Variational Method
(See CTDL 1148-1155, [Variational Method]
252-263, 295-307[Density Matrices])

Last time:
Quasi-Degeneracy $\rightarrow$ Diagonalize a part of infinite $H$

* sub-matrix : $H^{(0)} + H^{(1)}$
* corrections for effects of out-of-block elements: $H^{(2)}$
  (the Van Vleck transformation)
* diagonalize $H_{\text{eff}} = H^{(0)} + H^{(1)} + H^{(2)}$

Coupled H-O's 2 : 1 ($\omega_1 \approx 2\omega_2$) Fermi resonance example: polyads

1. Perturbation Theory vs. Variational Method
2. Variational Theorem
3. Stupid nonlinear variation
4. Linear Variation $\rightarrow$ new kind of secular Equation
5. Linear combined with nonlinear variation
6. Strategies for criteria of goodness — various kinds of variational calculations

1. Perturbation Theory vs. Variational Method

Perturbation Theory in effect uses $\infty$ basis set
goals: parametrically parsimonious fit model, $H_{\text{eff}}$
fit parameters (molecular constants) $\leftrightarrow$ parameters that define $V(x)$
order - sorting $\frac{H_{nk}^{(1)}}{E_n^{(0)} - E_k^{(0)}} < 1$ — errors less than this “mixing angle” times the previous order non-zero correction term
(n is in-block, k is out-of block) because diagonalization is $\infty$ order
(within block).

Variational Method

best possible estimate for lowest few $E_n$, $\psi_n$ (and properties derivable from these) using finite basis set and exact form of $H$. 
Vast majority of computer time in Chemistry is spent in variational calculations. Goal is numbers. Insight is secondary.

"Ab Initio" vs. "semi-empirical" or "fitting"

[intentionally bad basis set: Hückel, tight binding – qualitative behavior obtained by a fit to a few microscopic–like control parameters]

2. Variational Theorem

If \( \phi \) is approximation to eigenfunction of \( \hat{A} \) belonging to lowest eigenvalue \( a_0 \), then

\[
\alpha \equiv \frac{\langle \phi | \hat{A} | \phi \rangle}{\langle \phi | \phi \rangle} \geq a_0
\]

the variational Theorem

PROOF: eigenbasis (which we do not know – but know it must exist)

\[
\hat{A} | n \rangle = a_n | n \rangle
\]

expand \( \phi \) in eigenbasis of \( \hat{A} \), exploiting completeness

\[
| \phi \rangle = \sum_n | n \rangle \langle n | \phi \rangle
\]

completeness

\[
\langle \phi | \hat{A} | \phi \rangle = \sum_{n,n'} \langle \phi | n \rangle \langle n | \hat{A} | n' \rangle \langle n' | \phi \rangle = \sum_n \langle \phi | n \rangle^2 a_n
\]

eigenbasis

\[
\langle \phi | \phi \rangle = \sum_n \langle \phi | n \rangle \langle n | \phi \rangle = \sum_n \langle \phi | n \rangle^2
\]

subtract \( a_0 \) from both sides

\[
\alpha - a_0 = \frac{\sum_n (a_n - a_0) \langle n | \phi \rangle^2}{\sum_{n'} \langle n' | \phi \rangle^2} \geq 0
\]

again, all terms in both sums are \( \geq 0 \)
because, by definition of \( a_0 \), \( a_n \geq a_0 \) for all \( n \) and all terms in sum are \( \therefore \geq 0 \).

\[ \therefore \alpha \geq a_0. \quad \text{QED} \]

It is possible to perform a variational calculation for any \( A \), not limited to \( H \).

3. Stupid Nonlinear Variation

Use the wrong functional form or the wrong variational criterion to get poor results — illustrates that the variational function must have sufficient flexibility and the variational criterion must be as it is specified in the variational theorem, as opposed to a clever shortcut.

The H atom Schröd. Eq. \( (\ell = 0) \)

\[
H = -\frac{1}{2} \frac{1}{r^2} \frac{\partial}{\partial r} r^2 \frac{\partial}{\partial r} + \frac{1}{r}
\]

and we know

\[
\psi_{1s}(r) = \langle r | 1s \rangle = \pi^{-1/2} e^{-r}
\]

\[
E_{1s} = -1/2 \text{ au}
\]

\[ 1 \text{ au} = 219475 \text{ cm}^{-1} \]

but try \( \langle r | \phi \rangle = \left[ \frac{\xi^3}{2\pi} \right]^{1/2} \left( \xi r e^{-\xi r} \right) \)

normalized for all \( \xi \)

\[ \xi \] is a scale factor that controls overall size of \( \phi(r) \)

[actually this is the form of \( \psi_{2p}(r) \) which is necessarily orthogonal to \( \psi_{1s} \)! STUPID!]

\[
\left( \phi(0) = 0 \quad \text{but} \quad \psi_{1s}(0) = \pi^{-1/2} \right)
\]

\[
\varepsilon = \frac{\langle \phi | H | \phi \rangle}{\langle \phi | \phi \rangle} = \frac{4}{3} \left( \frac{\xi^2 - 3\xi}{8} \right)
\]

skipped a lot of algebra

minimize \( \varepsilon \):

\[
\frac{d\varepsilon}{d\xi} = 0 \quad \xi_{\min} = 3/2 \quad \varepsilon_{\min} = -3/8 \text{ au}
\]

FAILURE!

\[
\text{c.f. the true values: } E_{1s} = -1/2 \text{ au, } E_{2s} = -\frac{1}{8} \text{ au}
\]
Try something clever (but lazy):
What is the value of $\xi$ that maximizes $\langle \phi|1s\rangle$?

for the best variational $\xi = 3/2$, $C_{1s} = \langle \phi|\xi = 3/2|1s\rangle = 0.9775$
if we maximize $C_{1s}$ wrt. $\xi : \xi = 5/3 \rightarrow C_{1s} = 0.9826$ better?

but $\varepsilon = -0.370$ results, a poorer bound than $\xi = 3/2 \rightarrow \varepsilon = -0.375$

* need flexibility in $\phi$
* can improve on $\frac{d\varepsilon}{d\xi}$ by employing an alternative variational strategy

This was stupid anyway because we would never use the variational method when we already know the answer!

4. Linear Variation $\rightarrow$ Secular Equation

$$\phi = \sum_{n=1}^{N} c_n \chi_n$$

$$\langle \chi_n | H | \chi_{n'} \rangle = H_{nn'}$$
$$\langle \chi_n | \chi_{n'} \rangle = S_{nn'}$$

overlap integrals
(non-orthogonal basis sets are often convenient)

$$\varepsilon = \frac{\langle \phi | H | \phi \rangle}{\langle \phi | \phi \rangle} = \sum_{n,n'} c_n c_{n'} H_{nn'}$$
$$\sum_{m,m'} c_m c_{m'} S_{mm'}$$

rearrange this equation

to find minimum value of $\varepsilon$,
take $\frac{\partial}{\partial c_j}$ for each $j$ and require that

$$\frac{\partial \varepsilon}{\partial c_j} = 0 \text{ for each } j$$

linear variation!

because we are seeking to minimize $\varepsilon$ with respect to each $c_j$.

Find the global minimum of the $\varepsilon(c_1, c_2, \ldots c_N)$ hypersurface.

the only terms that survive $\frac{\partial}{\partial c_j}$ are those that include $c_j$. 

modified 10/9/02 10:21 AM
5.73 Lecture #18

\[ \varepsilon \sum_m c_m (S_{mj} + S_{jm}) = \sum_n c_n (H_{jn} + H_{nj}) \]

if \( \{ \chi_n \} \) are real \( S_{ij} = S_{ji}, \ H_{ij} = H_{ji} \)

\[ 0 = \sum_{n=1}^{N} c_n (H_{jn} - \varepsilon S_{jn}) \]

one such equation for each \( j \) (same set of unknown \( \{ c_n \} \))

N linear homogeneous equations in N unknown \( c_n \)'s
Non trivial \( \{ c_n \} \) only if \( |H - \varepsilon S| = 0 \)
(Not same form as \( |H - E| = 0 \))

The result is N special values of \( \varepsilon \) that satisfy this equation.

CTDL show: all N \( \varepsilon_n \) values are upper bounds to the lowest N \( E_n \)'s
and all \( \{ \phi_n \} \)'s are orthogonal!
(provided that they belong to different values of \( E_n \))

How to solve \( |H - \varepsilon S| = 0 \)

1. Diagonalize \( S \)
   \[ U^\dagger S U = \tilde{S} \quad \tilde{S}_{ij} = s_i \delta_{ij} \]
   (orthogonalize \( \{ \chi \} \) basis)

2. Normalize \( \tilde{S} \)
   \[
   \begin{pmatrix}
   \tilde{S}^{-1/2} \tilde{S} \tilde{S}^{-1/2}
   \end{pmatrix}^{\text{3 diagonal matrices}} = \mathbf{1} = \tilde{S} = T^\dagger S T
   \]
   where \( T = US^{-1/2} \)
   \[
   \begin{pmatrix}
   S^{-1/2} \\
   0
   \end{pmatrix} \leq \begin{pmatrix}
   \mathbf{1} \leq S^{-1/2} = \begin{pmatrix}
   s_1^{-1/2} & 0 & 0 \\
   0 & s_2^{-1/2} & 0 \\
   0 & 0 & \ddots
   \end{pmatrix}
   \end{pmatrix}
   \]

This is not an orthogonal transformation, but it does not destroy orthogonality because each function is only being multiplied by a constant.
3. Transform H to orthonormalized basis set

\[ H = S^{1/2} \left( U^T H U \right) S^{-1/2} \]

\[ U \] diagonalizes \( S \) not \( H \)

new secular equation

\[ \| H - \varepsilon S \| = 0 \quad \text{but} \quad S = 1 \]

\[ \| H - \varepsilon \mathbf{1} \| = 0 \quad \text{usual} \quad H \quad \text{diagonalized by usual procedure!} \]

5. Combine Linear and Nonlinear Variation

typically done in \textit{ab initio} electronic structure calculations

Basis set: \[ \chi_n(\xi_n r) \] linear variation where \( \varepsilon_n \) is a radial scale factor

\[ \psi = \sum_n \varepsilon_n \chi_n(\xi_n r) \] nonlinear variation

\[ S_{nn'}(\xi_n, \xi_n'), H_{nn'}(\xi_n, \xi_n') \]

0. pick arbitrary set of \( \{ \xi_i \} \)
1. calculate all \( H_{ij}(\xi_i, \xi_j) \) & \( S_{ij}(\xi_i, \xi_j) \)
2. Solve \( | H - \varepsilon S \| = 0 \)

\[ a. \quad S \rightarrow \tilde{S} \quad \text{diagonalize } S \quad \text{(orthogonalize)} \]

\[ b. \quad \left( \tilde{S} \right)^{-1/2} \quad \text{(normalize)} \]

\[ c. \quad H \rightarrow \tilde{H} \]

\[ d. \quad \text{diagonalize } \tilde{H} \]

nonlinear variation begins – find global minimum of \( \varepsilon_{\text{lowest}} \)
with respect to each \( \xi_i \)
3. change $\xi_1$ from $\xi_1^{(0)} \to \xi_1^{(1)} = \xi_1^{(0)} + \delta$

4. recalculate all integrals in $H$ and $S$ involving $\chi_1$

5. Solve $|H - \epsilon S| = 0$ to obtain a new set of $\{\epsilon_i\}$.
   Pick lowest $\epsilon_i$.

6. calculate $\frac{\partial \epsilon_{\text{lowest}}}{\partial \xi_1} = \frac{\epsilon_{\text{old}}^{\text{lowest}} - \epsilon_{\text{new}}^{\text{lowest}}}{\xi_1^{(0)} - \xi_1^{(1)}}$

7. repeat #3 – 6 for each $\xi_i$ (always looking only at lowest $\epsilon_i$)
   This defines a gradient on a multidimensional $\epsilon(\xi_1, \ldots, \xi_N)$ surface. We seek the minimum of this hypersurface. Take a step in direction of steepest descent by an amount determined by $|\partial \epsilon / \partial \xi_{\text{steepest}}|$ (small slope, small step; large slope, large step).

   This completes 1st iteration. All values of $\{\xi_i\}$ are improved.

8. Return to #3, iterate #3-7 until convergence is obtained.

   Nonlinear variations are much slower than linear variations.
   Typically use ENORMOUS LINEAR $\{\chi\}$ basis set.

   Contract this basis set by optimizing nonlinear parameters (exponential scale factors) in a SMALL BASIS SET to match the lowest $\{\phi\}$’s that had initially been expressed in large basis set.
6. Alternative Strategies

* rigorous variational minimization of $E_{\text{lowest}}$: “ab initio”
* constrain variational function to be orthogonal to specific subset of functions
  
  e.g. orthogonal to ground state – to get variational convergence.
  
  Applies only to higher members of specific symmetry class
  
  or orthogonal to core: frozen-core approximation.
  
  “Pseudopotentials” (use some observed energy levels to
  
  determine $Z_{\text{eff}}(r)$ of frozen core)

* least squares fitting

  minimize differences between a set of measured energy levels (or other
  properties) and a set of computed variational eigen-energies (or other
  properties computed from variational wavefunctions).

  $$\{\text{observed } E_n\} \leftrightarrow \{\text{parameters in } H_{\text{eff}}\}$$

  molecular constants

  $\downarrow$

  experimental $\psi$ ‘s in finite
  
  variational basis set

* semi-empirical model

  replace exact $\hat{H}$ by a grossly simplified form and restrict basis set to a simple
  form too.
  
  Then adjust parameters in $H$ to match some observed pattern of energy
  splittings. Use parameters to predict unobserved properties or use values of
  fit parameters to build insight.