

# Lecture I : The variational principle

## I. THE APPROXIMATE WAVEFUNCTION

In quantum mechanics, the main task is to solve the Schrödinger equation,

$$\hat{H}\psi = E\psi. \quad (1)$$

The Schrödinger equation is exactly solvable for a very narrow class of systems. In cases where the exact solution cannot be attained, the wavefunction may be approximated by a form that is easier to handle mathematically

$$\phi \approx \psi. \quad (2)$$

In most cases, we are interested in the ground state of the system, which we shall denote by  $\psi_0$  yielding the ground state energy,  $E_0$ . The excited states of the system will be denoted by  $\{\psi_1, \psi_2, \dots\}$  with corresponding energies  $\{E_1, E_2, \dots\}$ . In what follows, we will be interested in obtaining an approximation to  $\psi_0$

Unless we are very lucky, the approximate wavefunction  $\phi$  will no longer be an eigenvalue of the Hamiltonian operator,  $\hat{H}$ . The quality of the approximation is assessed based on how close the expectation value of  $\hat{H}$  for  $\psi$  given by

$$\tilde{E} = \frac{\langle \phi | \hat{H} | \phi \rangle}{\langle \phi | \phi \rangle} \quad (3)$$

comes to the actual energy eigenvalue,  $E_0$ . Assuming that the eigenstates of  $\hat{H}$  form a complete basis set, we may expand any other wavefunction of the system in terms of them. This observation applies also to our approximation,  $\phi$ . We therefore write down the following expansion

$$|\phi\rangle = \sum_n c_n |\psi_n\rangle \quad (4)$$

where  $c_n$  are the expansion coefficients. The eigenstates  $\{|\psi_n\rangle\}$  are assumed to be orthonormal. This assumption does not cause us to lose any generality because any complete set of eigenstates may be constructed to be orthonormal. Thus, the following property is satisfied

$$\langle \psi_n | \psi_m \rangle = \delta_{nm} \quad (5)$$

We make the further assumption that the eigenvalues,  $\{E_n\}$  are labeled in an increasing order, i.e.

$$E_0 \leq E_1 \leq E_2 \leq \dots \quad (6)$$

Substituting Eq. 4 into Eq. 3, we obtain the following

$$\tilde{E} = \frac{\langle \phi | \hat{H} | \phi \rangle}{\langle \phi | \phi \rangle} = \frac{\sum_{nm} c_n^* c_m \langle \psi_n | \hat{H} | \psi_m \rangle}{\sum_{nm} c_n^* c_m \langle \psi_n | \psi_m \rangle} \quad (7)$$

$$= \frac{\sum_{nm} c_n^* c_m E_n \langle \psi_n | \psi_m \rangle}{\sum_{nm} c_n^* c_m \langle \psi_n | \psi_m \rangle} = \frac{\sum_n |c_n|^2 E_n}{\sum_n |c_n|^2} \quad (8)$$

The last two equalities are a result of the fact that  $\{|\psi_n\rangle\}$  are eigenstates of  $\hat{H}$  and of orthonormality as stated in Eq. 5. If we now substitute all  $E_n$  in Eq. 8 with  $E_0$ , we have

$$\tilde{E} = \frac{\sum_n |c_n|^2 E_n}{\sum_n |c_n|^2} \leq \frac{\sum_n |c_n|^2 E_0}{\sum_n |c_n|^2} = E_0 \quad (9)$$

where Eq. 6 has been utilized. We thus arrive at the central result that makes the variational method possible and practical:

*Any approximation to the ground state wavefunction will yield an expectation value of the Hamiltonian that is greater than or equal to the ground state energy. Equality is satisfied only in the case that the approximate wavefunction is also a ground state wavefunction.*

In practice, the approximate wavefunction is written in terms of one or more parameters,

$$\phi = \phi(p_1, p_2, \dots, p_N). \quad (10)$$

The set of parameters that yield the best estimate to the ground state energy within the limits of the chosen form of  $\phi$  satisfies

$$\frac{\partial \tilde{E}(p_1, p_2, \dots, p_N)}{\partial p_1} = \frac{\partial \tilde{E}(p_1, p_2, \dots, p_N)}{\partial p_2} = \dots = \frac{\partial \tilde{E}(p_1, p_2, \dots, p_N)}{\partial p_N} = 0 \quad (11)$$

## II. VARIATION WITH RESPECT TO COEFFICIENTS (LINEAR VARIATION) AND LAGRANGE MULTIPLIERS

A very important class of electronic structure methods have variational origin. These methods make the initial assumption that the approximate wavefunction is a *sum* of functions satisfying intuitive properties and makes a variation over the expansion coefficients. Let us assume that the approximate wavefunction for a given system may be expanded in terms of a particular set of orbitals. Because we cannot work with an infinitely many number of such orbitals we truncate the sum and just consider the first  $N$  terms :

$$\phi(\vec{x}) = \sum_{i=1}^N c_i \chi_i(\vec{x}) \quad (12)$$

For each  $k$ , we would like the above expansion to satisfy the minimization condition, i.e.

$$\frac{\partial}{\partial c_k^*} \frac{\langle \phi | \hat{H} | \phi \rangle}{\langle \phi | \phi \rangle}. \quad (13)$$

In addition, we require the approximate wavefunction to remain normalized

$$\langle \phi | \phi \rangle = 1 \quad (14)$$

which then allows us to rewrite Eq. 13 as

$$\frac{\partial}{\partial c_k^*} \langle \phi | \hat{H} | \phi \rangle = 0 \quad (15)$$

We may in fact, satisfy both Eq. 13 and Eq. 14 by introducing a new quantity

$$K = \langle \phi | \hat{H} | \phi \rangle - \lambda [\langle \phi | \phi \rangle - 1] \quad (16)$$

and extending the minimization property to include the extra parameter  $\lambda$ ,

$$\frac{\partial K}{\partial c_k^*} = \frac{\partial K}{\partial \lambda} = 0 \quad (17)$$

Inserting Eq. 16 into Eq. 17 immediately yields

$$\langle \phi | \phi \rangle - 1 = 0 \quad (18)$$

and proves that minimizing  $K$  with respect to *all* of the parameters involved satisfies both of the conditions that we were aiming to satisfy. This method of introducing new variables into the problem to satisfy constraints is often used in classical and quantum mechanics. We may introduce as many variables into the problem as there are constraints. These variables are called *Lagrange multipliers*. The Lagrange multipliers are introduced as arbitrary parameters initially, however we shall see later on that they may correspond to physically meaningful quantities.

Inserting the expansion in Eq. 12 into Eq. 15 yields

$$\begin{aligned} & \frac{\partial}{\partial c_k^*} \left[ \langle \sum_i c_i \chi_i | \hat{H} | \sum_j c_j \chi_j \rangle - \lambda \left( \langle \sum_i c_i \chi_i | \sum_j c_j \chi_j \rangle - 1 \right) \right] \\ &= \frac{\partial}{\partial c_k^*} \sum_{i,j} c_i^* c_j [\langle \chi_i | \hat{H} | \chi_j \rangle - \lambda \langle \chi_i | \chi_j \rangle] \\ &= \sum_j c_j [\langle \chi_k | \hat{H} | \chi_j \rangle - \lambda \langle \chi_k | \chi_j \rangle] = 0 \end{aligned} \quad (19)$$

Rearranging Eq. 19, we obtain

$$\sum_j \langle \chi_k | \hat{H} | \chi_j \rangle c_j = \lambda \sum_j \langle \chi_k | \chi_j \rangle c_j \quad (20)$$

which we recognize immediately as a *generalized* eigenvalue equation

$$\mathbf{H} \cdot \mathbf{C} = \lambda \cdot \mathbf{S} \cdot \mathbf{C} \quad (21)$$

where  $\mathbf{H}$  and  $\mathbf{S}$  are the matrix representations of the Hamiltonian and overlap operators and their elements are defined by.

$$\begin{aligned} H_{nm} &= \langle \chi_n | \hat{H} | \chi_m \rangle \\ S_{nm} &= \langle \chi_n | \chi_m \rangle \end{aligned} \quad (22)$$

$\lambda$  in Eq. 21 corresponds to a diagonal matrix whose elements on the diagonal are the eigenvalues.

If we use  $N$  basis functions to expand the trial function  $\phi$ , Eq. 21 then gives  $N$  eigenvalues. But what do the eigenvalues correspond to? In order to see that let's sum both sides of Eq. 20 and isolate  $\lambda$  :

$$\lambda = \frac{\sum_{j,k} c_k^* c_j \langle \chi_k | \hat{H} | \chi_j \rangle}{\sum_{j,k} c_k^* c_j \langle \chi_k | \chi_j \rangle} = \frac{\langle \phi | \hat{H} | \phi \rangle}{\langle \phi | \phi \rangle}. \quad (23)$$

Eq. 23 implies that each of the  $N$  eigenvectors correspond to a series of expansion coefficients yielding different  $\phi$ 's and each  $\lambda$  corresponds to a different expectation value. The eigenvector corresponds to the smallest eigenvalue then corresponds to the best  $\phi$  and the smallest eigenvalue itself is the closest approximation to the ground state energy for the approximate form in Eq. 12.

### III. AN EXAMPLE : DELTA-FUNCTION POTENTIAL

As discussed above, one can imagine solving a variational problem adopting two different but equivalent strategies :

1. Assume a single wavefunction form having many parameters and minimize over the parameters
2. Assume a linear combination of wavefunctions without any parameters and minimize over the coefficients.

The second strategy is usually preferred in all but the simplest systems, both because the introduction of the variational parameters is more easily controlled and also because it results in an eigenvalue problem that has a straightforward solution. In the example below we are going to compare these two different families of solutions.

**Question 1 :** A very simple model for the potential seen by a hydrogen electron due to the nucleus is a single delta function  $V_H(x) = -\sqrt{\pi}\delta(x)$ . Make an estimate to the ground state energy assuming that the approximate wavefunction  $\phi_\alpha(x) = \exp(-\alpha x^2/2)$ .

**Solution :** First let us write the one-dimensional Hamiltonian using atomic coordinates ( $\hbar = m_e = 1$ ), which we'll see in detail later.

$$\hat{H} = \underbrace{-\frac{1}{2} \frac{d^2}{dx^2}}_{\hat{T}} + \underbrace{(-\sqrt{\pi}\delta(x))}_{\hat{V}} \quad (24)$$

The calculation of the energy requires the evaluation of a good deal of Gaussian-like integrals. We will therefore first remind ourselves of some simple identities :

$$\int_{-\infty}^{\infty} e^{-\beta x^2} dx = \sqrt{\pi} \beta^{-1/2} \quad (25)$$

$$\int_{-\infty}^{\infty} x^2 e^{-\beta x^2} dx = -\frac{\partial}{\partial \beta} \int_{-\infty}^{\infty} e^{-\beta x^2} dx = -\frac{\partial(\sqrt{\pi} \beta^{-1/2})}{\partial \beta} = \frac{1}{2} \sqrt{\pi} \beta^{-3/2} \quad (26)$$

Another identity that would make our life easier concerns the kinetic energy. For a given wavefunction, the kinetic energy may be written as :

$$T = -\frac{1}{2} \int_{-\infty}^{\infty} \phi^*(x) \frac{d^2}{dx^2} \phi(x) dx = \frac{1}{2} \int_{-\infty}^{\infty} |\phi'(x)|^2 dx \quad (27)$$

This result is obtained by employing integration by parts and taking the resulting surface term to zero.

The potential energy is much more easily calculated since it is just a delta function

$$V = -\sqrt{\pi} \int_{-\infty}^{\infty} \delta(x) e^{-\alpha x^2} dx = -\sqrt{\pi} \quad (28)$$

Finally putting everything together, we obtain the  $\alpha$ -dependent variational energy as

$$\tilde{E}(\alpha) = \frac{1}{4}\alpha - \alpha^{1/2}. \quad (29)$$

By taking the derivative with respect to  $\alpha$  of Eq. 29, we find that the value of  $\alpha$  which minimizes this energy is  $\alpha = 4$  and the corresponding minimum energy is  $E(\alpha = 1) = -1$ .  $\tilde{E}(\alpha)$  as a function of  $\alpha$  is displayed in Fig. 2

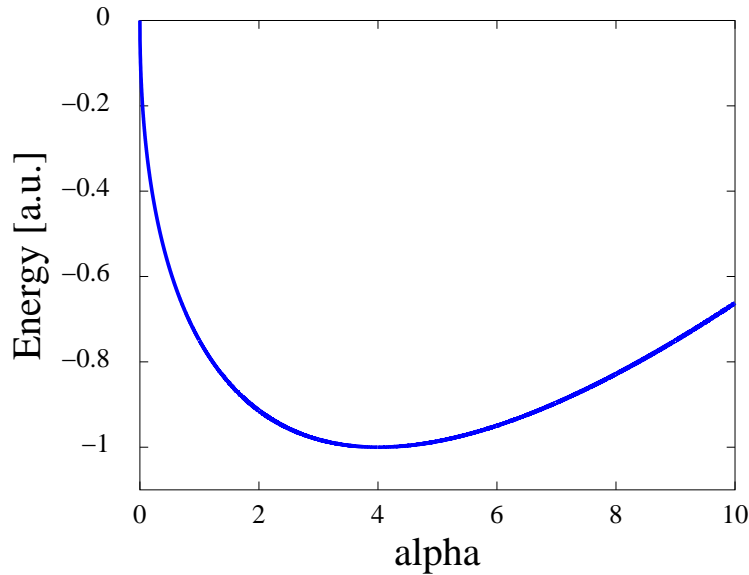


FIG. 1:  $\tilde{E}$  as a function of  $\alpha$ . The energy minimum is clearly seen at  $\alpha = 4$ .

**Question 2 :** Consider the same delta potential as in Question 1. This time, make an estimate to the ground state energy assuming that the approximate wavefunction is a linear combination of two Gaussian functions of different width,  $\phi(x) = c_1 e^{-x^2/2} + c_2 e^{-2x^2/2}$ .

**Solution :** In accordance with Eq. 21, we need to first construct the Hamiltonian and overlap matrices for the set of two functions in the linear expansion above,  $\phi_1(x) = e^{-x^2/2}$  and  $\phi_2(x) = e^{-2x^2/2}$ . Let's see how we construct the matrix elements with an example :

$$\begin{aligned} H_{11} &= T_{11} + V_{11} = \frac{1}{2} \langle \chi_1 | \hat{T} | \chi_1 \rangle + \langle \chi_1 | \hat{V} | \chi_1 \rangle \\ &= \int_{-\infty}^{\infty} x^2 e^{-x^2} dx - \int_{-\infty}^{\infty} \sqrt{\pi} \delta(x) e^{-x^2} dx = \frac{1}{2} \sqrt{\pi} - \sqrt{\pi} = -\frac{\sqrt{\pi}}{2} \\ S_{11} &= \langle \chi_1 | \chi_1 \rangle = \int_{-\infty}^{\infty} e^{-x^2} dx = \sqrt{\pi} \end{aligned} \quad (30)$$

When all the matrix elements are calculated we obtain the Hamiltonian and overlap matrices in the basis of  $\{\chi_1, \chi_2\}$

$$H = \sqrt{\pi} \begin{pmatrix} -3/4 & (2/27)^{1/2} - 1 \\ (2/27)^{1/2} - 1 & \sqrt{1/8} - 1 \end{pmatrix} \quad (31)$$

$$S = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \sqrt{\pi} \begin{pmatrix} 1 & (2/3)^{1/2} \\ (2/3)^{1/2} & \sqrt{1/2} \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} \quad (32)$$

Using these matrices, the generalized eigenvalue equation in Eq. 21 can be solved using algorithms available in software such as Matlab or may be hand-coded. Below is an excerpt from an Octave (Matlab clone for Linux) session.

```
octave:1> a=sqrt(2/27)
a = 0.27217
octave:2> b=sqrt(2/3)
b = 0.81650
octave:3> H=sqrt(pi)*[-3/4 a-1;a-1 sqrt(1/8)-1]
H =

    -1.3293    -1.2901
    -1.2901    -1.1458

octave:4> S=sqrt(pi)*[1 b;b sqrt(1/2)]
S =

    1.7725    1.4472
    1.4472    1.2533

octave:5> [AA,BB,Q,Z,V,W,lambda]=qz(H,S);
octave:6> V
V =

    0.13024   -0.87543
   -1.00000    1.00000

octave:7> lambda
lambda =

   -0.91825
    1.20935
```

The algorithm used in the above calculation is called *QZ decomposition* and the resulting eigenvalues are contained in the variable `lambda`. The ground state energy is given by the lower of the two eigenvalues. The coefficients  $c_1$  and  $c_2$  are the elements of the first column of `V`. The other column corresponds to the "excited state" of the variational problem. One must exercise caution in talking about the excited state here because although the ground state energy found using variational principle is an upper bound to the true ground state energy, there is no such guarantee for higher energy states.

The "ground" and "excited" states of this problem are plotted in Fig. ???. The Octave codes used to generate this plot can be found on the course Web site. Note that the integrals and derivatives in the Octave functions have been evaluated numerically rather than analytically.

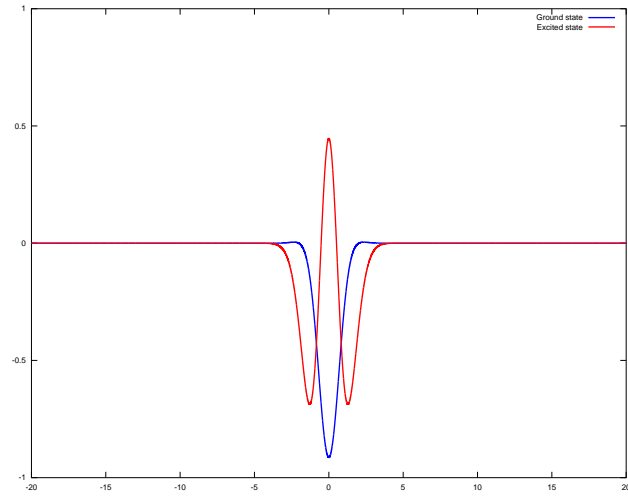


FIG. 2: Ground and excited states of the linear variation problem.