Introduction to Computational Chemistry

Frédéric Castet

f.castet@ism.u-bordeaux1.fr

http://blake.ism.u-bordeaux1.fr/~castet

Université Bordeaux 1

Summary

- The Schrödinger equation
- The Hartree-Fock-Roothaan method
- Semiempirical methods

THE SCHRÖDINGER EQUATION

THE SCHRÖDINGER EQUATION

The time-independent Schrödinger equation allows determining the *stationary states* of a system of particles (e.g. a molecule)

$$\hat{H}\Psi_n = E_n \Psi_n$$

Ψ_n is the total wavefunction of the system (electrons + nuclei)

Hamiltonian operator associated to the total energy

$$\hat{H} = \hat{H}_{nuc} + \hat{H}_{elec} + \hat{H}_{nuc/elec}$$

Born-Oppenheimer Approximation

Electronic and nuclei motions are uncoupled

$$\Psi_n = \Phi_{elec} \times \Phi_{nuc}$$

2 independent equations :

$$\hat{H}_{nuc}\Phi_{nuc} = E_{nuc}\Phi_{nuc}$$

vibration-rotation spectrum of the molecule

$$\hat{H}_{elec}\Phi_{elec} = E_{elec}\Phi_{elec}$$

Electronic states in the field of fixed nuclei

POTENTIAL ENERGY SURFACE

The time-independent Schrödinger equation allows determining the *stationary states* of a system of particles (e.g. a molecule)

Electronic **Schrödinger equation**

$$\hat{H}_{elec}\Phi_{elec} = E_{elec}\Phi_{elec}$$

Solution for different values of the nuclear coordinates (Q)

Energy $E_{el}(Q)$ of each electronic state

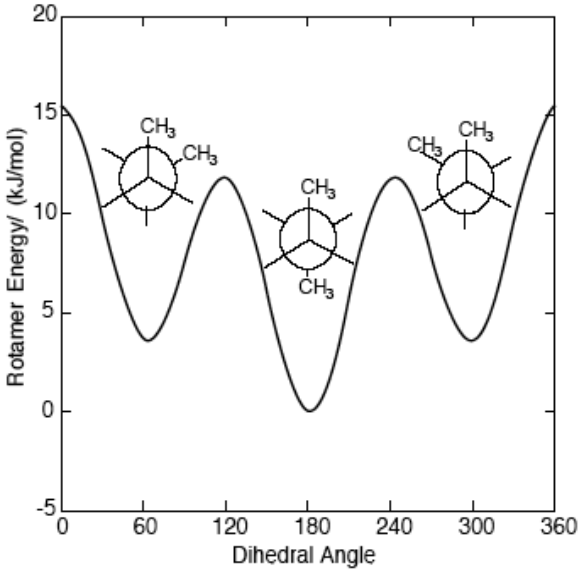
The operator describing the electrons/nuclei interactions depends on Q

$$\hat{H}_{nuc} = \sum_p -\frac{m_e}{2m_p}\Delta_p + \underbrace{\frac{1}{2}\sum_p\sum_q \frac{Z_p Z_q}{r_{pq}} + \boxed{\hat{E}_{el}(Q)}}_{V(Q)}$$

Potentielle energy of the molecule

POTENTIAL ENERGY SURFACE

V(Q)

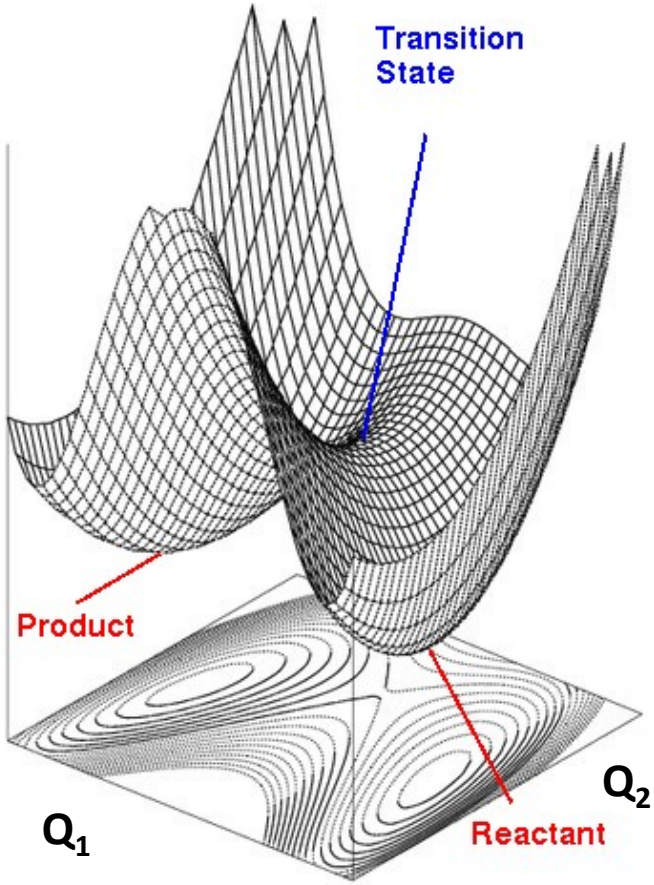


1 single coord. Q

Potential energy curve

Ex : Internal rotation interne of n-butane Q = dihedral angle

V(Q)



2 coord. Q

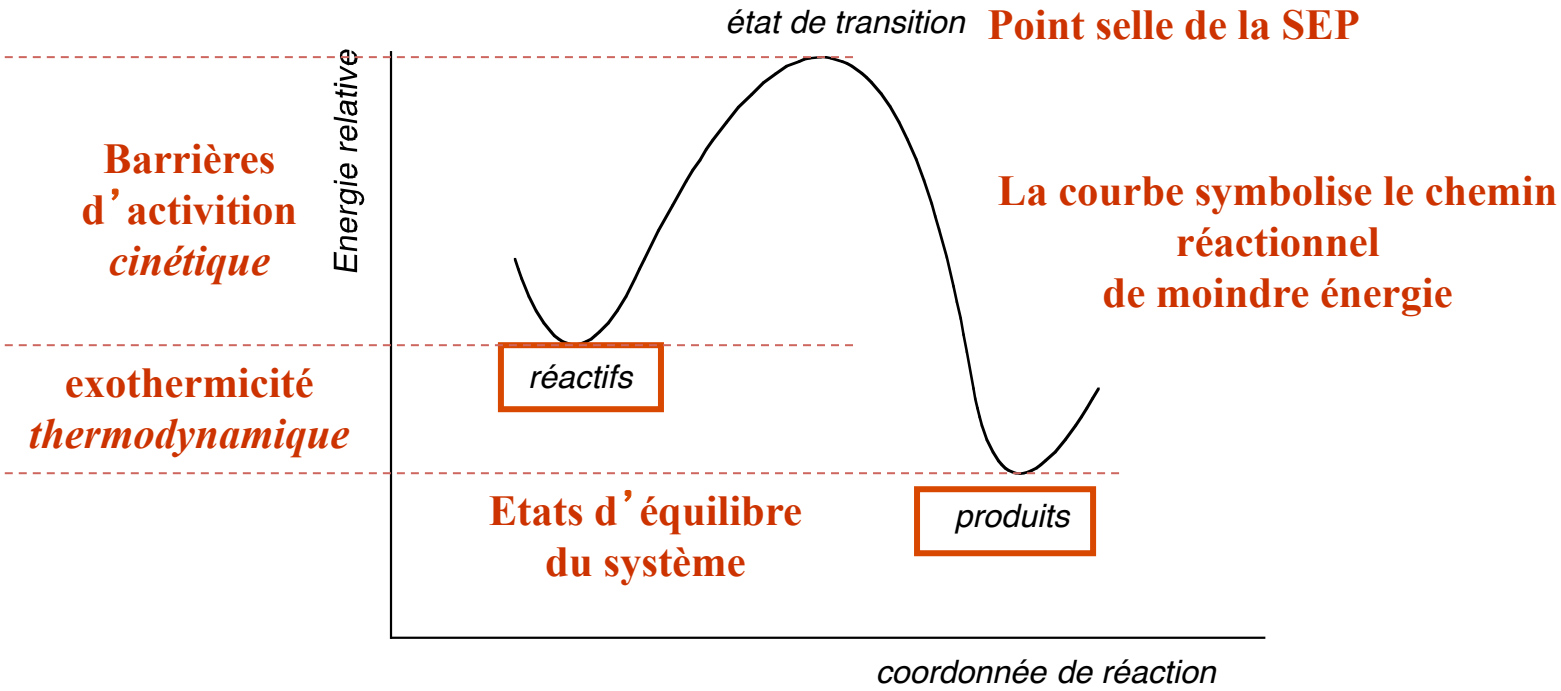
Potential energy surface

→ For N atoms, V depends on 3N–6 coord. Q

The PES is an hypersurface in a {3N–6}-dimension space

POTENTIAL ENERGY SURFACE

Simplified representation of the SEP of a chemical reaction



Déformation du système moléculaire (déplacements atomiques, rupture/formation de liaison…)

THE ELECTRONIC SCHRÖDINGER EQUATION

Aim of an electronic structure calculation

Solving the electronic Schrödinger equation for a given geometry of the molecule (i.e. a given point of the PES)

$$\hat{H}_{elec}\Phi_{elec} = E_{elec}\Phi_{elec}$$

Electronic Hamiltonian :

$$\hat{H}_{elec} = -\frac{1}{2}\sum_i \Delta_i - \frac{1}{2}\sum_i\sum_p \frac{Z_p}{r_{ip}} + \sum_i\sum_{j>i} \frac{1}{r_{ij}}$$

Electronic wavefunction

Electronic state in the electrostatic field of fixed nuclei

$$\Phi_{elec}(x_1, y_1, z_1, \ldots x_N, y_N, z_N)$$

Total energy of the molecule

$$V(Q) = E_{elec}(Q) + E_{nuc}$$

THE HARTREE-FOCK METHOD

THE HARTREE-FOCK-ROOTHAAN METHOD

Electronic Hamiltonian for a molecule with 2N electrons and P nuclei

$$\hat{H} = -\sum_{\mu=1}^{2N}\frac{1}{2}\Delta_\mu - \sum_{\mu=1}^{2N}\sum_{A=1}^{P}\frac{Z_A}{r_\mu} + \sum_{\mu<\upsilon}^{2N}\sum_{\upsilon=1}^{2N}\frac{1}{r_{\mu\upsilon}} = \sum_{\mu=1}^{2N}\hat{h}_\mu + \sum_{\mu<\upsilon}^{2N}\sum_{\upsilon=1}^{2N}\frac{1}{r_{\mu\upsilon}}$$

Single determinant wavefunction

$$\Psi = \frac{1}{\sqrt{(2N)!}}\begin{vmatrix} \psi_1(\mathbf{r}_1) & \psi_2(\mathbf{r}_1) & \dots & \psi_{2N}(\mathbf{r}_1) \\ \psi_1(\mathbf{r}_2) & \psi_2(\mathbf{r}_2) & \dots & \psi_{2N}(\mathbf{r}_2) \\ \dots & \dots & \dots & \dots \\ \psi_1(\mathbf{r}_{2N}) & \psi_2(\mathbf{r}_{2N}) & \dots & \psi_{2N}(\mathbf{r}_{2N}) \end{vmatrix}$$

Spin-orbitals

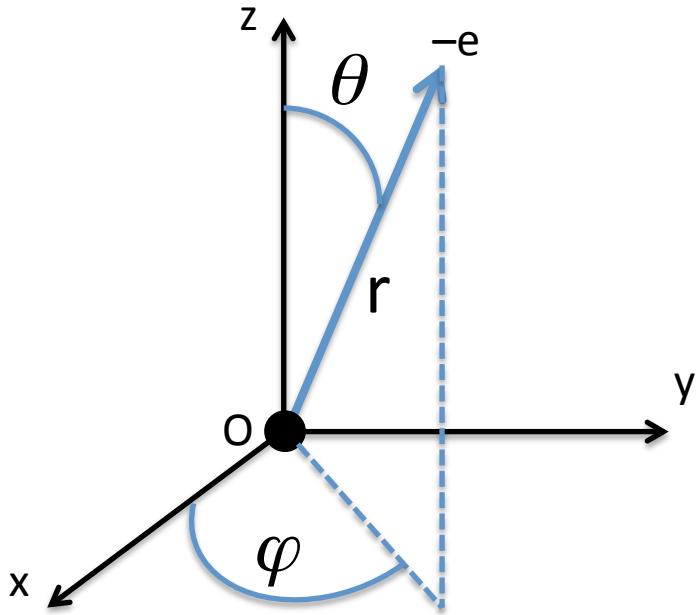
$$\psi_\mu(\mathbf{r}_\mu) = \phi_\mu(\mathbf{r}_\mu) \times \xi_\mu(\sigma_\mu)$$

Spin function α or β

Molecular orbital

$$\mathbf{r} = \{r, \theta, \varphi\}$$

$$d\mathbf{r} = dV = r^2 \sin\theta \, dr d\theta d\varphi$$



HF: search for the best *variational* wavefunction

THE HARTREE-FOCK-ROOTHAAN METHOD

Variational principle

$$\langle E \rangle = \int \Psi \hat{H} \Psi \, d\mathbf{r} \geq E_{exact}$$

1. Introduction of variational parameters $\Psi \rightarrow \Psi(\alpha_1 \ldots \alpha_M)$

2. Optimization of the parameters

$$\frac{\partial \langle E \rangle}{\partial \alpha_i} = 0 \quad \forall \, i$$ ***Minimization of the total energy***

Linear Combination of Atomic Orbitals (Roothaan)

$$\phi_i(\mathbf{r}) = \sum_{p=1}^{M} C_{pi} \chi_p(\mathbf{r})$$

- Basis set of known atomic functions
- Expansion coefficients = ***variational parameters***

HF: search for the best *orbitals*

THE HARTREE-FOCK-ROOTHAAN METHOD

Total energy

$$\langle E \rangle = \int \Psi \hat{H} \Psi \, d\mathbf{r}$$

$$\langle E \rangle = \sum_{i=1}^{N} 2I_i + \sum_{i=1}^{N}\sum_{j=1}^{N} \left(2J_{ij} - K_{ij}\right)$$

Sum over N doubly occupied molecular orbitals

$$\Psi = \frac{1}{\sqrt{(2N)!}} \begin{vmatrix} \psi_1(\mathbf{r}_1) & \psi_2(\mathbf{r}_1) & \dots & \psi_{2N}(\mathbf{r}_1) \\ \psi_1(\mathbf{r}_2) & \psi_2(\mathbf{r}_2) & \dots & \psi_{2N}(\mathbf{r}_2) \\ \dots & \dots & \dots & \dots \\ \psi_1(\mathbf{r}_{2N}) & \psi_2(\mathbf{r}_{2N}) & \dots & \psi_{2N}(\mathbf{r}_{2N}) \end{vmatrix}$$

HF integrals

$$I_i = \int \phi_i^*(\mathbf{r}_\mu) \hat{h} \phi_i(\mathbf{r}_\mu) d\mathbf{r}_\mu$$ 1-electron integrals

$$J_{ij} = \int \frac{\phi_i^*(\mathbf{r}_\mu)\phi_i(\mathbf{r}_\mu)\phi_j^*(\mathbf{r}_\nu)\phi_j(\mathbf{r}_\nu)}{r_{\mu\nu}} d\mathbf{r}_\mu d\mathbf{r}_\nu$$ Coulomb integrals

$$K_{ij} = \int \frac{\phi_i^*(\mathbf{r}_\mu)\phi_i(\mathbf{r}_\nu)\phi_j^*(\mathbf{r}_\nu)\phi_j(\mathbf{r}_\mu)}{r_{\mu\nu}} d\mathbf{r}_\mu d\mathbf{r}_\nu$$ Exchange integrals

THE HARTREE-FOCK-ROOTHAAN METHOD

Roothaan expansion

$$\phi_i(\mathbf{r}) = \sum_{p=1}^{M} C_{pi}\chi_p(\mathbf{r})$$

Integrals in the AO basis

$$I_i = \int \phi_i^*(\mathbf{r}_\mu)\hat{h}\phi_i(\mathbf{r}_\mu)d\mathbf{r}_\mu = \sum_{p=1}^{M}\sum_{q=1}^{M} C_{pi}C_{qi}h_{pq}$$

$$h_{pq} = \int \chi_p(\mathbf{r}_\mu)\hat{h}\chi_q(\mathbf{r}_\mu)d\mathbf{r}_\mu \Longrightarrow \textbf{M}^2 \textit{ integrals}$$

$$J_{ij} = \int \frac{\phi_i^*(\mathbf{r}_\mu)\phi_i(\mathbf{r}_\mu)\phi_j^*(\mathbf{r}_\nu)\phi_j(\mathbf{r}_\nu)}{r_{\mu\nu}} d\mathbf{r}_\mu d\mathbf{r}_\nu = \sum_{p=1}^{M}\sum_{q=1}^{M}\sum_{r=1}^{M}\sum_{s=1}^{M} C_{pi}C_{qi}C_{rj}C_{sj}(pq|rs)$$

$$(pq|rs) = \int \frac{\chi_p(\mathbf{r}_\mu)\chi_q(\mathbf{r}_\mu)\chi_r(\mathbf{r}_\nu)\chi_s(\mathbf{r}_\nu)}{r_{\mu\nu}} d\mathbf{r}_\mu\, d\mathbf{r}_\nu \Longrightarrow \textbf{M}^4 \textit{ integrals}$$

$$(pq|rs) = (qp|rs) = (pq|sr) = (qp|sr) = (rs|pq) = (sr|pq) = (rs|qp) = (sr|qp)$$

THE HARTREE-FOCK-ROOTHAAN METHOD

Energy minimization

$$\langle E \rangle = \sum_{i=1}^{N} 2I_i + \sum_{i=1}^{N}\sum_{j=1}^{N}\left(2J_{ij} - K_{ij}\right)$$

Minimisation →

N Fock equations

$$\hat{F}(\mathbf{r})\phi_i(\mathbf{r}) = \varepsilon_i\phi_i(\mathbf{r})$$

Fock operator

$$\hat{F}(\mathbf{r}_\mu) = \hat{h}(\mathbf{r}_\mu) + \sum_{i=1}^{N}\left[2\hat{J}_i(\mathbf{r}_\mu) - \hat{K}_i(\mathbf{r}_\mu)\right]$$

1-electron Fock operator

Describes electron μ in the mean electrostatic field of the other electrons

$$\hat{J}_i(\mathbf{r}_\mu)\phi_j(\mathbf{r}_\mu) = \left(\int \frac{\phi_i^*(\mathbf{r}_\nu)\phi_i(\mathbf{r}_\nu)}{r_{\mu\nu}} dV_\nu\right)\phi_j(\mathbf{r}_\mu)$$

Coulomb operator

$$\hat{K}_i(\mathbf{r}_\mu)\phi_j(\mathbf{r}_\mu) = \left(\int \frac{\phi_i^*(\mathbf{r}_\nu)\phi_j(\mathbf{r}_\nu)}{r_{\mu\nu}} dV_\nu\right)\phi_i(\mathbf{r}_\mu)$$

Exchange operator

The Fock operator depends on its own solutions **ϕ(r)**

↳ ***Iterative process until self-consistence***

THE HARTREE-FOCK-ROOTHAAN METHOD

Matrix form of the Fock equations

$$\hat{F}(\mathbf{r})\phi_i(\mathbf{r}) = \varepsilon_i\phi_i(\mathbf{r})$$

$$\text{with } \phi_i(\mathbf{r}) = \sum_{p=1}^{M} C_{pi}\chi_p(\mathbf{r})$$

Method of Linear Variations

$$\mathbf{FC} = \varepsilon\mathbf{SC}$$

W. Ritz, J. Reine Angew. Math. **135**, 1 (1909).
See A. Szabo and N. S. Ostlund, *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory*, McGraw-Hill, New York, 1989.

Overlap matrix S

Dimension M x M

Elements $S_{pq} = \int \chi_p^*(\mathbf{r}_\nu)\chi_q(\mathbf{r}_\nu)d\mathbf{r}_\nu$

Fock matrix F

Dimension M x M

Elements $F_{pq} = \int \chi_p^*(\mathbf{r}_\nu)\hat{F}(\mathbf{r}_\nu)\chi_q(\mathbf{r}_\nu)d\mathbf{r}_\nu = h_{pq} + \sum_r\sum_s D_{rs}\left[(pq|rs) - (ps|rq)\right]$

First-order density matrix D

Dimension M x M

Elements $D_{pq} = \sum_i n_i C_{pi}C_{qi}$

THE HARTREE-FOCK-ROOTHAAN METHOD

General solution of the HF matrix equations for non orthogonal basis sets

$$\mathbf{FC} = \mathbf{SC}\varepsilon$$

One seeks to obtain an eigenvalue equation of the form

$$\mathbf{F}'\mathbf{C}' = \mathbf{C}'\varepsilon$$

Löwdin orthogonalization

P. O. Löwdin *J. Chem. Phys*. **1950**, *18*, 365

$$\mathbf{F} \times \mathbf{S}^{-1/2} \times \mathbf{S}^{1/2} \times \mathbf{C} = \mathbf{S}^{1/2} \times \mathbf{S}^{1/2}\mathbf{C}\varepsilon$$

By multiplying by $S^{-1/2}$ on each side

$$\underbrace{\mathbf{S}^{-1/2} \times \mathbf{F} \times \mathbf{S}^{-1/2}}_{\mathbf{F}'} \times \underbrace{\mathbf{S}^{1/2} \times \mathbf{C}}_{\mathbf{C}'} = \underbrace{\mathbf{S}^{1/2}\mathbf{C}}_{\mathbf{C}'}\varepsilon$$

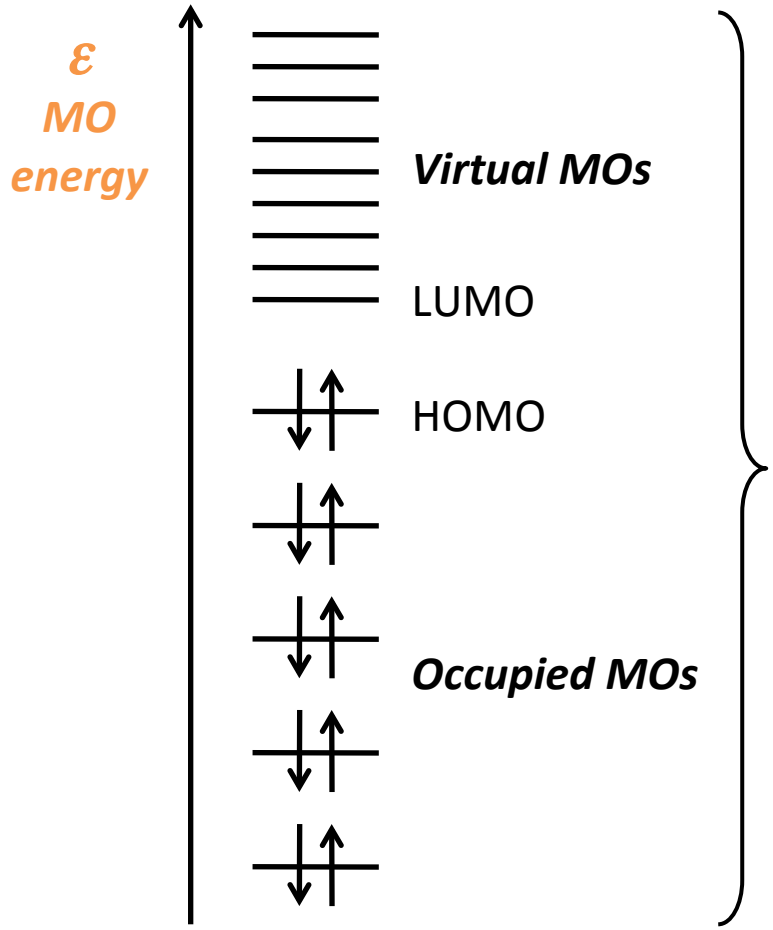
$$\Rightarrow \mathbf{F}'\mathbf{C}' = \mathbf{C}'\varepsilon$$

$\mathbf{C}'$ *= eigenvectors of* $\mathbf{F}'$

$$\mathbf{C} = \mathbf{S}^{-1/2}\mathbf{C}'$$

THE HARTREE-FOCK-ROOTHAAN METHOD

Solution of the HF equations



M basis functions
⇒ M MOs

Koopmans theorem: IP ≈ – ε

Electron Density and Related Properties

Total electron density

$$\rho(\mathbf{r}) = \sum_{i=1}^{M} n_i \phi_i^*(\mathbf{r}) \phi_i(\mathbf{r})$$

$$\int_{\text{space}} \rho(\mathbf{r}) d\mathbf{r} = \sum_{i=1}^{M} n_i \int_{\text{space}} \phi_i^*(\mathbf{r}) \phi_i(\mathbf{r}) d\mathbf{r} = \sum_{i=1}^{M} n_i = N$$

Mulliken AO populations

$$\rho(\mathbf{r}) = \sum_{i=1}^{M} n_i \sum_{p=1}^{M} \sum_{q=1}^{M} C_{pi} C_{qi} \chi_p(\mathbf{r}) \chi_q(\mathbf{r}) = \sum_{p=1}^{M} \sum_{q=1}^{M} D_{pq} \chi_p(\mathbf{r}) \chi_q(\mathbf{r}) \quad \textit{with} \quad D_{pq} = \sum_{i=1}^{M} n_i C_{pi} C_{qi}$$

$$\int_{\text{space}} \rho(\mathbf{r}) d\mathbf{r} = N = \sum_{p=1}^{M} \sum_{q=1}^{M} D_{pq} S_{pq} = \sum_{p=1}^{M} \left(D_{pp} + \sum_{q \neq p}^{M} D_{pq} S_{pq} \right) = \sum_{p=1}^{M} \left(Q_p \right)$$

$$Q_p = D_{pp} + \sum_{q \neq p}^{M} D_{pq} S_{pq}$$

Electron population in χ_p

Mulliken atomic charges

$$n_A = \sum_{p \in A} Q_p \quad \Rightarrow \quad \delta_A = Z_A - n_A$$

Net charge on atom A

ELECTRON DENSITY AND RELATED PROPERTIES

Dipole moment

$$\vec{\mu} = \int \Psi \hat{\vec{\mu}} \Psi d\mathbf{r} = \int \Psi \left(\hat{\vec{\mu}}_{elec} + \hat{\vec{\mu}}_{nuc} \right) \Psi d\mathbf{r}$$

$$with \quad \hat{\vec{\mu}}_{elec} = \sum_i -e\mathbf{r}_i \quad and \quad \hat{\vec{\mu}}_{nuc} = \sum_A eZ_A \vec{R}_A$$

$$\vec{\mu}_{nuc} = \sum_A \int \phi_i^*(\mathbf{r}) \; eZ_A \vec{R}_A \phi_i^*(\mathbf{r}) d\mathbf{r} = \sum_A eZ_A \vec{R}_A$$

$$\vec{\mu}_{elec} = -2e \sum_{i=1}^{occ} \int \phi_i^*(\mathbf{r}) \; \mathbf{r} \; \phi_i^*(\mathbf{r}) d\mathbf{r}$$

$$\vec{\mu}_{elec} = -2e \sum_{i=1}^{occ} C_{pi} C_{qi} \sum_{p=1}^{M} \sum_{q=1}^{M} \int \chi_p(\mathbf{r}) \; \mathbf{r} \; \chi_q(\mathbf{r}) d\mathbf{r}$$

$$= -e \sum_{p=1}^{M} \sum_{q=1}^{M} D_{pq} \underbrace{\int \chi_p(\mathbf{r}) \; \mathbf{r} \; \chi_q(\mathbf{r}) d\mathbf{r}}$$

Dipole integrals

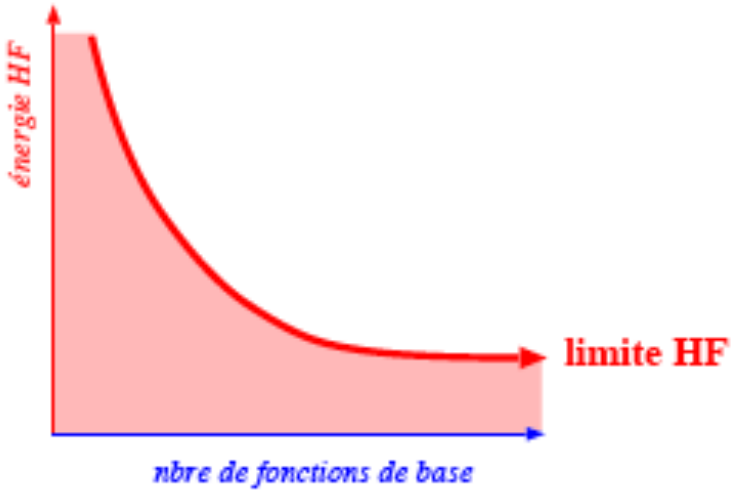
ATOMIC BASIS SETS

Roothaan expansion

$$\phi_i(\mathbf{r}) = \sum_{p=1}^{M} C_{pi}\chi_p(\mathbf{r})$$

Rigorously, the number of the basis functions should be infinite. In this case, the wavefunction corresponds to the lowest possible energy.

This asymptotic limit is referred to as ***the Hartree-Fock limit***.



In practice, one uses finite-size basis sets. The quality of the results depends on the size and on the nature of the basis set.

VARIOUS TYPES OF BASIS SETS

$$\phi_i(\mathbf{r}) = \sum_{p=1}^{M} C_{pi} \underbrace{\chi_p(\mathbf{r})}$$

Atomic basis functions

Slater-type orbitals (STO)

$$\chi^{STO}_{\alpha,n,l,m}(r,\theta,\varphi) = NY_{l,m}(\theta,\varphi) r^{n-1} e^{-\alpha r}$$

✓ ***STO are accurate because they are derived from the exact solutions of the Schrödinger equation for hydrogen-like systems***

✗ ***STO are not efficient for evaluating the 3- and 4-center integrals***

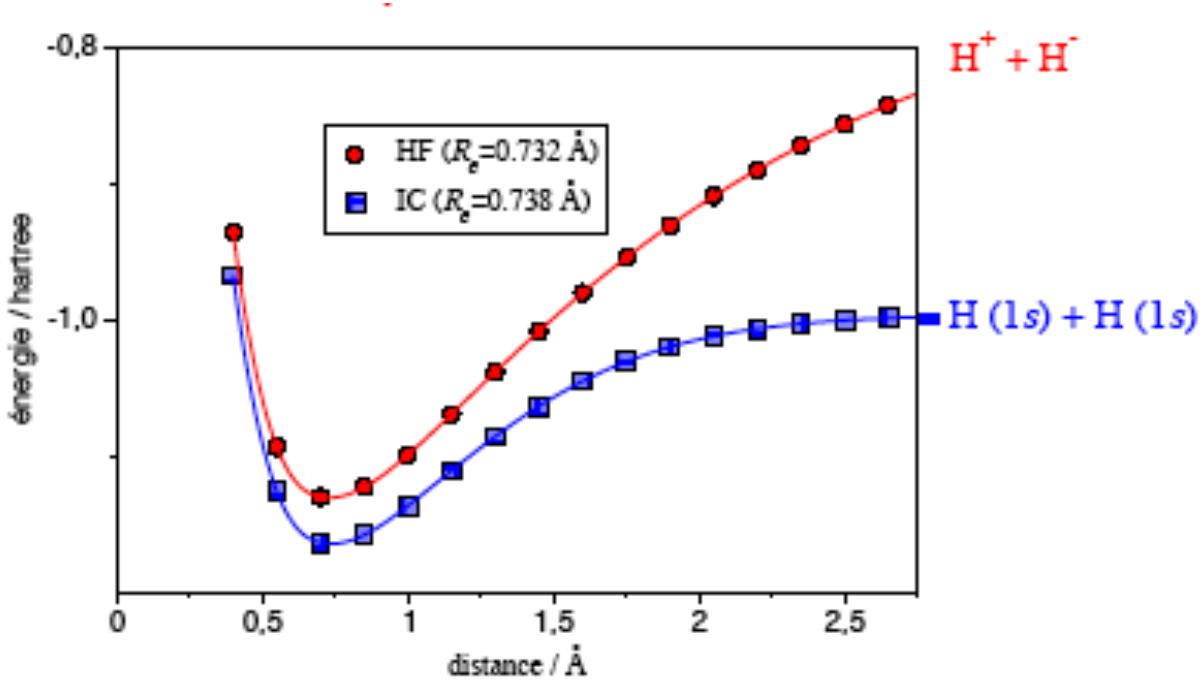
Gaussian-type orbitals (GTO)

$$\chi^{GTO}_{\alpha,n,l,m}(r,\theta,\varphi) = NY_{l,m}(\theta,\varphi) r^{2n-2-l} e^{-\alpha r^2}$$

✗ ***GTO are less accurate than STO. Several GTO are used to describe one AO (split-valence basis sets).***

✓ ***The use of GTO ensures analytical solutions for all integrals of the HF method***

LIMIT OF THE HF MODEL



Base 6-31+G(d,p)

exp	IC
R_e= 0.741 Å	R_e = 0.738 Å
D_e = 4.75 eV	D_e = 4.5 eV

(1hartree = 27.21 eV)

The HF model allows to describe the stability of the covalent bond. However, it is not able to describe its breacking or formation.

ELECTRON CORRELATION

The HF wavefunction does not include ***electron correlation***.

Each electron of the molecules feels an average electron density, independent of its position.

This lack of electron correlation is a consequence of writting the wavefunction as a single Slater determinant, in which electrons are affected to orbitals without possibility of moving from one MO to another.

This lack of flexibility is at the origin of the failure of HF for describing fine electronic effects.

Correlated models exist to overcome this drawback:

Multi-configurational wavefunctions (IC, MC-SCF)
Perturbative models (MP2...)
Density Functional Theory (DFT)
…

PRACTICAL LIMITATION OF AB INITIO IMPLEMENTATIONS

For quantitative results, it is often necessary to use quantum chemical methods including electron correlation, which leads to complex calculations.

Even the HF method has a significant computational cost :

Inventory of the integrals to be computed

$$I_i = \int \phi_i^*(\mathbf{r}_\mu)\hat{h}\phi_i(\mathbf{r}_\mu)d\mathbf{r}_\mu = \sum_{p=1}^{M}\sum_{q=1}^{M} C_{pi}C_{qi}h_{pq}$$

$$h_{pq} = \int \chi_p(\mathbf{r}_\mu)\hat{h}\chi_q(\mathbf{r}_\mu)d\mathbf{r}_\mu \Longrightarrow M^2 \textit{ integrals}$$

$$J_{ij} = \int \frac{\phi_i^*(\mathbf{r}_\mu)\phi_i(\mathbf{r}_\mu)\phi_j^*(\mathbf{r}_\nu)\phi_j(\mathbf{r}_\nu)}{r_{\mu\nu}} d\mathbf{r}_\mu d\mathbf{r}_\nu = \sum_{p=1}^{M}\sum_{q=1}^{M}\sum_{r=1}^{M}\sum_{s=1}^{M} C_{pi}C_{qi}C_{rj}C_{sj}\left(pq|rs\right)$$

$$\left(pq|rs\right) = \int \frac{\chi_p(\mathbf{r}_\mu)\chi_q(\mathbf{r}_\mu)\chi_r(\mathbf{r}_\nu)\chi_s(\mathbf{r}_\nu)}{r_{\mu\nu}} d\mathbf{r}_\mu\, d\mathbf{r}_\nu \Longrightarrow M^4 \textit{ integrals}$$

A HF calculation on the benzene molecule using a double-zeta basis set requires calculating M^4 = 18,974,736 (pq\|rs) integrals

SEMIEMPIRICAL MODELS

SEMI-EMPIRICAL MODELS: BASIC PRINCIPLES

Approximations on the integrals

a priori weak integrals are neglected in the HF equations

Empirical corrections

The integrals considered in the calculations are parameterized to reproduce either experimental results or results obtained using more sophisticated *ab initio* methods

Several families of methods

We can distinguish methods that are:

- **Restricted to all valence electrons**
  - Extended Hückel method, proposed by R. Hoffman
  - CNDO, INDO, NDDO, first introduced by J. Pople

  For an historical overview, see: Pople and Beveridge, *Approximate molecular orbital theory*, McGraw Hill, New York (1970)
  - INDO/S (ZINDO) dedicated to the simulation of electronic spectra
- **Restricted to π electrons**
  - Pariser-Parr-Pople (PPP) … Hubbard for physicists
  - Hückel … Tight-Binding for physicists

Parameterization

*All these methods include **empirical parameters** that have to be properly adjusted*

General Philosophy of NDDO Parameterization

M. Dewar (1918-1997)

Experimental data from a **small group of representative molecules** (*molecular basis set for parameterization*, MBSP) is used to derive the parameters values

The items of data used are **heats of formation, ionization potential, dipole moments, molecular geometries, heats of reaction, rotational and inversion barriers** (all gas phase)

Extensive experience has shown that the set of parameters values is **extendible beyond the limit of those molecules** in the MBSP

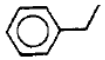
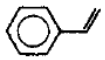
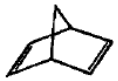
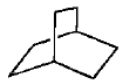
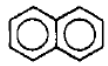
The different parameterizations depend on:

- The parameterization algorithm
- The molecules selected in the MBSP
- The targeted physico-chemical properties

MOLECULAR BASIS SET FOR PARAMETERIZATION

Parameterization algorithms optimize the parameters in such a way that the statistical error with respect to reference experimental data is minimized

Table II. Heats of Formation ΔH_f for Closed Shell Molecules

Molecule	ΔH_f, kcal/mol Calcd	Exptl	Error	Ref	Molecule	ΔH_f, kcal/mol Calcd	Exptl	Error	Ref
H_2	0.7	0.0	+0.7			8.6	7.2	+1.4	*a*
CH_4	−11.9	−17.9	+6.0	*a*					
C_2H_6	−19.7	−20.2	+0.5	*a*		37.5	35.3	+2.2	*a*
C_2H_4	15.3	12.5	+2.8	*a*					
C_2H_2	57.3	54.3	+3.0	*a*		33.6	43.9	−10.3	*a*
$CH_3CH_2CH_3$	−24.9	−24.8	−0.1	*a*					
$CH_3CH{=}CH_2$	4.9	4.9	0.0	*a*		69.1	51.9	+17.2	*a*
$CH_3C{\equiv}CH$	40.9	44.4	−3.5	*a*					
$CH_2{=}C{=}CH_2$	43.8	45.6	−1.8	*a*		33.7	44.3	−10.6	*a*
	−29.7	−30.4	+0.7	*a*					
	−26.8	−32.4	+5.6	*a*		28.6	30.9	−2.3	*a*
	0.3	−0.2	+0.5	*a*					
	−5.1	−3.0	−2.1	*a*		30.2	37.3	−7.1	*c*
	−4.0	−1.9	−2.1	*a*					
	−2.1	−4.3	+2.2	*a*		−10.3	−12.4	+2.1	*e*
	33.4	38.8	−5.4	*a*		62.6	59.7	+2.9	*c*
	28.9	26.0	+2.9	*a*					
	35.6	39.5	−3.9	*a*		−26.3	−24.1	−2.2	*e*
	24.4	34.7	−10.3	*a*					
	65.0	72.8	−7.8	*b*		38.1	36.1	+2.0	*a*
	102.2	113.0	−10.8	*b*					

M. J. S. Dewar and W. Thiel, *Journal of the American Chemical Society*, **1977**, *99*, 4907.

VALENCE-ELECTRON SEMI-EMPIRICAL MODELS

Common features of valence-electron semiempirical methods

MOs expansion

$$\phi_i = \sum_p C_{pi}\,\chi_p$$

*The MOs are linear combination of **valence** AOs*
*AOs are described using a **minimal basis set of Slater functions***

The core electrons are considered to be a part of the nuclei

ZDO (Zero Differential Overlap) approximation

The overlap integrals

$$S_{pq} = \int \chi_p(\nu)\chi_q(\nu)\mathrm{d}V_\nu$$

are neglected when solving the Fock equations

$$\mathbf{FC} = \varepsilon\mathbf{SC} \Rightarrow \mathbf{FC} = \varepsilon\mathbf{IC}$$

VALENCE-ELECTRON SEMI-EMPIRICAL MODELS

NDDO (Neglect of Diatomic Differential Overlap)

The most recent and complete level of semiempirical approximation

- 1- and 2-electron integrals are *parameterized*

$$h_{pq} = \int \chi_p(\nu)\hat{h}(\nu)\chi_q(\nu)\,d\mathbf{r}_\nu = (p\,|\,q)$$

$$(pq|rs) = \int \frac{\chi_p(\mathbf{r}_\mu)\chi_q(\mathbf{r}_\mu)\chi_r(\mathbf{r}_\nu)\chi_s(\mathbf{r}_\nu)}{r_{\mu\nu}}\,d\mathbf{r}_\mu\,d\mathbf{r}_\nu$$

- Most of 2-electron are neglected

one-center integrals $\left(p^A q^A \middle| r^A s^A\right)$

two-center integrals $\left(p^A q^A \middle| r^B s^B\right)$

three- and four-center integrals are neglected

Different methods have been developped on the basis of the NDDO approximation by using more and more precise parameterizations

MNDO (Dewar 1977), **AM1** (Dewar 1985), **PM3** (Stewart 1989), **PM6** (Stewart 2007)
PM7 (Stewart, under development)*… and many others!*

SEMI-EMPIRICAL PARAMETERS

MNDO and AM1 *differ only in the way the core-core repulsion is treated, and how the parameters are assigned*

MNDO

$$V_{AB} = Z_A^* Z_B^* (s^A s^A \mid s^B s^B)(1 + e^{-\alpha_A R_{AB}} + e^{-\alpha_B R_{AB}})$$

$$\text{if B=H: } V_{AH} = Z_A^* Z_H (s^A s^A \mid s^H s^H)\left(1 + \frac{e^{-\alpha_A R_{AH}}}{R_{AH}} + e^{-\alpha_H R_{AH}}\right)$$

Overstimation of the repulsion between nuclei at large distances

→ ***Weak interactions are unreliale (e.g. it does not predict hydrogen bonds)***

AM1

$$V_{AB} = Z_A^* Z_B^* (s^A s^A \mid s^B s^B)\left(1 + \exp(-\alpha_A R_{AB}) + \exp(-\alpha_B R_{AB})\right)$$

$$+ Z_A^* Z_B^* (s^A s^A \mid s^B s^B)\left(\sum_i K_{Ai} \exp\left(L_{Ai}(R_{AB} - M_{Bi}\right)^2 + \sum_j K_{Bj} \exp\left(L_{Bj}(R_{AB} - M_{Bj}\right)^2\right)$$

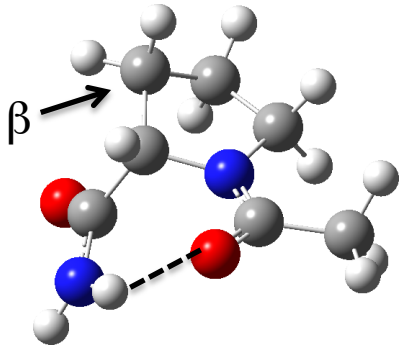
Use of "attractive" Gaussian functions at large distance

→ ***AM1 predicts hydrogen bonds with correct strengths***
(however the geometry can be wrong)

PM3: *Complete reparameterization of AM1 using an automatic process and a significantly larger MBSP*

PM6: *further modifications in core-core repulsion terms, and extension to transition metals (see J. Mol. Model. 2007, 1173.)*

A Representative Example: the N-Acetyl Proline Amide

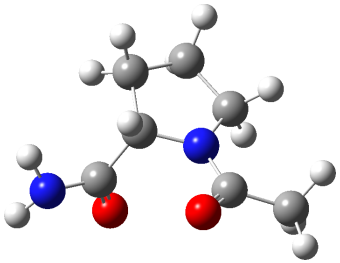


HF/6-31G(d,p)

H bond = 2.09 Å

Five membered proline ring in the envelope form, with atoms C-N-C-C approximately in one plane and the β carbon atom out of this plane (N-C-C-C_β = 29°)

Planar amide group

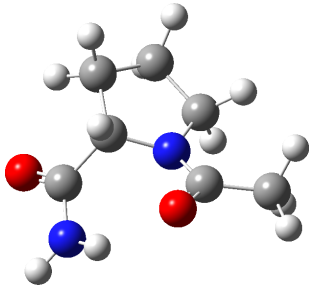


MNDO

✗ No H bond

✗ N-C-C-C_β = -1°

✗ Non-planar amide group

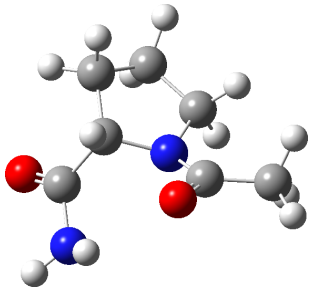


AM1

✓ H bond = 2.14 Å

✗ N-C-C-C_β = 1°

✓ Planar amide group

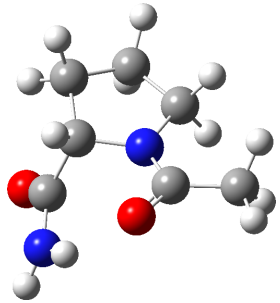


PM3

✓✗ H bond = 2.59 Å

✗ N-C-C-C_β = 0°

✗ Non-planar amide group



PM6

✓ H bond = 2.01 Å

✓ N-C-C-C_β = 11°

✓ Planar amide group

SEMIEMPIRICAL NDDO CALCULATION OF HEATS OF FORMATION

The Heats of Formation resulting from semiempirical calculations are calculated using empirical atomic Heats of Atomization

Example : $N + 3H \rightarrow NH_3$

1) The total AM1 energy of the system is first calculated. The **atomization energy** is then computed by subtracting the total energies of the atoms (as predicted by the model being used) in their stoichiometric ratios.

$$\Delta E^{atom}_{AM1} = \underbrace{E^{AM1}_{TOT}(NH_3)}_{\text{total AM1 energy}} - \underbrace{3E^{AM1}_{TOT}(H) - E^{AM1}_{TOT}(N)}_{\text{empirical parameters depending on the parametrization}}$$

($E^{AM1}(H) = -11.396$ eV !!)

2) The Heat of Formation ΔH_f of the system is then calculated using the **enthalpies of atomization of the atoms** and the atomization energy.

$$\Delta H^{atom}_{298K} = \underbrace{3 \times \Delta H^{atom}_{298}(H) + \Delta H^{atom}_{298}(N)}_{\text{experimental parameters not dependent on the parametrization}}$$

Then : $\Delta H_f = \Delta E^{atom}_{AM1} + \Delta H^{atom}_{298}$

Electron Density Within the NDDO Approximation

Mulliken AO populations

$$Q_p = \sum_i n_i \left(C_{pi}\right)^2$$

Occupation number of MO ϕ_i (n_i = 0, 1 or 2)

LCAO coefficient

Average number of electrons around atom A within the molecule:

$$n_A = \sum_{p \in A} Q_p = \sum_{p \in A} \sum_i n_i \left(C_{pi}\right)^2$$

Net atomic charges

$$\delta_A = Z_A - n_A$$

Number of valence electrons in atom A isolated

Electron Density Within the NDDO Approximation

Dipole moment

$$\vec{\mu}_0 = \left\langle \Psi_0 \left| \hat{\vec{\mu}} \right| \Psi_0 \right\rangle \qquad \hat{\vec{\mu}} = \sum_{i}^{\text{electrons}} -|e|\mathbf{r}_i + \sum_{A}^{\text{nuclei}} Z_A |e| \mathbf{R}_A$$

Rearranging the above expressions for a closed-shell system:

$$\vec{\mu} = 2\sum_{i}^{occ} \left\langle \phi_i \left| -e\mathbf{r} \right| \phi_i \right\rangle + \sum_{A}^{\text{nuclei}} eZ_A R_A$$

Developing the MOs over the AOs:

$$\vec{\mu} = -e\sum_{p,q}^{occ} D_{pq} \left\langle \chi_p \left| \mathbf{r} \right| \chi_q \right\rangle + e\sum_{A}^{\text{nuclei}} Z_A \mathbf{R}_A$$

with D_{pq} the elements of the first order density matrix in the AOs basis.

Dipole integrals in the NDDO approximation

$$\left\langle \chi_p^A \left| \mathbf{r} \right| \chi_q^B \right\rangle = 0 \qquad \left\langle \chi_p^A \left| \mathbf{r} \right| \chi_p^A \right\rangle = \mathbf{R}_A$$

The two-center dipole integrals are ignored. Only the integrals in which p and q are centered on the same atom are non zero

$$\left\langle \chi_p^A \left| \vec{r} \right| \chi_q^A \right\rangle = \frac{2n+1}{\sqrt{3}} \times \sqrt{2\alpha_p} \times \sqrt{2\alpha_q} \times \frac{\left(4\alpha_p\alpha_q\right)^n}{\left(\alpha_p + \alpha_q\right)^{2n+2}}$$

where n, α_p and α_q are the quantum number and the Slater exponents of the AOs χ_p and χ_q

Dipole moment expression

$$\vec{\mu} = \underbrace{e\sum_{A}^{\text{nuclei}} q_A R_A}_{\text{point charge contribution}} \underbrace{- 2e\sum_{A}^{\text{nuclei}} \sum_{p<q}^{p,q\in A} D_{pq} \left\langle \chi_p^A \left| \mathbf{r} \right| \chi_q^A \right\rangle}_{\text{hybrid contribution}}$$

with q_A the net charge of atom A

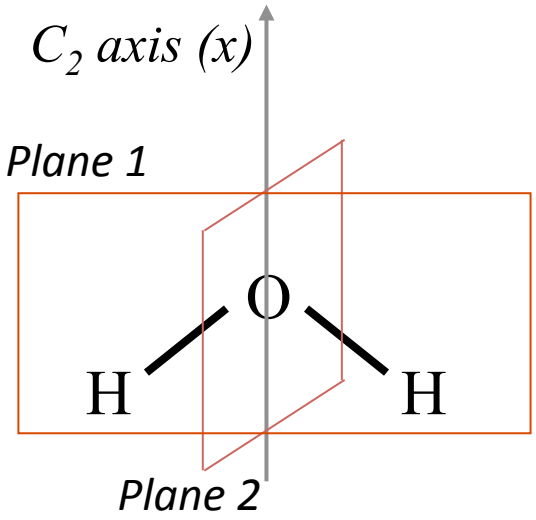
EXAMPLE OF NDDO CALCULATIONS: THE WATER MOLECULE

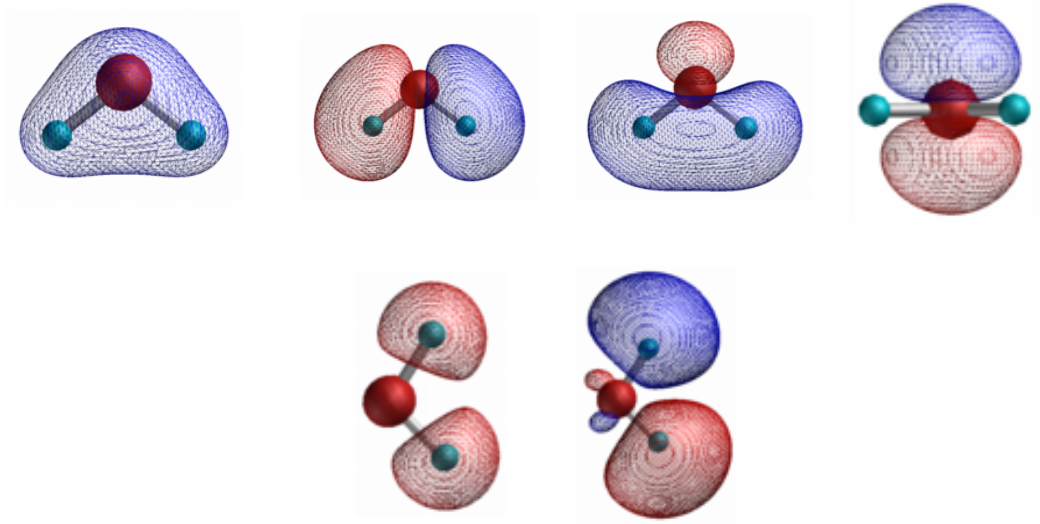
$$\phi_i = \sum_p C_{pi}\,\chi_p$$

O : $1s^2\,2s^2\,2p^4$: 4 AOs, 6 valence electrons
H : $1s^1$: 1 AO, 1 valence electron

AM1 RHF EIGENVECTORS AND EIGENVALUES

ROOT NO.		1	2	3	4	5	6
		-36.425	-18.199	-14.954	-12.464	4.419	6.191
		1 A1	**1 B2**	**2 A1**	**1 B1**	**3 A1**	**2 B2**
s O	1	.8955	.0000	.3599	.0000	-.2620	.0000
Px O	1	.0950	-.6049	-.4965	.0000	-.3573	-.5009
Py O	1	.1205	.4768	-.6299	.0000	-.4534	.3948
Pz O	1	.0000	.0000	.0000	1.0000	.0000	.0000
s H	2	.2955	-.4510	-.3371	.0000	.5469	.5446
s H	3	.2955	.4510	-.3371	.0000	.5469	-.5446





EXAMPLE OF NDDO CALCULATIONS: THE WATER MOLECULE

MNDO

NET ATOMIC CHARGES AND DIPOLE CONTRIBUTIONS

ATOM NO.	TYPE	CHARGE	ATOM ELECTRON DENSITY
1	O	-0.3255	6.3255
2	H	0.1628	0.8372
3	H	0.1628	0.8372

δ_A n_A

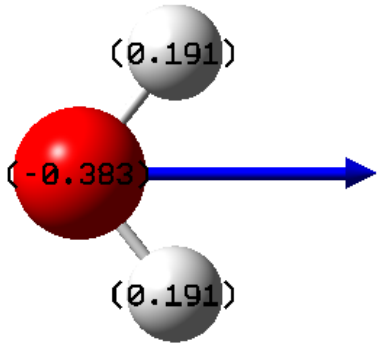
DIPOLE (DEBYE)	X	Y	Z	TOTAL
POINT-CHG.	0.524	0.706	0.000	0.879
HYBRID	0.539	0.725	0.000	0.904
SUM	1.063	1.431	0.000	1.783

AM1

NET ATOMIC CHARGES AND DIPOLE CONTRIBUTIONS

ATOM NO.	TYPE	CHARGE	ATOM ELECTRON DENSITY
1	O	-0.3826	6.3826
2	H	0.1913	0.8087
3	H	0.1913	0.8087

DIPOLE (DEBYE)	X	Y	Z	TOTAL
POINT-CHG.	0.677	0.859	0.000	1.094
HYBRID	0.475	0.602	0.000	0.767
SUM	1.152	1.461	0.000	1.861



Experimental dipole value: 1.847 D (*J. Chem. Phys*. **1973**, *59*, 2254)

How Accurate is a Semiempirical Calculation?

Overall performance of MNDO, AM1 and PM3

Reported average errors (From *Introduction to Computational Chemistry* - Frank Jensen)

Table 3.1 Average heat of formation error in kcal/mol (number of compounds)

Compounds:	MNDO	AM1	PM3
H, C, N, O (276)	18.5	10.5	7.9
F (133)	84.2	49.5	11.2
Si (78)	22.9	20.8	14.2
All normal valent (607)	24.3	14.8	11.2
Hypervalent (106)	104.5	62.3	17.3
All (713)	46.2	27.6	11.6

Table 3.2 Average errors in bond distances (Å)

Bonds to:	MNDO	AM1	PM3
H	0.015	0.006	0.005
C	0.002	0.002	0.002
N	0.015	0.014	0.012
O	0.017	0.011	0.006
F	0.023	0.017	0.011
Si	0.030	0.019	0.045

Angles (degrees)

MNDO	4.3
AM1	3.3
PM3	3.9

Ionization potentials (eV)

MNDO	0.78
AM1	0.61
PM3	0.57

Dipole moments (D)

MNDO	0.45
AM1	0.35
PM3	0.38

AM1 contains more adjustable parameters than MNDO
All parameters are fully optimized in PM3
→ **Errors decrease in the order MNDO > AM1 > PM3**

Caution! Considering average errors can be misleading.
For specific (classes of) compounds the ordering may be different.
A semiempirical method should always be validated before use.

Some General Conclusions on Semiempirical Methods

Semiempirical methods can provide results accurate enough to be useful, particularly for organic molecules with computation requirements low enough to make them convenient on personal computers.

These methods are generally good for predicting molecular geometry and energetics.

Semiempirical methods can be used for predicting vibrational modes and transition structures, but do so less reliably than ab initio methods.

Semiempirical calculations generally give poor results for van der Waals, dispersion intermolecular forces and polarisabilities, due to the lack of diffuse basis functions.

Semiempirical calculations generally give poor results for metal-containing compounds.

Computational Chemistry: A Practical Guide for Applying Techniques to Real-World Problems.
David C. Young, 2001, John Wiley & Sons, Inc.