

Chapter 4

Some important tools of theory

4.1 Perturbation theory and the variational method

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 can not be solved exactly. To illustrate how desperate this situation is, I note that neither of the following two Schrödinger equations have ever been solved exactly (meaning analytically):

- (i) The Schrödinger equation for the two electrons moving about the He nucleus:

$$\left[-\frac{\hbar^2}{2m_e} \nabla_1^2 - \frac{\hbar^2}{2m_e} \nabla_2^2 - \frac{2e^2}{r_1} - \frac{2e^2}{r_2} + \frac{e^2}{r_{1,2}} \right] \psi = E\psi. \quad (4.1)$$

- (ii) The Schrödinger equation for the two electrons moving in an H_2 molecule even if the locations of the two nuclei (labeled A and B) are held clamped:

$$\left[-\frac{\hbar^2}{2m_e} \nabla_1^2 - \frac{\hbar^2}{2m_e} \nabla_2^2 - \frac{e^2}{r_{1,A}} - \frac{e^2}{r_{2,A}} - \frac{e^2}{r_{1,B}} - \frac{e^2}{r_{2,B}} + \frac{e^2}{r_{1,2}} \right] \psi = E\psi. \quad (4.2)$$

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 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 as starting points.

But, once one has decided on a reasonable “starting point” 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.

4.1.1 Perturbation theory

In this method, one has available a set of equations for generating a sequence of approximations to the true energy E and true wave function ψ . I will now briefly outline the derivation of these working equations for you. First, one decomposes the true Hamiltonian H into a so-called zeroth order part H^0 (this is the Hamiltonian of the model problem one has chosen to use to represent the real system) and the difference $(H - H^0)$ which is called the perturbation and often denoted V :

$$H = H^0 + V. \quad (4.3)$$

The fundamental assumption of perturbation theory is that the wave functions and energies can be expanded in a Taylor series involving various powers of the perturbation. That is, one expands the energy E and the wave function ψ into zeroth, first, second, etc., order pieces which form the unknowns in this method:

$$E = E^0 + E^1 + E^2 + E^3 + \dots, \quad (4.4)$$

$$\psi = \psi^0 + \psi^1 + \psi^2 + \psi^3 + \dots. \quad (4.5)$$

Next, one substitutes these expansions for E of H and of 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. 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:

$$H^0\psi^0 = E^0\psi^0, \quad (4.6)$$

$$H^0\psi^1 + V\psi^0 = E^0\psi^1 + E^1\psi^0, \quad (4.7)$$

$$H^0\psi^2 + V\psi^1 = E^0\psi^2 + E^1\psi^1 + E^2\psi^0. \quad (4.8)$$

The zeroth order equation simply instructs us to solve the zeroth order Schrödinger equation to obtain the zeroth order wave function ψ^0 and its zeroth order energy E^0 . 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^K in terms of the complete set of wave functions

of the zeroth order problem $\{\psi_j^0\}$. This means that one must be able to 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 (e.g., excited states if ψ^0 is the ground state) zeroth order states $\{\psi_j^0\}$. For example, expanding ψ^1 in this manner gives

$$\psi^1 = \sum_j C_j^1 \psi_j^0. \quad (4.9)$$

Now, the unknowns in the first order equation become E^1 and the C_j^1 expansion coefficients. Substituting this expansion into $H^0\psi^1 + V\psi^0 = E^0\psi^1 + E^1\psi^0$ and solving for these unknowns produces the following final first order working equations:

$$E^1 = \langle \psi^0 | V | \psi^0 \rangle. \quad (4.10)$$

$$\psi^1 = \sum_j \psi_j^0 \{ \langle \psi^0 | V | \psi_j^0 \rangle / (E^0 - E_j^0) \}. \quad (4.11)$$

where the index J is restricted such that ψ_j^0 not equal the state ψ^0 you are interested in. These are the fundamental working equations of first order 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 in terms of coefficients multiplying various other zeroth order wave functions ψ_j^0 .

An analogous approach is used to solve the second and higher order equations. Although modern quantum mechanics does indeed use 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 attention. The expression for the second order energy correction is found to be

$$E^2 = \sum_j \langle \psi^0 | V | \psi_j^0 \rangle^2 / (E^0 - E_j^0), \quad (4.12)$$

where again, the index J is restricted as noted above. Let's now consider an example problem that illustrates how perturbation theory is used.

Example problem for perturbation theory

As we discussed earlier, an electron moving in a 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 $V = e\varepsilon(x - \frac{L}{2})$ 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.

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.

$$\begin{aligned}
 V &= e\varepsilon \left(x - \frac{L}{2} \right), & \Psi_n^{(0)} &= \left(\frac{2}{L} \right)^{\frac{1}{2}} \sin\left(\frac{n\pi x}{L} \right), \\
 E_n^{(0)} &= \frac{\hbar^2 \pi^2 n^2}{2mL^2}, \\
 E_{n=1}^{(1)} &= \left\langle \Psi_{n=1}^{(0)} \left| V \right| \Psi_{n=1}^{(0)} \right\rangle = \left\langle \Psi_{n=1}^{(0)} \left| e\varepsilon \left(x - \frac{L}{2} \right) \right| \Psi_{n=1}^{(0)} \right\rangle \\
 &= \left(\frac{2}{L} \right) \int_0^L \sin^2\left(\frac{\pi x}{L} \right) e\varepsilon \left(x - \frac{L}{2} \right) dx \\
 &= \left(\frac{2e\varepsilon}{L} \right) \int_0^L \sin^2\left(\frac{\pi x}{L} \right) x dx - \left(\frac{2e\varepsilon}{L} \right) \frac{L}{2} \int_0^L \sin^2\left(\frac{\pi x}{L} \right) dx.
 \end{aligned}$$

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

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

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

$$\begin{aligned}
 \int_0^L \sin^2\left(\frac{\pi x}{L} \right) dx &= \frac{L}{\pi} \int_0^\pi \sin^2 \theta d\theta \\
 \int_0^\pi \sin^2 \theta d\theta &= -\frac{1}{4} \sin(2\theta) + \frac{\theta}{2} \Big|_0^\pi = \frac{\pi}{2}.
 \end{aligned}$$

Making all of these appropriate substitutions we obtain:

$$\begin{aligned}
 E_{n=1}^{(1)} &= \left(\frac{2e\varepsilon}{L} \right) \left(\frac{L^2}{4} - \frac{L}{2} \frac{L}{\pi} \right) = 0, \\
 \Psi_{n=1}^{(1)} &= \frac{\left\langle \Psi_{n=2}^{(0)} \left| e\varepsilon \left(x - \frac{L}{2} \right) \right| \Psi_{n=1}^{(0)} \right\rangle \Psi_{n=2}^{(0)}}{E_{n=1}^{(0)} - E_{n=2}^{(0)}}, \\
 \Psi_{n=1}^{(1)} &= \frac{\left(\frac{2}{L} \right) \int_0^L \sin\left(\frac{2\pi x}{L} \right) e\varepsilon \left(x - \frac{L}{2} \right) \sin\left(\frac{\pi x}{L} \right) dx}{\frac{\hbar^2 \pi^2}{2mL^2(1^2 - 2^2)}} \left(\frac{2}{L} \right)^{\frac{1}{2}} \sin\left(\frac{2\pi x}{L} \right).
 \end{aligned}$$

The two integrals in the numerator need to be evaluated:

$$\int_0^L x \sin\left(\frac{2\pi x}{L} \right) \sin\left(\frac{\pi x}{L} \right) dx, \quad \text{and} \quad \int_0^L \sin\left(\frac{2\pi x}{L} \right) \sin\left(\frac{\pi x}{L} \right) dx.$$

Using the integral $\int_x \cos(ax) dx = \frac{1}{a^2} \cos(ax) + \frac{x}{a} \sin(ax)$, and the integral $\int \cos(ax) dx = \frac{1}{a} \sin(ax)$, we obtain the following:

$$\begin{aligned}
 \int_0^L \sin\left(\frac{2\pi x}{L} \right) \sin\left(\frac{\pi x}{L} \right) dx &= \frac{1}{2} \left[\int_0^L \cos\left(\frac{\pi x}{L} \right) dx - \int_0^L \cos\left(\frac{3\pi x}{L} \right) dx \right] \\
 &= \frac{1}{2} \left[\frac{L}{\pi} \sin\left(\frac{\pi x}{L} \right) \Big|_0^L - \frac{L}{3\pi} \sin\left(\frac{3\pi x}{L} \right) \Big|_0^L \right] = 0.
 \end{aligned}$$

$$\begin{aligned}
\int_0^L x \sin\left(\frac{2\pi x}{L}\right) \sin\left(\frac{\pi x}{L}\right) dx &= \frac{1}{2} \left[\int_0^L x \cos\left(\frac{\pi x}{L}\right) dx - \int_0^L x \cos\left(\frac{3\pi x}{L}\right) dx \right] \\
&= \frac{1}{2} \left[\left(\frac{L^2}{\pi^2} \cos\left(\frac{\pi x}{L}\right) + \frac{Lx}{\pi} \sin\left(\frac{\pi x}{L}\right) \right) \Big|_0^L \right. \\
&\quad \left. - \left(\frac{L^2}{9\pi^2} \cos\left(\frac{3\pi x}{L}\right) + \frac{Lx}{3\pi} \sin\left(\frac{3\pi x}{L}\right) \right) \Big|_0^L \right] \\
&= \frac{L^2}{2\pi^2} (\cos \pi - \cos 0) + \frac{L^2}{2\pi} \sin \pi - 0 \\
&= \frac{-2L^2}{2\pi^2} - \frac{-2L^2}{18\pi^2} = \frac{L^2}{9\pi^2} - \frac{L^2}{\pi^2} = -\frac{8L^2}{9\pi^2}.
\end{aligned}$$

Making all of these appropriate substitutions we obtain

$$\begin{aligned}
\Psi_{n=1}^{(1)} &= \frac{\left(\frac{2}{L}\right) (e\mathcal{E}) \left(-\frac{8L^2}{9\pi^2} - \frac{L}{2}(0)\right)}{\frac{-3\hbar^2\pi^2}{2mL^2}} \left(\frac{2}{L}\right)^{\frac{1}{2}} \sin\left(\frac{2\pi x}{L}\right) \\
&= \frac{32mL^3 e\mathcal{E}}{27\hbar^2\pi^4} \left(\frac{2}{L}\right)^{\frac{1}{2}} \sin\left(\frac{2\pi x}{L}\right).
\end{aligned}$$

Now, let's compute the induced dipole moment caused by the polarization of the electron density due to the electric field effect using the equation $\mu_{\text{induced}} = -e \int \Psi^*(x - \frac{L}{2}) \Psi dx$ with Ψ now 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.

$$\begin{aligned}
\mu_{\text{induced}} &= -e \int \Psi^* \left(x - \frac{L}{2}\right) \Psi dx, \quad \text{where } \Psi = (\Psi_1^{(0)} + \Psi_1^{(1)}). \\
\mu_{\text{induced}} &= -e \int_0^L (\Psi_1^{(0)} + \Psi_1^{(1)})^* \left(x - \frac{L}{2}\right) (\Psi_1^{(0)} + \Psi_1^{(1)}) dx \\
&= -e \int_0^L \Psi_1^{(0)*} \left(x - \frac{L}{2}\right) \Psi_1^{(0)} dx - e \int_0^L \Psi_1^{(0)*} \left(x - \frac{L}{2}\right) \Psi_1^{(1)} dx \\
&\quad - e \int_0^L \Psi_1^{(1)*} \left(x - \frac{L}{2}\right) \Psi_1^{(0)} dx - e \int_0^L \Psi_1^{(1)*} \left(x - \frac{L}{2}\right) \Psi_1^{(1)} dx.
\end{aligned}$$

The first integral is zero (see the evaluation of this integral for $E_1^{(1)}$ above). The fourth integral is neglected since it is proportional to ε^2 . The second and third integrals are the same and are combined to give

$$\mu_{\text{induced}} = -2e \int_0^L \Psi_1^{(0)*} \left(x - \frac{L}{2}\right) \Psi_1^{(1)} dx.$$

Substituting $\Psi_1^{(0)} = \left(\frac{2}{L}\right)^{\frac{1}{2}} \sin\left(\frac{\pi x}{L}\right)$ and $\Psi_1^{(1)} = \frac{32mL^3 e \varepsilon}{27 \hbar^2 \pi^4} \left(\frac{2}{L}\right)^{\frac{1}{2}} \sin\left(\frac{2\pi x}{L}\right)$, we obtain

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

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

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

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 $\alpha = \left.\frac{\partial \mu}{\partial \varepsilon}\right|_{\varepsilon=0}$. Using our induced moment result above, we then find

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

Notice that this finding suggests that the larger the box (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 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 :

$$E(\{\alpha_J\}) = \langle \psi^0 | H | \psi^0 \rangle / \langle \psi^0 | \psi^0 \rangle. \quad (4.13)$$

The optimal values of the α_J parameters are determined by making

$$dE/d\alpha_J = 0, \quad (4.14)$$

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).

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

$$H\Psi_K = E_K\Psi_K. \quad (4.15)$$

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 :

$$\psi^0 = \sum_K C_K \Psi_K. \quad (4.16)$$

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:

$$E = \langle \psi^0 | H | \psi^0 \rangle / \langle \psi^0 | \psi^0 \rangle = \left\langle \sum_K C_K \Psi_K \left| H \right| \sum_L C_L \Psi_L \right\rangle / \left\langle \sum_K C_K \Psi_K \left| \sum_L C_L \Psi_L \right. \right\rangle. \quad (4.17)$$

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)

$$\langle \Psi_K | \Psi_L \rangle = \delta_{KL}, \quad (4.18)$$

the above expression reduces to

$$\begin{aligned} E &= \sum_K \langle C_K \Psi_K | H | C_K \Psi_K \rangle / \left(\sum_K \langle C_K \Psi_K | C_K \Psi_K \rangle \right) \\ &= \sum_K |C_K|^2 E_K / \sum_K |C_K|^2. \end{aligned} \quad (4.19)$$

One of the basic properties of the kind of Hamiltonian 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, 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

$$E \geq \sum_K |C_K|^2 E_0 / \sum_K |C_K|^2 = E_0. \quad (4.20)$$

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.

Example variational problem

Suppose you are given a trial wave function of the form

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

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

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

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

$$\begin{aligned} W &= \left(Z_c^2 - 2ZZ_c + \frac{5}{8}Z_c \right) \frac{e^2}{a_0}, \\ \frac{dW}{dZ_c} &= \left(2Z_c - 2Z + \frac{5}{8} \right) \frac{e^2}{a_0} = 0, \\ 2Z_c - 2Z + \frac{5}{8} &= 0, \\ 2Z_c &= 2Z - \frac{5}{8}, \\ Z_c &= Z - \frac{5}{16} = Z - 0.3125 \end{aligned}$$

(n.b., 0.3125 represents the shielding factor of one 1s electron to the other).

Now, using this optimal Z_c in our energy expression gives

$$\begin{aligned} W &= Z_c \left(Z_c - 2Z + \frac{5}{8} \right) \frac{e^2}{a_0} \\ &= \left(Z - \frac{5}{16} \right) \left[\left(Z - \frac{5}{16} \right) - 2Z + \frac{5}{8} \right] \frac{e^2}{a_0} \end{aligned}$$

$$\begin{aligned}
 &= \left(Z - \frac{5}{16}\right) \left(-Z + \frac{5}{16}\right) \frac{e^2}{a_0} \\
 &= -\left(Z - \frac{5}{16}\right) \left(Z - \frac{5}{16}\right) \frac{e^2}{a_0} = -\left(Z - \frac{5}{16}\right)^2 \frac{e^2}{a_0} \\
 &= -(Z - 0.3125)^2 (27.21) \text{ eV}
 \end{aligned}$$

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

Is this energy “any good”? The total energies of some two-electron atoms and ions have been experimentally determined to be:

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

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 percentage error in our estimate for each ion.

Z	Atom	Experimental (eV)	Calculated (eV)	% Error
1	H ⁻	-14.35	-12.86	10.38
2	He	-78.98	-77.46	1.92
3	Li ⁺	-198.02	-196.46	0.79
4	Be ⁺²	-371.5	-369.86	0.44
5	B ⁺³	-599.3	-597.66	0.27
6	C ⁺⁴	-881.6	-879.86	0.19
7	N ⁺⁵	-1218.3	-1216.48	0.15
8	O ⁺⁶	-1609.5	-1607.46	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 dissociation into a hydrogen atom and an electron. 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.

4.2 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 section, 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.

4.2.1 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 reflections, σ_v , σ'_v , σ''_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. 4.1).

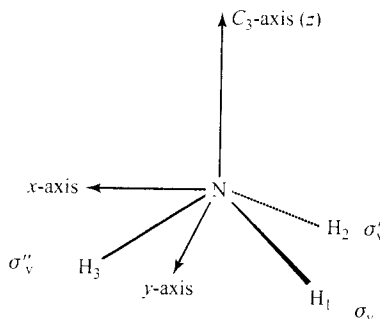


Figure 4.1 Ammonia molecule and its symmetry elements.

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

- (i) A combination rule is defined through which two group elements are combined to give a result which 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).
- (ii) 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).
- (iii) 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.
- (iv) 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 a 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.

4.2.2 Matrices as group representations

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

- (i) 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 m.o. case; they are unit vectors along the x , y , and z directions located on each of the atoms, and represent displacements along each of these directions, in the vibration/rotation case.

- (ii) 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 m.o. 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:

$$\begin{array}{ll}
 (S_N, S_1, S_2, S_3) & \xrightarrow{E} (S_N, S_1, S_2, S_3) \\
 & \xrightarrow{C_3} (S_N, S_3, S_1, S_2) \\
 & \xrightarrow{C_3^2} (S_N, S_2, S_3, S_1) \\
 & \xrightarrow{\sigma_v} (S_N, S_1, S_3, S_2) \\
 & \xrightarrow{\sigma_v''} (S_N, S_3, S_2, S_1) \\
 & \xrightarrow{\sigma_v'} (S_N, S_2, S_1, S_3)
 \end{array} \quad (4.21)$$

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

$$D^{(4)}(C_3) \begin{bmatrix} S_N \\ S_1 \\ S_2 \\ S_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} S_N \\ S_1 \\ S_2 \\ S_3 \end{bmatrix} = \begin{bmatrix} S_N \\ S_3 \\ S_1 \\ S_2 \end{bmatrix}. \quad (4.22)$$

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

$$\begin{aligned}
 D^{(4)}(C_3^2) &= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \end{bmatrix}, & D^{(4)}(E) &= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \\
 D^{(4)}(\sigma_v) &= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}, & D^{(4)}(\sigma'_v) &= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}, \\
 D^{(4)}(\sigma''_v) &= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.
 \end{aligned} \tag{4.23}$$

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

$$\sigma_v C_3 = \sigma'_v, \quad \text{or,} \quad \begin{array}{ccccc} & S_1 & & S_3 & & S_3 \\ & & \xrightarrow{C_3} & & \xrightarrow{\sigma_v} & \\ \sigma_v C_3 = \sigma'_v & & S_2 & S_3 & S_1 & S_2 & S_2 & S_1 \end{array} \tag{4.24}$$

Note that this same relationship is carried by the matrices:

$$\begin{aligned}
 D^{(4)}(\sigma_v)D^{(4)}(C_3) &= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \\
 &= D^{(4)}(\sigma'_v).
 \end{aligned} \tag{4.25}$$

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

$$\begin{aligned}
 D^{(4)}(C_3)D^{(4)}(\sigma_v) &= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \\
 &= D^{(4)}(\sigma''_v).
 \end{aligned} \tag{4.26}$$

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.

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 $\text{Tr}(D)$:

$$\text{Tr}(D) = \sum_i D_{ii} = \chi. \quad (4.27)$$

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

$$\chi(E) = 4, \quad (4.28)$$

$$\chi(C_3) = \chi(C_3^2) = 1, \quad (4.29)$$

$$\chi(\sigma_v) = \chi(\sigma_v') = \chi(\sigma_v'') = 2. \quad (4.30)$$

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 them. 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 those 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.

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:

$$\begin{array}{lll} E & C_3 & C_3^2 \\ S_N \rightarrow S_N, & S_N \rightarrow S_N, & S_N \rightarrow S_N, \\ \sigma_v & \sigma_v' & \sigma_v'' \\ S_N \rightarrow S_N, & S_N \rightarrow S_N, & S_N \rightarrow S_N. \end{array} \quad (4.31)$$

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

$$\begin{array}{lll} D^{(1)}(E) = 1, & D^{(1)}(C_3) = 1, & D^{(1)}(C_3^2) = 1, \\ D^{(1)}(\sigma_v) = 1, & D^{(1)}(\sigma_v') = 1, & D^{(1)}(\sigma_v'') = 1. \end{array} \quad (4.32)$$

Again we have

$$\begin{aligned} D^{(1)}(\sigma_v)D^{(1)}(C_3) &= 1 \cdot 1 = D^{(1)}(\sigma_v''), \\ \text{and } D^{(1)}(C_3)D^{(1)}(\sigma_v) &= 1 \cdot 1 = D^{(1)}(\sigma_v'). \end{aligned} \quad (4.33)$$

These six 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.

4.2.3 Reducible and irreducible representations

A reducible representation

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

$$\begin{array}{|c|ccc|} \hline 1 & 0 & 0 & 0 \\ \hline 0 & \boxed{\text{3} \times \text{3 matrix}} & & \\ \hline 0 & & & \\ \hline 0 & & & \\ \hline \end{array}$$

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 $D^{(3)}$.

$$\begin{aligned} D^{(3)}(E) &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, & D^{(3)}(C_3) &= \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}, & D^{(3)}(C_3^2) &= \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}. \\ D^{(3)}(\sigma_v) &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}, & D^{(3)}(\sigma_v') &= \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}, & D^{(3)}(\sigma_v'') &= \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \end{aligned} \quad (4.34)$$

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) ; also note that these characters are equal to those of $D^{(4)}$ minus those of $D^{(1)}$.

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:

$$T_1 = S_1 + S_2 + S_3. \quad (4.35)$$

$$T_2 = 2S_1 - S_2 - S_3. \quad (4.36)$$

$$T_3 = S_2 - S_3. \quad (4.37)$$

(Don't worry about how we construct 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,

$$\begin{aligned} (T_1, T_2, T_3) &\xrightarrow{\sigma_v} (T_1, T_2, -T_3), & (T_1, T_2, T_3) &\xrightarrow{E} (T_1, T_2, T_3), \\ (T_1, T_2, T_3) &\xrightarrow{\sigma'_v} (S_3 + S_2 + S_1, 2S_3 - S_2 - S_1, S_2 - S_1) \\ &= \left(T_1, -\frac{1}{2}T_2 - \frac{3}{2}T_3, -\frac{1}{2}T_2 + \frac{1}{2}T_3 \right); \\ (T_1, T_2, T_3) &\xrightarrow{\sigma''_v} (S_2 + S_1 + S_3, 2S_2 - S_1 - S_3, S_1 - S_3) \\ &= \left(T_1, -\frac{1}{2}T_2 + \frac{3}{2}T_3, \frac{1}{2}T_2 + \frac{1}{2}T_3 \right), \\ (T_1, T_2, T_3) &\xrightarrow{C_3} (S_3 + S_1 + S_2, 2S_3 - S_1 - S_2, S_1 - S_2) \\ &= \left(T_1, -\frac{1}{2}T_2 - \frac{3}{2}T_3, \frac{1}{2}T_2 - \frac{1}{2}T_3 \right), \\ (T_1, T_2, T_3) &\xrightarrow{C_3^2} (S_2 + S_3 + S_1, 2S_2 - S_3 - S_1, S_3 - S_1) \\ &= \left(T_1, -\frac{1}{2}T_2 + \frac{3}{2}T_3, -\frac{1}{2}T_2 - \frac{1}{2}T_3 \right). \end{aligned} \quad (4.38)$$

So the matrix representations in the new T_i basis are

$$\begin{aligned} D^{(3)}(E) &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, & D^{(3)}(C_3) &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & -\frac{1}{2} & -\frac{3}{2} \\ 0 & +\frac{1}{2} & -\frac{1}{2} \end{bmatrix}, \\ D^{(3)}(C_3^2) &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & -\frac{1}{2} & +\frac{3}{2} \\ 0 & -\frac{1}{2} & -\frac{1}{2} \end{bmatrix}, & D^{(3)}(\sigma_v) &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix}, \\ D^{(3)}(\sigma'_v) &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & -\frac{1}{2} & -\frac{3}{2} \\ 0 & -\frac{1}{2} & +\frac{1}{2} \end{bmatrix}, & D^{(3)}(\sigma''_v) &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & -\frac{1}{2} & +\frac{3}{2} \\ 0 & +\frac{1}{2} & +\frac{1}{2} \end{bmatrix}. \end{aligned} \quad (4.39)$$

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:

$$\begin{aligned} D^{(2)}(E) &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, & D^{(2)}(C_3) &= \begin{bmatrix} -\frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} \end{bmatrix}, & D^{(2)}(C_3^2) &= \begin{bmatrix} -\frac{1}{2} & \frac{\sqrt{3}}{2} \\ -\frac{\sqrt{3}}{2} & -\frac{1}{2} \end{bmatrix}, \\ D^{(2)}(\sigma_v) &= \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, & D^{(2)}(\sigma'_v) &= \begin{bmatrix} -\frac{1}{2} & -\frac{\sqrt{3}}{2} \\ -\frac{\sqrt{3}}{2} & \frac{1}{2} \end{bmatrix}, & D^{(2)}(\sigma''_v) &= \begin{bmatrix} -\frac{1}{2} & \frac{\sqrt{3}}{2} \\ -\frac{\sqrt{3}}{2} & \frac{1}{2} \end{bmatrix}. \end{aligned} \quad (4.40)$$

The characters can be obtained by summing diagonal elements:

$$\chi(E) = 2, \quad \chi(2C_3) = -1, \quad \chi(3\sigma_v) = 0. \quad (4.41)$$

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:

$$\begin{aligned} R_z &\xrightarrow{E} R_z, & R_z &\xrightarrow{C_3} R_z, & R_z &\xrightarrow{C_3^2} R_z, \\ R_z &\xrightarrow{\sigma_v} -R_z, & R_z &\xrightarrow{\sigma'_v} -R_z, & R_z &\xrightarrow{\sigma''_v} -R_z. \end{aligned} \quad (4.42)$$

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

$$\begin{aligned} D^{(1)}(E) &= 1, & D^{(1)}(C_3) &= 1, & D^{(1)}(C_3^2) &= 1, \\ D^{(1)}(\sigma_v) &= -1, & D^{(1)}(\sigma'_v) &= -1, & D^{(1)}(\sigma''_v) &= -1. \end{aligned} \quad (4.43)$$

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 can not 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.

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_2	1	1	1
E	2	-1	0
$2A_1 \oplus E$	4	1	2

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 $D^{(4)}$

$$\begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ E & C_3 & & \sigma_v & & \end{bmatrix} \cdot \begin{bmatrix} 4 & E \\ 1 & C_3 \\ 1 & \\ 2 & \sigma_v \\ 2 & \\ 2 & \end{bmatrix} = 4 + 1 + 1 + 2 + 2 + 2 = 12. \quad (4.44)$$

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

$$n_\Gamma = \frac{1}{g} \sum_R \chi_\Gamma(R) \chi_{\text{red}}(R), \quad (4.45)$$

where g is the order of the group and $\chi_\Gamma(R)$ are the characters of the Γ th irreducible representation.

Commonly used bases

We could take *any* set of functions as a basis for a group representation. Commonly used sets include: 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 $-xy, yz, xz, x^2 - y^2, z^2$, 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 the Appendix.

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.

4.2.4 Another example

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 $C_3, C_3^2, \sigma_v, \sigma'_v, \sigma''_v$ 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 the Appendix) 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 $C_3, C_3^2, \sigma_v, \sigma'_v, \sigma''_v$ and the identity operation map $2p_x$ and $2p_y$ among one

another. Specifically,

$$C_3 \begin{bmatrix} 2p_x \\ 2p_y \end{bmatrix} = \begin{bmatrix} \cos 120^\circ & -\sin 120^\circ \\ \sin 120^\circ & \cos 120^\circ \end{bmatrix} \begin{bmatrix} 2p_x \\ 2p_y \end{bmatrix}, \quad (4.46)$$

$$C_3^2 \begin{bmatrix} 2p_x \\ 2p_y \end{bmatrix} = \begin{bmatrix} \cos 240^\circ & -\sin 240^\circ \\ \sin 240^\circ & \cos 240^\circ \end{bmatrix} \begin{bmatrix} 2p_x \\ 2p_y \end{bmatrix}, \quad (4.47)$$

$$E \begin{bmatrix} 2p_x \\ 2p_y \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 2p_x \\ 2p_y \end{bmatrix}. \quad (4.48)$$

$$\sigma_v \begin{bmatrix} 2p_x \\ 2p_y \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 2p_x \\ 2p_y \end{bmatrix}. \quad (4.49)$$

$$\sigma_v' \begin{bmatrix} 2p_x \\ 2p_y \end{bmatrix} = \begin{bmatrix} +\frac{1}{2} & +\frac{\sqrt{3}}{2} \\ +\frac{\sqrt{3}}{2} & -\frac{1}{2} \end{bmatrix} \begin{bmatrix} 2p_x \\ 2p_y \end{bmatrix}, \quad (4.50)$$

$$\sigma_v'' \begin{bmatrix} 2p_x \\ 2p_y \end{bmatrix} = \begin{bmatrix} +\frac{1}{2} & -\frac{\sqrt{3}}{2} \\ -\frac{\sqrt{3}}{2} & -\frac{1}{2} \end{bmatrix} \begin{bmatrix} 2p_x \\ 2p_y \end{bmatrix}. \quad (4.51)$$

The 2×2 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 ($D^{(IR)}$) for that particular operation and for this particular irreducible representation (IR). For example,

$$\begin{bmatrix} +\frac{1}{2} & +\frac{\sqrt{3}}{2} \\ +\frac{\sqrt{3}}{2} & -\frac{1}{2} \end{bmatrix} = D^{(E)}(\sigma_v'). \quad (4.52)$$

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.

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 $\int 2p_x^* C_3 2p_x d\tau$ and $\int 2p_y^* C_3 2p_y d\tau$. 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,

$$\sum_i \int \phi_i^* S\phi_i d\tau = \chi(S). \quad (4.53)$$

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

$$\chi(E) = 2, \quad \chi(C_3) = \chi(C_3^2) = -1, \quad \chi(\sigma_v) = \chi(\sigma'_v) = \chi(\sigma''_v) = 0. \quad (4.54)$$

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 the Appendix. 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_{xz} 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 IRs. 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.

4.2.5 Projection 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:

$$P_{A_1} = \sum_S \chi_{A_1}(S)S, \quad (4.55)$$

$$P_E = \sum_S \chi_E(S)S, \quad (4.56)$$

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

$$\begin{aligned} P_{A_1} 1s_{H_1} &= 1s_{H_1} + 1s_{H_2} + 1s_{H_3} + 1s_{H_2} + 1s_{H_3} + 1s_{H_1} \\ &= 2(1s_{H_1} + 1s_{H_2} + 1s_{H_3}) = \phi_{A_1}, \end{aligned} \quad (4.57)$$

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

$$P_E 1s_{H_1} = 2 \times 1s_{H_1} - 1s_{H_2} - 1s_{H_3} \equiv \phi_{E,1}, \quad (4.58)$$

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 $1s_{H_3}$:

$$P_E 1s_{H_2} = 2 \cdot 1s_{H_2} - 1s_{H_1} - 1s_{H_3} \equiv \phi_{E,2}, \quad (4.59)$$

$$P_E 1s_{H_3} = 2 \cdot 1s_{H_3} - 1s_{H_1} - 1s_{H_2} = \phi_{E,3}. \quad (4.60)$$

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:

$$\phi_{E,3} = -(\phi_{E,1} + \phi_{E,2}). \quad (4.61)$$

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 $\phi'_E = a\phi_{E,1} + b\phi_{E,3}$ is orthogonal to $\phi_{E,1}$: $\int \phi'_E \phi_{E,1} d\tau = 0$. 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.

4.2.6 Summary

Let us now summarize what we have learned. 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). 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

$$\chi(S) = \sum_i n_i \chi_i(S), \quad (4.62)$$

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

$$n_i = \frac{1}{g} \sum_S \chi(S) \chi_i(S), \quad (4.63)$$

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:

$$n_{A_1} = \frac{1}{6}(3 \cdot 1 + 2 \cdot 0 \cdot 1 + 3 \cdot 1 \cdot 1) = 1. \quad (4.64)$$

$$n_{A_2} = \frac{1}{6}(3 \cdot 1 + 2 \cdot 0 \cdot 1 + 3 \cdot 1 \cdot (-1)) = 0. \quad (4.65)$$

$$n_E = \frac{1}{6}(3 \cdot 2 + 2 \cdot 0 \cdot (-1) + 3 \cdot 1 \cdot 0) = 1. \quad (4.66)$$

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

$$P_i = \sum_j \chi_i(S)S \quad (4.67)$$

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 of 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:

$$n_{A_1} = \frac{1}{6}[1 \cdot 1 \cdot 12 + 2 \cdot 1 \cdot 0 + 3 \cdot 1 \cdot 2] = 3, \quad (4.68)$$

$$n_{A_2} = \frac{1}{6}[1 \cdot 1 \cdot 12 + 2 \cdot 1 \cdot 0 + 3 \cdot (-1) \cdot 2] = 1, \quad (4.69)$$

$$n_E = \frac{1}{6}[1 \cdot 2 \cdot 12 + 2 \cdot (-1) \cdot 0 + 3 \cdot 0 \cdot 2] = 4. \quad (4.70)$$

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.

4.2.7 Direct product representations

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

$$S(\phi_1\phi_2\phi_3\cdots\phi_N) = (S\phi_1)(S\phi_2)(S\phi_3)\cdots(S\phi_N). \quad (4.71)$$

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 short-cut 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:

$$\chi(S) = \prod_i \chi_i(S) = (\chi_{A_1}(S))^2 (\chi_{A_2}(S))^2 (\chi_E(S))^2. \quad (4.72)$$

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 $a_1^2 a_2^2 e_x^m e_{x'}^{m'}$, where m and m' denote the occupation (1 or 0) of the two degenerate orbitals e_x and $e_{x'}$. 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 10 of my *QMIC* text. 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.

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

$$\int \psi_a^* \mathbf{V} \psi_b d\tau = V_{a,b} \quad (4.73)$$

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

$$\sum_S \chi_A(S) S \psi_a^* \mathbf{V} \psi_b \quad (4.74)$$

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.

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:

$$\int \psi_a^* \mathbf{r} \psi_b d\tau. \quad (4.75)$$

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

$$\mathbf{r} = e \sum_n Z_n \mathbf{R}_n - e \sum_j \mathbf{r}_j. \quad (4.76)$$

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 into the $\pi^* b_1$ 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:

$$a_2 \times b_1 = b_2, \quad (4.77)$$

$$a_2 \times b_2 = b_1, \quad (4.78)$$

$$a_2 \times a_1 = a_2. \quad (4.79)$$

There is no component of a_1 symmetry in the triple direct product, so the integral vanishes. This allows us to conclude that the $n \rightarrow \pi^*$ excitation in formaldehyde is electric dipole forbidden.

4.2.8 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. 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.