Lecture 18
Post-Hartree-Fock Methods

Configuration interaction; coupled cluster methods; perturbation theory

Prof. David A. Kofke
CE 500 – Modeling Potential-Energy Surfaces
Department of Chemical & Biological Engineering
University at Buffalo

University at Buffalo
The State University of New York

Electron correlation is essential for accurate electronic structure, yet is neglected by HF

- Summary of HF formulas

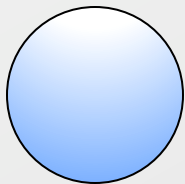$$\hat{F}\varphi_i = \varphi_i \varepsilon_i \quad \hat{F} = \hat{h} + \left(\hat{C} - \hat{X}\right)$$

$$\hat{C}\varphi_i(\tau_1) = \sum_j \int \frac{1}{r_{12}} \varphi_j^*(\tau_2)\varphi_j(\tau_2)\varphi_i(\tau_1)d\tau_2$$

$$\hat{X}\varphi_i(\tau_1) = \sum_j \int \frac{1}{r_{12}} \varphi_j(\tau_1)\varphi_j^*(\tau_2)\varphi_i(\tau_2)d\tau_2$$

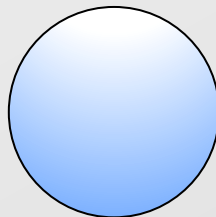- Two applications using infinite basis-set limit

Hydrogen atom
1 electron

Helium atom
2 electrons

E_{HF} = −0.500 00 au
E_{exact} = −0.500 00 au

E_{HF} = −2.861 68 au
E_{exact} = −2.903 72 au
Error = 0.04204 au
        = 26.4 kcal/mol ~ 13 000 K

Roothaan-Hall HF treatment usually yields more molecular orbitals than needed

Coefficient matrix, C

- Method for solution of HF equation by expressing molecular orbitals in terms of basis functions

$$\varphi_i(\mathbf{r}) = \sum_\mu \chi_\mu(\mathbf{r}) C_{\mu,i}$$

- Roothaan-Hall equation specifies coefficients $C_{\mu,i}$ by minimizing HF expectation energy

$$
\begin{array}{c}
\begin{matrix} \chi_1 \\ \chi_2 \\ . \\ . \\ \chi_N \end{matrix}
\begin{bmatrix}
C_{1,1} & C_{1,2} & \ldots & C_{1,N} \\
C_{2,1} & C_{2,2} & \ldots & C_{2,N} \\
\vdots & \vdots & C_{\mu,i} & \vdots \\
C_{N,1} & C_{N,2} & \ldots & C_{N,N}
\end{bmatrix} \\
\phi_1 \quad \phi_2 \quad \ldots \quad \phi_N
\end{array}
$$

number of molecular orbitals is equal to number of basis functions, which is usually more than the number of electron pairs. Lowest-energy orbitals are filled; remainder are "virtual orbitals"

$$FC = SC\varepsilon$$

Fock matrix   overlap matrix   diagonal eigenvalue matrix

A wavefunction with electron correlation can be constructed by summing SDs using virtual orbitals

- Sum Slater determinants to ensure antisymmetry

$$\Psi(1, 2, 3, \ldots, N) = \sum_{I=0}^{D} \Phi_I(1, 2, 3, \ldots, N) c_I$$

Number of determinants

Correlated wavefunction

Different Slater determinants

Coefficients to be determined

- Φ_0 is the usual HF "reference" determinant

- Other Φ_I are formed from it by swapping in virtual orbitals

- Reference is SD formed from N lowest-energy orbitals

$$\Phi_0 = |\varphi_1, \varphi_2, \ldots, \varphi_i, \ldots, \varphi_j, \ldots, \varphi_N|$$

- A single orbital substitution with virtual orbital yields new SD

  – Replace orbital i with virtual orbital a

$$\Phi_i^a = |\varphi_1, \varphi_2, \ldots, \varphi_a, \ldots, \varphi_j, \ldots, \varphi_N|$$

- Orbital a will be orthogonal to all other orbitals

  – So new SD will be orthogonal to Φ_0

    - Easy creation of a set of orthonormal antisymmetric wavefunctions

  – With V virtual orbitals, can create $N \times V$ single-substitution SDs

    - Single substitutions, or "singles"

5

$$\Phi_0 = |\varphi_1, \varphi_2, \ldots, \varphi_i, \ldots, \varphi_j, \ldots, \varphi_N|$$

- Double substitutions, or "doubles" use two virtual orbitals

$$\Phi_{ij}^{ab} = |\varphi_1, \varphi_2, \ldots, \varphi_a, \ldots, \varphi_b, \ldots, \varphi_N|$$

  – $N(N-1)/2 \times V(V-1)/2$ possible choices

  – All normalized and mutually orthogonal

- Triples, quadruples, etc. follow in a similar manner

- General form of new Slater determinant sum

$$\Psi = \Phi_0 c_0 + \sum_{i,a} \Phi_i^a c_i^a + \sum_{\substack{i<j \\ a<b}} \Phi_{ij}^{ab} c_{ij}^{ab} + \sum_{\substack{i<j<k \\ a<b<c}} \Phi_{ijk}^{abc} c_{ijk}^{abc} + \cdots + \text{ up to } N\text{-tuples}$$

The coefficients of the SD sum are determined using the variation principle

- Minimize energy expectation $\left\langle \Psi | \hat{H} | \Psi \right\rangle$

- For simplicity writing the SD sum as $\Psi = \sum_I \Phi_I c_I$

- We need to evaluate the matrix elements $\left\langle \Phi_I | \hat{H} | \Phi_J \right\rangle$

- Minimization $\rightarrow$ eigenvalue equation involving $\hat{H}$ matrix
  - Each eigenvector gives coefficients c_I for one wavefunction
  - Eigenvalues are electronic state energies

- Each SD in the basis represents an electronic configuration
  - Method is called *Configuration Interaction* (full CI (FCI) goes to *N*-tuples)

7

Hamiltonian matrix involves calculation of ERIs, but Slater-Condon rules simplify things

Diagonal element for some determinant Φ

$$\langle \Phi | \widehat{H} | \Phi \rangle = \sum_{r \, \text{in} \, \Phi} h_{rr} + \frac{1}{2} \sum_{r,s \, \text{in} \, \Phi} ([rr \mid ss] - [rs \mid sr])$$

1-electron integrals

Electron-repulsion integrals

Off-diagonal for single substitution

$$\langle \Phi | \widehat{H} | \Phi_p^q \rangle = h_{pq} + \sum_{r \, \text{in} \, \Phi} ([pq \mid rr] - [pr \mid rq])$$

Off-diagonal for double substitution

$$\langle \Phi | \widehat{H} | \Phi_{pr}^{qs} \rangle = [pq \mid rs] - [ps \mid rq]$$

- For more than two substitutions in Φ, element is zero

- Also: $\langle \Phi_0 | \widehat{H} | \Phi_p^q \rangle = 0$ Brillouin theorem

The Hamiltonian matrix has large parts that are very sparse

	Ψ_{HF}	Ψ_i^a	Ψ_{ij}^{ab}	Ψ_{ijk}^{abc}
Ψ_{HF}	E_{HF}	**0**	dense	**0**
Ψ_i^a	**0**	dense	sparse	very sparse
Ψ_{ij}^{ab}	dense	sparse	sparse	extremely sparse
Ψ_{ijk}^{abc}	**0**	very sparse	extremely sparse	extremely sparse

Cramer, video 4.02

9

CISD

Issues of *size consistency*

$$E(\text{A+B}) = E(\text{A}) + E(\text{B}) \text{ (for noninteracting A\&B)}$$

and *size extensivity*

$$\lim_{N \to \infty} E^{\text{corr}}(N \cdot A)/N = \text{const}$$

arise with TCI, related to inconsistent level of approximation for collected vs. separate molecules

MCSCF (multiconfiguration SCF) and MRCI (multireference CI) are compromise methods between simple TCI and FCI

Cramer, video 4.02

10

Method/basis	No. of basis functions	Total E/au
HF/STO-3G	2	−1.116 706
HF/cc-pVDZ	10	−1.128 711
HF/cc-pVTZ	28	−1.132 959
HF/cc-pVQZ	60	−1.133 458
HF/cc-pV5Z	110	−1.133 607
⋮	⋮	HF limit
FCI/STO-3G	2	−1.137 274
FCI/cc-pVDZ	10	−1.163 403
FCI/cc-pVTZ	28	−1.172 335
FCI/cc-pVQZ	60	−1.173 796
FCI/cc-pV5Z	110	−1.174 223
Exact[b]		−1.174 476
MP2/cc-pV5Z	110	−1.167 262
CCSD/cc-pV5Z	110	−1.174 223

STO-3G is a minimal GTO basis: one 1s function per hydrogen contracted from three GTOs, to fit a 1s STO with exponent 1.24. The other basis sets are polarized, correlation-consistent valence double-(D) to quintuple- (5)ζ GTO basis sets developed specifically for post-HF calculations. The number of contracted basis functions is listed. MP2 is discussed in Section 22.3.

J. S. Sims and S. A. Hagstrom, 'High precision variational calculations for the Born-Oppenheimer energies of the ground state of the hydrogen molecule', J. Chem. Phys. 124 (2006) 094101.

2 electrons!

Note that energy of each H atom is −0.5 au, so binding energy is tabulated value +1

Error: 25 kcal/mol ~ 13,000 K

Error: 6.9 kcal/mol ~ 3500 K

Error: 0.16 kcal/mol ~ 80 K

Error: 4.5 kcal/mol ~ 2300 K

Error: 0 kcal/mol

J. Autschbach

11

Coupled Cluster (CC) methods group SD excitations to allow focus on more highly correlated electrons

- General N-tuple SD sum

$$\Psi = \Phi_0 c_0 + \sum_{i,a} \Phi_i^a c_i^a + \sum_{\substack{i<j \\ a<b}} \Phi_{ij}^{ab} c_{ij}^{ab} + \sum_{\substack{i<j<k \\ a<b<c}} \Phi_{ijk}^{abc} c_{ijk}^{abc} + \cdots + \text{ up to } N\text{-tuples}$$

- Define operators

Cluster amplitudes, TBD

$$\hat{\tau} = \sum_m^N \tau_m \qquad \hat{\tau}_1 = \sum_{i,a} \Phi_i^a t_i^a \qquad \hat{\tau}_2 = \sum_{\substack{i<j \\ a<b}} \Phi_{ij}^{ab} t_{ij}^{ab} \qquad \text{etc.}$$

- Given a complete 1-particle basis, exact wavefunction of N-electron system generated from a reference Φ_0 via

$$e^{\hat{\tau}} \Phi_0 = \Phi_0 + \hat{\tau} \Phi_0 + \frac{1}{2!} \hat{\tau}^2 \Phi_0 + \cdots$$

$$= \Phi_0 + \hat{\tau}_1 \Phi_0 + \hat{\tau}_2 \Phi_0 + \hat{\tau}_1 \hat{\tau}_2 \Phi_0 + \frac{1}{2} \hat{\tau}_1^2 \Phi_0 + \cdots$$

Unlike TCI, truncation of CC framework can maintain size extensivity

- Full CC is just as unwieldy as FCI

- Cluster operator is truncated to make approach tractable
  - E.g., CCSD (coupled cluster singles doubles)

$$\hat{\tau} = \hat{\tau}_1 + \hat{\tau}_2$$

- Using exponential of operator makes approach size extensive
  - E.g., in CCSD, triple and quadruple excitations are still present, via $\hat{\tau}_2^2$, $\hat{\tau}_2\hat{\tau}_1^2$, $\hat{\tau}_1^4$

Solution of CC equations does not rely on the variation principle

- Schrödinger equation for CC wavefunction $\widehat{H} e^{\hat{\tau}} \Phi_0 = e^{\hat{\tau}} \Phi_0 E$

- Apply inverse of exponential operator to each side
$$e^{-\hat{\tau}} \widehat{H} e^{\hat{\tau}} \Phi_0 = \Phi_0 E$$

- Energy from: $\left\langle \Phi_0 \mid e^{-\hat{\tau}} \widehat{H} e^{\hat{\tau}} \mid \Phi_0 \right\rangle = \langle \Phi_0 \mid \Phi_0 \rangle E = E$

- Cluster amplitudes (t coefficients defining τ operators) use, for example $\left\langle \Phi_{ij}^{ab} \mid e^{-\hat{\tau}} \widehat{H} e^{\hat{\tau}} \mid \Phi_0 \right\rangle = 0$ (similar for other SDs)
  - Leads to set of nonlinear equations for amplitudes in terms of 1- and 2-electron integrals
  - Solved iteratively

Perturbation theory treats an intractable system by focusing on how it differs from a tractable one

- Very general approach that is used in every field of science and engineering

- Sometimes the difference is exactly what's wanted

- In molecular modeling, there are many examples
  - With/without electric or magnetic field
  - With/without spin coupling
  - Movement of an atom to get a force constant
  - With/without electron correlation ⬅ our application

- Series for Hamiltonian, wavefunction, and energy

$$\widehat{H}(\lambda) = \sum_n \frac{1}{n!} \frac{d^n \widehat{H}}{d\lambda^n} \lambda^n = \widehat{H}^{(0)} + \widehat{H}^{(1)}\lambda + \widehat{H}^{(2)}\lambda^2 + \widehat{H}^{(3)}\lambda^3 + \mathscr{O}(\lambda^4)$$

Often truncate H here without approximation, defining $H^{(1)}$

$$\psi(\lambda) = \sum_n \frac{1}{n!} \frac{d^n \psi}{d\lambda^n} \lambda^n = \psi^{(0)} + \psi^{(1)}\lambda + \psi^{(2)}\lambda^2 + \psi^{(3)}\lambda^3 + \mathscr{O}(\lambda^4)$$

$$E(\lambda) = \sum_n \frac{1}{n!} \frac{d^n E}{d\lambda^n} \lambda^n = E^{(0)} + E^{(1)}\lambda + E^{(2)}\lambda^2 + E^{(3)}\lambda^3 + \mathscr{O}(\lambda^4)$$

$$\widehat{H}\Phi = \Phi E$$

$$\sum_{i,j} H^{(i)}\lambda^i \Phi^{(j)}\lambda^j$$

$$= \sum_{k,l} \Phi^{(k)}\lambda^k E^{(l)}\lambda^l$$

```
Hser = Series[H[λ], {λ, 0, 3}] /. {Derivative[n_][H][_] :> n! H_n, H[_] :> H_0}
Phiser = Series[Φ[λ], {λ, 0, 3}] /. {Derivative[n_][Φ][_] :> n! Φ_n, Φ[_] :> Φ_0}
eser = Series[e[λ], {λ, 0, 3}] /. {Derivative[n_][e][_] :> n! e_n, e[_] :> e_0}
Hser Phiser - Phiser eser;
CoefficientList[%, λ] // TableForm
```

$$H_0 + H_1\lambda + H_2\lambda^2 + H_3\lambda^3 + O[\lambda]^4$$

$$\Phi_0 + \Phi_1\lambda + \Phi_2\lambda^2 + \Phi_3\lambda^3 + O[\lambda]^4$$

$$e_0 + e_1\lambda + e_2\lambda^2 + e_3\lambda^3 + O[\lambda]^4$$

```
//TableForm=
 -e_0 Φ_0 + H_0 Φ_0
 -e_1 Φ_0 + H_1 Φ_0 - e_0 Φ_1 + H_0 Φ_1
 -e_2 Φ_0 + H_2 Φ_0 - e_1 Φ_1 + H_1 Φ_1 - e_0 Φ_2 + H_0 Φ_2
 -e_3 Φ_0 + H_3 Φ_0 - e_2 Φ_1 + H_2 Φ_1 - e_1 Φ_2 + H_1 Φ_2 - e_0 Φ_3 + H_0 Φ_3
```

- Collected terms

$$\lambda^0 : H_0\Phi_0 = e_0\Phi_0$$

$$\lambda^1 : -e_1\Phi_0 - e_0\Phi_1 + H_1\Phi_0 + H_0\Phi_1 = 0$$

$$\lambda^2 : -e_2\Phi_0 - e_1\Phi_1 - e_0\Phi_2 + H_2\Phi_0 + H_1\Phi_1 + H_0\Phi_2 = 0$$

$$\lambda^3 : -e_3\Phi_0 - e_2\Phi_1 - e_1\Phi_2 - e_0\Phi_3 + H_3\Phi_0 + H_2\Phi_1 + H_1\Phi_2 + H_0\Phi_3 = 0$$

To get first-order energy correction, project λ^1 equation on Φ_0

$$H_1\Phi_0 + H_0\Phi_1 = e_1\Phi_0 + e_0\Phi_1$$

$$\langle \Phi_0 \mid H_1 \mid \Phi_0 \rangle + \langle \Phi_0 \mid H_0 \mid \Phi_1 \rangle = \langle \Phi_0 \mid e_1 \mid \Phi_0 \rangle + \langle \Phi_0 \mid e_0 \mid \Phi_1 \rangle$$

$$\langle \Phi_0 \mid H_1 \mid \Phi_0 \rangle + \langle H_0\Phi_0 \mid \Phi_1 \rangle = e_1\langle \Phi_0 \mid \Phi_0 \rangle + e_0\langle \Phi_0 \mid \Phi_1 \rangle$$

$$\langle \Phi_0 \mid H_1 \mid \Phi_0 \rangle + e_0\langle \Phi_0 \mid \Phi_1 \rangle = e_1\langle \Phi_0 \mid \Phi_0 \rangle + e_0\langle \Phi_0 \mid \Phi_1 \rangle$$

$$\langle \Phi_0 \mid H_1 \mid \Phi_0 \rangle = e_1$$

- Result is known as *Hellmann-Feynman theorem*

- Equation for the 2$^{\text{nd}}$-order term can be written

Moller-Plesset (MP) theory uses perturbation theory to introduce electron correlation

$$\widehat{H}^{(1)} \equiv \widehat{H} - \widehat{H}^{(0)} = \frac{1}{2} \sum_{m \neq n} \frac{1}{r_{m,n}} - \sum_{m} \hat{v}_{\mathrm{HF}}(m)$$

SCF HF 2-electron energy

- First-order theory (MP1) yields just the HF energy

- Second-order theory uses CI wavefunctions to represent perturbations to $\Phi^{(0)}$. Result for correlation energy is:

$$E_{\mathrm{MP2}}^{(2)} = \sum_{\substack{i<j \\ a<b}} \frac{|[ia \mid jb] - [ib \mid ja]|^2}{\varepsilon_i + \varepsilon_j - \varepsilon_a - \varepsilon_b}$$

i,j,a,b are the swapped orbitals, with energies ε_k

Performance of methods is balanced against computational cost

Post-HF levels: Price/Performance

$$HF < MP2 \sim MP3 \sim CCD < CISD$$

$$< MP4SDQ \sim QCISD \sim CCSD < MP4 < QCISD(T) \sim CCSD(T) < \ldots$$

Scaling behavior	Method(s)
N^4	HF
N^5	MP2
N^6	MP3, CISD, MP4SDQ, CCSD, QCISD
N^7	MP4, CCSD(T), QCISD(T)
N^8	MP5, CISDT, CCSDT
N^9	MP6
N^{10}	MP7, CISDTQ, CCSDTQ

Cramer, video 4.03

20

Calendar

Suggested Reading/Viewing

- Autschbach, Chapter 20, 22.1-22.4

- Cramer
    - 4.02 https://www.youtube.com/watch?v=5BZxa6fZHZU
    - 4.03 https://www.youtube.com/watch?v=np81k16E4I0
    - 4.05 https://www.youtube.com/watch?v=n3D1c8zV-x0