The Variation Principle

The variation theorem states that given a system with a Hamiltonian $H$, then if $\varphi$ is any normalised, well-behaved function that satisfies the boundary conditions of the Hamiltonian, then

$$<\varphi|H|\varphi \geq E_0$$

(1)

where $E_0$ is the true value of the lowest energy eigenvalue of $H$. This principle allows us to calculate an upper bound for the ground state energy by finding the trial wavefunction $\varphi$ for which the integral is minimised (hence the name; trial wavefunctions are varied until the optimum solution is found). Let us first verify that the variational principle is indeed correct.

We first define an integral

$$I = <\varphi|-E_0|\varphi>$$

$$= <\varphi|H|\varphi > - <\varphi|E_0|\varphi >$$

$$= <\varphi|H|\varphi > - E_0 <\varphi|\varphi >$$

(�ince $\varphi$ is normalised)

If we can prove that $I \geq 0$ then we have proved the variation theorem.

Let $\psi_i$ and $E_i$ be the true eigenfunctions and eigenvalues of $H$, so $H \psi_i = E_i \psi_i$. Since the eigenfunctions $\psi_i$ form a complete basis set for the space spanned by $H$, we can expand any wavefunction $\varphi$ in terms of the $\psi_i$ (so long as $\varphi$ satisfies the same boundary conditions as $\psi_i$).

$$\varphi = \sum_k a_k \psi_k$$

Substituting this function into our integral $I$ gives

$$I = <\sum_k a_k \psi_k | H-E_0 | \sum_j a_j \psi_j >$$

$$= <\sum_k a_k \psi_k | \sum_j (H-E_0) a_j \psi_j >$$

If we now use $H \psi = E \psi$, we obtain

$$I = <\sum_k a_k \psi_k | \sum_j a_j (E_j-E_0) \psi_j >$$

$$= \sum_k \sum_j a_k^* a_j (E_j-E_0) <\psi_k | \psi_j >$$

$$= \sum_k \sum_j a_k^* a_j (E_j-E_0) \delta_{jk}$$

We now perform the sum over $j$, losing all terms except the $j=k$ term, to give

$$I = \sum_k a_k^* a_k (E_k-E_0)$$

$$= \sum_k |a_k|^2 (E_k-E_0)$$

Since $E_0$ is the lowest eigenvalue, $E_k-E_0$ must be positive, as must $|a_k|^2$. This means that all terms in the sum are non-negative and $I \geq 0$ as required.

For wavefunctions that are not normalised, the variational integral becomes:
\[
\langle \phi | H | \phi \rangle \geq E_0
\]

**Linear variation method**

A special type of variation widely used in the study of molecules is the so-called linear variation function, a linear combination of \( n \) linearly independent functions \( f_1, f_2, \ldots, f_n \) (often atomic orbitals) that satisfy the boundary conditions of the problem, i.e. \( \phi = \Sigma_i c_i f_i \). The coefficients \( c_i \) are parameters to be determined by minimising the variational integral. In this case, we have:

\[
\langle \phi | H | \phi \rangle = \Sigma \sum c_i^* c_j H_{ij} \quad \text{where } H_{ij} \text{ is the Hamiltonian matrix element.}
\]

\[
\langle \phi | \phi \rangle = \Sigma \sum c_i^* c_j S_{ij} \quad \text{where } S_{ij} \text{ is the overlap matrix element.}
\]

The variational energy is therefore

\[
E = \frac{\Sigma \Sigma c_i^* c_j H_{ij}}{\Sigma \Sigma c_i^* c_j S_{ij}}
\]

which rearranges to give

\[
E \Sigma \Sigma c_i^* c_j S_{ij} = \Sigma \Sigma c_i^* c_j H_{ij}
\]

We want to minimise the energy with respect to the linear coefficients \( c_i \), requiring that \( \frac{\partial E}{\partial c_i} = 0 \) for all \( i \).

Differentiating both sides of the above expression gives,

\[
\frac{\partial E}{\partial c_k} \Sigma \Sigma c_i^* c_j S_{ij} + 2 E \Sigma \Sigma [\frac{\partial c_i^*}{\partial c_k} c_j + \frac{\partial c_j}{\partial c_k} c_i^*] S_{ij} = \Sigma \Sigma [\frac{\partial c_i^*}{\partial c_k} c_j + \frac{\partial c_j}{\partial c_k} c_i^*] H_{ij}
\]

Since \( \frac{\partial c_i^*}{\partial c_k} = \delta_{ik} \) and \( S_{ij} = S_{ji}, H_{ij} = H_{ji} \), we have

\[
\frac{\partial E}{\partial c_k} \Sigma \Sigma c_i^* c_j S_{ij} + 2 E \Sigma S_{ik} = 2 \Sigma c_i H_k
\]

When \( \frac{\partial E}{\partial c_k} = 0 \), this gives

\[
\Sigma_i c_i (H_{ik} - ES_{ik}) = 0 \quad \text{for all } k \quad \text{SECULAR EQUATIONS}
\]

We therefore 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, i.e.
\[ |H_{ik} - ES_{ik}| = 0 \]

The resulting algebraic equation 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 \( f_i \) in the trial wavefunction \( \phi \). To obtain the approximate wavefunction for a particular state, we substitute the appropriate energy into the secular equations and solve for the coefficients \( c_i \).

NB: Using this method it is possible to find all the coefficients \( c_1 \ldots c_k \) in terms of one coefficient; normalising the wavefunction provides the absolute values for the coefficients.

**Huckel Molecular Orbital Theory**

Huckel theory is a simplified version of the linear variation method which can be applied to conjugated \( \pi \) systems. We assume \( \sigma-\pi \) separability, which means that the nuclei, electrons of the atomic inner shells and localised \( \sigma \) bonds provide an effective field in which the remaining \( \pi \) electrons move. We then average the effective field.

\[ H_{\pi} = \sum_{i=1}^{n_{\pi}} H_{\text{eff}}(i) \]

The next simplification is to approximate the \( \pi \) molecular orbitals as linear combinations of atomic orbitals. In a minimal basis set calculation of a planar conjugated hydrocarbon, the only atomic orbitals with \( \pi \) symmetry are the 2p orbitals on carbon. The \( \pi \) orbitals can therefore be written

\[ \sigma_i = \sum_{j=1}^{n} c_{ij} f_j \]

where \( f_{ij} \) is a 2p\( \pi \) orbital on the jth carbon atom. The optimum values for the n lowest \( p \) orbitals satisfy the secular equations

\[ \sum_{k=1}^{n} (H_{jk}^{\text{eff}} - S_{jk}E_i)c_{ij} = 0 \]

in which the \( E_i \)'s are the roots of the secular determinant

\[ |H_{jk}^{\text{eff}} - S_{jk}E_i| = 0 \]

The key assumptions in Huckel theory involve the integrals \( H \) and \( S \). We assume that \( H_{ii} \) has the same value for every carbon atom in the molecule, and also for carbon atoms in different planar hydrocarbons. The integral \( H_{ij} \) is assumed to have the same value for any two carbons bonded to each other and to vanish for two non-bonded atoms. The overlap integral \( S_{ij} \) is assumed to be equal to zero unless \( i=j \), in which case it will be equal to unity since the atomic orbitals \( f_i \) are normalised. This can be summarised as:

\[ H_{ii} = \langle f_i | H^{\text{eff}} | f_i \rangle = \alpha \]
\[ H_{ij} = \langle f_i | H^{\text{eff}} | f_j \rangle = \beta \quad \text{for atoms } i, j \text{ bonded} \]
\[ H_{ij} = \langle f_i | H^{\text{eff}} | f_j \rangle = 0 \quad \text{for atoms } i, j \text{ not bonded} \]
\[ S_{ij} = <f_i|f_j> = \delta_{ij} \]

\( \alpha \) and \( \beta \) are called the Coulomb and bond (sometimes resonance) integrals. Since non-bonded carbon atoms are usually well separated, the third assumption of Huckel theory is reasonable, but assuming an overlap integral of zero for bonded atoms is a poor approximation (in practice, the theory is often modified to improve on this assumption).

As an example, consider butadiene, \( H_2C=CH-CH=CH_2 \). For the purposes of Huckel theory, only the connectivity of the carbon framework is important; no distinction is made between the cis- and trans-conformations. The Huckel assumptions give:

\[
\begin{align*}
H_{11} &= H_{22} = H_{33} = H_{44} = \alpha \\
H_{12} &= H_{23} = H_{34} = \beta \\
H_{13} &= H_{14} = H_{24} = 0
\end{align*}
\]

The secular equations are:

\[
\begin{pmatrix}
\alpha - E & \beta & 0 & 0 \\
\beta & \alpha - E & \beta & 0 \\
0 & \beta & \alpha - E & \beta \\
0 & 0 & \beta & \alpha - E
\end{pmatrix}
\begin{pmatrix}
f_1 \\
f_2 \\
f_3 \\
f_4
\end{pmatrix}
= 0
\]

Solving the secular determinant gives the four roots \( E_1 = \alpha - 1.618\beta \), \( E_2 = \alpha - 0.618\beta \), \( E_3 = \alpha + 0.618\beta \) and \( E_4 = \alpha + 1.618\beta \). Substituting each of these energies into the secular equations in turn and solving for the coefficients gives the normalised molecular orbitals:

\[
\begin{align*}
\phi_1 &= 0.372f_1 + 0.602f_2 + 0.602f_3 + 0.372f_4 \\
\phi_2 &= 0.602f_1 + 0.372f_2 - 0.372f_3 - 0.602f_4 \\
\phi_3 &= 0.602f_1 - 0.372f_2 - 0.372f_3 + 0.602f_4 \\
\phi_4 &= 0.372f_1 - 0.602f_2 + 0.602f_3 - 0.372f_4
\end{align*}
\]

The molecular orbital coefficients can be used to determine a range of molecular properties. As an example, one quantity that is sometimes quoted in relation to Huckel theory is the \( \pi \) electron or charge density on an atom. The charge density on atom \( j \) is given by \( q_j = \sum n_i |c_{ij}|^2 \). While this is not a true electron density, measuring neither charge nor probability density, but merely the number of \( \pi \) electrons in the vicinity of atom \( i \), it can sometimes be used to give at least a rough indication of the reactivity of particular atoms with nucleophiles and electrophiles.