Variational Method

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

Last time:

Quasi-Degeneracy → Diagonalize a part of infinite **H**

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

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 → 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 ∞ basis set

goals: parametrically parsimonious fit model, $\mathbf{H}^{eff}$

fit parameters (molecular constants) ↔ parameters that define V(x)

$$\text{order - sorting } \frac{H^{(1)}_{nk}}{E^{(0)}_n - E^{(0)}_k} < 1$$

— 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 ∞ order (within block).

Variational Method

best possible estimate for lowest few E_n, ψ_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–<u>like</u>
control parameters]

2. Variational Theorem

not necessarily normalized

any observable

If ϕ is approximation to eigenfunction of $\hat{\mathbf{A}}$
belonging to lowest eigenvalue a_0, then

$$\alpha \equiv \frac{\langle \phi|\mathbf{A}|\phi\rangle}{\langle \phi|\phi\rangle} \geq a_0$$

the variational Theorem

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

$$\mathbf{A}|n\rangle = a_n|n\rangle$$

expand ϕ in eigenbasis of **A**, exploiting completeness

$$|\phi\rangle = \sum_n |n\rangle\langle n|\phi\rangle$$

completeness

$$\langle \phi|\mathbf{A}|\phi\rangle = \sum_{n,n'} \langle \phi|n\rangle \underbrace{\langle n|\mathbf{A}|n'\rangle}_{a_n \delta_{nn'}} \langle n'|\phi\rangle = \sum_n |\langle \phi|n\rangle|^2 a_n$$

eigenbasis

$$\langle \phi|\phi\rangle = \sum_n \langle \phi|n\rangle\langle n|\phi\rangle = \sum_n |\langle \phi|n\rangle|^2$$

$$\alpha \equiv \frac{\langle \phi|\mathbf{A}|\phi\rangle}{\langle \phi|\phi\rangle} = \frac{\sum_n a_n |\langle n|\phi\rangle|^2}{\sum_{n'} |\langle n'|\phi\rangle|^2}$$

all terms in both sums are ≥ 0

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 ≥ 0

because, by definition of a_0, $a_n \geq a_0$ for all n and all terms in sum are $\therefore \geq 0$.

$$\therefore \alpha \geq a_0. \qquad \text{QED} \qquad \begin{pmatrix}\text{but useless because we}\\ \text{can' t know } a_n \text{ or } \langle n|\phi\rangle\end{pmatrix}$$

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 Schr. Eq. ($\ell = 0$)

$$\mathbf{H} = \underbrace{-\frac{1}{2}\frac{1}{r^2}\frac{\partial}{\partial r}r^2\frac{\partial}{\partial r}}_{T} \quad \underbrace{-\frac{1}{r}}_{V}$$

and we know $\begin{cases}\psi_{1s}(r) = \langle r|1s\rangle = \pi^{-1/2}e^{-r}\\ E_{1s} = -1/2\ au\end{cases}$ $\qquad \left[1\ au = 219475\ cm^{-1}\right]$

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

ξ is a scale factor that controls overall size of ϕ(r)

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

$$\left(\phi(0) = 0 \qquad \textit{but} \qquad \psi_{1s}(0) = \pi^{-1/2}\right)$$

$$\varepsilon = \frac{\langle\phi|\mathbf{H}|\phi\rangle}{\langle\phi|\phi\rangle} = \frac{4}{3}\left(\frac{\xi^2 - 3\xi}{8}\right)$$ skipped a lot of algebra

minimize ε: $\quad \frac{d\varepsilon}{d\xi} = 0 \quad \xi_{min} = 3/2 \rightarrow \varepsilon_{min} = -3/8\ \ au$

FAILURE! $\qquad \left[\text{c.f. the true values: } E_{1s} = -1/2\ \text{au},\ E_{2s} = -\frac{1}{8}\text{au}\right]$

Try somethng clever (but lazy):
What is the value of ξ that maximizes $\langle\phi|$ 1s$\rangle$?

for the best variational $\xi = 3/2$, $C_{1s} \equiv \langle\phi(\xi = 3/2)|1s\rangle = 0.9775$

if we maximize C_{1s} *wrt.* ξ : $\xi = 5/3 \rightarrow C_{1s} = 0.9826$ better?

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

* need flexibility in ϕ

* can't 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 → Secular Equation

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

KEY TOPIC for this lecture

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

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

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

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

rearrange this equation

$$\varepsilon \sum_{m,m'} c_m c_{m'} S_{mm'} = \sum_{n,n'} c_n c_{n'} H_{nn'}$$

to find minimum value of ε,
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 ε 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.

$$\varepsilon\sum_m c_m\left(S_{mj}+S_{jm}\right)=\sum_n c_n\left(H_{jn}+H_{nj}\right)$$

These are all of the surviving terms (i.e. those that include j). Each j term appears twice in both sums, once as a bra and once as a ket.

if $\{\chi_n\}$ are real $\quad S_{ij}=S_{ji},\ H_{ij}=H_{ji}$

$$0=\sum_{n=1}^{N} c_n\left(H_{jn}-\varepsilon S_{jn}\right)$$

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 $|\mathbf{H}-\varepsilon\mathbf{S}|=0$
(Not same form as $|\mathbf{H}-\mathbb{1}E|=0$)

The result is N special values of ε that satisfy this equation.

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

How to solve $|\mathbf{H}-\varepsilon\mathbf{S}|=0$

1. Diagonalize **S**

   $\mathbf{U}^\dagger\mathbf{S}\mathbf{U}=\tilde{\mathbf{S}} \qquad \tilde{S}_{ij}=s_i\delta_{ij}$

   (orthogonalize $\{\chi\}$ basis)

2. Normalize $\tilde{\mathbf{S}}$

   $$\underbrace{\left(\tilde{\mathbf{S}}\right)^{-1/2}\tilde{\mathbf{S}}\left(\tilde{\mathbf{S}}\right)^{-1/2}}_{\text{3 diagonal matrices}}=\mathbf{1}\equiv\tilde{\tilde{\mathbf{S}}}=\mathbf{T}^\dagger\mathbf{S}\mathbf{T} \qquad \text{where } \mathbf{T}=\mathbf{U}\mathbf{S}^{-1/2}$$

   $$\left(\tilde{\mathbf{S}}^{-1/2}\right)^{\leq}=\tilde{\mathbf{S}}^{-1/2}=\begin{pmatrix} s_1^{-1/2} & 0 & 0\\ 0 & s_2^{-1/2} & 0\\ 0 & 0 & \ddots\end{pmatrix}$$

   unitary

   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

$$\tilde{\tilde{\mathbf{H}}} = \underbrace{\mathbf{S}^{-1/2}\left(\mathbf{U}^{\dagger}\right)}_{\mathbf{T}^{\dagger}}\mathbf{H}\underbrace{\left(\mathbf{U}\right)\mathbf{S}^{-1/2}}_{\mathbf{T}}$$

U diagonalizes **S** not **H**

new secular equation

$$\left|\tilde{\tilde{\mathbf{H}}} - \varepsilon\tilde{\tilde{\mathbf{S}}}\right| = 0 \quad \text{but} \quad \tilde{\tilde{\mathbf{S}}} = \mathbf{1}$$

$$\left|\tilde{\tilde{\mathbf{H}}} - \varepsilon\mathbf{1}\right| = 0 \quad \text{usual} \quad \tilde{\tilde{\mathbf{H}}}$$

diagonalized by usual procedure!

---

5. Combine Linear and Nonlinear Variation

typically done in *ab initio* electronic structure calculations

Basis set: $\chi_n(\xi_n r)$

$$\psi = \sum_n c_n \chi_n(\xi_n r)$$

linear variation where ε_n is a radial scale factor

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 $|\mathbf{H} - \varepsilon\mathbf{S}| = 0$

   a. $\mathbf{S} \rightarrow \tilde{\mathbf{S}}$ diagonalize **S** (orthogonalize)

   b. $\left(\tilde{\mathbf{S}}\right)^{-1/2}$ (normalize)

   c. $\mathbf{H} \rightarrow \tilde{\tilde{\mathbf{H}}}$

   d. diagonalize $\tilde{\tilde{\mathbf{H}}}$

nonlinear variation begins – find global minimum of $\varepsilon_{\text{lowest}}$ with respect to each ξ_i

3. change ξ_1 from $\xi_1^{(0)} \rightarrow \xi_1^{(1)} = \xi_1^{(0)} + \delta$
4. recalculate all integrals in **H** and **S** involving χ_1
5. Solve $|\mathbf{H} - \varepsilon\mathbf{S}| = 0$ to obtain a new set of $\{\varepsilon_i\}$.
   Pick lowest ε_i.
6. calculate $\dfrac{\partial\varepsilon_{\text{lowest}}}{\partial\xi_1} = \dfrac{\varepsilon_{\text{lowest}}^{\text{old}} - \varepsilon_{\text{lowest}}^{\text{new}}}{\xi_1^{(0)} - \xi_1^{(1)}}$

7. repeat #3 – 6 for each ξ_i (always looking only at lowest ε_i)
   This defines a gradient on a multidimensional $\varepsilon(\xi_1,\ldots\xi_N)$ surface. We seek the minimum of this hypersurface. Take a step in direction of steepest descent by an amount determined by $|\partial\varepsilon/\partial\xi_{\text{steepest}}|$ (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_{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 $Z^{eff}(r)$ 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).

$$\{\text{observed } E_n\} \leftrightarrow \{\text{parameters in } \mathbf{H}^{eff}\}$$

molecular constants

$\Downarrow$

experimental ψ 's in finite variational basis set

* semi-empirical model

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