6.3. The Rayleigh–Ritz Variational Method

So far our approximation schemes have depended on us knowing the solution to the unperturbed problem which typically describes the major aspects of the system. Of course one can seldom solve such a problem, for example in a multi-electron atom or molecular physics, we cannot ignore all the inter-electron interactions or even spin. The variational method does not rely on solving exactly any simpler Hamiltonian in order to approximate the energy levels.

The mathematical problem underlying the variational method is to find the conditions under which variations in the arbitrary state $|\psi, |\psi\rangle$ leave the functional

$$L[\psi, |\psi\rangle] = \langle \psi | H | \psi\rangle - E \langle \psi | \psi \rangle$$

stationary with Hamilton operators and $E$.
The stationary conditions are found by taking the variational derivatives of $H$ with respect to $|\phi\rangle$ and $|\alpha\rangle$. That is, suppose we allow $|\phi\rangle$ and $|\alpha\rangle$ to vary from infinitesimally in an independent way.

$|\phi\rangle = |\phi_0\rangle + a|\alpha\rangle$

$|\alpha\rangle = |\alpha_0\rangle + b|\beta\rangle$

with $a$, $b$ infinitesimal complex numbers and $|\alpha\rangle$ and $|\beta\rangle$ arbitrary states.

Then

$L[\phi,\alpha] = (\langle\phi_0 | + a^* \langle \alpha |) H (|\alpha_0\rangle + b|\beta\rangle)$

$- E (\langle\phi_0 | + a^* \langle \alpha | (|\alpha_0\rangle + b|\beta\rangle)$

$= L[\phi_0,\alpha_0] + a^* \langle \alpha | (H - E) |\alpha_0\rangle$

$+ b \langle\phi_0 | (H - E) |\beta\rangle$

$+ a^* b L[\alpha_0,\beta]$
Then the variation of $L$ is zero,

$$\delta L = \langle \psi_0 | \mathcal{H} \psi_0 \rangle - \langle \psi_0 | \mathcal{H} \psi_0 \rangle$$

$$= a^* \langle \alpha | (H - E) \psi_0 \rangle + b \langle \psi_0 | (H - E) | \beta \rangle$$

$$= 0$$

if $(H - E) \psi_0 = 0$ and

$$\langle \psi_0 | (H - E) = 0 \Rightarrow (H^+ - E^*) \psi_0 = 0$$

Since $a, b$ are independent,

Then $L$ is stationary if $| \psi_0 \rangle$ is the right eigenvector of $H^+$ and $| \psi_0 \rangle$ is the left eigenvector of $H$ and $E^+$ is the (complex) eigenvalue $E$. If $H = H^+$, as in our case (most of the time),

Then $E = E^*$ and $| \psi_0 \rangle = | \chi_0 \rangle$.

Then $\delta L | \psi_0 \rangle = \langle \chi | (H - E) | \psi_0 \rangle$. Of course, we need only consider one of the eigenvalues, $\delta L | \psi_0 \rangle$, until we evaluate $L$ at the stationary state $| \psi_0 \rangle$; it is not zero, i.e., $| \psi_0 \rangle$ is arbitrary.
The idea of the method then is to choose any trial states $|\psi\rangle$ and $|\phi\rangle$ (guided by physical intuition) which in general depend on certain parameters which we vary until $\langle \phi | L | \psi \rangle$ is stationary. We then can extract approximate values for $E$, the energy levels of the system from this. This approach is encapsulated in the theorem:

**Variational Theorem:** If $H = H^\dagger$ and $E_0$ is the ground state (lowest energy eigenvalue of $H$), then for any state $|\psi\rangle$ the inequality is true

$$E_0 \leq \frac{\langle \psi H | \psi \rangle}{\langle \psi | \psi \rangle}$$

**Proof:** The proof is straightforward. The eigenstates of $H$

$$H |\psi_n\rangle = E_n |\psi_n\rangle$$
The equation is complete in the form of \( \sum_{n} C_n |1_{2n}\rangle \), where we expand any \( |2\rangle \) as \( |2\rangle = \sum_{n} C_n |1_{2n}\rangle \). For orthonormal \( |1_{2n}\rangle \), \( C_n = \langle 1_{2n} | 2 \rangle \) then
\[
\langle 2 | 1_{14} | 2 \rangle = \sum_{n} E_n |C_n|^2
\]
Since \( E_0 \) is the lowest value of \( E_n \),
\[
\langle 2 | 1_{14} | 2 \rangle \geq E_0 \sum_{n} |C_n|^2
\]
Since \( \langle 2 | 2 \rangle = \sum_{n} |C_n|^2 \) we have
\[
\frac{\langle 2 | 1_{14} | 2 \rangle}{\langle 2 | 2 \rangle} \geq E_0
\]
Clearly the equality holds only when \( C_n = 0 \), \( n \neq 0 \). But this implies \( \langle 1_{2} \rangle = C_0 |1_{20}\rangle \) the true ground state.
So the equality is true iff $|2\rangle = c|\chi_0\rangle$.

The expectation value of the Hamiltonian above is called the energy functional

$$E_{\Sigma_4} = \frac{\langle 2 | H | 2 \rangle}{\langle 2 | 2 \rangle}.$$ 

Its lower bound is $E_0 \leq E_{\Sigma_4}$, for any $\Psi$. Note that

$$\langle 2 | 2 \rangle E_{\Sigma_4} = L \Sigma_4 + E \langle 2 | 2 \rangle$$

or

$$L \Sigma_4 = (E_{\Sigma_4} - E) \langle 2 | 2 \rangle$$

defines

$$E_{\Sigma_4} \geq (E_0 - E) \langle 2 | 2 \rangle.$$

Further at the stationary points of $L \Sigma_4$ $E = E_n$ the eigenvalues, thus, for $\Sigma L = 0 \Rightarrow E_{\Sigma_{4,n}} = E_n$ as is clear from its definition.
The variation of the energy functional is found from

\[ \Delta E = \frac{2a_{1} + a \langle \chi_{1} \rangle H \langle \chi_{1} | a \rangle + a \langle \chi_{1} | a \rangle}{\langle \chi_{1} | a \rangle + a \langle \chi_{1} | a \rangle + a \langle \chi_{1} | a \rangle + a \langle \chi_{1} | a \rangle} \]

with \( a \) arbitrary and \( \Delta \) infinitesimal.

\[ \Delta E = \frac{\langle \chi_{1} | H | \chi_{1} \rangle}{\langle \chi_{1} | \chi_{1} \rangle} \left[ 1 - a \frac{\langle \chi_{1} | a \rangle}{\langle \chi_{1} | a \rangle} - a \frac{\langle \chi_{1} | a \rangle}{\langle \chi_{1} | a \rangle} \right] \]

\[ + a \frac{\langle \chi_{1} | H | \chi_{1} \rangle}{\langle \chi_{1} | \chi_{1} \rangle} \]

\[ = \langle \chi_{0} | H | \chi_{0} \rangle + \frac{a}{\langle \chi_{0} | \chi_{0} \rangle} \langle \chi_{0} | (H - 3 E_{0}) | \chi_{0} \rangle \]

\[ + \frac{a}{\langle \chi_{0} | \chi_{0} \rangle} \langle \chi_{0} | (H - 3 E_{0}) | \chi_{0} \rangle \cdot \]

So far \( SE \equiv E \langle \chi_{0} \rangle - 3 E_{0} \chi_{0} = 0 \)

\[ \Rightarrow \quad H \chi_{0} = E E_{0} \chi_{0} \]

Thus \( \chi_{0} \rangle \) is an eigenstate of \( H \)

and \( E E_{0} \) is the eigenvalue.

Thus the stationary points of \( \frac{\partial H}{\partial \mathbf{E}} \) are

the stationary points of \( \mathbf{E} E_{0} \).
\[ \delta \mathcal{E} 3 \mathcal{I} = 0 \iff \delta \mathcal{E} 3 \mathcal{I} \mathcal{J} = 0 \ \text{and both conditions imply} \ \mathcal{H} \mathcal{I} = \mathcal{E} \mathcal{I} \mathcal{J} , \ \mathcal{I} \mathcal{J} \ \text{is a solution to the Schrödinger equation.} \\
(\text{Note: We can use this as the dynamical postulate for quantum mechanical matter through the Schrödinger equation} \ \delta \mathcal{E} 3 \mathcal{I} = 0 .) \\

So equivalently we can evaluate \( \mathcal{E} \mathcal{I} \mathcal{J} \) for any trial state \( \mathcal{I} \mathcal{J} \) which in practice depends on various parameters. The parameters can be varied and \( \mathcal{E} \mathcal{I} \mathcal{J} \) can be made stationary. As with \( \mathcal{E} \mathcal{I} \mathcal{J} \mathcal{K} \).

The values of \( \mathcal{E} \mathcal{I} \mathcal{J} = \mathcal{E} \) is then an approximation to the eigenvalues of \( \mathcal{H} \). For the lowest \( \mathcal{E} \mathcal{I} \mathcal{J} \) this gives an approximation to the ground state \( \mathcal{E} \) with.

\[ \mathcal{E} \mathcal{I} \mathcal{J} \geq \mathcal{E} \mathcal{0} . \]

The space of trial states over which we vary \( \mathcal{E} \mathcal{I} \mathcal{J} \) is smaller than all if so energy \( \mathcal{E} \) is only an approximation to the true energy \( \mathcal{E} \mathcal{0} . \)
Suppose we consider excited states $\psi_n > E_n > E_0$.

Then if we know the exact eigenstates $\psi_n >$ below energy $E_n$, we can form the trial states

$$|\phi_n> = |n> - \sum_{i=0}^{n-1} \langle \psi_i|n><\psi_i|\psi_n>$$

$$\mathcal{E}[\phi_n] = \frac{\langle \phi_n|H|\phi_n>}{\langle \phi_n|\phi_n>}$$

$$= \sum_{n=1}^{\infty} E_n |c_n|^2 - \frac{\sum_{n=1}^{\infty} 1|c_n|^2}{\sum_{n=1}^{\infty} 1|c_n|^2}$$

$$\mathcal{E}[\phi_n] \geq E_n,$$  \text{The equality holding iff } \phi_n = c \psi_n.

Thus each excited state is a stationary point and with the above trial functions we can approximate the excited energy eigenvalues by the variational equation

$$\mathcal{D}[\mathcal{E}[\phi_n] = 0, \text{Then } \mathcal{E}[\phi_n] \approx E_n.$$
Of course we must know exactly the lower-lying eigenstates in order to estimate $E_n$. The approximate eigenstates are all are really illusory and so each of those errors will compound when estimating $E_n$. Hence only low-lying excited states are generally determined with any degree of accuracy with this method.

Finally we can estimate the error in the energy eigenvalue when approximated by $|\phi_0\rangle$ which differs from the true eigenstate $|\phi\rangle$ by a small amount

$$|\Delta\rangle = |\phi_0\rangle + \alpha |\phi\rangle \quad , \quad \alpha \ll 1.$$  

We choose $|\phi_0\rangle$ so that $\Delta E[|\phi_0\rangle] = 0$ even through it is not the exact eigenstate

$$\Delta E = E_0 \rightarrow 0 = \Delta E[|\phi\rangle + O(1\alpha^2)]$$

Thus $E_0 = \Delta E[|\phi\rangle + O(1\alpha^2)]$. The error in the energy is second order in smallness even though the error in the energy eigenstate is first order in smallness.
Example: Let's return to the example of Helium-like atoms and approximate the ground state energy by the variational method. As a trial wave function we use the zeroth-order perturbation theory wavefunction

$$
\psi_{0}(\vec{r}_1, \vec{r}_2) \equiv \langle \vec{r}_1, \vec{r}_2 | \psi_0 \rangle = \frac{1}{\pi a^3} e^{-\frac{\vec{r}_1^2 + \vec{r}_2^2}{a}}
$$

But now let "\( a \)" be a variational parameter rather than \( a = a_0 \).

So since

$$
\langle \psi_0 | \psi_0 \rangle = \int d^3r_1 d^3r_2 \left| \langle \vec{r}_1, \vec{r}_2 | \psi_0 \rangle \right|^2 = 1
$$

we have

$$
E [\psi_0] = \langle \psi_0 | H | \psi_0 \rangle
= \int d^3r_1 d^3r_2 \psi_0^*(\vec{r}_1, \vec{r}_2) H \psi_0 (\vec{r}_1, \vec{r}_2)
$$

where Recall

$$
H(\vec{r}_1, \vec{r}_2) = \frac{\hbar^2}{2m} (\nabla_1^2 + \nabla_2^2) - \frac{Ze^2}{r_1} - \frac{Ze^2}{r_2}
$$

$$
H' = \frac{e^2}{|\vec{r}_1 - \vec{r}_2|} + H'
$$
Recall the \( \psi_n = 2^{l} r_1^{l_1} (r_2) 2^{l_2} r_1^{l_2} (r_2) \)
with \( r_1, r_2 \) the ground state eigenfunction for an electron in the Coulomb field of charge \( Z'e \) with \( Z' = \frac{a_0}{a} \).

Hence the total energy of such an electron is
\[
- Z'^2 E_H = - Z'^2 \frac{M e^4}{2 \hbar^2} = - Z'^2 \left( 13.6 \text{eV} \right)
\]
which, recalling from our study of H-atom,
\[
\langle 2^{l_1}_1^{l_1} | \frac{\hat{P}^2}{2m} | 2^{l_1} \rangle = + Z'^2 E_H
\]
\[
\langle 2^{l_1} | \frac{Z'e^2}{r} | 2^{l_1} \rangle = - 2Z'^2 E_H
\]
Now the H-atom potential has \( - \frac{Z e^2}{r} \), so
\[
\langle 2^{l_1} | - \frac{Z e^2}{r} | 2^{l_1} \rangle = \frac{Z}{Z'} \left( - 2Z'^2 E_H \right)
\]
\[
= - 2Z Z' E_H.
\]

Finally, we recall the perturbation calculation of \( \langle \psi_n | H' | \psi_n \rangle \) page 862,
\[
\langle \psi_n | H' | \psi_n \rangle = \frac{5}{4} Z' E_H
\]
Putting this altogether we get
\[ E \left[ \psi_n \right] = \langle \psi_n | H | \psi_n \rangle \]
\[ = 2E_H \left[ Z'^2 - 2Z'(Z - \frac{5}{16}) \right] \]
recalling that \( Z' = \frac{a_0}{a} \). Hence we vary \( E \left[ \psi_n \right] \) as a function of \( a \) or equivalently \( Z' \) to minimize it.
\[ \frac{\Delta E \left[ \psi_n \right]}{\Delta Z'} = 0 = \left[ 2Z' - 2(Z - \frac{5}{16}) \right] 2E_H \]
\[ \Rightarrow Z' = Z - \frac{5}{16} \]

At this minimum value we find
\[ E \left[ \psi_n \right] = E_{\text{var}} = -2(Z - \frac{5}{16})^2 E_H \]
\[ Z' = Z - \frac{5}{16} \]

This is the variational method's approximation of the He-atom ground state energy
\[ E_{\text{var}} = -2(Z - \frac{5}{16})^2 E_H. \]
Recalling our table on page 862:

<table>
<thead>
<tr>
<th>Z</th>
<th>$E_{RS}$ (eV)</th>
<th>$E_{exp}$ (eV)</th>
<th>$E_{var}$ (eV)</th>
</tr>
</thead>
<tbody>
<tr>
<td>He</td>
<td>2</td>
<td>-74</td>
<td>-78.6</td>
</tr>
<tr>
<td>Li+</td>
<td>3</td>
<td>-193</td>
<td>-197.1</td>
</tr>
<tr>
<td>Be++</td>
<td>4</td>
<td>-365.5</td>
<td>-370.0</td>
</tr>
</tbody>
</table>

Then $E_{var}$ is closer to $E_{exp}$ than $E_{RS}$ in fact,

$$E_{var} = E_{RS} - \frac{25}{128}E_H,$$

$$= 2.6 \text{ eV}$$

Of course, $E_{var} \geq E_{exp}$ as the variational theorem requires.

The variational trial state $|\psi_n\rangle$ is interpreted as the state for 2 independent electrons moving in a Coulomb field with charge $Z'e$, $eZ' = (Z-\frac{Z}{2})e$ is smaller than the nuclear charge $Ze$. The difference $Z'e$ represents the screening of the nuclear charge felt by each $e^-$ due to the presence of the other electron.