

## 22 The Variational Principle

1. We have so far dealt with particle in a box, hydrogen atom and harmonic oscillator. These were problems that can be solved analytically. However, all other chemical problems (with more than one electron) are problems that cannot be solved exactly and approximate methods are necessary to treat such problems. For the rest of the course we will deal with some basic ideas that are required to understand these approximate methods. The first such idea is the variational principle which we will study here.
2. We want to solve the time-independent Schrödinger Equation

$$H |\psi_n\rangle = E_n |\psi_n\rangle \quad (22.1)$$

for some given Hamiltonian operator. If we left multiply both sides by  $\langle\psi_n|$ :

$$\begin{aligned} \langle\psi_n| H |\psi_n\rangle &= \langle\psi_n| E_n |\psi_n\rangle \\ \frac{\langle\psi_n| H |\psi_n\rangle}{\langle\psi_n| \psi_n\rangle} &= E_n \end{aligned} \quad (22.2)$$

3. The whole purpose of the time-independent Schrödinger Equation is to find the  $|\psi_n\rangle$ . If we know this we know the energy from the above expression. In most cases we can find an expression that approximates  $|\psi_n\rangle$ . (We will study a couple of such approximations for molecular systems later in this course.) Let us say  $|\phi\rangle$  is an approximation to  $|\psi_n\rangle$ . Therefore associated with this approximation is an approximate energy given by:

$$E'[\phi] = \frac{\langle\phi|H|\phi\rangle}{\langle\phi|\phi\rangle} \quad (22.3)$$

what are the properties of such an approximate energy expression?

4. Let us say we can write  $|\phi\rangle = |\psi_n\rangle + \alpha|\chi\rangle$ . Using this:

$$\begin{aligned} (H - E_n)|\psi_n\rangle &= 0 \\ (H - E_n)[|\phi\rangle - \alpha|\chi\rangle] &= 0 \\ (H - E_n)|\phi\rangle &= \alpha(H - E_n)|\chi\rangle \end{aligned} \quad (22.4)$$

Now what is  $E'[\phi] - E_n$ ?

$$\begin{aligned} E'[\phi] - E_n &= \frac{\langle \phi | (H - E_n) | \phi \rangle}{\langle \phi | \phi \rangle} = \frac{\langle \phi | (H - E_n) | \alpha \chi \rangle}{\langle \phi | \phi \rangle} \\ &= \frac{\langle \alpha \chi | (H - E_n) | \alpha \chi \rangle}{\langle \phi | \phi \rangle} \\ &= |\alpha|^2 \frac{\langle \chi | (H - E_n) | \chi \rangle}{\langle \phi | \phi \rangle} \end{aligned} \quad (22.5)$$

where we have used  $\{(H - E_n) | \phi \rangle\}^\dagger = \langle \phi | (H - E_n)$  since  $(H - E_n)$  is Hermitian.

5. The last expression in Eq. (22.5) tells us that  $E'[\phi]$  differs from  $E_n$  by an amount that is second order in the variation in  $\alpha |\chi\rangle$ . If  $\alpha$  is small,  $\phi$  is close to  $\psi$  and  $E'[\phi]$  will be close to  $E_n$ .
6. And once  $\phi$  is close to  $\psi$ , the energy is stationary with respect to small variations in  $\phi$ . For small  $\alpha$ ,  $|\alpha|^2$  changes even slower. This is what is meant when we say that the energy is second order in  $\phi$ . (Notice that the numerator in Eq. (22.5) is also quadratic in  $\phi$ .)

7. Let us say we know the eigenfunctions  $\{\psi_i\}$  of  $H$  and say these have eigenvalues  $\{E_i\}$ . Since  $H$  is a Hermitian operator, these eigenfunctions form a complete set and hence we can expand the function  $\phi$  using this complete set:

$$\phi = \sum_i c_i \psi_i \quad (22.6)$$

where  $c_j = \langle \psi_j | \phi \rangle$ , since  $\{\psi_i\}$  are complete and orthonormal. (To check this left multiply the above equation by  $\langle \psi_j |$  to see what you get.) Then the energy of the function  $\phi$  is:

$$\begin{aligned} E'[\phi] &= \frac{\langle \phi | H | \phi \rangle}{\langle \phi | \phi \rangle} = \frac{\langle \phi | H \sum_i c_i |\psi_i\rangle}{\langle \phi | \phi \rangle} \\ &= \frac{\langle \phi | \sum_i c_i E_i |\psi_i\rangle}{\langle \phi | \phi \rangle} = \frac{\sum_i c_i E_i \langle \phi | \psi_i \rangle}{\langle \phi | \phi \rangle} \\ &= \frac{\sum_i |c_i|^2 E_i}{\langle \phi | \phi \rangle} \end{aligned} \quad (22.7)$$

Now, the denominator:

$$\begin{aligned} \langle \phi | \phi \rangle &= \sum_j c_j^* \langle \psi_j | \sum_i c_i |\psi_i\rangle = \sum_i \sum_j c_j^* c_i \langle \psi_j | \psi_i \rangle \\ &= \sum_i \sum_j c_j^* c_i \delta_{i,j} = \sum_i |c_i|^2 \end{aligned} \quad (22.8)$$

Therefore,

$$E'[\phi] = \frac{\sum_i |c_i|^2 E_i}{\sum_i |c_i|^2} \quad (22.9)$$

Now the energies are ordered such that  $E_1 \leq E_2 \leq E_3 \leq \dots \leq E_n \leq \dots$ . The Eq. (22.9) is a weighted average of these energies. **Hence, the lowest value that  $E'[\phi]$  can have is when  $c_1 = 1$  and all other  $c_i = 0$ , that is when  $\phi$  is the ground state  $\psi_1$ .** For all other  $\phi$ ,  $E'[\phi] > E_1$ , hence

$$E'[\phi] \geq E_1 \quad (22.10)$$

and its minimum value is when  $\phi \equiv \psi_1$ . This is the variational principle.

8. Notice that we had chosen our “basis” functions as the eigenstates of the Hamiltonian. If we had not done that we would have found that the lowest energy that we can get for  $E'$ , given a finite basis set, is always greater than or equal to the ground state energy. To put it in a different way, the lowest value of energy  $E'$  obtained from any finite basis forms an *upper bound* to the ground state energy. That is the ground state is either below or equal to the lowest value of  $E'$ .

9. Now we shall see how we can use the variational principle. *The important point that the variational principle conveys to us is that the ground state energy is the minimum value of all possible energies that any vector can have.* We will see how this works below. Let us expand a “trial” function  $\phi$  using some complete set of vectors  $|f_i\rangle$ , that is  $|\phi\rangle = a_i |f_i\rangle$ . But let us now assume that  $|f_i\rangle$  do not form an orthonormal set. That is  $\langle f_i | f_j \rangle \neq \delta_{i,j}$ . [So  $\{|f_i\rangle\}$  are our coordinate system, and we are trying to find  $|\phi\rangle = a_i |f_i\rangle$ , represented in that coordinate system. And in our case the coordinate system is not necessarily an orthogonal coordinate systems. So, what space are we in again?]

Then

$$\begin{aligned}
 E'[\phi] &= \frac{\langle \phi | H | \phi \rangle}{\langle \phi | \phi \rangle} \\
 &= \frac{\sum_j a_j^* \langle f_j | H \sum_i a_i | f_i \rangle}{\sum_j a_j^* \langle f_j | \sum_i a_i | f_i \rangle} \\
 &= \frac{\sum_i \sum_j a_i a_j^* \langle f_j | H | f_i \rangle}{\sum_i \sum_j a_i a_j^* \langle f_j | f_i \rangle} = \frac{\sum_i \sum_j a_i a_j^* H_{i,j}}{\sum_i \sum_j a_i a_j^* S_{i,j}} \quad (22.11)
 \end{aligned}$$

where we have defined:  $H_{i,j} = \langle f_j | H | f_i \rangle$  and  $S_{i,j} = \langle f_j | f_i \rangle$ . These are called “matrix elements”. Why? Consider  $H_{i,j} = \langle f_j | H | f_i \rangle$ , for each  $i$  and  $j$  we have one number. If there are  $N$  basis functions  $f_i$ , then we have an  $N \times N$  matrix. Similarly for  $S_{i,j} = \langle f_j | f_i \rangle$ . Using these definitions we can rewrite the above equation as:

$$\sum_i \sum_j a_i a_j^* [H_{i,j} - E' S_{i,j}] = 0 \quad (22.12)$$

10. OK. So far we have just rewritten the expression for  $E'[\phi]$ . Now we want to ask what does the variational principle really mean? We said the minimum value of energy that we can get from any function  $\phi$  is when it is equal to the ground eigenstate. This is the *minimum* value of energy. Hence, from calculus  $\frac{\partial E'}{\partial a_k} = 0$ . Using this in Eq. (22.12):

$$\begin{aligned}
 \frac{\partial}{\partial a_k} \sum_i \sum_j a_i a_j^* [H_{i,j} - E' S_{i,j}] &= \sum_i \sum_j \delta_{i,k} a_j^* [H_{i,j} - E' S_{i,j}] + \\
 &\quad \sum_i \sum_j a_i \delta_{j,k} [H_{i,j} - E' S_{i,j}] + \\
 &\quad \sum_i \sum_j a_i a_j^* \left[ -\frac{\partial E'}{\partial a_k} S_{i,j} \right] \\
 &= \sum_j a_j^* [H_{k,j} - E' S_{k,j}] + \\
 &\quad \sum_i a_i [H_{i,k} - E' S_{i,k}] \\
 &= 0 \qquad (22.13)
 \end{aligned}$$

which is true for every  $k$ . We are free to change the  $i \rightarrow j$  in the second term of the last equation. Then we note that since  $H$  is Hermitian,  $H_{i,k}^* = H_{k,i}$ .

11. If we restrict ourselves to real functions and real values of  $a_i$  then Eq. (22.13) reduces to:

$$\frac{\partial}{\partial a_k} \sum_i \sum_j a_i a_j^* [H_{i,j} - E' S_{i,j}] = 2 \sum_i a_i [H_{i,k} - E' S_{i,k}] = 0 \quad (22.14)$$

Equation (22.14) is a system of equations. There is one equation for each  $k$ . And each equation is a linear equation in the set of *unknowns*  $\{a_i\}$ . Certainly one solution to this system of equations is  $a_1 = a_2 = a_3 = \dots = 0$ . This called the *trivial solution* and in this case is completely meaningless since they  $\phi$  would be zero. For such a set of equations to have a “meaningful” solution, it is required that the determinant:

$$|H - E'S| = 0 \quad (22.15)$$

Solving this determinant gives us  $E'$  which can be used to determine the eigenstates of the Hamiltonian. This equation is called the secular equation. Let us do a couple of examples. (I will do this on the board.)

Can we now see what we really did in

<http://www.indiana.edu/%7Essiweb/C561/PDFfiles/Diag-handout.pdf?>

[If you recall, this was a handout given to you during the first mid-term exam to help you diagonalize a matrix.]