7.2: Linear Variational Method and the Secular Determinant

🧠 Learning Objectives

- Understand how the variational method can be expanded to include trial wavefunctions that are a linear combination of functions with coefficients that are the parameters to be varied.
- To be able to construct secular equations to solve the minimization procedure intrinsic to the variational method approach.
- To map the secular equations into the secular determinant
- To understand how the Linear Combination of Atomic Orbital (LCAO) approximation is a specific application of the linear variational method.

A special type of variation widely used in the study of molecules is the so-called **linear variation function**, where the trial wavefunction is a linear combination of N linearly independent functions (often atomic orbitals) that not the eigenvalues of the Hamiltonian (since they are not known). For example

$$|\psi_{trial}\rangle = \sum_{j=1}^{N} a_j |\phi_j\rangle \tag{7.2.1}$$

and

$$\langle\psi_{trial}| = \sum_{j=1}^{N} a_j^* \langle\phi_j| \tag{7.2.2}$$

In these cases, one says that a 'linear variational' calculation is being performed.

📌 Linear Variational Basis Functions

The set of functions $\{\phi_j\}$ are called the 'linear variational' basis functions and are nothing more than members of a set of functions that are convenient to deal with. However, they are typically not arbitrary and are usually selected to address specific properties of the system:

- to obey all of the boundary conditions that the exact state $|\psi_{trial}\rangle$ obeys,
- to be functions of the the same coordinates as $|\psi_{trial}\rangle$,
- to be of the same symmetry as $|\psi_{trial}\rangle$, and
- to be convenient to evaluate Hamiltonian terms elements $\langle\phi_i|H|\phi_j\rangle$.

Beyond these conditions, nothing other than effort can limit the selection and number of such basis functions in the expansions in Equations 7.2.1 and 7.2.2.

As discussed in Section 7.1, the variational energy for a generalized trial wavefunction is

$$E_{trial} = \frac{\langle\psi_{trial}|\hat{H}|\psi_{trial}\rangle}{\langle\psi_{trial}|\psi_{trial}\rangle} \tag{7.2.3}$$

Substituting Equations 7.2.1 and 7.2.2 into Equation 7.2.3 involves addressing the numerator and denominator individually. For the numerator, the integral can be expanded thusly:

$$\langle\psi_{trial}|H|\psi_{trial}\rangle = \sum_i^N \sum_j^N a_i^* a_j \langle\phi_i|H|\phi_j\rangle. \tag{7.2.4}$$

$$= \sum_{i,j}^{N,N} a_i^* a_j \langle\phi_i|H|\phi_j\rangle. \tag{7.2.5}$$

We can rewrite the following integral in Equation 7.2.5 as a function of the basis elements (not the trial wavefunction) as

$$H_{ij} = \langle\phi_i|H|\phi_j\rangle$$

So the numerator of the right side of Equation 7.2.3 becomes

$$\langle \psi_{trial} | H | \psi_{trial} \rangle = \sum_{i,\,j}^{N,\,N} a_i^* a_j H_{ij} \tag{7.2.6}$$

Similarly, the denominator of the right side of Equation 7.2.3 can be expanded

$$\langle \psi_{trial} | \psi_{trial} \rangle = \sum_{i,\,j}^{N,\,N} a_i^* a_j \langle \phi_i | \phi_j \rangle \tag{7.2.7}$$

We often simplify the integrals on the right side of Equation 7.2.7 as

$$S_{ij} = \langle \phi_i | \phi_j \rangle$$

where S_{ij} are **overlap integrals** between the different $\{\phi_j\}$ basis functions. Equation 7.2.7 is thus expressed as

$$\langle \psi_{trial} | \psi_{trial} \rangle = \sum_{i,\,j}^{N,\,N} a_i^* a_j S_{ij} \tag{7.2.8}$$

> ☰ Orthonormality of Basis Functions
>
> There is no explicit rule that the $\{\phi_j\}$ functions have to be **orthogonal** or **normalized** functions, although they often are selected that way for convenience. Therefore, *a priori*, S_{ij} does not have to be δ_{ij}.

Substituting Equations 7.2.6 and 7.2.8 into the variational energy formula (Equation 7.2.3) results in

$$E_{trial} = \frac{\displaystyle\sum_{i,\,j}^{N,\,N} a_i^* a_j H_{ij}}{\displaystyle\sum_{i,\,j}^{N,\,N} a_i^* a_j S_{ij}} \tag{7.2.9}$$

For such a trial wavefunction as Equation 7.2.1, the variational energy depends quadratically on the 'linear variational' a_j coefficients. These coefficients can be varied just like the parameters in the trial functions of Section 7.1 to find the optimized trial wavefunction ($|\psi_{trial}\rangle$) that approximates the true wavefunction ($|\psi\rangle$) that we cannot analytically solve for.

Minimizing the Variational Energy

The expression for variational energy (Equation 7.2.9) can be rearranged

$$E_{trial} \sum_{i,\,j}^{N,\,N} a_i^* a_j S_{ij} = \sum_{i,\,j}^{N,\,N} a_i^* a_j H_{ij} \tag{7.2.10}$$

The optimum coefficients are found by searching for minima in the variational energy landscape spanned by varying the $\{a_i\}$ coefficients (Figure 7.2.1).
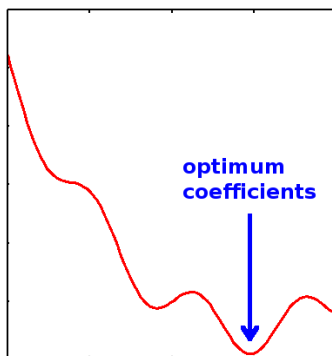
Figure 7.2.1 : Linear Variational Principle in action. Two (or more) wavefunctions are mixed by linear combination. The coefficients a_1, a_2 determine the weight each of them is given. The optimum coefficients are found by searching for minima in the potential landscape spanned by a_1 and a_2. (CC BY-SA-NC; Rudolf Winter; Aberystwyth University).

We want to minimize the energy with respect to the linear coefficients $\{a_i\}$, which requires that

$$\frac{\partial E_{trial}}{\partial a_i} = 0$$

for all i.

Differentiating both sides of Equation 7.2.10 for the kth coefficient gives,

$$\frac{\partial E_{trial}}{\partial a_k} \sum_{i,j}^{N,N} a_i^* a_j S_{ij} + E_{trial} \sum_i \sum_j \left[\frac{\partial a_i^*}{\partial a_k} a_j + \frac{\partial a_j}{\partial a_k} a_i^* \right] S_{ij} = \sum_{i,j}^{N,N} \left[\frac{\partial a_i^*}{\partial a_k} a_j + \frac{\partial a_j}{\partial a_k} a_i^* \right] H_{ij} \qquad (7.2.11)$$

Since the coefficients are independent

$$\frac{\partial a_i^*}{\partial a_k} = \delta_{ik}$$

and

$$S_{ij} = S_{ji}$$

and also since the Hamiltonian is a Hermitian Operator (see below)

$$H_{ij} = H_{ji}$$

then Equation 7.2.11 simplifies to

$$\frac{\partial E_{trial}}{\partial a_k} \sum_i \sum_j a_i^* a_j S_{ij} + 2 E_{trial} \sum_i a_i S_{ik} = 2 \sum_i a_i H_{ik} \qquad (7.2.12)$$

At the minimum variational energy, when

$$\frac{\partial E_{trial}}{\partial a_k} = 0$$

then Equation 7.2.12 gives

$$\sum_i^N a_i \left(H_{ik} - E_{trial} S_{ik} \right) = 0 \qquad (7.2.13)$$

for all k. The equations in 7.2.13 are called the **Secular Equations**.

📌 Hermitian Operators

Hermitian operators are operators that satisfy the general formula

$$\langle \phi_i | \hat{A} | \phi_j \rangle = \langle \phi_j | \hat{A} | \phi_i \rangle \qquad (7.2.14)$$

> If that condition is met, then $\hat{A}$ is a Hermitian operator. For any operator that generates a real eigenvalue (e.g., observables), then that operator is Hermitian. The Hamiltonian $\hat{H}$ meets the condition of a Hermitian operator. Equation 7.2.14 can be rewriten as
>
> $$A_{ij} = A_{ji}^*$$
>
> where
>
> $$A_{ij} = \langle \phi_i | \hat{A} | \phi_j \rangle$$
>
> and
>
> $$A_{ji} = \langle \phi_j | \hat{A} | \phi_i \rangle$$
>
> Therefore, when applied to the Hamiltonian operator
>
> $$H_{ij}^* = H_{ji}.$$

If the functions $\{|\phi_j\rangle\}$ are orthonormal, then the overlap matrix S reduces to the unit matrix (one on the diagonal and zero every where else) and the Secular Equations in Equation 7.2.13 reduces to the more familiar Eigenvalue form:

$$\sum_i^N H_{ij} a_j = E_{trial} a_i. \tag{7.2.15}$$

Hence, the secular equation, in either form, have as many eigenvalues E_i and eigenvectors $\{C_{ij}\}$ as the dimension of the H_{ij} matrix as the functions in $|\psi_{trail}\rangle$ (Example 7.2.1). It can also be shown that between successive pairs of the eigenvalues obtained by solving the secular problem at least one exact eigenvalue must occur (i.e., $E_{i+1} > E_{exact} > E_i$, for all i). This observation is referred to as '*the bracketing theorem*'.
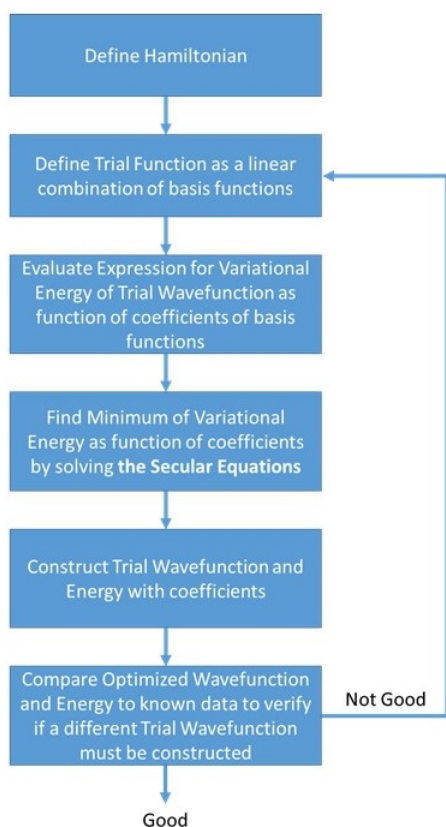

Figure 7.2.2 : Simplified algorithmic flowchart of the linear Variational Method approximation. Compared to the similar algorithmic flowchart for the "normal" Variational Method approximation.

Variational methods, in particular the linear variational method, are the most widely used approximation techniques in quantum chemistry. To implement such a method one needs to know the Hamiltonian H whose energy levels are sought and one needs to construct a trial wavefunction in which some 'flexibility' exists (e.g., as in the linear variational method where the a_j coefficients can be varied). This tool will be used to develop several of the most commonly used and powerful molecular orbital methods in chemistry.

The Secular Determinant

From the secular equations with an orthonormal functions (Equation), we have k simultaneous secular equations in k unknowns. These equations can also be written in matrix notation, and for a non-trivial solution (i.e. $c_i \neq 0$ for all i), the determinant of the secular matrix must be equal to zero.

$$|H_{ik} - ES_{ik}| = 0 \qquad (7.2.16)$$

> ⏻ **Properties of Determinants**
>
> - The determinant is a **real** number, it is **not** a matrix.
> - The determinant can be a negative number.
> - It is **not** associated with absolute value at all except that they both use vertical lines.
> - The determinant **only** exists for square matrices (2×2, 3×3, ..., $n \times n$). The determinant of a 1×1 matrix is that single value in the determinant.
> - The inverse of a matrix will exist only if the determinant is not zero.

> ⏻ **Expanding Determinants**
>
> The determinant can be evaluated using an expansion method involving minors and cofactors. Before we can use them, we need to define them. It is the product of the elements on the main diagonal minus the product of the elements off the main diagonal. In the case of a 2×2 matrix, the specific formula for the determinant is
>
> $$|A| = \begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc.$$
>
> Similarly, suppose we have a 3×3 matrix A, and we want the specific formula for its determinant $|A|$:
>
> $$|A| = \begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix} = a\begin{vmatrix} e & f \\ h & i \end{vmatrix} - b\begin{vmatrix} d & f \\ g & i \end{vmatrix} + c\begin{vmatrix} d & e \\ g & h \end{vmatrix}$$
>
> $$= aei + bfg + cdh - ceg - bdi - afh.$$

To solve Equation , the determinate should be expanded and then set to zero. That generates a polynomial (called a **characteristic equation**) that can be directly solved with linear algebra methods or numerically.

> ✔ **Example 7.2.1 : A Simple Two Component Basis Set**
>
> If $|\psi_{trial}\rangle$ is a linear combination of two functions. In math terms,
>
> $$|\psi_{trial}\rangle = \sum_{n=1}^{N=2} a_n |f_n\rangle = a_1|\phi_1\rangle + a_2|\phi_2\rangle$$
>
> then the secular determinant (Equation), in matrix formulation would look like this
>
> $$\begin{vmatrix} H_{11} - E_{trial}S_{11} & H_{12} - E_{trial}S_{12} \\ H_{12} - E_{trial}S_{12} & H_{22} - E_{trial}S_{22} \end{vmatrix} = 0$$
>
> **Solution**
>
> Solving the secular equations is done by finding E_{trial} and putting the value into the expansion of the secular determinant
>
> $$a_1^2 H_{11} + 2a_1 a_2 H_{12} + a_2^2 H_{22} = 0$$

and

$$a_1(H_{12} - E_{trial}S_{12}) + a_2(H_{22} - E_{trial}S_{22}) = 0$$

Equation 7.2.16 can be solved to obtain the energies E. When arranged in order of increasing energy, these provide approximations to the energies of the first k states (each having an energy higher than the true energy of the state by virtue of the variation theorem). To find the energies of a larger number of states we simply use a **greater number** of basis functions $\{\phi_i\}$ in the trial wavefunction (Example 7.2.1). To obtain the approximate wavefunction for a particular state, we substitute the appropriate energy into the secular equations and solve for the coefficients a_i.

Using this method it is possible to find all the coefficients $a_1 \ldots a_k$ in terms of one coefficient; normalizing the wavefunction provides the absolute values for the coefficients.

> ✔ Example 7.2.2 : Linear Combination of Atomic Orbitals (LCAO) Approximation
>
> Trial wavefunctions that consist of linear combinations of simple functions
>
> $$|\psi(r)\rangle = \sum_i a_i |\phi_i(r)\rangle$$
>
> form the basis of the Linear Combination of Atomic Orbitals (LCAO) method introduced by Lennard and Jones and others to compute the energies and wavefunctions of atoms and molecules. The functions $\{|\phi_i\rangle\}$ are selected so that matrix elements can be evaluated analytically. Two basis sets of atomic orbitals functions can be used: Slater type and Gaussian type:
>
> Slater orbitals using Hydrogen-like wavefunctions
>
> $$|\phi_i\rangle = Y_l^m(\theta, \phi)e^{-\alpha r}$$
>
> and Gaussian orbitals of the form
>
> $$|\phi_i\rangle = Y_l^m(\theta, \phi)e^{-\alpha r^2}$$
>
> are the most widely used forms, where $Y_l^m(\theta, \phi)$ are the spherical harmonics that represent the angular part of the atomic orbitals. Gaussian orbitals form the basis of many quantum chemistry computer codes.

Dec 23, 2018, 11:08 PM

Because Slater orbitals give exact results for Hydrogen, we will use Gaussian orbitals to test the LCAO method on Hydrogen, following S.F. Boys, Proc. Roy. Soc. A 200, 542 (1950) and W.R. Ditchfield, W.J. Hehre and J.A. Pople, J. Chem. Phys. Rev. 52, 5001 (1970) with the basis set. Because products of Gaussians are also Gaussian, the required matrix elements are easily computed.

The linear variational method is used extensively in molecular orbitals of molecules and further examples will be postponed until that discussion in Chapters 9.