Molecular Simulation I

Quantum Chemistry

\[ E = \frac{\langle \Psi | H | \Psi \rangle}{\langle \Psi | \Psi \rangle} \]

Classical Mechanics

\[ U = E_{\text{bond}} + E_{\text{angle}} + E_{\text{torsion}} + E_{\text{non-bond}} \]

Jeffry D. Madura
Department of Chemistry & Biochemistry
Center for Computational Sciences
Duquesne University
Quantum Chemistry

• Molecular Orbital Theory
  – Based on a wave function approach
  – Schrödinger equation

• Density Functional Theory
  – Based on the total electron density
  – Hohenberg – Kohn theorem

• Semi-empirical
  – Some to most integrals parameterized
  – MNDO, AM1, EHT

• Empirical
  – All integrals are parameterized
  – Huckel method
The Beginning...

- Schrödinger equation

\[ H \Psi = E \Psi \]

- Hamiltonian operator

\[ H = -\frac{\hbar^2}{2m} \left( \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \right) + V(r) = -\frac{\hbar^2}{2m} \nabla^2 + V(r) \]

- Wave function (\( \Psi \))
  - characterizes the particles motion
  - various properties of the particle can be derived
Quantum Chemistry

• Start with Schrödinger’s equation

\[ H\Psi = E\Psi \]

• Make some assumptions
  – Born-Oppenheimer approximation
  – Linear combination of atomic orbitals

\[ \Psi = c_a \varphi_a + c_b \varphi_b \]

• Apply the variational method

\[ E \leq \frac{\int \Psi^* H \Psi d\tau}{\int \Psi^* \Psi d\tau} \]
LCAO

• A practical and common approach to solving the Hartree-Fock equations is to write each spin orbital as a linear combination of single electron orbitals (LCAO)

\[ \psi_i = \sum_{\nu=1}^{K} c_{vi} \phi_{\nu} \]

– the \( \phi_{\nu} \) are commonly called \textit{basis functions} and often correspond to \textit{atomic orbitals}
– K basis functions lead to K molecular orbitals
– the point at which the energy is not reduced by the addition of basis functions is known as the Hartree-Fock limit
Basis Sets

• Slater type orbitals (STO)

\[ R_{nl}(r) = (2\zeta)^{n+1/2} \left[(2n)!\right]^{-1/2} r^{n-1} e^{-\zeta r} \]

• Gaussian type orbitals (GTO)
  – functional form

\[ x^a y^b z^c e^{-a r^2} \]
  – zeroth-order Gaussian function

\[ g_s(\alpha, r) = \left(\frac{2\alpha}{\pi}\right)^{3/4} e^{-\alpha r^2} \]
Property of Gaussian functions is that the product of two Gaussians can be expressed as a single Gaussian, located along the line joining the centers of the two Gaussians:

\[ e^{-a_m r_m^2} e^{-a_n r_n^2} = e^{-\frac{a_m a_n}{a_m + a_n} r_{mn}^2} e^{-a r_c^2} \]
• STO vs. GTO
• Gaussian expansion
  – the coefficient
  – the exponent
  – uncontracted or primitive and contracted
  – s and p exponents in the same shell are equal

\[ \phi_\mu = \sum_{i=1}^{L} d_{i\mu} \phi_i(\alpha_{i\mu}) \]

• Minimal basis set
  – STO-NG

• Double zeta basis set
  – linear combination of a ‘contracted’ function and a ‘diffuse’ function.

• Split valence
  – 3-21G, 4-31G, 6-31G
• Polarization
  – to solve the problem of non-isotropic charge distribution.
  – 6-31G*, 6-31G**

• Diffuse functions
  – fulfill as deficiency of the basis sets to describe significant amounts of electron density away from the nuclear centers. (e.g. anions, lone pairs, etc.)
  – 3-21+G, 6-31++G
Roothaan-Hall equations

• The recasting of the integro-differential equations into matrix form.
  – The Hartree-Fock equation is written as
    \[
    F_i (1) \sum_{v=1}^{K} c_{vi} \phi_v (1) = \varepsilon_i \sum_{v=1}^{K} c_{vi} \phi_v (1)
    \]
  – This is transformed to give a Fock matrix (for closed-shell systems)
    \[
    F_{\mu\nu} = H^\text{core}_{\mu\nu} + \sum_{\lambda=1}^{K} \sum_{\sigma=1}^{K} P_{\lambda\sigma} \left( (\mu\nu | \lambda\sigma) - \frac{1}{2} (\mu\lambda | \nu\sigma) \right)
    \]
  – where
    \[
    P_{\lambda\sigma} = 2 \sum_{i=1}^{N/2} c_{\lambda i} c_{\sigma i}
    \]
– The energy is

\[ E = \frac{1}{2} \sum_{\mu=1}^{K} \sum_{\nu=1}^{K} P_{\mu \nu} \left( H^\text{core}_\nu + F_{\mu \nu} \right) \]

– the electron density is expressed as

\[ \rho(r) = \sum_{\mu=1}^{K} \sum_{\nu=1}^{K} P_{\mu \nu} \phi_{\mu}(r) \phi_{\nu}(r) \]

• In matrix form the Roothaan-Hall equation is written as

\[ FC = SCE \]
Solving the Roothaan-Hall Equation

- Common scheme for solving the Roothaan-Hall equations is
  - calculate the integrals to form the Fock matrix, $F$
  - calculate the overlap matrix, $S$
  - diagonalize $S$
  - form $S^{-1/2}$
  - guess or calculate an initial density matrix, $P$
  - Form the Fock matrix using the integrals and density matrix
  - solve the secular equation $|F'-EI|=0$ to give the eigenvalues $E$
    and the eigenvectors $C'$ by diagonalizing $F'$
  - calculate the molecular orbital coefficients, $C$, from $C=S^{-1/2}C'$
  - calculate a new density matrix, $P$, from matrix $C$
  - check for convergence
RHF vs. UHF

- **Restricted Hartree-Fock (RHF)**
  - closed-shell molecules

- **Restricted Open-shell Hartree-Fock (ROHF)**
  - combination of singly and doubly occupied molecular orbitals.

- **Unrestricted Hartree-Fock (UHF)**
  - open-shell molecules
  - Pople and Nesbet: one set of molecular orbitals for \( \alpha \) spin and another for the \( \beta \) spin.
UHF and RHF Dissociation Curves for $H_2$
Electron Correlation

• The most significant drawback to HF theory is that it fails to adequately represent electron correlation.

\[ E_{corr} = E^{NR} - E^{HF} \]

• Configuration Interactions
  – excited states are included in the description of an electronic state

• Many Body Perturbation Theory
  – based upon Rayleigh-Schrödinger perturbation theory
Configuration Interaction

• The CI wavefunction is written as

$$\Psi = c_0 \Psi_0 + c_1 \Psi_1 + c_2 \Psi_2 + \cdots$$

– where $\Psi_0$ is the HF single determinant
– where $\Psi_1$ is the configuration derived by replacing one of the occupied spin orbitals by a virtual spin orbital
– where $\Psi_2$ is the configuration derived by replacing one of the occupied spin orbitals by a virtual spin orbital

• The system energy is minimized in order to determine the coefficients, $c_0, c_1, \text{etc.}$, using a linear variational approach
Many Body Perturbation Theory

- Based upon perturbation concepts \( H = H_0 + V \)
- The correction to the energies are
  \[
  E_i^{(0)} = \int \Psi_i^{(0)} H_0 \Psi_i^{(0)} d\tau \\
  E_i^{(1)} = \int \Psi_i^{(0)} V \Psi_i^{(0)} d\tau \\
  E_i^{(2)} = \int \Psi_i^{(0)} V \Psi_i^{(1)} d\tau \\
  E_i^{(3)} = \int \Psi_i^{(0)} V \Psi_i^{(2)} d\tau
  \]
  - Perturbation methods are size independent
  - these methods are not variational
Geometry Optimization

• Derivatives of the energy

\[ E(x_f) = E(x) + \sum_i \frac{\partial E(x)}{\partial x_i} (x_{i_f} - x_i) + \frac{1}{2} \sum_i \sum_j \frac{\partial^2 E(x)}{\partial x_i \partial x_j} (x_{i_f} - x_i)(x_{j_f} - x_j) + \cdots \]

– the first term is set to zero
– the second term can be shown to be equivalent to a force
– the third term can be shown to be equivalent to a force constant
• Internal coordinate, Cartesian coordinate, and redundant coordinate optimization
  – choice of coordinate set can determine whether a structure reaches a minimum/maximum and the speed of this convergence.
  – Internal coordinates are defined as bond lengths, bond angles, and torsions. There are $3N-6$ ($3N-5$) such degrees of freedom for each molecule. Chemists work in this world. Z-matrix...
  – Cartesian coordinates are the standard x, y, z coordinates. Programs often work in this world.
  – Redundant coordinates are defined as the number of coordinates larger than $3N-6$. 
Frequency Calculation

• The second derivatives of the energy with respect to the displacement of coordinate yields the force constants.

• These force constants in turn can be used to calculate frequencies.
  – All real frequencies (positive force constants): local minimum
  – One imaginary frequency (one negative force constant): saddle point, a.k.a. transition state.

• From vibrational analysis can compute thermodynamic data
Molecular Properties

- Charges
  - Mulliken
  - Löwdin
  - electrostatic fitted (ESP)

- Bond orders

- Bonding
  - Natural Bond Analysis
  - Bader’s AIM method

- Molecular orbitals and total electron density

- Dipole Moment

- Energies
  - ionization and electron affinity
Energies

• Koopman’s theorem
  – equating the energy of an electron in an orbital to the energy required to remove the electron to the corresponding ion.
    • ‘frozen’ orbitals
    • lack of electron correlation effects
Dipole Moments

• The electric multipole moments of a molecule reflect the distribution of charge.
  – Simplest is the dipole moment
    \[ \mu = \sum_i q_i r_i \]
  – nuclear component
    \[ \mu_{\text{nuclear}} = \sum_{A=1}^{M} Z_A R_A \]
  – electronic
    \[ \mu_{\text{electronic}} = \sum_{\mu=1}^{K} \sum_{\nu=1}^{K} P_{\mu\nu} \int d\tau \phi_\mu (-r) \phi_\nu \]
Molecular Orbitals and Total Electron Density

- Electron density at a point $r$
  \[ \rho (r) = 2 \sum_{i=1}^{N/2} |\psi_i (r)|^2 = \sum_{\mu=1}^{K} P_{\mu\mu} \phi_\mu (r) \phi_\mu (r) + 2 \sum_{\mu=1}^{K} \sum_{\nu=\mu+1}^{K} P_{\mu\nu} \phi_\mu (r) \phi_\nu (r) \]

- Number of electrons is
  \[ N = 2 \sum_{i=1}^{N/2} \int dr |\psi_i (r)|^2 = \sum_{\mu=1}^{K} P_{\mu\mu} + 2 \sum_{\mu=1}^{K} \sum_{\nu=\mu+1}^{K} P_{\mu\nu} S_{\mu\nu} \]

- Molecular orbitals
  - HOMO
  - LUMO
Bonding

• Natural Bond Analysis
  – a way to describe N-electron wave functions in terms of localized orbitals that are closely tied to chemical concepts.

• Bader
  – F. W. Bader’s theory of ‘atoms in molecules’.
  – This method provides an alternative way to partition the electrons among the atoms in a molecule.
  – Gradient vector path
  – bond critical points
  – charges are relatively invariant to the basis set
Bond Orders

- Wiberg

\[ W_{AB} = \sum_{\mu \text{ on } A} \sum_{\nu \text{ on } B} |P_{\mu \nu}|^2 \]

- Mayer

\[ B_{AB} = \sum_{\mu \text{ on } A} \sum_{\nu \text{ on } B} (PS)_{\mu \nu} (PS)_{\nu \mu} \]

- Bond orders can be computed for intermediate structures which can be useful way to describe similarity of the TS to the reactants or to the products.
Charges

• Mulliken

\[ q_A = Z_A - \sum_{\mu=1; \text{on } A}^K P_{\mu\mu} - \sum_{\mu=1; \text{on } A}^K \sum_{\nu=1; \nu \neq \mu}^K P_{\mu\nu} S_{\mu\nu} \]

• Löwdin

– atomic orbitals are transformed to an orthogonal set, along with the mo coefficients

\[ \phi'_\mu = \sum_{\nu=1}^K \left( S^{-1/2} \right)_{\nu\mu} \phi_\nu \]

\[ q_A = Z_A - \sum_{\mu=1; \mu \text{ on } A}^K \left( S^{1/2} P^{1/2} \right)_{\mu\mu} \]
• Electrostatic potentials
  – the electrostatic potential at a point \( r \), \( \phi(r) \), is defined as the work done to bring a unit positive charge from infinity to the point.
  – the electrostatic interaction energy between a point charge \( q \) located at \( r \) and the molecule equals \( q\phi(r) \).
  – there is a nuclear part and electronic part

\[
\phi_{\text{nucl}}(r) = \sum_{A=1}^{M} \frac{Z_A}{|r - R_A|} \quad \phi_{\text{elec}}(r) = -\int \frac{dr' \rho(r')}{|r' - r|}
\]

\[
\phi(r) = \phi_{\text{nucl}}(r) + \phi_{\text{elec}}(r)
\]
Water Example

Geometry

Z-MATRIX (ANGSTROMS AND DEGREES)

<table>
<thead>
<tr>
<th>CD</th>
<th>Cent</th>
<th>Atom</th>
<th>N1</th>
<th>Length/X</th>
<th>N2</th>
<th>Alpha/Y</th>
<th>N3</th>
<th>Beta/Z</th>
<th>J</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>H</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>O</td>
<td>1</td>
<td>0.989400</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>100.028</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>H</td>
<td>2</td>
<td>0.989400</td>
<td>1</td>
<td>100.028</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Energy

E(RHF) = -74.9658952265 A.U.

Population Analysis

Population analysis using the SCF density.

Alpha occ. eigenvalues -- -20.25226 -1.25780 -0.59411 -0.45987 -0.39297
Alpha virt. eigenvalues -- 0.58175 0.69242

Condensed to atoms (all electrons):

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>H</td>
<td>0.626190</td>
<td>0.253760</td>
</tr>
<tr>
<td>2</td>
<td>O</td>
<td>0.253760</td>
<td>7.823081</td>
</tr>
<tr>
<td>3</td>
<td>H</td>
<td>-0.045250</td>
<td>0.253760</td>
</tr>
</tbody>
</table>

Total atomic charges:

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>H</td>
<td>0.165300</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>O</td>
<td>-0.330601</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>H</td>
<td>0.165300</td>
<td></td>
</tr>
</tbody>
</table>
Limitations, Strengths & Reliability

• Limitations
  – Requires more CPU time
  – Can treat smaller molecules
  – Calculations are more complex
  – Have to worry about electronic configuration

• Strengths
  – No experimental bias
  – Can improve a calculation in a logical manner (e.g. basis set, level of theory, …)
  – Provides information on intermediate species, including spectroscopic data
  – Can calculate novel structures
  – Can calculate any electronic state

• Reliability
  – The mean deviation between experiment and theory for heavy-atom bond lengths in two-heavy-atom hydrides drops from 0.082 Å for the RHF/STO-3G level of theory to just 0.019 Å for MP2/6-31G(d).
  – Heats of hydrogenation of a range of saturated and unsaturated systems are calculated sufficiently well at the Hartree-Fock level of theory with a moderate basis set (increasing the basis set from 6-31G(d) to 6-31G(d,p) has little effect on the accuracy of these numbers).
  – Inclusion of electron correlation is mandatory in order to get good agreement between experiment and theory for bond dissociation energies (MP2/6-31G(d,p) does very well for the one-heavy-atom hydrides).