1 + S_3, 2S_2 - S_1 - S_3, S_1 - S_3) = (T_1, -1/2 T_2 + 3/2 T_3, 1/2 T_2 + 1/2 T_3);$$

$$(T_1, T_2, T_3) \xrightarrow{C_3} (S_3 + S_1 + S_2, 2S_3 - S_1 - S_2, S_1 - S_2) = (T_1, -1/2 T_2 - 3/2 T_3, 1/2 T_2 - 1/2 T_3);$$

$$(T_1, T_2, T_3) \xrightarrow{C_3^2} (S_2 + S_3 + S_1, 2S_2 - S_3 - S_1, S_3 - S_1) = (T_1, -1/2 T_2 + 3/2 T_3, -1/2 T_2 - 1/2 T_3).$$

So the matrix representations in the new T_i basis are:

$$D^{(3)}(E) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad D^{(3)}(C_3) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1/2 & -3/2 \\ 0 & 1/2 & -1/2 \end{pmatrix};$$

$$D^{(3)}(C_3^2) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1/2 & 3/2 \\ 0 & -1/2 & -1/2 \end{pmatrix} \quad D^{(3)}(\sigma_v) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix};$$

$$D^{(3)}(\sigma_v') = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1/2 & -3/2 \\ 0 & -1/2 & 1/2 \end{pmatrix} \quad D^{(3)}(\sigma_v'') = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1/2 & 3/2 \\ 0 & 1/2 & 1/2 \end{pmatrix}.$$

3. Reduction of the Reducible Representation

These six matrices can be verified to multiply just as the symmetry operations do; thus they form another three-dimensional representation of the group. We see that in the T_i basis the matrices are block diagonal. This means that the space spanned by the T_i functions, which is the same space as the S_j span, forms a reducible representation that can be decomposed into a one dimensional space and a two dimensional space (via formation of the T_i functions). Note that the characters (traces) of the matrices are not changed by the change in bases.

The one-dimensional part of the above reducible three-dimensional representation is seen to be the same as the totally symmetric representation we arrived at before, $D^{(1)}$. The two-dimensional representation that is left can be shown to be irreducible ; it has the following matrix representations:

$$D^{(2)}(E) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad D^{(2)}(C_3) = \begin{pmatrix} -1/2 & -3/2 \\ 1/2 & -1/2 \end{pmatrix} \quad D^{(2)}(C_3^2) = \begin{pmatrix} -1/2 & 3/2 \\ -1/2 & -1/2 \end{pmatrix}$$

$$D^{(2)}(\sigma_v) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad D^{(2)}(\sigma_v') = \begin{pmatrix} -1/2 & -3/2 \\ -1/2 & 1/2 \end{pmatrix} \quad D^{(2)}(\sigma_v'') = \begin{pmatrix} -1/2 & -3/2 \\ -1/2 & 1/2 \end{pmatrix}$$

The characters can be obtained by summing diagonal elements:

$$\chi(E) = 2, \chi(2C_3) = -1, \chi(3\sigma_v) = 0.$$

4. Rotations as a Basis

Another one-dimensional representation of the group can be obtained by taking rotation about the Z-axis (the C_3 axis) as the object on which the symmetry operations act:

$$\begin{array}{ccc} R_Z \xrightarrow{E} R_Z & R_Z \xrightarrow{C_3} R_Z & R_Z \xrightarrow{C_3^2} R_Z; \\ R_Z \xrightarrow{\sigma_v} -R_Z & R_Z \xrightarrow{\sigma_{v''}} -R_Z & R_Z \xrightarrow{\sigma_{v'}} -R_Z. \end{array}$$

In writing these relations, we use the fact that reflection reverses the sense of a rotation. The matrix representations corresponding to this one-dimensional basis are:

$$\begin{array}{ccc} D^{(1)}(E) = 1 & D^{(1)}(C_3) = 1 & D^{(1)}(C_3^2) = 1; \\ D^{(1)}(\sigma_v) = -1 & D^{(1)}(\sigma_{v''}) = -1 & D^{(1)}(\sigma_{v'}) = -1. \end{array}$$

These one-dimensional matrices can be shown to multiply together just like the symmetry operations of the C_{3v} group. They form an irreducible representation of the group (because it is one-dimensional, it cannot be further reduced). Note that this one-dimensional representation is not identical to that found above for the 1s N-atom orbital, or the T_1 function.

5. Overview

We have found three distinct irreducible representations for the C_{3v} symmetry group; two different one-dimensional and one two dimensional representations. Are there any more? An important theorem of group theory shows that the number of irreducible representations of a group is equal to the number of classes. Since there are three classes of operation (i.e., E, C_3 and σ_v), we have found all the irreducible representations of the C_{3v} point group. There are no more.

The irreducible representations have standard names; the first $D^{(1)}$ (that arising from the T_1 and $1s_N$ orbitals) is called A_1 , the $D^{(1)}$ arising from R_Z is called A_2

and $D^{(2)}$ is called E (not to be confused with the identity operation E). We will see shortly where to find and identify these names.

Thus, our original $D^{(4)}$ representation was a combination of two A_1 representations and one E representation. We say that $D^{(4)}$ is a direct sum representation: $D^{(4)} = 2A_1 \oplus E$. A consequence is that the characters of the combination representation $D^{(4)}$ can be obtained by adding the characters of its constituent irreducible representations.

	E	$2C_3$	$3\sigma_v$
A_1	1	1	1
A_1	1	1	1
E	2	-1	0
$2A_1 \oplus E$	4	1	2

6. How to Decompose Reducible Representations in General

Suppose you were given only the characters (4,1,2). How can you find out how many times A_1 , E, and A_2 appear when you reduce $D^{(4)}$ to its irreducible parts? You want to find a linear combination of the characters of A_1 , A_2 and E that add up (4,1,2). You can treat the characters of matrices as vectors and take the dot product of A_1 with $D^{(4)}$

$$\begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ E & C_3 & C_3^2 & \sigma_v & \sigma_{v'} & \sigma_{v''} \end{pmatrix} \begin{pmatrix} 4 & E \\ 1 & C_3 \\ 1 & C_3^2 \\ 2 & \sigma_v \\ 2 & \sigma_{v'} \\ 2 & \sigma_{v''} \end{pmatrix} = 4 + 1 + 1 + 2 + 2 + 2 = 12.$$

The vector (1,1,1,1,1,1) is not normalized; hence to obtain the component of (4,1,1,2,2,2) along a unit vector in the (1,1,1,1,1,1) direction, one must divide by the norm of (1,1,1,1,1,1); this norm is 6. The result is that the reducible representation contains $12/6 = 2$ A_1 components. Analogous projections in the E and A_2 directions give components of 1 and 0, respectively. In general, to determine the number n_Γ of times irreducible representation Γ appears in the reducible representation with characters χ_{red} , one calculates

$$n_\Gamma = \frac{1}{g} \sum_R \chi_\Gamma(R) \chi_{red}(R),$$

where g is the order of the group (i.e.. the number of operations in the group; six in our example) and $\chi_\Gamma(R)$ are the characters of the Γ^{th} irreducible representation.

7. Commonly Used Bases

We could take any set of functions as a basis for a group representation. Commonly used sets include: Cartesian displacement coordinates (x,y,z) located on the atoms of a polyatomic molecule (their symmetry treatment is equivalent to that involved in treating a set of p orbitals on the same atoms), quadratic functions such as d orbitals - xy,yz,xz,x²-y²,z², as well as rotations about the x, y and z axes. The transformation properties of these very commonly used bases are listed in the character tables shown in Section 4.4.

8. Summary

The basic idea of symmetry analysis is that any basis of orbitals, displacements, rotations, etc. transforms either as one of the irreducible representations or as a direct sum (reducible) representation. Symmetry tools are used to first determine how the basis transforms under action of the symmetry operations. They are then used to decompose the resultant representations into their irreducible components.

4.3.6. More Examples

1. The 2p Orbitals of Nitrogen

For a function to transform according to a specific irreducible representation means that the function, when operated upon by a point-group symmetry operator, yields a linear combination of the functions that transform according to that irreducible representation. For example, a $2p_z$ orbital (z is the C_3 axis of NH_3) on the nitrogen atom belongs to the A_1 representation because it yields unity times itself when C_3 , C_3^2 , σ_v , σ_v' , σ_v'' or the identity operation act on it. The factor of 1 means that $2p_z$ has A_1 symmetry since the characters (the numbers listed opposite A_1 and below E , $2C_3$, and $3\sigma_v$ in the C_{3v} character table shown in Section 4.4) of all six symmetry operations are 1 for the A_1 irreducible representation.

The $2p_x$ and $2p_y$ orbitals on the nitrogen atom transform as the E representation since C_3 , C_3^2 , σ_v , σ_v' , σ_v'' and the identity operation map $2p_x$ and $2p_y$ among one another. Specifically,

$$C_3 \begin{pmatrix} 2p_x \\ 2p_y \end{pmatrix} = \begin{pmatrix} \cos 120^\circ & -\sin 120^\circ \\ \sin 120^\circ & \cos 120^\circ \end{pmatrix} \begin{pmatrix} 2p_x \\ 2p_y \end{pmatrix}$$

$$C_3^2 \begin{pmatrix} 2p_x \\ 2p_y \end{pmatrix} = \begin{pmatrix} \cos 240^\circ & -\sin 240^\circ \\ \sin 240^\circ & \cos 240^\circ \end{pmatrix} \begin{pmatrix} 2p_x \\ 2p_y \end{pmatrix}$$

$$E \begin{pmatrix} 2p_x \\ 2p_y \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 2p_x \\ 2p_y \end{pmatrix}$$

$$\sigma_v \begin{pmatrix} 2p_x \\ 2p_y \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 2p_x \\ 2p_y \end{pmatrix}$$

$$\sigma_{v'} \begin{pmatrix} 2p_x \\ 2p_y \end{pmatrix} = \begin{pmatrix} 1/2 & \frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & -1/2 \end{pmatrix} \begin{pmatrix} 2p_x \\ 2p_y \end{pmatrix}$$

$$\sigma_{v''} \begin{pmatrix} 2p_x \\ 2p_y \end{pmatrix} = \begin{pmatrix} 1/2 & \frac{-\sqrt{3}}{2} \\ \frac{-\sqrt{3}}{2} & -1/2 \end{pmatrix} \begin{pmatrix} 2p_x \\ 2p_y \end{pmatrix}.$$

The 2 x 2 matrices, which indicate how each symmetry operation maps $2p_x$ and $2p_y$ into some combinations of $2p_x$ and $2p_y$, are the representation matrices ($D^{(\text{IR})}$) for that particular operation and for this particular irreducible representation (IR). For example,

$$\begin{pmatrix} 1/2 & \frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & -1/2 \end{pmatrix} = D^{(\text{E})}(\sigma_{v'})$$

This set of matrices have the same characters as the $D^{(2)}$ matrices obtained earlier when the T_i displacement vectors were analyzed, but the individual matrix elements are different because we used a different basis set (here $2p_x$ and $2p_y$; above it was T_2 and T_3). This illustrates the invariance of the trace to the specific representation; the trace only depends on the space spanned, not on the specific manner in which it is spanned.

2. A Short-Cut

A short-cut device exists for evaluating the trace of such representation matrices (that is, for computing the characters). The diagonal elements of the representation matrices are the projections along each orbital of the effect of the

symmetry operation acting on that orbital. For example, a diagonal element of the C_3 matrix is the component of $C_3 2p_y$ along the $2p_y$ direction. More rigorously, it is $\int 2p_y^* C_3 2p_y d\tau$. Thus, the character of the C_3 matrix is the sum of $\int 2p_y^* C_3 2p_y d\tau$ and $\int 2p_x^* C_3 2p_x d\tau$. In general, the character χ of any symmetry operation S can be computed by allowing S to operate on each orbital ϕ_i , then projecting $S\phi_i$ along ϕ_i (i.e., forming $\int \phi_i^* S\phi_i d\tau$, and summing these terms,

$$\sum_i \int \phi_i^* S\phi_i d\tau = \chi(S).$$

If these rules are applied to the $2p_x$ and $2p_y$ orbitals of nitrogen within the C_{3v} point group, one obtains

$$\chi(E) = 2, \chi(C_3) = \chi(C_3^2) = -1, \chi(\sigma_v) = \chi(\sigma_v'') = \chi(\sigma_v') = 0.$$

This set of characters is the same as $D^{(2)}$ above and agrees with those of the E representation for the C_{3v} point group. Hence, $2p_x$ and $2p_y$ belong to or transform as the E representation. This is why (x,y) is to the right of the row of characters for the E representation in the C_{3v} character table shown in Section 4.4. In similar fashion, the C_{3v} character table (please refer to this table now) states that $d_{x^2-y^2}$ and d_{xy} orbitals on nitrogen transform as E , as do d_{xy} and d_{yz} , but d_{z^2} transforms as A_1 .

Earlier, we considered in some detail how the three $1s_H$ orbitals on the hydrogen atoms transform. Repeating this analysis using the short-cut rule just described, the traces (characters) of the 3×3 representation matrices are computed by allowing E , $2C_3$, and $3\sigma_v$ to operate on $1s_{H_1}$, $1s_{H_2}$, and $1s_{H_3}$ and then computing the component of the resulting function along the original function. The resulting characters are $\chi(E) = 3$, $\chi(C_3) = \chi(C_3^2) = 0$, and $\chi(\sigma_v) = \chi(\sigma_v') = \chi(\sigma_v'') = 1$, in agreement with what we calculated before.

Using the orthogonality of characters taken as vectors we can reduce the above set of characters to $A_1 + E$. Hence, we say that our orbital set of three $1s_H$

orbitals forms a reducible representation consisting of the sum of A_1 and E IR's. This means that the three $1s_H$ orbitals can be combined to yield one orbital of A_1 symmetry and a pair that transform according to the E representation.

4.3.7. Projector Operators: Symmetry Adapted Linear Combinations of Atomic Orbitals

To generate the above A_1 and E symmetry-adapted orbitals, we make use of so-called symmetry projection operators P_E and P_{A_1} . These operators are given in terms of linear combinations of products of characters times elementary symmetry operations as follows:

$$P_{A_1} = \sum_S \chi_{A_1}(S)S$$

$$P_E = \sum_S \chi_E(S)S$$

where S ranges over $C_3, C_3^2, \sigma_v, \sigma_v'$ and σ_v'' and the identity operation. The result of applying P_{A_1} to say $1s_{H_1}$ is

$$\begin{aligned} P_{A_1} 1s_{H_1} &= 1s_{H_1} + 1s_{H_2} + 1s_{H_3} + 1s_{H_2} + 1s_{H_3} + 1s_{H_1} \\ &= 2(1s_{H_1} + 1s_{H_2} + 1s_{H_3}) = \phi_{A_1}, \end{aligned}$$

which is an (unnormalized) orbital having A_1 symmetry. Clearly, this same ϕ_{A_1} orbital would be generated by P_{A_1} acting on $1s_{H_2}$ or $1s_{H_3}$. Hence, only one A_1 orbital exists. Likewise,

$$P_E 1s_{H_1} = 2 \cdot 1s_{H_1} - 1s_{H_2} - 1s_{H_3} = \phi_{E,1}$$

which is one of the symmetry adapted orbitals having E symmetry. The other E orbital can be obtained by allowing P_E to act on $1s_{H_2}$ or $1s_{H_3}$:

$$P_E 1s_{H_2} = 2 \cdot 1s_{H_2} - 1s_{H_1} - 1s_{H_3} \equiv \phi_{E,2}$$

$$P_E 1s_{H_3} = 2 \cdot 1s_{H_3} - 1s_{H_1} - 1s_{H_2} = \phi_{E,3}.$$

It might seem as though three orbitals having E symmetry were generated, but only two of these are really independent functions. For example, $\phi_{E,3}$ is related to $\phi_{E,1}$ and $\phi_{E,2}$ as follows:

$$\phi_{E,3} = -(\phi_{E,1} + \phi_{E,2}).$$

Thus, only $\phi_{E,1}$ and $\phi_{E,2}$ are needed to span the two-dimensional space of the E representation. If we include $\phi_{E,1}$ in our set of orbitals and require our orbitals to be orthogonal, then we must find numbers a and b such that $\phi'_E = a\phi_{E,2} + b\phi_{E,3}$ is orthogonal to $\phi_{E,1}$: $\int \phi_E^* \phi_E d\tau = 0$. A straightforward calculation gives $a = -b$ or $\phi'_E = a(1s_{H_2} - 1s_{H_3})$ which agrees with what we used earlier to construct the T_i functions in terms of the S_j functions.

4.3.8. Summary

Let us now summarize what we have learned thus far about point group symmetry. Any given set of atomic orbitals $\{\phi_i\}$, atom-centered displacements, or rotations can be used as a basis for the symmetry operations of the point group of the molecule. The characters $\chi(S)$ belonging to the operations S of this point group within any such space can be found by summing the integrals $\int \phi_i^* S\phi_i d\tau$ over all the atomic orbitals (or corresponding unit vector atomic displacements or rotations). The

resultant characters will, in general, be reducible to a combination of the characters of the irreducible representations $\chi_i(S)$. To decompose the characters $\chi(S)$ of the reducible representation to a sum of characters $\chi_i(S)$ of the irreducible representation

$$\chi_{\text{red}}(S) = \sum_k \int \phi_k^* S \phi_k d\tau = \sum_i n_i \chi_i(S),$$

it is necessary to determine how many times, n_i , the i -th irreducible representation occurs in the reducible representation. The expression for n_i is

$$n_i = \frac{1}{g} \sum_R \chi_i(R) \chi_{\text{red}}(R)$$

in which g is the order of the point group- the total number of symmetry operations in the group (e.g., $g = 6$ for C_{3v}).

For example, the reducible representation $\chi(E) = 3$, $\chi(C_3) = 0$, and $\chi(\sigma_v) = 1$ formed by the three $1s_H$ orbitals discussed above can be decomposed as follows:

$$n_{A_1} = 1/6 (3 \cdot 1 + 2 \cdot 0 \cdot 1 + 3 \cdot 1 \cdot 1) = 1,$$

$$n_{A_2} = 1/6 (3 \cdot 1 + 2 \cdot 0 \cdot 1 + 3 \cdot 1 \cdot -1) = 0,$$

$$n_E = 1/6 (3 \cdot 2 + 2 \cdot 0 \cdot -1 + 3 \cdot 1 \cdot 0) = 1.$$

These equations state that the three $1s_H$ orbitals can be combined to give one A_1 orbital and, since E is degenerate, one pair of E orbitals, as established above. With knowledge of the n_i , the symmetry-adapted orbitals can be formed by allowing the projectors

$$P_i = \sum_S \chi_i(S) S$$

to operate on each of the primitive atomic orbitals. How this is carried out was illustrated for the $1s_H$ orbitals in our earlier discussion. These tools allow a symmetry decomposition of any set of atomic orbitals into appropriate symmetry-adapted orbitals.

Before considering other concepts and group-theoretical machinery, it should once again be stressed that these same tools can be used in symmetry analysis of the translational, vibrational and rotational motions of a molecule. The twelve motions of NH_3 (three translations, three rotations, six vibrations) can be described in terms of combinations of displacements of each of the four atoms in each of three (x,y,z) directions. Hence, unit vectors placed on each atom directed in the x, y, and z directions form a basis for action by the operations $\{S\}$ of the point group. In the case of NH_3 , the characters of the resultant 12×12 representation matrices form a reducible representation in the C_{2v} point group: $\chi(E) = 12, \chi(C_3) = \chi(C_3^2) = 0, \chi(\sigma_v) = \chi(\sigma_v') = \chi(\sigma_v'') = 2$. For example under σ_v , the H_2 and H_3 atoms are interchanged, so unit vectors on either one will not contribute to the trace. Unit z-vectors on N and H_1 remain unchanged as well as the corresponding y-vectors. However, the x-vectors on N and H_1 are reversed in sign. The total character for σ_v' the H_2 and H_3 atoms are interchanged, so unit vectors on either one will not contribute to the trace. Unit z-vectors on N and H_1 remain unchanged as well as the corresponding y-vectors. However, the x-vectors on N and H_1 are reversed in sign. The total character for σ_v is thus $4 - 2 = 2$. This representation can be decomposed as follows:

$$n_{A_1} = 1/6 (1 \cdot 1 \cdot 12 + 2 \cdot 1 \cdot 0 + 3 \cdot 1 \cdot 2) = 3,$$

$$n_{A_2} = 1/6 (1 \cdot 1 \cdot 12 + 2 \cdot 1 \cdot 0 + 3 \cdot -1 \cdot 2) = 1,$$

$$n_E = 1/6 (1 \cdot 2 \cdot 12 + 2 \cdot -1 \cdot 0 + 3 \cdot 0 \cdot 2) = 4.$$

From the information on the right side of the C_{3v} character table, translations of all

four atoms in the z, x and y directions transform as $A_1(z)$ and $E(x,y)$, respectively, whereas rotations about the z(R_z), x(R_x), and y(R_y) axes transform as A_2 and E. Hence, of the twelve motions, three translations have A_1 and E symmetry and three rotations have A_2 and E symmetry. This leaves six vibrations, of which two have A_1 symmetry, none have A_2 symmetry, and two (pairs) have E symmetry. We could obtain symmetry-adapted vibrational and rotational bases by allowing symmetry projection operators of the irreducible representation symmetries to operate on various elementary Cartesian (x,y,z) atomic displacement vectors centered on the four atoms.

4.3.9 Direct Product Representations

1. Direct Products in N-Electron Wave functions

We now turn to the symmetry analysis of orbital products. Such knowledge is important because one is routinely faced with constructing symmetry-adapted N-electron configurations that consist of products of N individual spin orbitals, one for each electron. A point-group symmetry operator S, when acting on such a product of orbitals, gives the product of S acting on each of the individual orbitals

$$S(\phi_1\phi_2\phi_3\dots\phi_N) = (S\phi_1) (S\phi_2) (S\phi_3) \dots (S\phi_N).$$

For example, reflection of an N-orbital product through the σ_v plane in NH_3 applies the reflection operation to all N electrons.

Just as the individual orbitals formed a basis for action of the point-group operators, the configurations (N-orbital products) form a basis for the action of these same point-group operators. Hence, the various electronic configurations can be treated as functions on which S operates, and the machinery illustrated earlier for decomposing orbital symmetry can then be used to carry out a symmetry analysis of configurations.

Another shortcut makes this task easier. Since the symmetry adapted individual orbitals $\{\phi_i, i = 1, \dots, M\}$ transform according to irreducible representations,

the representation matrices for the N-term products shown above consist of products of the matrices belonging to each ϕ_i . This matrix product is not a simple product but what is called a direct product. To compute the characters of the direct product matrices, one multiplies the characters of the individual matrices of the irreducible representations of the N orbitals that appear in the electron configuration. The direct-product representation formed by the orbital products can therefore be symmetry-analyzed (reduced) using the same tools as we used earlier.

For example, if one is interested in knowing the symmetry of an orbital product of the form $a_1^2 a_2^2 e^2$ (note: lower case letters are used to denote the symmetry of electronic orbitals, whereas capital letters are reserved to label the overall configuration's symmetry) in C_{3v} symmetry, the following procedure is used. For each of the six symmetry operations in the C_{2v} point group, the product of the characters associated with each of the six spin orbitals (orbital multiplied by α or β spin) is formed

$$\chi(S) = \prod_i \chi_i(S) = (\chi_{A_1}(S))^2 (\chi_{A_2}(S))^2 (\chi_E(S))^2.$$

In the specific case considered here, $\chi(E) = 4$, $\chi(2C_3) = 1$, and $\chi(3\sigma_v) = 0$. Notice that the contributions of any doubly occupied non-degenerate orbitals (e.g., a_1^2 , and a_2^2) to these direct product characters $\chi(S)$ are unity because for all operators $(\chi_k(S))^2 = 1$ for any one-dimensional irreducible representation. As a result, only the singly occupied or degenerate orbitals need to be considered when forming the characters of the reducible direct-product representation $\chi(S)$. For this example this means that the direct-product characters can be determined from the characters $\chi_E(S)$ of the two active (i.e., non-closed-shell) orbitals - the e^2 orbitals. That is, $\chi(S) = \chi_E(S) \cdot \chi_E(S)$.

From the direct-product characters $\chi(S)$ belonging to a particular electronic configuration (e.g., $a_1^2 a_2^2 e^2$), one must still decompose this list of characters into a sum of irreducible characters. For the example at hand, the direct-product characters

$\chi(S)$ decompose into one A_1 , one A_2 , and one E representation. This means that the e^2 configuration contains A_1 , A_2 , and E symmetry elements. Projection operators analogous to those introduced earlier for orbitals can be used to form symmetry-adapted orbital products from the individual basis orbital products of the form $a_1^2 a_2^2 e_x^m e_y^{m'}$, where m and m' denote the occupation (1 or 0) of the two degenerate orbitals e_x and e_y . In Appendix III of Electronic Spectra and Electronic Structure of Polyatomic Molecules, G. Herzberg, Van Nostrand Reinhold Co., New York, N.Y. (1966) the resolution of direct products among various representations within many point groups are tabulated.

When dealing with indistinguishable particles such as electrons, it is also necessary to further project the resulting orbital products to make them antisymmetric (for Fermions) or symmetric (for Bosons) with respect to interchange of any pair of particles. This step reduces the set of N -electron states that can arise. For example, in the above e^2 configuration case, only 3A_2 , 1A_1 , and 1E states arise; the 3E , 3A_1 , and 1A_2 possibilities disappear when the antisymmetry projector is applied. In contrast, for an $e^1 e'^1$ configuration, all states arise even after the wave function has been made antisymmetric. The steps involved in combining the point group symmetry with permutational antisymmetry are illustrated in Chapter 6 of this text as well as in Chapter 10 of my QMIC text.

2. Direct Products in Selection Rules

Two states ψ_a and ψ_b that are eigenfunctions of a Hamiltonian \mathbf{H}_0 in the absence of some external perturbation (e.g., electromagnetic field or static electric field or potential due to surrounding ligands) can be "coupled" by the perturbation \mathbf{V} only if the symmetries of \mathbf{V} and of the two wave functions obey a so-called selection rule. In particular, only if the coupling integral

$$\int \psi_a^* V \psi_b d\tau = V_{a,b}$$

is non-vanishing will the two states be coupled by \mathbf{V} .

The role of symmetry in determining whether such integrals are non-zero can be demonstrated by noting that the integrand, considered as a whole, must contain a component that is invariant under all of the group operations (i.e., belongs to the totally symmetric representation of the group) if the integral is to not vanish. In terms of the projectors introduced above we must have

$$\sum_S \chi_A(S) S[\psi_a^* S \psi_b]$$

not vanish. Here the subscript A denotes the totally symmetric representation of whatever point group applies. The symmetry of the product $\psi_a^* \mathbf{V} \psi_b$ is, according to what was covered earlier, given by the direct product of the symmetries of ψ_a^* of \mathbf{V} and of ψ_b . So, the conclusion is that the integral will vanish unless this triple direct product contains, when it is reduced to its irreducible components, a component of the totally symmetric representation.

Another way to state the above result, and a way this is more often used in practice, is that the integral $\int \psi_a^* V \psi_b d\tau$ will vanish unless the symmetry of the direct product $V\psi_b$ matches the symmetry of ψ_a^* . Only when these symmetries match will the triple direct product contain a non-zero component of the totally symmetric representation. This is very much the same as what we saw earlier in this Chapter when we discussed how angular momentum coupling could limit which states contribute to the second-order perturbation theory energy. The angular momenta of \mathbf{V} and of ψ_b , when coupled, must have a component that matches the angular momentum of ψ_a .

To see how this result is used, consider the integral that arises in formulating the interaction of electromagnetic radiation with a molecule within the electric-dipole approximation:

$$\int \psi_a^* r \psi_b d\tau .$$

Here, \mathbf{r} is the vector giving, together with e , the unit charge, the quantum mechanical dipole moment operator

$$\mathbf{r} = e \sum_n Z_n \mathbf{R}_n - e \sum_i \mathbf{r}_i,$$

where Z_n and \mathbf{R}_n are the charge and position of the n^{th} nucleus and \mathbf{r}_j is the position of the j^{th} electron. Now, consider evaluating this integral for the singlet $n \rightarrow \pi^*$ transition in formaldehyde. Here, the closed-shell ground state is of 1A_1 symmetry and the singlet excited state, which involves promoting an electron from the non-bonding b_2 lone pair orbital on the Oxygen atom into the anti-bonding π^* b_1 orbital on the CO moiety, is of 1A_2 symmetry ($b_1 \times b_2 = a_2$). The direct product of the two wave function symmetries thus contains only a_2 symmetry. The three components (x, y, and z) of the dipole operator have, respectively, b_1 , b_2 , and a_1 symmetry. Thus, the triple direct products give rise to the following possibilities:

$$a_2 \times b_1 = b_2,$$

$$a_2 \times b_2 = b_1,$$

$$a_2 \times a_1 = a_2.$$

There is no component of a_1 symmetry in the triple direct product, so the integral vanishes. The alternative way of reaching this same conclusion is to notice that the direct product of the symmetries of the π^* b_1 orbital and the b_2 lone pair orbital is a_2 ($b_1 \times b_2 = a_2$), which does not match the symmetry of any component of the dipole operator. Either route allows us to conclude that the $n \rightarrow \pi^*$ excitation in formaldehyde is electric dipole forbidden.

4.3.10 Overview

We have shown how to make a symmetry decomposition of a basis of atomic orbitals (or Cartesian displacements or orbital products) into irreducible representation components. This tool is very helpful when studying spectroscopy and when constructing the orbital correlation diagrams that form the basis of the Woodward-Hoffmann rules that play useful roles in predicting whether chemical reactions will have energy barriers in excess of thermodynamic barriers. We also learned how to form the direct-product symmetries that arise when considering configurations consisting of products of symmetry-adapted spin orbitals. Finally, we learned how the direct product analysis allows one to determine whether or not integrals of products of wave functions with operators between them vanish. This tool is of utmost importance in determining selection rules in spectroscopy and for determining the effects of external perturbations on the states of the species under investigation.

4.4 Character Tables

Point Group Character Tables

C_1	E
A	1

C_s	E	σ_h		
A'	1	1	x,y,R _z	x^2,y^2,z^2,xy
A''	1	-1	z,R _x ,R _y	yz,xz

C_i	E	i		
A _g	1	1	R _x ,R _y ,R _z	x^2,y^2,z^2,xy,xz,yz

A_u	1	-1			x,y,z	
-------	---	----	--	--	---------	--

C_2	E	C_2				
A	1	1			z,R_z	x^2,y^2,z^2,xy
B	1	-1			x,y,R_x,R_y	yz,xz

D_2	E	$C_2(z)$	$C_2(y)$	$C_2(x)$		
A	1	1	1	1		x^2,y^2,z^2
B_1	1	1	-1	-1	z,R_z	xy
B_2	1	-1	1	-1	y,R_y	xz
B_3	1	-1	-1	1	x,R_x	yz

D_3	E	$2C_3$	$3C_2$			
A_1	1	1	1			x^2+y^2,z^2
A_2	1	1	-1	z,R_z		
E	2	-1	0	$(x,y)(R_x,R_y)$		$(x^2-y^2,xy)(xz,yz)$

D ₄	E	2C ₄	C ₂ (=C ₄ ²)	2C ₂ '	2C ₂ "		
A ₁	1	1	1	1	1		x ² +y ² ,z ²
A ₂	1	1	1	-1	-1	z,R _Z	
B ₁	1	-1	1	1	-1		x ² -y ²
B ₂	1	-1	1	-1	1		xy
E	2	0	-2	0	0	(x,y)(R _x ,R _y)	(xz,yz)

C _{2v}	E	C ₂	σ _v (xz)	σ _v '(yz)		
A ₁	1	1	1	1	z	x ² ,y ² ,z ²
A ₂	1	1	-1	-1	R _Z	xy
B ₁	1	-1	1	-1	x,R _y	xz
B ₂	1	-1	-1	1	y,R _x	yz

C _{3v}	E	2C ₃	3σ _v		
A ₁	1	1	1	z	x ² +y ² ,z ²
A ₂	1	1	-1	R _Z	
E	2	-1	0	(x,y)(R _x ,R _y)	(x ² -y ² ,xy)(xz,yz)

C _{4v}	E	2C ₄	C ₂	2σ _v	2σ _d		
A ₁	1	1	1	1	1	z	x ² +y ² ,z ²
A ₂	1	1	1	-1	-1	R _Z	
B ₁	1	-1	1	1	-1		x ² -y ²
B ₂	1	-1	1	-1	1		xy
E	2	0	-2	0	0	(x,y)(R _x ,R _y)	(xz,yz)

C_{2h}	E	C_2	i	σ_h		
A_g	1	1	1	1	R_z	x^2, y^2, z^2, xy
B_g	1	-1	1	-1	R_x, R_y	xz, yz
A_u	1	1	-1	-1	z	
B_u	1	-1	-1	1	x, y	

D _{2h}	E	C ₂ (z)	C ₂ (y)	C ₂ (x)	i	σ(xy)	σ(xz)	σ(yz)		
A _g	1	1	1	1	1	1	1	1		x ² ,y ² ,z ²
B _{1g}	1	1	-1	-1	1	1	-1	-1	R _z	xy
B _{2g}	1	-1	1	-1	1	-1	1	-1	R _y	xz
B _{3g}	1	-1	-1	1	1	-1	-1	1	R _x	yz
A _u	1	1	1	1	-1	-1	-1	-1		
B _{1u}	1	1	-1	-1	-1	-1	1	1	z	
B _{2u}	1	-1	1	-1	-1	1	-1	1	y	
B _{3u}	1	-1	-1	1	-1	1	1	-1	x	

D _{3h}	E	2C ₃	3C ₂	σ _h	2S ₃	3σ _v		
A ₁ '	1	1	1	1	1	1		x ² +y ² ,z ²
A ₂ '	1	1	-1	1	1	-1	R _z	
E'	2	-1	0	2	-1	0	(x,y)	(x ² -y ² ,xy)
A ₁ "	1	1	1	-1	-1	-1		
A ₂ "	1	1	-1	-1	-1	1	z	
E"	2	-1	0	-2	1	0	(R _x ,R _y)	(xz,yz)

D _{4h}	E	2C ₄	C ₂	2C ₂ '	2C ₂ "	i	2S ₄	σ _h	2σ _v	2σ _d		
A _{1g}	1	1	1	1	1	1	1	1	1	1		x ² +y ² ,z ²
A _{2g}	1	1	1	-1	-1	1	1	1	-1	-1	R _z	
B _{1g}	1	-1	1	1	-1	1	-1	1	1	-1		x ² -y ²
B _{2g}	1	-1	1	-1	1	1	-1	1	-1	1		xy
E _g	2	0	-2	0	0	2	0	-2	0	0	(R _x ,R _y)	(xz,yz)
A _{1u}	1	1	1	1	1	-1	-1	-1	-1	-1		

$$\begin{array}{l}
 A_{2u} \\
 B_{1u} \\
 B_{2u} \\
 E_u
 \end{array}
 \left| \begin{array}{cccccccccc}
 1 & 1 & 1 & -1 & -1 & -1 & -1 & -1 & 1 & 1 \\
 1 & -1 & 1 & 1 & -1 & -1 & 1 & -1 & -1 & 1 \\
 1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 \\
 2 & 0 & -2 & 0 & 0 & -2 & 0 & 2 & 0 & 0
 \end{array} \right| \begin{array}{l}
 z \\
 \\
 \\
 (x,y)
 \end{array}$$

D _{6h}	E	2C ₆	2C ₃	C ₂	3C ₂ '	3C ₂ ''	i	2S ₃	2S ₆	σ _h	3σ _d	3σ _v		
A _{1g}	1	1	1	1	1	1	1	1	1	1	1	1		x ² +y ² ,z ²
A _{2g}	1	1	1	1	-1	-1	1	1	1	1	-1	-1	R _z	
B _{1g}	1	-1	1	-1	1	-1	1	-1	1	-1	1	-1		
B _{2g}	1	-1	1	-1	-1	1	1	-1	1	-1	-1	1		
E _{1g}	2	1	-1	-2	0	0	2	1	-1	-2	0	0	(R _x ,R _y)	(xz,yz)
E _{2g}	2	-1	-1	2	0	0	2	-1	-1	2	0	0		(x ² -y ² ,xy)
A _{1u}	1	1	1	1	1	1	-1	-1	-1	-1	-1	-1		
A _{2u}	1	1	1	1	-1	-1	-1	-1	-1	-1	1	1	z	
B _{1u}	1	-1	1	-1	1	-1	-1	1	-1	1	-1	1		
B _{2u}	1	-1	1	-1	-1	1	-1	1	-1	1	1	-1		
E _{1u}	2	1	-1	-2	0	0	-2	-1	1	2	0	0	(x,y)	
E _{2u}	2	-1	-1	2	0	0	-2	1	1	-2	0	0		

D _{2d}	E	2S ₄	C ₂	2C ₂ '	2σ _d		
A ₁	1	1	1	1	1		x ² +y ² ,z ²
A ₂	1	1	1	-1	-1	R _z	
B ₁	1	-1	1	1	-1		x ² -y ²
B ₂	1	-1	1	-1	1	z	xy
E	2	0	-2	0	0	(x,y)(R _x ,R _y)	(xz,yz)

D _{3d}	E	2C ₃	3C ₂	i	2S ₆	3σ _d		
A _{1g}	1	1	1	1	1	1		x ² +y ² ,z ²
A _{2g}	1	1	-1	1	1	-1	R _z	
E _g	2	-1	0	2	-1	0	(R _x ,R _y)	(x ² -y ² ,xy)(xz,yz)
A _{1u}	1	1	1	-1	-1	-1		

A _{2u}	1	1	-1	-1	-1	1	z
E _u	2	-1	0	-2	1	0	(x,y)

S ₄	E	S ₄	C ₂	S ₄ ³		
A	1	1	1	1	R _z	x ² +y ² ,z ²
B	1	-1	1	-1	z	x ² -y ² ,xy
E	{1 1}	{i -i}	{-1 -1}	{-i i}	(x,y)(R _x ,R _y)	(xz,yz)

T	E	4C ₃	4C ₃ ²	3C ₂		ε=exp(2πi/3)
A	1	1	1	1		x ² +y ² +z ²
E	{1 1}	{ε ε*}	{ε* ε}	{1 -1}		(2z ² -x ² -y ² ,x ² -y ²)
T	3	0	0	-1	(R _x ,R _y ,R _z)(x,y,z)	(xy,xz,yz)

T _h	E	4C ₃	4C ₃ ²	3C ₂	i	4S ₆	4S ₆ ⁵	3σ _h		ε=exp(2πi/3)
A _g	1	1	1	1	1	1	1	1		x ² +y ² +z ²
A _u	1	1	1	1	-1	-1	-1	-1		
E _g	{1 1}	{ε ε*}	{ε* ε}	{1 1}	{1 1}	{ε ε*}	{ε* ε}	{1 1}		(2z ² -x ² -y ² , x ² -y ²)
E _u	{1 1}	{ε ε*}	{ε* ε}	{1 1}	{-1 -1}	{-ε -ε*}	{-ε* -ε}	{-1 -1}		
T _g	3	0	0	-1	1	0	0	-1	(R _x ,R _y ,R _z)	
T _u	3	0	0	-1	-1	0	0	1	(x,y,z)	(xy,xz,yz)

T _d	E	8C ₃	3C ₂	6S ₄	6σ _d		
A ₁	1	1	1	1	1		x ² +y ² +z ²
A ₂	1	1	1	-1	-1		
E	2	-1	2	0	0		(2z ² -x ² -y ² , x ² -y ²)
T ₁	3	0	-1	1	-1	(R _x , R _y , R _z)	
T ₂	3	0	-1	-1	1	(x, y, z)	(xy, xz, yz)

O	E	6C ₄	3C ₂	8C ₃	6C ₂		
			(=C ₄ ²)				
A ₁	1	1	1	1	1		x ² +y ² +z ²
A ₂	1	-1	1	1	-1		
E	2	0	2	-1	0		(2z ² -x ² -y ² , x ² -y ²)
T ₁	3	1	-1	0	-1	(R _x , R _y , R _z)(x, y, z)	
T ₂	3	-1	-1	0	1		(xy, xz, yz)

O _h	E	8C ₃	6C ₂	6C ₄	3C ₂	i	6S ₄	8S ₆	3σ _h	6σ _d		
					(=C ₄ ²)							
A _{1g}	1	1	1	1	1	1	1	1	1	1		x ² +y ² +z ²
A _{2g}	1	1	-1	-1	1	1	-1	1	1	-1		
E _g	2	-1	0	0	2	2	0	-1	2	0		(2z ² -x ² -y ² , x ² -y ²)
T _{1g}	3	0	-1	1	-1	3	1	0	-1	-1	(R _x , R _y , R _z)	
T _{2g}	3	0	1	-1	-1	3	-1	0	-1	1		(xy, xz, yz)
A _{1u}	1	1	1	1	1	-1	-1	-1	-1	-1		
A _{2u}	1	1	-1	-1	1	-1	1	-1	-1	1		

E_u	2	-1	0	0	2	-2	0	1	-2	0	(x,y,z)
T_{1u}	3	0	-1	1	-1	-3	-1	0	1	1	
T_{2u}	3	0	1	-1	-1	-3	1	0	1	-1	

$C_{\infty v}$	E	$2C_{\infty\Phi}$...	$\infty\sigma_v$		
$A_1 \equiv \Sigma^+$	1	1	...	1	z	x^2+y^2, z^2
$A_2 \equiv \Sigma^-$	1	1	...	-1	R_z	
$E_1 \equiv \Pi$	2	$2\cos\Phi$...	0	$(x,y)(R_x,R_y)$	(xz,yz)
$E_2 \equiv \Delta$	2	$2\cos 2\Phi$...	0		(x^2-y^2, xy)
$E_3 \equiv \Phi$	2	$2\cos 3\Phi$...	0		
...		

$D_{\infty h}$	E	$2C_{\infty}\Phi$...	$\infty\sigma_v$	i	$2S_{\infty}\Phi$...	∞C_2		
Σ_g^+	1	1	...	1	1	1	...	1		x^2+y^2, z^2
Σ_g^-	1	1	...	-1	1	1	...	-1	R_z	
Π_g	2	$2\cos\Phi$...	0	2	$-2\cos\Phi$...	0	(R_x, R_y)	(xz, yz)
Δ_g	2	$2\cos 2\Phi$...	0	2	$2\cos 2\Phi$...	0		(x^2-y^2, xy)
...		
Σ_u^+	1	1	...	1	-1	-1	...	-1	z	
Σ_u^-	1	1	...	-1	-1	-1	...	1		
Π_u	2	$2\cos\Phi$...	0	-2	$2\cos\Phi$...	0	(x, y)	
Δ_u	2	$2\cos 2\Phi$...	0	-2	$-2\cos 2\Phi$...	0		
...		

4.5 Time Dependent Perturbation Theory

When dealing with the effects of external perturbations (e.g., applied fields, collisions with other species), one needs to have a way to estimate the probabilities and rates of transitions among states of the system of interest induced by these perturbations. Time-dependent perturbation theory (TDPT) offers a framework within which such estimates can be achieved.

In deriving the working equations of TDPT, one begins with the time-dependent Schrödinger equation

$$i\hbar \frac{\partial \Psi}{\partial t} = [H^0 + V(t)]\Psi$$

in which H^0 is the Hamiltonian for the system whose transitions are to be probed, and $V(t)$ is the perturbation caused by the external field or the collision. The wave function that solves this equation is expanded in an order-by-order manner as in conventional perturbation theory

$$\Psi = \psi^0(r) \exp(-it \frac{E^0}{\hbar}) + \psi^1 + \dots$$

Here ψ^0 is the eigenfunction of H^0 from which transitions to other eigenstates (denoted ψ_f^0) of H^0 are being considered. Because, in the absence of the external perturbation $V(t)$, the states of H^0 are known to vary with time as $\exp(-it \frac{E_f^0}{\hbar})$, this component of the time dependence of the total wave function is included in the above expansion. Then, the first-order correction ψ^1 is expanded in terms of the complete set of states $\{\psi_f^0\}$ after which the expansion coefficients $\{C_f^1(t)\}$ become the unknowns to be solved for

$$\psi^1 = \sum_f \psi_f^0(r) \exp(-it \frac{E_f^0}{\hbar}) C_f^1(t).$$

It should be noted that this derivation treats the zeroth-order states $\{\psi^0$ and $\psi_f^0\}$ as eigenfunctions of H^0 . However, in most practical applications of TDPT, $\{\psi^0$ and $\psi_f^0\}$ are not known exactly and, in fact, are usually approximated by using variational or perturbative methods (e.g., to treat differences between HF mean-field and true Coulombic interactions among electrons). So, the derivation of TDPT that we are pursuing assumes the $\{\psi^0$ and $\psi_f^0\}$ are exact eigenfunctions. When the final TDPT working equations are thus obtained, one usually substitutes perturbative or variational approximations to $\{\psi^0$ and $\psi_f^0\}$ into these equations.

Substituting the order-by-order expansion into the Schrödinger equation gives, for the left- and right-hand sides,

$$i\hbar \frac{\partial \Psi}{\partial t} = E^0 \psi^0 \exp(-it \frac{E^0}{\hbar}) + \sum_f \{E_f^0 \psi_f^0(r) \exp(-it \frac{E_f^0}{\hbar}) C_f^1(t) - i\hbar \psi_f^0(r) \exp(-it \frac{E_f^0}{\hbar}) \frac{dC_f^1(t)}{dt}\}$$

and

$$[H^0 + V(t)]\Psi = E^0\psi^0 \exp(-it\frac{E^0}{\hbar}) + \sum_f E_f^0\psi_f^0(r) \exp(-it\frac{E_f^0}{\hbar})C_f^1(t) + V(t)\psi^0 \exp(-it\frac{E^0}{\hbar}),$$

respectively, through first-order. Multiplying each of these equations on the left by the complex conjugate of a particular ψ_f^0 and integrating over the variables that H^0 depends on produces the following equation for the unknown first-order coefficients

$$-i\hbar\frac{dC_f^1(t)}{dt} = \langle \psi_f^0 | V(t) | \psi^0(r) \rangle \exp(-it\frac{(E^0 - E_f^0)}{\hbar}).$$

The states ψ^0 and ψ_f^0 can be different electronic states, vibrational states, or rotational states. In Chapter 15 of my book *Quantum Mechanics in Chemistry* referred to in Chapter 1, I treat each of these types of transitions in detail. In the present discussion, I will limit myself to the general picture of TDPT, rather than focusing on any of these particular forms of spectroscopic transitions.

To proceed further, one needs to say something about how the perturbation $V(t)$ depends on time. In the most common application of TDPT, the perturbation is assumed to consist of a term that depends on spatial variables (denoted $v(\mathbf{r})$) multiplied by a time-dependent factor of sinusoidal character. An example of such a perturbation is provided by the electric dipole potential

$$V(t) = \mathcal{E} \cdot [e \sum_n Z_n R_n - e \sum_i r_i] \cos(\omega t)$$

characterizing photons of frequency ω interacting with the nuclei and electrons of a molecule. $\mathcal{E} \cdot [e \sum_n Z_n R_n - e \sum_i r_i]$ is the spatial part $v(\mathbf{r})$ and $\cos(\omega t)$ is the time-dependence.

To allow for the possibility that photons over a range of frequencies may impinge on the molecules, we can proceed with the derivation for photons of a given

frequency ω and, after obtaining our final result, average over a distribution of frequencies characterized by a function $f(\omega)$ giving the number of photons with frequencies between ω and $\omega + d\omega$. For perturbations that do not vary in a sinusoidal manner (e.g., a perturbation arising from a collision with another molecule), the derivation follows a different path at this point (later in this Chapter, I offer an example). Because spectroscopic time-dependent perturbations are extremely common in chemistry, we will focus much of our attention to this class of perturbations in this Chapter. For the reader who wishes a more complete and diverse treatment of TDPT as applied to chemistry, I suggest the text *Radiation and Noise in Quantum Electronics*, W. H. Louisell, R. E. Krieger, Pub., Huntington, N. Y. (1977) as well as my text *Quantum Mechanics in Chemistry* mentioned above.

To proceed deriving the working equations of TDPT, the above expression for $V(t)$ is inserted into the differential equation for the expansion coefficients and the equation is integrated from an initial time t_i to a final time t_f . These times describe when the external perturbation is first turned on and when it is turned off, respectively. For example, a laser whose photon intensity profile is described by $f(\omega)$ might be pulsed on from t_i to t_f , and one wants to know what fraction of the molecules initially in ψ^0 have undergone transitions to each of the ψ_f^0 . Alternatively, the molecules may be flowing in a stream that passes through a laser light source that is continually on, entering the laser beam at t_i and exiting from the laser beam at t_f . In either case, the molecules would be exposed to the photons from t_i until t_f . The result of integrating the differential equation is

$$C_f^1(t) = \frac{-1}{2i\hbar} \int_{t_0}^{t_f} \langle \psi_f^0 | v(r) | \psi^0(r) \rangle [\exp(i\omega t) + \exp(-i\omega t)] \exp(-it \frac{(E^0 - E_f^0)}{\hbar}) dt$$

$$= \frac{-1}{2i\hbar} \int_{t_0}^{t_f} \langle \psi_f^0 | v(r) | \psi^0(r) \rangle [\exp(i(\omega + \omega_{f,0})t) + \exp(-i(\omega - \omega_{f,0})t)] dt$$

$$= \frac{-1}{2i\hbar} \langle \psi_f^0 | v(r) | \psi^0(r) \rangle \left[\frac{\exp(i(\omega + \omega_{f,0})t_f) - \exp(i(\omega + \omega_{f,0})t_i)}{i(\omega + \omega_{f,0})} + \frac{\exp(-i(\omega - \omega_{f,0})t_f) - \exp(-i(\omega - \omega_{f,0})t_i)}{i(\omega - \omega_{f,0})} \right]$$

where the transition frequencies $\omega_{f,0}$ are defined by

$$\omega_{f,0} = \frac{E_f^0 - E^0}{\hbar}$$

and t is the time interval $t_f - t_i$.

Now, if the frequency ω is close to one of the transition frequencies, the term with $(\omega - \omega_{f,0})$ in the denominator will be larger than the term containing $(\omega + \omega_{f,0})$. Of course, if ψ_f^0 has a higher energy than ψ^0 , so one is studying stimulate emission spectroscopy, $\omega_{f,0}$ will be negative, in which case the term containing $(\omega + \omega_{f,0})$ will dominate. In on-resonance absorption spectroscopy conditions, the above expression for the first-order coefficients reduces to

$$C_f^1(t) = \frac{-1}{2i\hbar} \langle \psi_f^0 | v(r) | \psi^0(r) \rangle \left[\frac{\exp(-i(\omega - \omega_{f,0})t_f) - \exp(-i(\omega - \omega_{f,0})t_i)}{i(\omega - \omega_{f,0})} \right].$$

The modulus squared of this quantity gives a measure of the probability of observing the system in state ψ_f^0 after being subjected to the photons of frequency ω for a length of time t .

$$\begin{aligned} |C_f^1(t)|^2 &= \frac{|\langle \psi_f^0 | v(r) | \psi^0(r) \rangle|^2}{4\hbar^2} \left[\frac{2\{1 - \cos((\omega - \omega_{f,0})t)\}}{(\omega - \omega_{f,0})^2} \right] \\ &= \frac{|\langle \psi_f^0 | v(r) | \psi^0(r) \rangle|^2}{\hbar^2} \frac{\sin^2(1/2(\omega - \omega_{f,0})t)}{(\omega - \omega_{f,0})^2}. \end{aligned}$$

The function $\frac{\sin^2(1/2(\omega - \omega_{f,0})t)}{(\omega - \omega_{f,0})^2}$ is plotted in Fig. 4.4 for a given value of t as a function of ω . It is sharply peaked around $\omega = \omega_{f,0}$, decays rapidly as $|\omega - \omega_{f,0}|$ increases, and displays recurrences of smaller and smaller intensity when $(\omega - \omega_{f,0})t$ passes through multiples of π .

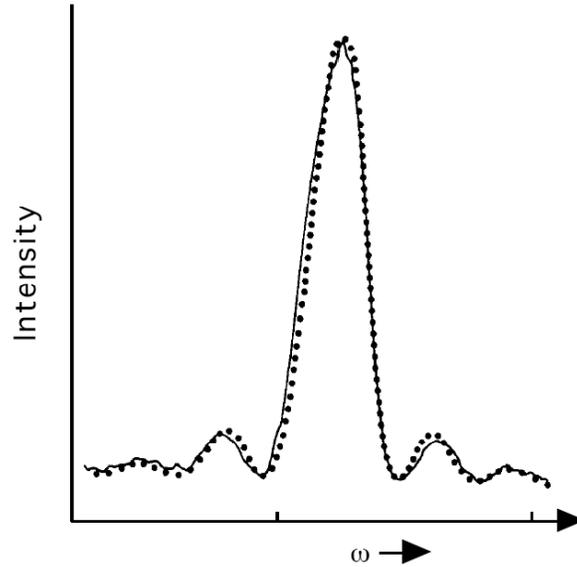


Figure 4.4 Plot of $\frac{\sin^2(1/2(\omega - \omega_{f,0})t)}{(\omega - \omega_{f,0})^2}$ vs ω for a given value of t .

At larger values of t , the main peak in the plot of this function becomes narrower and higher such that, in the $t \rightarrow \infty$ limit, the area under this plot approaches $\pi/2$:

$$Area = \int \frac{\sin^2(1/2(\omega - \omega_{f,0})t)}{(\omega - \omega_{f,0})^2} d\omega = t \frac{\pi}{2}.$$

The importance of this observation about the area under the plot shown in Fig. 4.4 can be appreciated by returning to our result

$$|C_f^1(t)|^2 = \frac{|\langle \psi_f^0 | v(r) | \psi^0(r) \rangle|^2}{\hbar^2} \frac{\sin^2(1/2(\omega - \omega_{f,0})t)}{(\omega - \omega_{f,0})^2}$$

and introducing the fact that the photon source used to induce the transitions being studied most likely is not perfectly monochromatic. If it is characterized, as suggested earlier, by a distribution of frequencies $f(\omega)$ that is broader than the width of the large central peak in Fig. 4.4 (n.b., this will be true if the time duration t is long enough), then when we average $|C_f^1(t)|^2$ over $f(\omega)$ to obtain a result that directly relates to this kind of experiment, we obtain

$$\begin{aligned} \int_{-\infty}^{\infty} f(\omega) |C_f^1(t)|^2 d\omega &= \frac{|\langle \psi_f^0 | v(r) | \psi^0(r) \rangle|^2}{\hbar^2} \int_{-\infty}^{\infty} f(\omega) \frac{\sin^2(1/2(\omega - \omega_{f,0})t)}{(\omega - \omega_{f,0})^2} d\omega \\ &= \frac{\pi |\langle \psi_f^0 | v(r) | \psi^0(r) \rangle|^2 t}{2\hbar^2} f(\omega_{f,0}) = \langle |C_f^1|^2 \rangle \end{aligned}$$

We are allowed to write the integral over ω as ranging from $-\infty$ to $+\infty$ because the function shown in Fig. 4.4 is so sharply peaked around $\omega_{f,0}$ that extending the range of integration makes no difference. We are allowed to factor the $f(\omega)$ out of the integral as $f(\omega_{f,0})$ by assuming the light source's distribution function $f(\omega)$ is very smoothly varying (i.e., not changing much) in the narrow range of frequencies around $\omega_{f,0}$ where the function in Fig. 4.4 is sharply peaked.

The result of this derivation of TDPT is the above expression for the average probability of observing a transition from state ψ^0 to state ψ_f^0 . This probability is seen to grow linearly with the time duration over which the system is exposed to the light source. Because we carried out this derivation within first-order perturbation theory, we should trust this result only under conditions where the effects of the perturbation are small. In the context of the example considered here, this means only for short times. That is, we should view

$$\frac{\pi |\langle \psi_f^0 | v(r) | \psi^0(r) \rangle|^2 t}{2\hbar^2} f(\omega_{f,0}) = \langle |C_f^1|^2 \rangle$$

as expressing the short-time estimate of the probability of a transition from ψ^0 to ψ_f^0 and

$$\text{Rate} = \frac{\pi |\langle \psi_f^0 | v(r) | \psi^0(r) \rangle|^2}{2\hbar^2} f(\omega_{f,0})$$

(obtained as $\frac{d \langle |C_f^1|^2 \rangle}{dt}$) as expressing the initial rate of such transitions within the first-order TDPT approximation.

It should be noted that the rate expression given above will not be valid if the time duration t of the perturbation does not obey $\omega_{f,0} t \gg \pi$; only when this condition is met can the function shown in Fig. 4.4 be integrated to generate a probability prediction that grows linearly with time. So, one has to be careful when using pulsed lasers of very short duration to not employ the simplified rate expression given above (e.g., 1 eV corresponds to a frequency of ca. $2.4 \times 10^{14} \text{ s}^{-1}$, so to study an electronic transition of this energy, one needs to use a light source of duration significantly longer than 10^{-14} s to make use of the simplified result).

The working equations of TDPT, given above, allow one to estimate (because this is a first-order theory) the rates of transitions from one quantum state to another induced by a perturbation whose spatial dependence is characterized by $v(\mathbf{r})$ and whose time dependence is sinusoidal. The same kind of coupling matrix elements $\langle \psi_f^0 | v(r) | \psi^0(r) \rangle$ as we experienced in time-independent PT govern the selection rules and intensities for these transitions, so there is no need to repeat how symmetry can be used to analyze these integrals.

Before closing this treatment of TDPT, it is useful to address a few issues that were circumvented in the derivation presented above.

1. In some cases, one is interested in transitions from a particular initial state $\psi^0(r)$ into a manifold of states that exist in a continuum having energies between E_f^0 and $E_f^0 + dE_f^0$. This occurs, for example, when treating photoionization of a neutral or photodetachment of an anion; here the ejected electron exists in a continuum wave

function whose density of states $\rho(E_f^0)$ is given by the formulas discussed in Chapter 2. In such cases, the expression given above for the rate is modified by summing over all final states having energies within E_f^0 and $E_f^0 + dE_f^0$. Returning to the earlier expression

$$\int \rho(E_f^0) \frac{|\langle \psi_f^0 | v(r) | \psi^0(r) \rangle|^2}{\hbar^2} \int_{-\infty}^{\infty} f(\omega) \frac{\sin^2(1/2(\omega - \omega_{f,0})t)}{(\omega - \omega_{f,0})^2} d\omega dE_f^0$$

using $dE_f^0 = \hbar d\omega_{f,0}$, and assuming the matrix elements $\langle \psi_f^0 | v(r) | \psi^0(r) \rangle$ do not vary significantly within the narrow range between E_f^0 and $E_f^0 + dE_f^0$, one arrives at a rate expression of

$$Rate = \frac{\pi |\langle \psi_f^0 | v(r) | \psi^0(r) \rangle|^2}{2\hbar} f(\omega_{f,0}) \rho(E_f^0)$$

which is much like we obtained earlier but now contains the density of states $\rho(E_f^0)$. In some experiments, one may not have only a single state $\psi^0(r)$ that can absorb light of a given frequency ω ; in such a situation, attenuation of the light source at this frequency can occur through absorptions from many initial states $\psi^0(r)$ into all possible final states ψ_f^0 whose energy differs from that of the initial state by $\hbar\omega$. In this case, the correct expression for the total rate of absorption of photons of energy $\hbar\omega$ is obtained by averaging the above result over the probabilities P_i of the system being in various initial states (which we label ψ_i^0):

$$Rate = \sum_i P_i \frac{\pi |\langle \psi_f^0 | v(r) | \psi_i^0(r) \rangle|^2}{2\hbar} f(\omega_{f,i}) \rho(E_f^0) \delta(\omega - \omega_{f,i}).$$

Here the $\delta(\omega - \omega_{f,i})$ function guarantees that only states ψ_i^0 and ψ_f^0 whose energies differ by $\hbar\omega$ are permitted to enter the sum. The nature of the initial-state probability P_i

depends on what kind of experiment is being carried out. P_i might be a Boltzmann distribution if the initial states are in thermal equilibrium, for example.

2. In Fig. 4.4 the function $\frac{\sin^2(1/2(\omega - \omega_{f,0})t)}{(\omega - \omega_{f,0})^2}$ is plotted for one value of t as a function

of ω . There also appear in this figure, dots that represent experimental data. These data were obtained by allowing a stream of HCN molecules to flow through a laser beam of width L with the laser frequency tuned to ω . From the flow velocity v of the HCN stream and the laser beam width L , one can determine the duration over which the molecules were exposed to the light source $t = L/v$. After the molecules exited the laser beam, they were probed to determine whether they were in an excited state. This experiment was repeated for various values of the frequency ω . The population of excited states was then plotted as a function of ω to obtain the data plotted in Fig. 4.4. This experiment is described in the text *Molecules and Radiation*, J. I. Steinfeld, MIT Press, Cambridge, Mass. (1981). This kind of experiment provided direct proof of the oscillatory frequency dependence observed in the population of excited states as predicted in our derivation of TDPT.

3. To give an example of how one proceeds in TDPT when the perturbation is not oscillatory in time, let us consider an atom located at the origin of our coordinate system that experiences a collision with an ion of charge χ whose trajectory is described in Fig. 4.5.

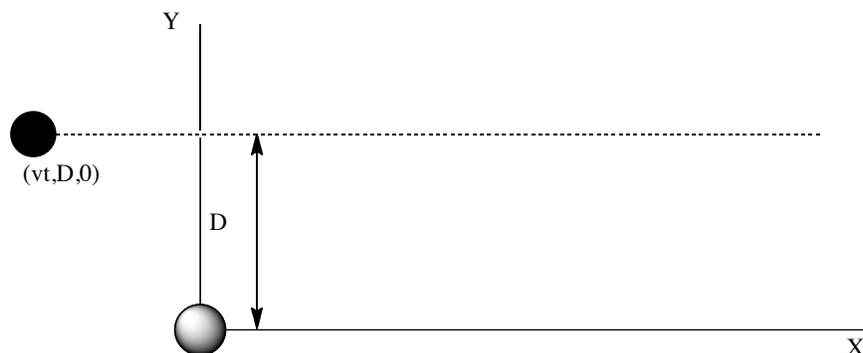


Figure 4.5 An atom (at the origin $X = Y = Z = 0$) undergoing a collision with an ion of charge χ moving along the X -axis with constant velocity v .

As an approximation, we assume

- (1) that the ion moves in a straight line: $X = vt$, $Y = D$, $Z = 0$, characterized by an impact parameter D and a velocity v (this would be appropriate if the ion were moving so fast that it would not be deflected by interactions with the atom),
- (2) that the perturbation caused by the ion on the electrons of the atom at the origin can be represented by

$$-\sum_{i=1}^N \frac{\chi}{|r_i - R|}$$

where \mathbf{r}_i is the position of the i^{th} electron in the atom and $\mathbf{R} = (vt, D, 0)$ is the position of the ion. The time dependence of the perturbation arises from the motion of the ion along the X-axis.

Writing the distance $|r_i - R|$ as

$$|r_i - R| = \sqrt{(x_i - vt)^2 + (y_i - D)^2 + z_i^2}$$

and expanding in inverse powers of $\sqrt{D^2 + (vt)^2}$ we can express the ion-atom interaction potential as

$$-\sum_{i=1}^N \frac{\chi}{|r_i - R|} = \sum_{i=1}^N \left[\frac{-\chi}{\sqrt{D^2 + (vt)^2}} + \frac{-\chi(vtx_i + Dy_i + r_i^2)}{(D^2 + (vt)^2)^{3/2}} + \dots \right].$$

The first term contains no factors dependent on the atom's electronic coordinates, so it plays no role in causing electronic transitions. In the second term, the factor r_i^2 can be neglected compared to the $vtx_i + Dy_i$ terms because the ion is assumed to be somewhat distant from the atom's valence electrons.

To derive an equation for the probability of the atom undergoing a transition from $\psi^0(r)$ to ψ_f^0 , one returns to the TDPT expression

$$-i\hbar \frac{dC_f^1(t)}{dt} = \langle \psi_f^0 | V(t) | \psi^0(r) \rangle \exp(-it \frac{(E^0 - E_f^0)}{\hbar})$$

and substitutes the above expression for the perturbation to obtain

$$\frac{dC_f^1(t)}{dt} = \frac{-1}{i\hbar} \langle \psi_f^0 | \sum_{i=1}^N \frac{-\chi(vtx_i + Dy_i)}{(D^2 + (vt)^2)^{3/2}} | \psi^0(r) \rangle \exp(-it \frac{(E^0 - E_f^0)}{\hbar}).$$

This is the equation that must be solved to evaluate C_f^1 by integrating from $t = -\infty$ to $t = +\infty$ (representing the full collision with the ion starting far to the left on the X-axis and proceeding far to the right).

There are two limiting cases in which the solution is straightforward. First, if the time duration of the collision (i.e., the time over which the ion is close to the atom) D/v is long compared to $\omega_{f,0}$ where

$$\omega_{f,0} = \frac{(E^0 - E_f^0)}{\hbar},$$

then the integrand will oscillate repeatedly during the time D/v as a result of which the integral

$$C_f^1 = \int_{-\infty}^{\infty} \frac{dC_f^1(t)}{dt} dt$$

will be vanishingly small. So, in this so-called adiabatic case (i.e., with the ion moving slowly relative to the oscillation frequency $\omega_{f,0}$), electronic transitions should not be expected. In the other limit $\omega_{f,0}D/v \ll 1$, the factor $\exp(-it \frac{(E^0 - E_f^0)}{\hbar})$ will remain approximately equal to unity, so the integration needed reduces to

$$C_f^1 = \frac{-1}{i\hbar} \int_{-\infty}^{\infty} \langle \psi_f^0 | \sum_{i=1}^N \frac{-\chi(vtx_i + Dy_i)}{(D^2 + (vt)^2)^{3/2}} | \psi^0(r) \rangle dt .$$

The integral involving vtx_i vanishes because vt is odd and the remainder of the integrand is an even function of t . The integral involving Dy_i can be performed by trigonometric substitution ($vt = D \tan(\theta)$) so the denominator reduces to $D^3 (1+(\sin\theta/\cos\theta)^2)^{3/2} = D^3/(\cos\theta)^3$ and gives

$$C_f^1 = \frac{-2\chi}{i\hbar v D} \langle \psi_f^0 | \sum_{i=1}^N y_i | \psi^0(r) \rangle .$$

This result suggests that the probability of a transition

$$|C_f^1|^2 = \frac{4\chi^2}{\hbar^2 v^2 D^2} |\langle \psi_f^0 | \sum_{i=1}^N y_i | \psi^0(r) \rangle|^2$$

should vary as the square of the ion's charge and inversely with the speed of the collision. Of course, this result can not be trusted if the speed v is too low because, then the condition $\omega_{f0}D/v \ll 1$ will not hold. This example shows how one must re-derive the equations of TDPT when dealing with perturbations whose time-dependence is not sinusoidal.

4.6 Chapter Summary

In this Chapter, you should have learned about the following things:

1. Rayleigh-Schrödinger perturbation theory with several example applications.
2. The variational method for optimizing trial wave functions.
3. The use of point group symmetry.
4. Time dependent perturbation theory, primarily for sinusoidal perturbations characteristic of electromagnetic radiation.