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 example: the 2 : 1 ($\omega_1 \approx 2\omega_2$) Fermi resonance polyads

1. Perturbation Theory vs. Variational Method: non-orthognal → **S** (overlap matrix)
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 effectively uses ∞ basis set
non-degenerate: diagonalize $\mathbf{H}^{eff}$
quasi-degenerate: non-diagonal $\mathbf{H}^{eff}$ (model with quantum number scaling)
goals: parametrically parsimonious fit model, $\mathbf{H}^{eff}$
fit parameters (molecular constants) ↔ parameters that define V(x)

order-sorting $\dfrac{H_{nk}^{(1)}}{E_n^{(0)} - E_k^{(0)}} < 1$ — errors smaller than this "mixing angle" times the previous order non–zero correction term

(n is in-block, k is out-of block) because diagonalization is to ∞ order (within block).

Variational Method

best possible estimate for lowest few E_n, ψ_n (and properties derivable from these) using a finite (often huge) basis set and the 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 $\hat{\mathbf{A}}$
that belongs to the 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 that it must exist)

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

expand $|\phi\rangle$ 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 for **A**

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

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

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 $\sum_n |\langle n | \phi \rangle|^2 = \sum_{n'} |\langle n' | \phi \rangle|^2$

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

$\therefore \quad \alpha \geq a_0$. QED (but useless because we do not know a_n or $\langle n|\phi\rangle$ in advance)

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 Schrödinger Equation ($\ell = 0$)

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

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

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 at $\xi = 1$ is necessarily orthogonal to ψ_{1s}!
STUPID!

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

$$\varepsilon = \frac{\langle\phi|\mathbf{H}|\phi\rangle}{\langle\phi|\phi\rangle} = \frac{4}{3}\left(\frac{\xi^2 - 3\xi}{8}\right)$$ 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! $\left[c.f.\text{ the true values: } E_{1s} = -1/2\; au, E_{2s} = -\frac{1}{8}au\right]$

[insufficiently flexible variational function.

exact result (usually not known)

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

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

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

No. The value of ε = –0.370 results from maximizing C_{1s}, which is a poorer bound to ε than obtained by minimizing ε.
$\xi = 3/2 \rightarrow \ \varepsilon \ = -0.375$

* need flexibility in ϕ

* can't improve on $\frac{d\varepsilon}{d\xi} = 0$ 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

KEY TOPIC for this lecture

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

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

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

distributed Gaussians

overlap integrals
(non-orthogonal basis sets are often convenient)
also not necessarily normalized

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

Rearrange this equation.

To find the minimum value of ε,

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

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!

We are seeking to minimize ε with respect to each c_j. Find the global minimum of the e(c_1,c_2,…,c_N) hypersurface.

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

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

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

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

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.

We get 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 the same simple form as $|\mathbf{H}-\mathbf{1}\mathrm{E}|=0$, but we can deal with this extra computational complexity.)

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

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

How to solve $|\mathbf{H}-\varepsilon\mathbf{S}|=0$:
series of transformations on both **S** and **H**

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 $\tilde{\mathbf{S}}$

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

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

unitary

This is not an orthogonal transformation of $\tilde{\mathbf{S}}$, but it does not destroy orthogonality because each eigenfunction of $\tilde{\mathbf{S}}$ is only being multiplied by a constant.

3. Transform **H** to orthonormalized basis set

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

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

Obtain a new secular equation:

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

thus $\left|\tilde{\tilde{\mathbf{H}}} - \varepsilon\mathbf{1}\right| = 0$ by which $\tilde{\tilde{\mathbf{H}}}$ is diagonalized by the usual procedure

---

5. <u>Combine Linear and Nonlinear Variation</u>

typically done in *ab initio* electronic structure calculations

Basis set: $\chi_n(\xi_n r)$ ↓ linear variation in $\{\chi_n\}$, but where ξ_n is a radial scale factor, one for each χ_n

$\psi = \sum_n c_n \chi_n(\xi_n r)$ ↑ nonlinear variation

get $S_{nn'}(\xi_n, \xi_{n'}), H_{nn'}(\xi_n, \xi_{n'})$

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

and now the nonlinear variation begins — find global minimum of $\varepsilon_{\text{lowest}}$ with respect to 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* $|\mathbf{H}\text{-}\varepsilon\mathbf{S}| = 0$ to obtain a new set of $\{\varepsilon_i\}$.
   Pick lowest ε_i.
6. *calculate* $\dfrac{\partial \varepsilon_{lowest}}{\partial \xi_1} = \dfrac{\varepsilon_{lowest}^{old} - \varepsilon_{lowest}^{new}}{\xi_1^{(0)} - \xi_1^{(1)}}$

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

* Large linear variation to get the linear combination of ϕ's that have the lowest energies;

* define a set of functions $\{\psi_i(\xi_i x)\}$ that contain non-linear scale parameters;

* perform a least squares fit of the $\{\xi\}$ to match the lowest few energy states from the linear variation;

* ortho-normalize the small set of $\psi_i(\xi_i x)$ functions and use them in a linear variation, thereby replacing the many-component functions from the massive linear variation by few-component functions (contracted basis set) from the hybrid linear variation.

6. Alternative Strategies

* rigorous variational minimization of E_{lowest}: "*ab initio*"
* constrain variational functions to be orthogonal to specific subset of variationally optimized functions

e.g. orthogonal to ground state – to get variational convergence. Applied to higher energy 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 to truncated **H** i.e. $\mathbf{H}^{eff}$

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
⇓
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 **H** to match some observed pattern of energy splittings. Confirm by using parameters to predict unobserved properties. Use values of fit parameters to build insight.

Never-ending battle between accuracy and insight!

5.73 Quantum Mechanics I
Fall 2018