Computational Material Science
Part I

Ito Chao (趙奕姼)
Institute of Chemistry
Academia Sinica

Molecular Orbital Theory

How to get the system energy?

Shrödinger equation: $\boldsymbol{H}\psi = \boldsymbol{E}\psi$

$$\int \psi H \psi d\tau = \int \psi E \psi d\tau$$

$$\Rightarrow \qquad E = \frac{\int \psi H \psi d\tau}{\int \psi \psi d\tau}$$

H: Hamiltonian operator ψ: wavefunction E: system energy

$$\boldsymbol{H}\psi = \boldsymbol{E}\psi$$

- What is in the Hamiltonian?

$$H = -\sum_{i}\frac{\hbar^2}{2m_e}\nabla_i^2 - \sum_{k}\frac{\hbar^2}{2m_k}\nabla_k^2 - \sum_{i}\sum_{k}\frac{e^2 Z_k}{r_{ik}} + \sum_{i<j}\frac{e^2}{r_{ij}} + \sum_{k<l}\frac{e^2 Z_k Z_l}{r_{kl}}$$

Operator for kinetic energy of electron

For kinetic energy of nucleus

For attraction of the electrons to the nuclei

For the interelectronic repulsions

For the internuclear repulsions

Under the ***Born-Oppenheimer approximation*** (electrons respond to the motion of nucleus instantaneously, so assume fixed nuclear geometry when electronic energy is evaluated), this term is zero!

➢ The variational principle

For any normalized , acceptable function Φ

$$\frac{\int \Phi H \Phi d\mathbf{r}}{\int \Phi^2 d\mathbf{r}} \geq E_0$$

where E_0 is the lowest eigenvalue of *H*

⇒ We can judge the quality of wave functions that we arbitrarily guess by their associated energies: the lower the better!

⇒ No need to guess from orthonormal functions!

- The proof of variational principle

Φ is represented in terms of the complete orthonormal set of eigenfunctions $\{\psi_i\}$ of *H*

$$\Phi = \sum_i c_i \Psi_i$$

$$\int \Psi_i \Psi_j d\mathbf{r} = \delta_{ij}$$

$$\int \Psi_j H \Psi_i d\mathbf{r} = \int \Psi_j E_i \Psi_i d\mathbf{r}$$

$$\int \Psi_j H \Psi_i d\mathbf{r} = E_i \delta_{ij}$$

$$\begin{aligned}
\int \Phi^2 d\mathbf{r} = 1 &= \int \sum_i c_i \Psi_i \sum_j c_j \Psi_j d\mathbf{r} \\
&= \sum_{ij} c_i c_j \int \Psi_i \Psi_j d\mathbf{r} \\
&= \sum_{ij} c_i c_j \delta_{ij} \\
&= \sum_i c_i^2
\end{aligned}$$

$$\begin{aligned}\int \Phi H \Phi d\mathbf{r} &= \int \left(\sum_i c_i \Psi_i\right) H \left(\sum_j c_j \Psi_j\right) d\mathbf{r} \\ &= \sum_{ij} c_i c_j \int \Psi_i H \Psi_j d\mathbf{r} \\ &= \sum_{ij} c_i c_j E_j \delta_{ij} \\ &= \sum_i c_i^2 E_i\end{aligned}$$

E_0 is the lowest eigenvalue of H

$$\int \Phi H \Phi d\mathbf{r} - E_0 \int \Phi^2 d\mathbf{r} = \sum_i \underset{>0}{c_i^2} \underset{\geq 0}{(E_i - E_0)}$$

$$\Rightarrow \quad \int \Phi H \Phi d\mathbf{r} - E_0 \int \Phi^2 d\mathbf{r} \geq 0$$

$$\Rightarrow \quad \frac{\int \Phi H \Phi d\mathbf{r}}{\int \Phi^2 d\mathbf{r}} \geq E_0$$

- Construction of trial wave functions : *In the molecular orbital approach, the wavefunction of an MO is the linear combination of atomic orbitals*

$$\phi = \sum_{i=1}^{N} a_i \varphi_i$$

ϕ is MO; φ is AO

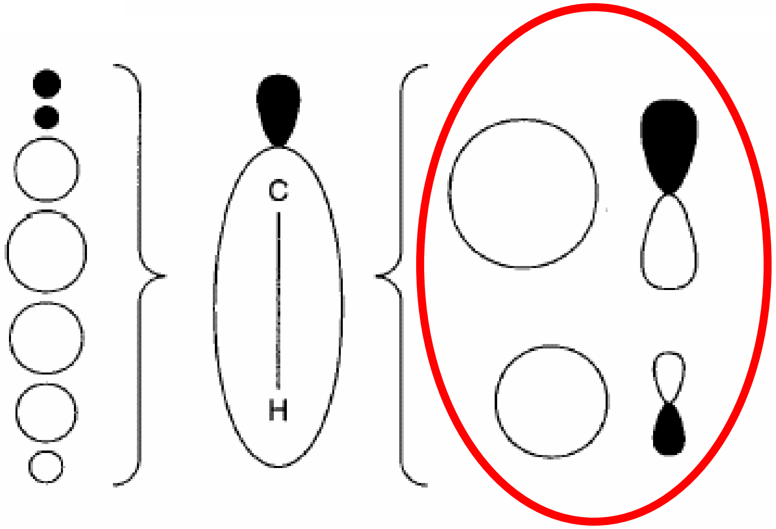


Figure 4.1 Two different basis sets for representing a C–H σ bonding orbital with the size of the basis functions roughly illustrating their weight in the hybrid MO. The set on the right is the more chemically intuitive since all basis functions are centered on the atoms. Note, however, that the use of a p function to polarize the hydrogen density goes beyond a purely minimalist approach. The set on the left is composed entirely of s functions distributed along the bond. Such a basis set may seem odd in concept, but is quite capable of accurately representing the electron density in space. Indeed, the basis set on the left would have certain computational advantages, chief among them the greater simplicity of working with s functions than with p functions

- *Let's use the multi atoms – one electron wavefunction to evaluate the energy => the* ***secular equation***

1.

$$E = \frac{\int \left(\sum_i a_i \varphi_i\right) H \left(\sum_j a_j \varphi_j\right) d\mathbf{r}}{\int \left(\sum_i a_i \varphi_i\right)\left(\sum_j a_j \varphi_j\right) d\mathbf{r}}$$

$$= \frac{\sum_{ij} a_i a_j \int \varphi_i H \varphi_j d\mathbf{r}}{\sum_{ij} a_i a_j \int \varphi_i \varphi_j d\mathbf{r}}$$

$$= \frac{\sum_{ij} a_i a_j H_{ij}}{\sum_{ij} a_i a_j S_{ij}}$$

2. Choose *a* to minimize *E*

$$\frac{\partial E}{\partial a_k} = 0 \qquad \forall k$$

$$\sum_{i=1}^{N} a_i (H_{ki} - E S_{ki}) = 0 \qquad \forall k$$

3. N equations involve N unknowns (*a*)
- has solution when determinant formed from coefficients of the unknowns (H – ES) = 0

$$\begin{vmatrix} H_{11} - ES_{11} & H_{12} - ES_{12} & \cdots & H_{1N} - ES_{1N} \\ H_{21} - ES_{21} & H_{22} - ES_{22} & \cdots & H_{2N} - ES_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ H_{N1} - ES_{N1} & H_{N2} - ES_{N2} & \cdots & H_{NN} - ES_{NN} \end{vmatrix} = 0$$

=> N roots for *E*, and each *E* will have a set of *a*

$$\phi = \sum_{i=1}^{N} a_i \varphi_i \quad \rightarrow \quad E = \frac{\int \left(\sum_i a_i \varphi_i\right) H \left(\sum_j a_j \varphi_j\right) d\mathbf{r}}{\int \left(\sum_i a_i \varphi_i\right) \left(\sum_j a_j \varphi_j\right) d\mathbf{r}}$$

$$\rightarrow \quad \frac{\partial E}{\partial a_k} = 0 \qquad \forall k \quad \rightarrow \quad \sum_{i=1}^{N} a_i (H_{ki} - ES_{ki}) = 0 \qquad \forall k$$

Eq (4.20)

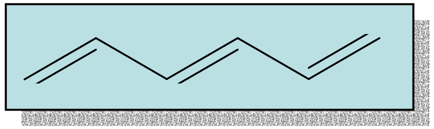
$$\rightarrow \quad \begin{vmatrix} H_{11} - ES_{11} & H_{12} - ES_{12} & \cdots & H_{1N} - ES_{1N} \\ H_{21} - ES_{21} & H_{22} - ES_{22} & \cdots & H_{2N} - ES_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ H_{N1} - ES_{N1} & H_{N2} - ES_{N2} & \cdots & H_{NN} - ES_{NN} \end{vmatrix} = 0$$

Summary

1. Select a set of N basis functions.
2. For that set of basis functions, determine all N^2 values of both H_{ij} and S_{ij}.
3. Form the secular determinant, and determine the N roots E_j of the secular equation.
4. For each of the N values of E_j, solve the set of linear Eqs. (4.20) in order to determine the basis set coefficients a_{ij} for that MO.

=> Orthogonal MOs determined

➢ **Hückel Theory** for conjugated systems

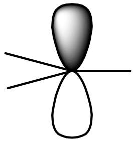


Principles:

1. Only p orbital for π bonds are considered (originally for carbon atoms)

2. Overlap matrix is defined by

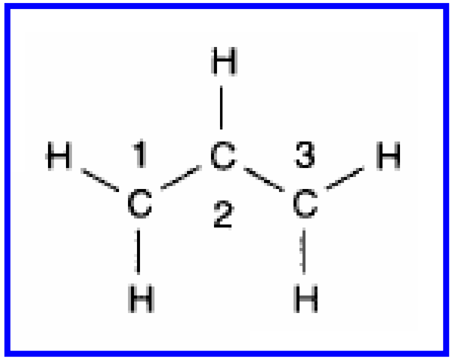
$$S_{ij} = \delta_{ij}$$ **=> meaning no overlap between neighboring atoms!!**

3. *Hii* = -IP of methyl radical (singly occupied 2p )
= α

4. *Hij* = β for neighboring atoms

5. *Hij* = 0 for nonneighboring atoms

Application to the allyl system:



$$\begin{vmatrix} H_{11} - ES_{11} & H_{12} - ES_{12} & \cdots & H_{1N} - ES_{1N} \\ H_{21} - ES_{21} & H_{22} - ES_{22} & \cdots & H_{2N} - ES_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ H_{N1} - ES_{N1} & H_{N2} - ES_{N2} & \cdots & H_{NN} - ES_{NN} \end{vmatrix} = 0$$

Sij = δ *ij*

Hii = α

Hij = β for neighboring atoms

Hij = 0 for nonneighboring atoms

$$\begin{vmatrix} \alpha - E & \beta & 0 \\ \beta & \alpha - E & \beta \\ 0 & \beta & \alpha - E \end{vmatrix} = 0$$

$$(\alpha - E)^3 + (\beta^2 \cdot 0) + (0 \cdot \beta^2) - [0 \cdot (\alpha - E) \cdot 0] - \beta^2(\alpha - E) - (\alpha - E)\beta^2 = 0$$

$$E = \alpha + \sqrt{2}\beta, \quad \alpha, \quad \alpha - \sqrt{2}\beta$$

$$\sum_{i=1}^{N} a_i (H_{ki} - E S_{ki}) = 0 \qquad \forall k$$

$\Downarrow$ Plug in solved E

$$a_1[\alpha - (\alpha + \sqrt{2}\beta) \cdot 1] + a_2[\beta - (\alpha + \sqrt{2}\beta) \cdot 0] + a_3[0 - (\alpha + \sqrt{2}\beta) \cdot 0] = 0$$

$$a_1[\beta - (\alpha + \sqrt{2}\beta) \cdot 0] + a_2[\alpha - (\alpha + \sqrt{2}\beta) \cdot 1] + a_3[\beta - (\alpha + \sqrt{2}\beta) \cdot 0] = 0$$

$$a_1[0 - (\alpha + \sqrt{2}\beta) \cdot 0] + a_2[\beta - (\alpha + \sqrt{2}\beta) \cdot 0] + a_3[\alpha - (\alpha + \sqrt{2}\beta) \cdot 1] = 0$$

$\Downarrow$

$$a_2 = \sqrt{2} a_1$$

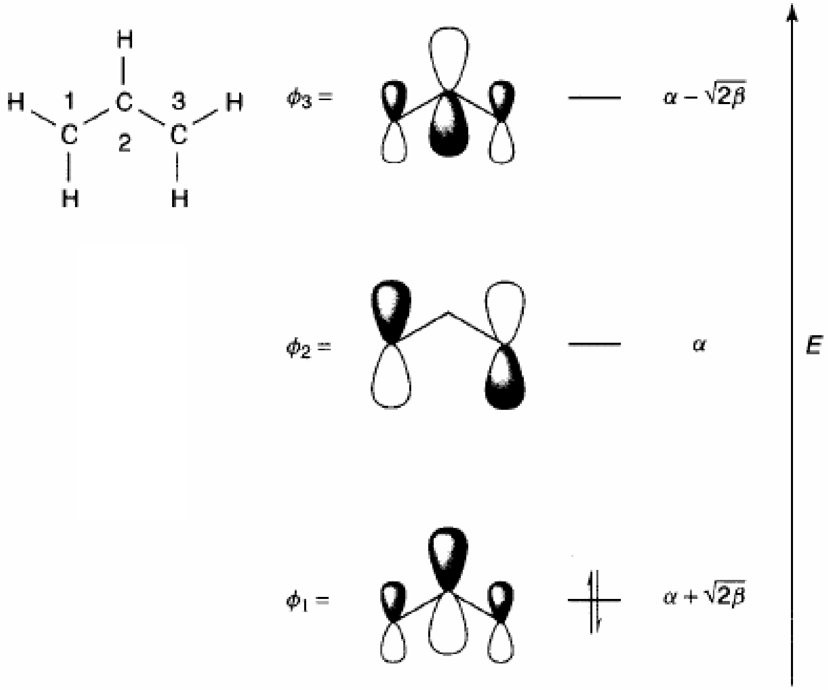
$$a_3 = a_1$$

$\Downarrow$ After normalization

$$a_{11} = \frac{1}{2}, \quad a_{21} = \frac{\sqrt{2}}{2}, \quad a_{31} = \frac{1}{2}$$

$\Downarrow$ Get wavefunction

$$\varphi_1 = \frac{1}{2} \mathrm{p}_1 + \frac{\sqrt{2}}{2} \mathrm{p}_2 + \frac{1}{2} \mathrm{p}_3$$



$$a_{13} = \frac{1}{2}, \quad a_{23} = -\frac{\sqrt{2}}{2}, \quad a_{33} = \frac{1}{2}$$

$$a_{12} = \frac{\sqrt{2}}{2}, \quad a_{22} = 0, \quad a_{32} = -\frac{\sqrt{2}}{2}$$

$$a_{11} = \frac{1}{2}, \quad a_{21} = \frac{\sqrt{2}}{2}, \quad a_{31} = \frac{1}{2}$$

Figure 4.2 Hückel MOs for the allyl system. One p_C orbital per atom defines the basis set. Combinations of these 3 AOs create the 3 MOs shown. The electron occupation illustrated corresponds to the allyl cation. One additional electron in ϕ_2 would correspond to the allyl radical, and a second (spin-paired) electron in ϕ_2 would correspond to the allyl anion

In the literature, coefficients are often labeled as c, rather than a.
You will see charge (q) and bond order (p) are represented as followed:

$$q_i = \sum_{k}^{\text{all MOs}} n_k c_{ik}^2 \qquad p_{ij} = \sum_{k}^{\text{all MOs}} n_k c_{ik} c_{jk}$$

Correlation between bond order and bond length

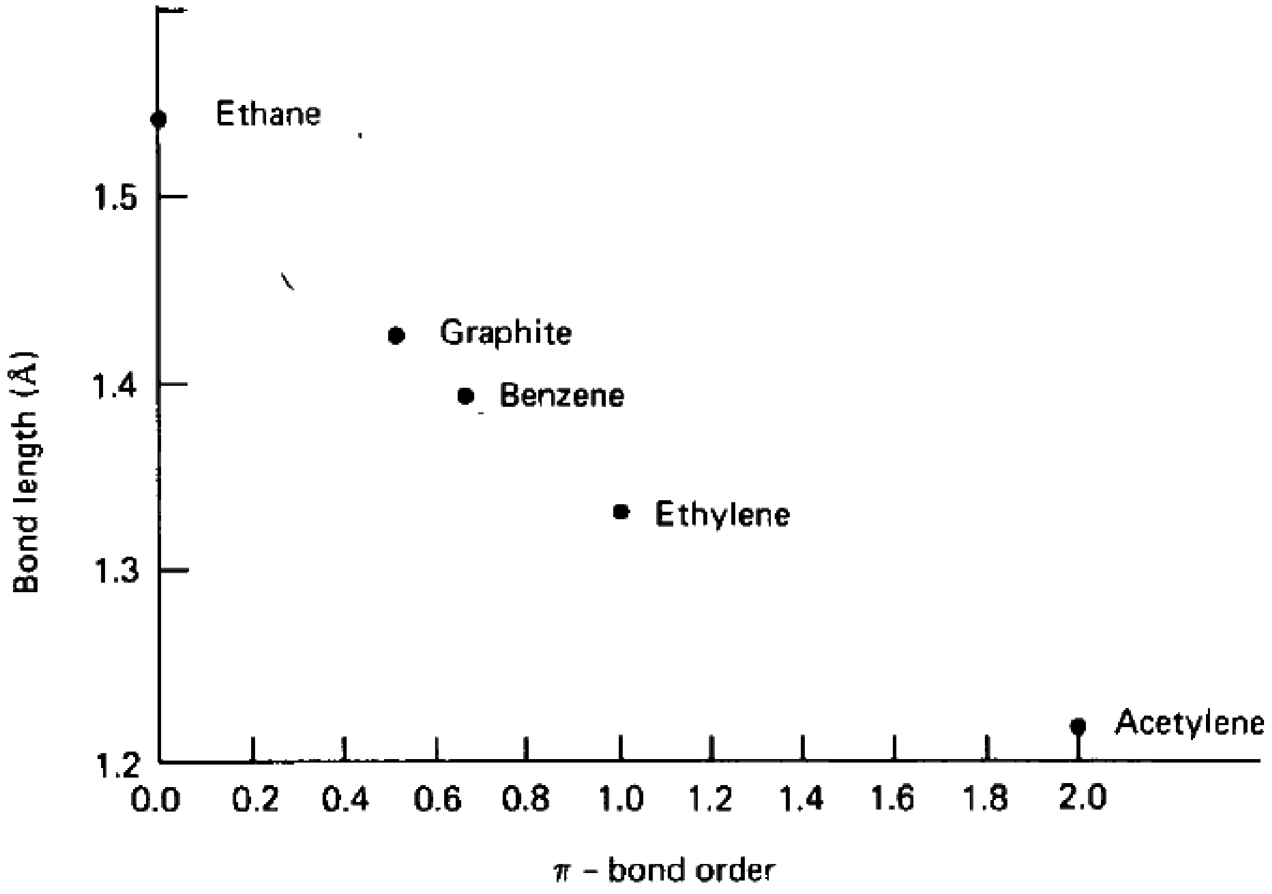
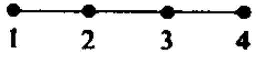
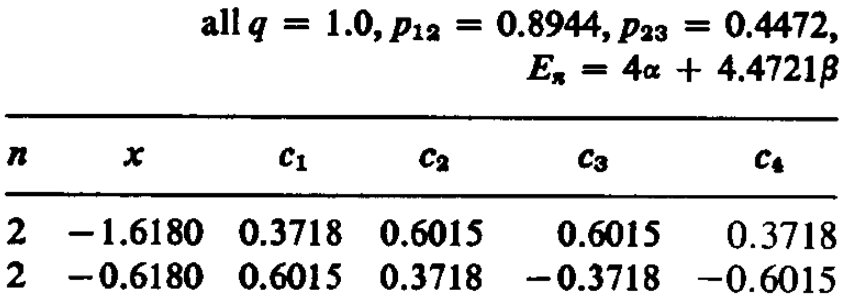


FIG. 8-14 π-Bond order versus bond length for some simple unsaturated hydrocarbons.

Bond order and wavefunction of butadiene

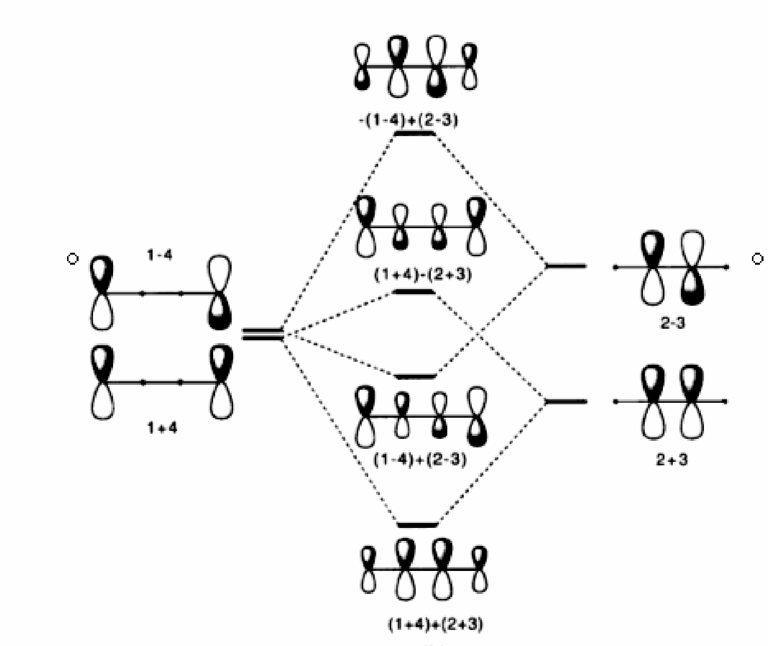
Butadiene (alternant) C_4H_6





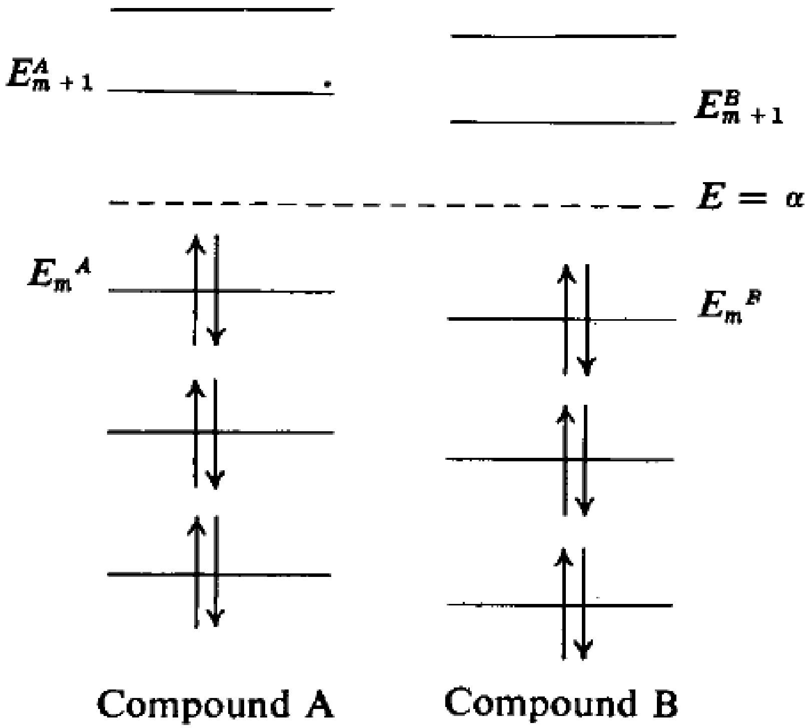
all $q = 1.0$, $p_{12} = 0.8944$, $p_{23} = 0.4472$,
$E_\pi = 4\alpha + 4.4721\beta$

n	x	c_1	c_2	c_3	c_4
2	−1.6180	0.3718	0.6015	0.6015	0.3718
2	−0.6180	0.6015	0.3718	−0.3718	−0.6015



Remember we talked about the shape of the wavefunctions and the smaller bond order between carbon 2 and 3?

Correlation between orbital energy and oxidation potential



(XI)

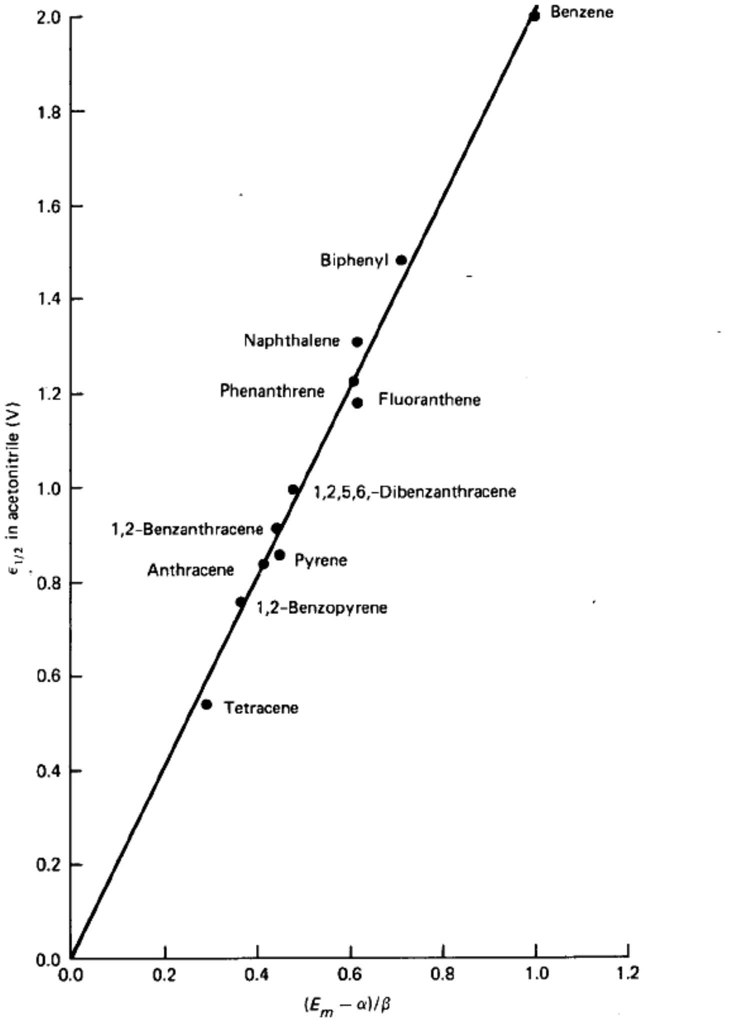


FIG. 8-20 Oxidation potentials in acetonitrile solution versus energy of HOMO (in units of β). ($\epsilon_{1/2}$ from Lund [10].)

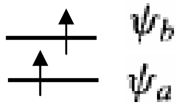
- Many-electron wave functions => *the **Slater Determinants***

The wavefunction has to

1. contain spin information
2. change sign when two electrons are interchanged
   (Pauli principle: the total wavefunction must be antisymmetric with respect to the interchange of any pair of electrons)

$$\begin{aligned} P_{ij}\Psi[\mathbf{q}_1(1), \ldots, \mathbf{q}_i(i), \ldots, \mathbf{q}_j(j), \ldots, \mathbf{q}_N(N)] \\ = \Psi[\mathbf{q}_1(1), \ldots, \mathbf{q}_j(i), \ldots, \mathbf{q}_i(j), \ldots, \mathbf{q}_N(N)] \\ = -\Psi[\mathbf{q}_1(1), \ldots, \mathbf{q}_i(i), \ldots, \mathbf{q}_j(j), \ldots, \mathbf{q}_N(N)] \end{aligned}$$

For



$$^3\Psi_{\mathrm{HP}} = \psi_a(1)\alpha(1)\psi_b(2)\alpha(2)$$

$$\begin{aligned} P_{12}[\psi_a(1)\alpha(1)\psi_b(2)\alpha(2)] &= \psi_b(1)\alpha(1)\psi_a(2)\alpha(2) \\ &\neq -\psi_a(1)\alpha(1)\psi_b(2)\alpha(2) \end{aligned}$$

DOES NOT FULFILL REQUIREMENTS!

A slight variation will fulfill the requirements….

$$ {}^{3}\Psi_{\mathrm{SD}} = \frac{1}{\sqrt{2}}[\psi_a(1)\alpha(1)\psi_b(2)\alpha(2) - \psi_a(2)\alpha(2)\psi_b(1)\alpha(1)] $$

In another equivalent form:

$$ {}^{3}\Psi_{\mathrm{SD}} = \frac{1}{\sqrt{2}}\begin{vmatrix} \psi_a(1)\alpha(1) & \psi_b(1)\alpha(1) \\ \psi_a(2)\alpha(2) & \psi_b(2)\alpha(2) \end{vmatrix} $$

Pauli exclusion principle: no two electrons can have the same set of quantum numbers. Slater determinant fulfill the principle: when electrons 1 and 2 are both in the same orbital ($\Psi_a\alpha$), Slater determinant = 0.

For the general case (use χ to express both the spatial and spin parts of the equation)

$$\Psi_{\mathrm{SD}} = \frac{1}{\sqrt{N!}}\begin{vmatrix} \chi_1(1) & \chi_2(1) & \cdots & \chi_N(1) \\ \chi_1(2) & \chi_2(2) & \cdots & \chi_N(2) \\ \vdots & \vdots & \ddots & \vdots \\ \chi_1(N) & \chi_2(N) & \cdots & \chi_N(N) \end{vmatrix}$$

For abbreviation:

$$\Psi_{\mathrm{SD}} = |\chi_1\chi_2\chi_3\cdots\chi_N\rangle$$

$$\Psi_{\mathrm{SD}} = |\psi_1^2\chi_3\cdots\chi_N\rangle$$

When α and β are in the same MO

The build-in electron correlation for like spin in Slater determinants

-- Fermi hole (like spin will avoid each other)

Consider two electron integral of like spin: (α and α)

$$
\begin{aligned}
&\int {}^{3}\Psi_{\mathrm{SD}} \frac{1}{r_{12}} {}^{3}\Psi_{\mathrm{SD}} d\mathbf{r}_1 d\omega_1 d\mathbf{r}_2 d\omega_2 \\
&= \frac{1}{2}\Big[\int |\psi_a(1)|^2 |\alpha(1)|^2 \frac{1}{r_{12}} |\psi_b(2)|^2 |\alpha(2)|^2 d\mathbf{r}_1 d\omega_1 d\mathbf{r}_2 d\omega_2 \\
&\quad - 2\int \psi_a(1)\psi_b(1) |\alpha(1)|^2 \frac{1}{r_{12}} \psi_b(2)\psi_a(2) |\alpha(2)|^2 d\mathbf{r}_1 d\omega_1 d\mathbf{r}_2 d\omega_2 \\
&\quad + \int |\psi_a(2)|^2 |\alpha(2)|^2 \frac{1}{r_{12}} |\psi_b(1)|^2 |\alpha(1)|^2 d\mathbf{r}_1 d\omega_1 d\mathbf{r}_2 d\omega_2 \Big] \\
&= \frac{1}{2}\Big[\underline{\int |\psi_a(1)|^2 \frac{1}{r_{12}} |\psi_b(2)|^2 d\mathbf{r}_1 d\mathbf{r}_2} \\
&\quad - 2\underline{\int \psi_a(1)\psi_b(1) \frac{1}{r_{12}} \psi_b(2)\psi_a(2) d\mathbf{r}_1 d\mathbf{r}_2} \\
&\quad + \underline{\int |\psi_a(2)|^2 \frac{1}{r_{12}} |\psi_b(1)|^2 d\mathbf{r}_1 d\mathbf{r}_2}\Big] \\
&= \frac{1}{2}\left(J_{ab} - 2\int \psi_a(1)\psi_b(1) \frac{1}{r_{12}} \psi_a(2)\psi_b(2) d\mathbf{r}_1 d\mathbf{r}_2 + J_{ab}\right) \\
&= J_{ab} - \underline{K_{ab}}
\end{aligned}
$$

***J*: Coulomb integral**

***K*: exchange integral**

Consider two electron integral of unlike spin: (α and β)

$$\Psi_{\mathrm{SD}} = \frac{1}{\sqrt{2}}[\psi_a(1)\alpha(1)\psi_b(2)\beta(2) - \psi_a(2)\alpha(2)\psi_b(1)\beta(1)]$$

$$\begin{aligned}
&\int \Psi_{\mathrm{SD}} \frac{1}{r_{12}} \Psi_{\mathrm{SD}} d\mathbf{r}_1 d\omega_1 d\mathbf{r}_2 d\omega_2 \\
&= \frac{1}{2}\Bigg[\int |\psi_a(1)|^2 \, |\alpha(1)|^2 \, \frac{1}{r_{12}} \, |\psi_b(2)|^2 \, |\beta(2)|^2 \, d\mathbf{r}_1 d\omega_1 d\mathbf{r}_2 d\omega_2 \\
&\quad - 2\int \psi_a(1)\psi_b(1)\alpha(1)\beta(1)\frac{1}{r_{12}}\psi_b(2)\psi_a(2)\alpha(2)\beta(2) d\mathbf{r}_1 d\omega_1 d\mathbf{r}_2 d\omega_2 \\
&\quad + \int |\psi_a(2)|^2 \, |\alpha(2)|^2 \, \frac{1}{r_{12}} \, |\psi_b(1)|^2 \, |\beta(1)|^2 \, d\mathbf{r}_1 d\omega_1 d\mathbf{r}_2 d\omega_2\Bigg] \\
&= \frac{1}{2}\Bigg[\int |\psi_a(1)|^2 \, \frac{1}{r_{12}} \, |\psi_b(2)|^2 \, d\mathbf{r}_1 d\mathbf{r}_2 \\
&\quad - 2 \cdot 0 \\
&\quad + \int |\psi_a(2)|^2 \, \frac{1}{r_{12}} \, |\psi_b(1)|^2 \, d\mathbf{r}_1 d\mathbf{r}_2\Bigg] \\
&= \frac{1}{2}(J_{ab} + J_{ab}) \\
&= J_{ab}
\end{aligned}$$

Only Coulomb integral left…

➢ The Hartree-Fock (HF) self-consistent field method

HF MOs are individually determined as eigenfunctions of a set of one-electron operators (each electron interacts with the static field of all of the other electrons)

1. Use Slater determinantal wave function
2. The one-electron Fock operator for RHF

$$f_i = -\frac{1}{2}\nabla_i^2 - \sum_k^{\text{nuclei}} \frac{Z_k}{r_{ik}} + V_i^{\text{HF}}\{j\}$$

3. Solve for various roots of E from secular equation

$$\begin{vmatrix} F_{11} - ES_{11} & F_{12} - ES_{12} & \cdots & F_{1N} - ES_{1N} \\ F_{21} - ES_{21} & F_{22} - ES_{22} & \cdots & F_{2N} - ES_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ F_{N1} - ES_{N1} & F_{N2} - ES_{N2} & \cdots & F_{NN} - ES_{NN} \end{vmatrix} = 0$$

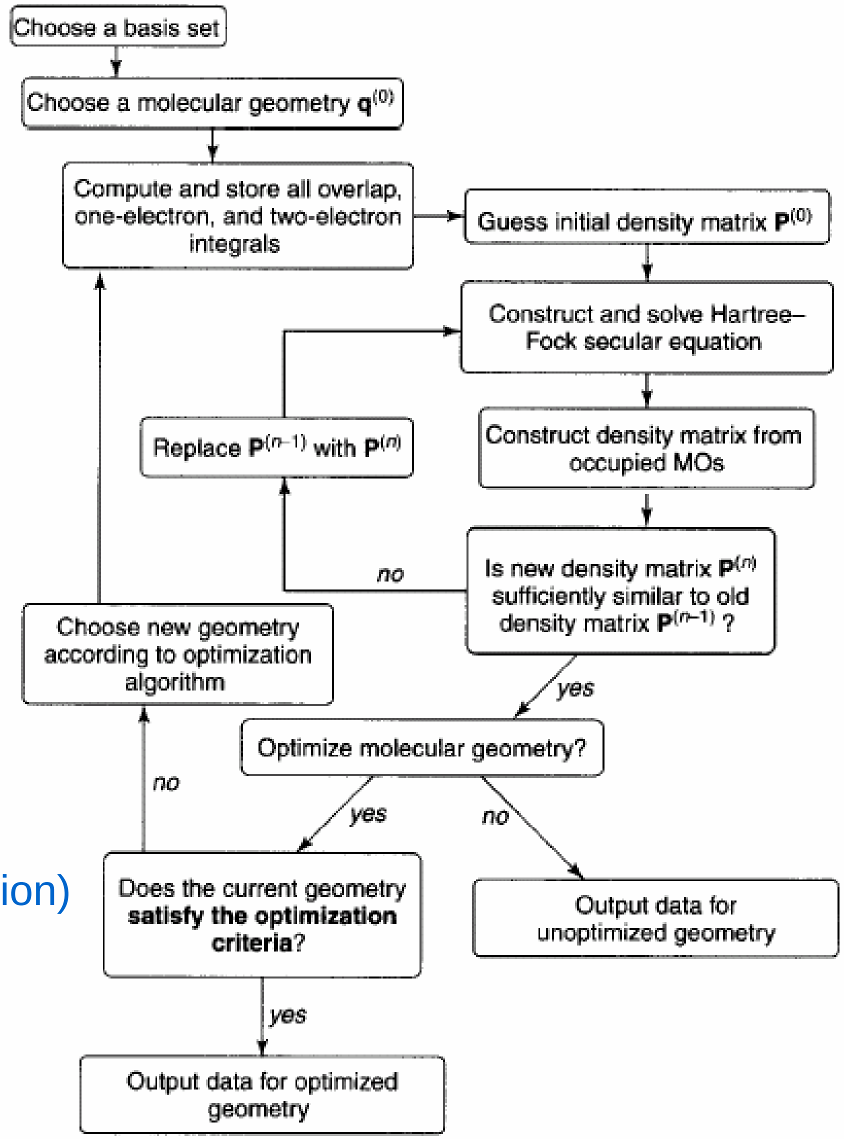
$$F_{\mu\nu} = \left\langle \mu \left| -\frac{1}{2}\nabla^2 \right| \nu \right\rangle - \sum_k^{\text{nuclei}} Z_k \left\langle \mu \left| \frac{1}{r_k} \right| \nu \right\rangle + \sum_{\lambda\sigma} P_{\lambda\sigma} \left[(\mu\nu|\lambda\sigma) - \frac{1}{2}(\mu\lambda|\nu\sigma) \right]$$

$$\langle \mu | g | \nu \rangle = \int \phi_\mu (g \phi_\nu) d\mathbf{r}.$$

$$(\mu\nu|\lambda\sigma) = \iint \phi_\mu(1)\phi_\nu(1)\frac{1}{r_{12}}\phi_\lambda(2)\phi_\sigma(2)d\mathbf{r}(1)d\mathbf{r}(2)$$

$$P_{\lambda\sigma} = 2\sum_i^{\text{occupied}} a_{\lambda i}a_{\sigma i}$$

* remember electrons are occupying molecular orbitals and each MO is a linear combination of atomic orbitals…
μ, ν represent different atoms; for electron 1
λ, σ represent different atoms; for electron 2



(Optimization calculation)

(The so-called single point calculation)

Figure 4.3 Flow chart of the HF SCF procedure. Note that data for an unoptimized geometry is referred to as deriving from a so-called 'single-point calculation'

- Basis set

Ψ = Slater determinant formed from the individual occupied MO; each MO is expressed as a linear combination of basis functions

For the general case (use χ to express both the spatial and spin parts of the equation)

$$\Psi_{\mathrm{SD}} = \frac{1}{\sqrt{N!}} \begin{vmatrix} \chi_1(1) & \chi_2(1) & \cdots & \chi_N(1) \\ \chi_1(2) & \chi_2(2) & \cdots & \chi_N(2) \\ \vdots & \vdots & \ddots & \vdots \\ \chi_1(N) & \chi_2(N) & \cdots & \chi_N(N) \end{vmatrix}$$

PREREQUISITE of basis set: wavefunction should have large amplitude in regions of space where the electron probability density is large…use atomic orbital?

- **Slater-type orbitals (STO) and Gaussian-type orbitals (GTO)**

$$\phi(x, y, z; \alpha, i, j, k) = \left(\frac{2\alpha}{\pi}\right)^{3/4}\left[\frac{(8\alpha)^{i+j+k}\, i!\, j!\, k!}{(2i)!(2j)!(2k)!}\right]^{1/2} x^i y^j z^k e^{-\alpha(x^2+y^2+z^2)}$$

$$\varphi(r, \theta, \phi; \zeta, n, l, m) = \frac{(2\zeta)^{n+1/2}}{[(2n)!]^{1/2}} r^{n-1} e^{-\zeta r} Y_l^m(\theta, \phi)$$

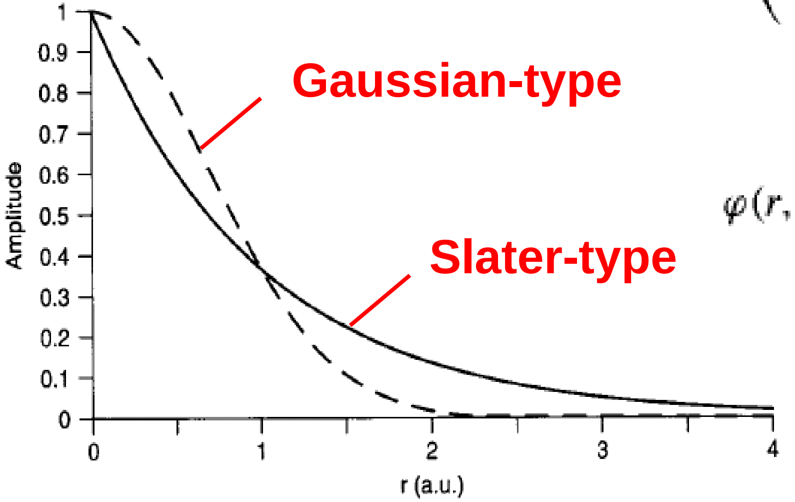


Figure 6.1 Behavior of e^x where $x = r$ (solid line, STO) and $x = r^2$ (dashed line, GTO)

- Slater-type orbital (e^{-r}) has no analytical solution for four index integral
- Boys proposed Gaussian-type AO-like orbitals

$i + j + k = 0$ => s
$i = 1$ => p_x; $j = 1$ => p_y; $k = 1$ => p_z
$i + j + k = 2$ => d Cartesian d functions: x^2, y^2, z^2, xy, xz, yz
Canonical d functions: xy, xz, yz, x^2-y^2, $3z^2 - r^2$

- **Problems of GTO:**

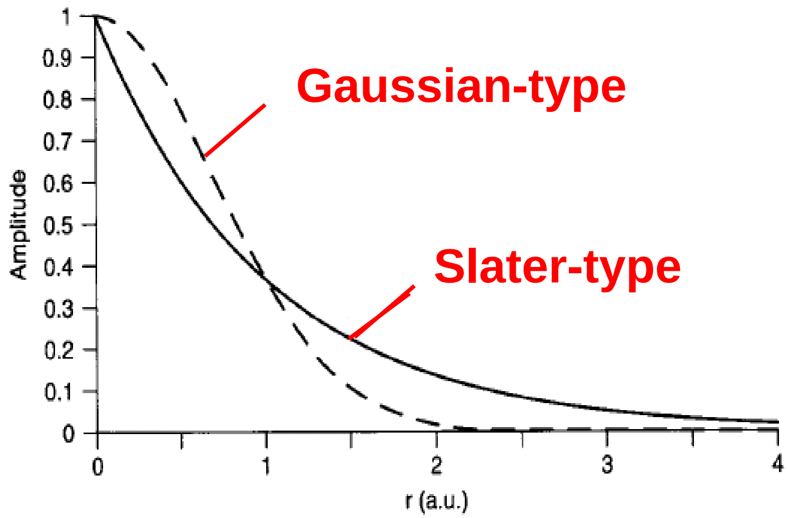


Figure 6.1 Behavior of e^x where $x = r$ (solid line, STO) and $x = r^2$ (dashed line, GTO)

1) No cusp at atomic center,
2) Too rapid a reduction in
   amplitude with distance
3) No radial nodal behavior

Solution: linear combination of GTOs to reproduce STOs

number of gaussian functions

exponent decides width

$$\underbrace{\varphi(x, y, z; \{\alpha\}, i, j, k)}_{\text{Basis function}} = \sum_{a=1}^{M} c_a \underbrace{\phi(x, y, z; \alpha_a, i, j, k)}_{\text{Gaussian function}}$$

chosen to optimize shape and ensure normalization

- **STO-*MG*** (Slater-Type Orbital approximated by *M* Gaussians)

$$\varphi(x,y,z;\{\alpha\},i,j,k)=\sum_{a=1}^{M} c_a \underbrace{\phi(x,y,z;\alpha_a,i,j,k)}_{\text{primitive gaussian}}$$

In a contracted basis function, c_a are fixed.

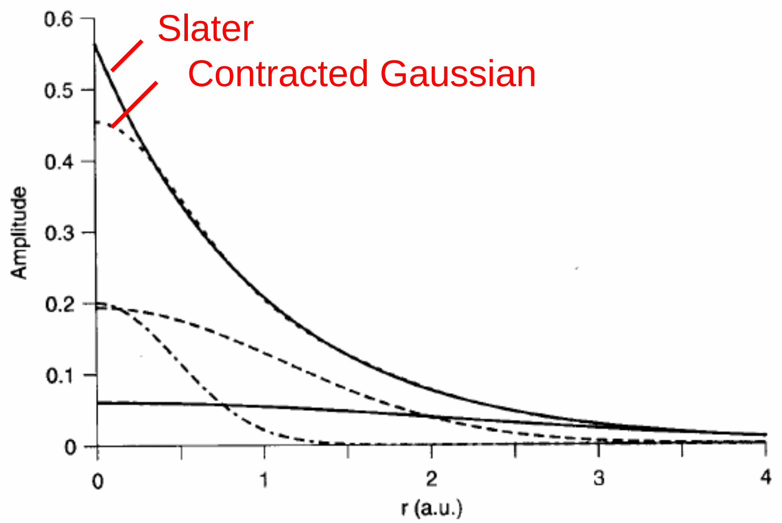


Figure 6.2 The radial behavior of various basis functions in atom-centered coordinates. The bold solid line at top is the STO ($\zeta = 1$) for the hydrogen 1s function; for the one-electron H system, it is also the exact solution of the Schrödinger equation. Nearest it is the contracted STO-3G 1s function (- - - - - -) optimized to match the STO. It is the sum of a set of one each tight (-· -· -· -·), medium (– – –), and loose (——) Gaussian functions shown below. The respective Gaussian primitive exponents α are 2.227660, 0.405771, and 0.109818, and the associated contraction coefficients c are 0.154329, 0.535328, and 0.444635. Note that from 0.5 to 4.0 a.u., the STO-3G orbital matches the correct orbital closely. However, near the origin there is a notable difference and, were the plot to extend to very large r, it would be apparent that the decay of the STO-3G orbital is more rapid than the correct orbital, in analogy to Figure 6.1

STO-3G more often used than other STO-*MG*

Notation for contraction:
(6s3p/3s)/[2s1p/1s]

H atom

1st row atom

( ) shows the number of primitive function; [ ] shows the number of contracted function.

A 1st row atom has 1s 2s 2p orbitals 3G for each type, finally 2 sets of S orbital and 1 set of p orbital contracted

- Single-ζ, Multiple-ζ, and Split-Valence basis sets

  - STO-3G is single-ζ : only one basis function for each AO

    Exponent for s and p are the same to save computing time (called sp basis function). Advantage: (ss|ss) = (ss|sp) = (ss|pp) = (sp|sp) for the radial part.

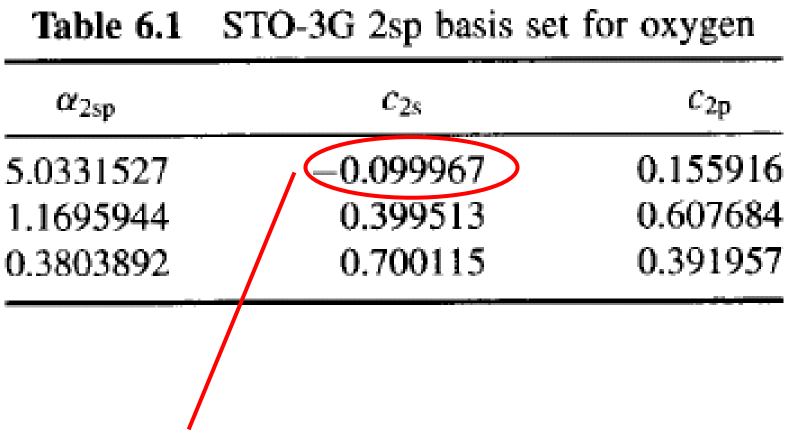


Table 6.1 STO-3G 2sp basis set for oxygen

α_{2sp}	c_{2s}	c_{2p}
5.0331527	−0.099967	0.155916
1.1695944	0.399513	0.607684
0.3803892	0.700115	0.391957

Negative sign to get nodal characteristics

Disadvantage of STO-3G: not flexible enough to represent electron population in molecules

- **Multiple-ζ** : “decontract” to increase flexibility of wavefunction

Ex: double-ζ : 2 functions for each AO
triple-ζ: 3 functions for each AO

Compare STO-3G and double-ζ with three gaussian functions

STO-3G: 3 primitive Gaussian contracted as 1 function
double-ζ: 2 functions--the first function being a contraction of 2 primitive
the second function being normalized third primitive

=> more coefficients have to be determined;
size of secular equation larger for double-ζ

$$\phi = \sum_{i=1}^{N} a_i \varphi_i \;\Rightarrow\; E = \frac{\int \left(\sum_i a_i \varphi_i\right) H \left(\sum_j a_j \varphi_j\right) d\mathbf{r}}{\int \left(\sum_i a_i \varphi_i\right) \left(\sum_j a_j \varphi_j\right) d\mathbf{r}} \;\Rightarrow\; \frac{\partial E}{\partial a_k} = 0 \qquad \forall k$$

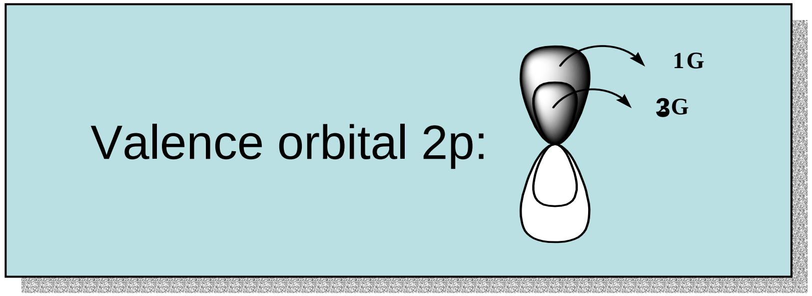
$$\Rightarrow\; \sum_{i=1} a_i (H_{ki} - E S_{ki}) = 0 \qquad \forall k$$

$$\Rightarrow\; \begin{vmatrix} H_{11} - ES_{11} & H_{12} - ES_{12} & \cdots & H_{1N} - ES_{1N} \\ H_{21} - ES_{21} & H_{22} - ES_{22} & \cdots & H_{2N} - ES_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ H_{N1} - ES_{N1} & H_{N2} - ES_{N2} & \cdots & H_{NN} - ES_{NN} \end{vmatrix} = 0$$

- **Split-valence** (valence multiple-ζ) : “decontract” to increase flexibility of the valence electrons (core electrons are less involved in bonding, still use contracted functions)

  Ex: 6-31G (by Pople)

  6 primitive Gaussian functions for contracted core function
  3 primitive Gaussian functions for contracted valence function
  1 primitive for uncontracted valence function



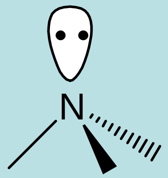
Valence orbital 2p:

For O atom: 1s 2s $2p_x$ $2p_y$ $2p_z$

STO-3G 1 + 1 + 1 + 1 + 1 = 5 (number of basis function)
6-31G 1 + 2 + 2 + 2 + 2 = 9
6-311G 1 + 3 + 3 + 3 + 3 = 13

- **Polarization Functions and Diffusion Functions**

Inversion barrier of pyramidal NH_3
5.8 kcal/mol



HF calc. with atomic centered basis set of s and p functions predicts planar geometry! *Not enough mathematical flexibility for molecules.*

=> *adding basis functions not centered on atomic center (**floating gaussian orbitals FLOGO**)* -less used

=> *adding basis functions corresponding to one quantum number of higher angular momentum than the valence orbitals (**polarization functions**)*

First row atom C, N, O, ….: add d orbitals
H: add p orbitals

Pyramidal geometry obtained from HF calc. Although some error in energy due to lack of electron correlation

Ex: d orbital for better bonding of O

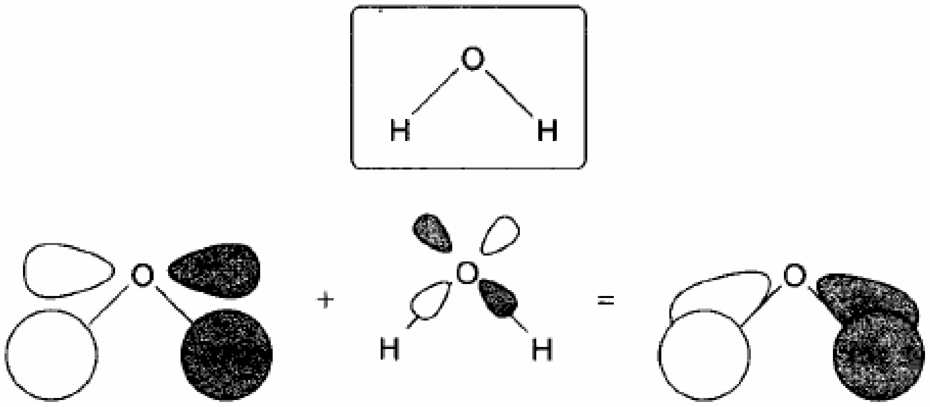


Figure 6.3 The MO formed by interaction between the antisymmetric combination of H 1s orbitals and the oxygen p_x orbital (see also Figure 6.7). Bonding interactions are enhanced by mixing a small amount of O d_{xz} character into the MO

Ex: d orbitals needed to describe hypervalent bonding situations in phophates, sulfoxides…

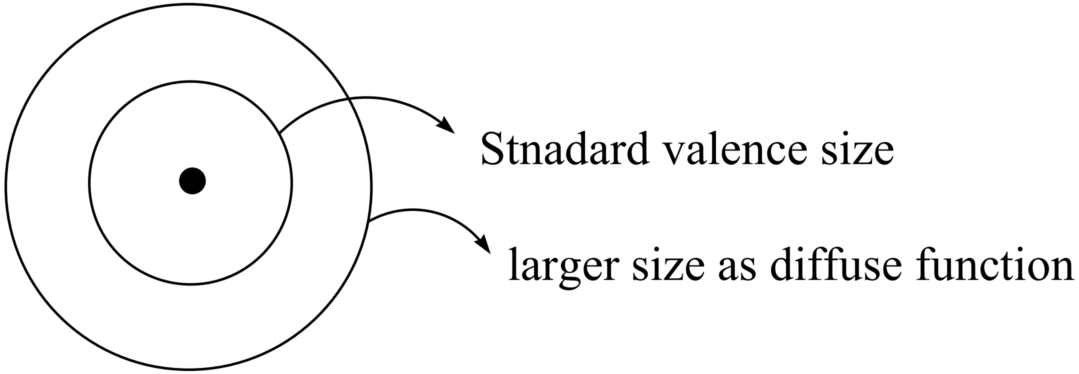
Highest energy MOs of anion, highly excited states, loose supermolecular complexes

=> Spatially diffuse!

OH^-

=> *adding basis functions that are more diffuse* (***diffuse functions***)

Ex: additional one s and one set of p functions having small exponents



- **Pople's nomenclature**

Ex: 6-31+G* or 6-31+G(d)

*polarization functions on nonhydrogen (**heavy**) atom*

*diffuse function on nonhydrogen (**heavy**) atom*

Ex: 6-31++G** or 6-31+G(d,p)

*polarization functions on **hydrogen** atom*

*second + means diffuse function on **hydrogen** atom*

Ex: 6-311++G(3df,2pd)

2 sets of p functions and 1 set of d function

3 sets of d functions, 1 set of f functions on heavy atom

Ex: 3-21G(*)

** in parenthesis means no d functions for 1st row atoms (C, N, O…)*

- **Other conventions**

MIDI (by Huzinaga)

MIDI! Or MIDIX (polarization functions on non-carbon heavy atoms)

MIDIY (polarization functions on hydrogen atom as well)

MIDIX+

cc-pV*n*Z (correlation-consistent polarized valence zeta; by Dunning)

n: D (double zeta), T (triple zeta), Q (quadruple zeta)

D: d on heavy atoms; p on H atom

T: 1 set of f and 2 sets of d on heavy;
1 set of d and 2 sets of p on H

Ex: aug-cc-pV6Z (very expensive!)

1 i, 2 h, 3g, 4 f, 5 d, 6 p and 6 s for heavy atoms; in addition to core functions

aug (augmented) means diffuse functions used

- The general impression you should get: Pople’s basis set => popular; MIDI => not expensive; cc-pV*n*Z => expensive

- **The HF limit**

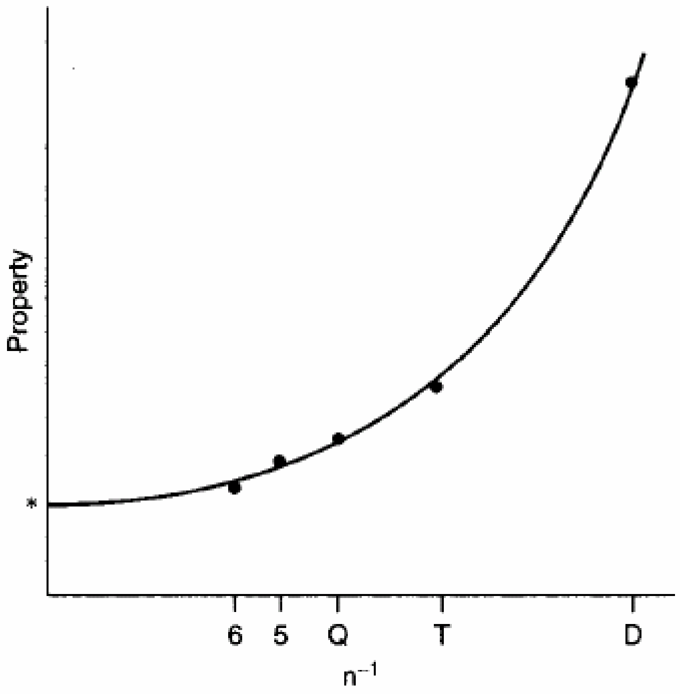


Figure 6.4 Use of an extrapolation procedure to estimate the expectation value for some property at the HF limit. The abscissa is marked off as n^{-1} in cc-pVnZ notation (see page 162). Note the sensitivity of the limiting value, which is to say the ordinate intercept, that might be expected based on the use of different curve-fitting procedures

Too large to calculate! Not practical for most case.

- **The additive assumption**

$$
\begin{aligned}
E[\mathrm{HF/6\text{-}311{+}{+}G(d,p)}] \approx{} & E[\mathrm{HF/6\text{-}31G}] \\
& + \{E[\mathrm{HF/6\text{-}31G(d,p)}] - E[\mathrm{HF/6\text{-}31G}]\} \\
& + \{E[\mathrm{HF/6\text{-}311G}] - E[\mathrm{HF/6\text{-}31G}]\} \\
& + \{E[\mathrm{HF/6\text{-}31{+}{+}G}] - E[\mathrm{HF/6\text{-}31G}]\}
\end{aligned}
$$

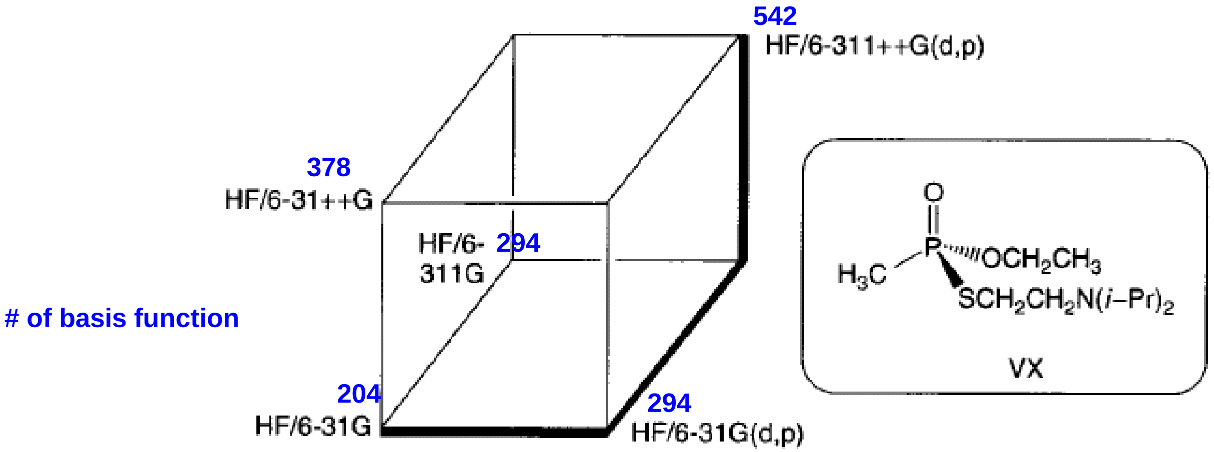


Figure 6.5 The chemical warfare agent VX and a conceptual illustration of the additivity concept embodied in Eq. (6.5). Each boldface line in the additivity cube represents one line on the r.h.s. of the equation

$(204^4 + 378^4 + 294^4 + 294^4)/542^4 = 0.43.$ *Less than half the time of evalution at the HF/6-311++G(d,p) level!*

- **Effective Core Potential (ECP)**

Replace the core electrons with analytical functions to represent the combined nuclear-electronic core to the remaining electrons

- reduce the # of electrons in calculation
- fold in the relativistic effects of heavy elements
- small-core (scale back to the next lower shell) and large-core (exclude only the outermost shell) approaches both present in the literature

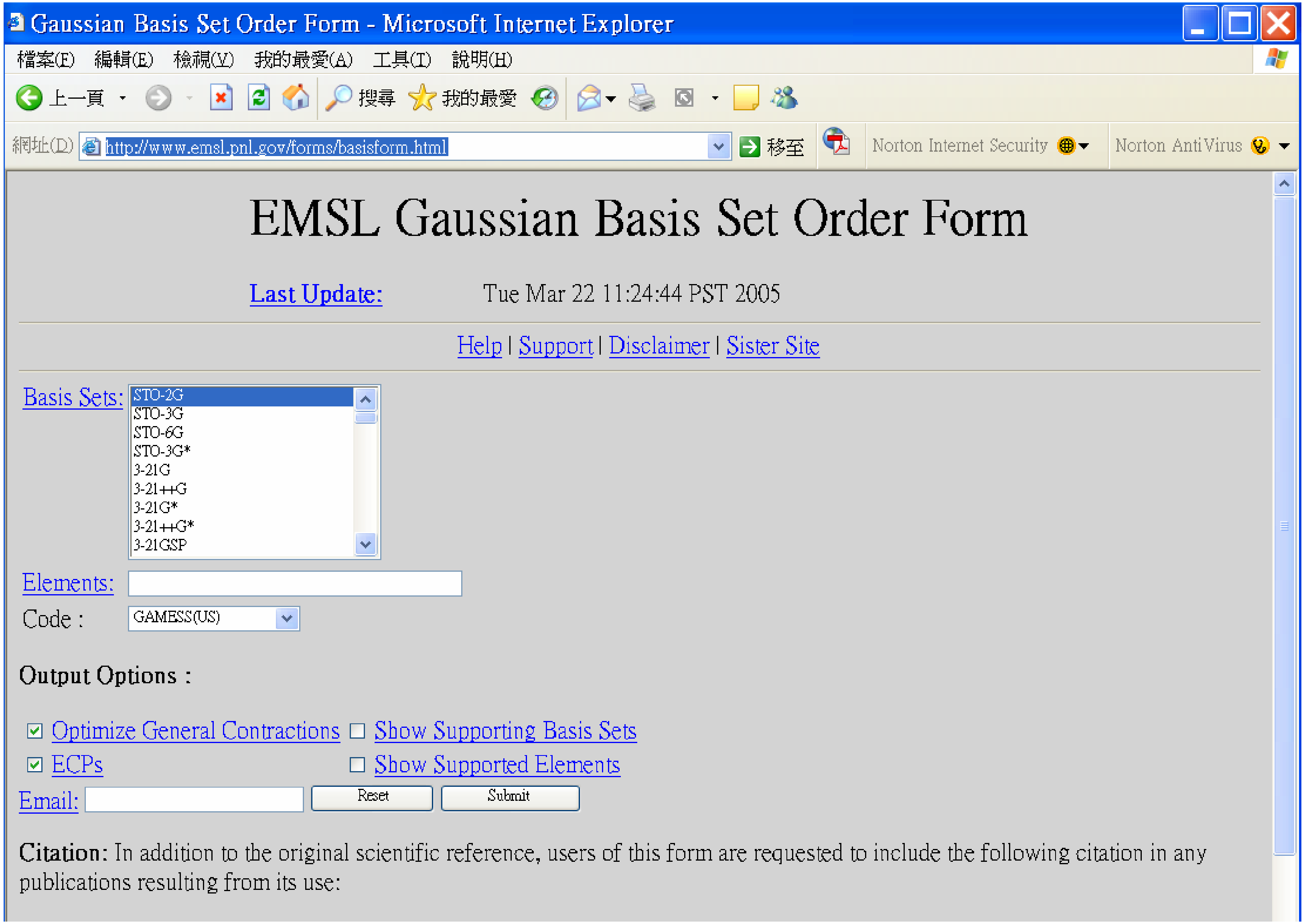
Commonly used ones:

By Hay and Wadt; LANL ECPs (Los Alamos National Laboratory)
non-relativistic for the first row of transition metals
Check optimized polarization and diffuse functions for LANL double-zeta basis set for p block elements
By Stevens
By Dolg; Stuttgart-Dresden ECPs

More recent ones:

By Lavallo and Klobukowski; for 1st row transition metals
By Dyall; for 4p, 5p, and 6p elements; correlation-consistent

- **A good electronic source of basis sets**



http://www.emsl.pnl.gov/forms/basisform.html

EMSL Gaussian Basis Set Order Form

Last Update: Tue Mar 22 11:24:44 PST 2005

Help | Support | Disclaimer | Sister Site

Basis Sets:
- STO-2G
- STO-3G
- STO-6G
- STO-3G*
- 3-21G
- 3-21++G
- 3-21G*
- 3-21++G*
- 3-21GSP

Elements:

Code : GAMESS(US)

Output Options :

☑ Optimize General Contractions ☐ Show Supporting Basis Sets

☑ ECPs ☐ Show Supported Elements

Email:

Citation: In addition to the original scientific reference, users of this form are requested to include the following citation in any publications resulting from its use:

Semiempirical Implementations of MO Theory

- N^4 integrals in HF; too many for large molecules. Simplification needed! Replace 2 electron or 1 electron integrals with parameters.
- Spirit of simplifications: make up for electron correlation with parameterization against experimental structural and energetic data
- Consider individual terms in the secular equation

$$\begin{vmatrix} H_{11} - ES_{11} & H_{12} - ES_{12} & \cdots & H_{1N} - ES_{1N} \\ H_{21} - ES_{21} & H_{22} - ES_{22} & \cdots & H_{2N} - ES_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ H_{N1} - ES_{N1} & H_{N2} - ES_{N2} & \cdots & H_{NN} - ES_{NN} \end{vmatrix} = 0$$

- Extended Hückel Theory

1. All core e^- ignored (same in all modern semiempirical methods)
2. Slater-type orbital (STO)

$$\varphi(r,\theta,\phi;\zeta,n,l,m) = \frac{(2\zeta)^{n+1/2}}{[(2n)!]^{1/2}} r^{n-1} e^{-\zeta r} Y_l^m(\theta,\phi)$$

MnO_4^- : 1 x 2s; 3 x 2p orbitals for each O atom
1 x 4s; 3 x 4p; 5 x 3d orbitals for Mn atom
=> 25 STO basis functions

3. Calculate overlap matrix elements
4. $H_{\mu\mu}$ parameterized (based on negative of valence-shell ionization potential; -IP. For 1s of $H_{\mu\mu}$ = -13.6 eV)
   ex: Fenske-Hall effective Hamiltonian: adjusted as a function of partial atomic charge)

5. Approximate $H_{\mu\nu} = ½ C_{\mu\nu} (H_{\mu\mu} + H_{\nu\nu}) S_{\mu\nu}$ (by Wolfsberg and Helmholz)

C: empirical constant (typically 1.75); S: overlap integral

- Use all the above approximations to construct secular determinant
- Matrix elements not depend on the final MO; not iterative
 => very fast
- In Fenske-Hall $H_{\mu\mu}$ adjusted as a function of partial atomic charge; iterative
- Insufficiently accurate for PES, so not useful for optimizing geometry (experimental geometry used); another drawback: no spin
- OK to generate qualitatively correct MO; explain the trends of UV

- CNDO (Complete Neglect of Differential Overlap) Formalism

1. One STO per valence orbital

2. In the secular determinant,overlap matrix elements are defined by

$$S_{\mu\nu} = \delta_{\mu\nu}$$

=> different basis function, no overlap

3. For two-electron integrals,

$$(\mu\nu|\lambda\sigma) = \delta_{\mu\nu}\delta_{\lambda\sigma}(\mu\mu|\lambda\lambda)$$

$$(\mu\mu|\lambda\lambda) = \gamma_{AB}$$ A,B represent the atoms

γ_{AA} and γ_{AB} can be calculated or parameterized.

As γ only depends on atoms, no distinction about s or p orbitals! (ss|ss) = (ss|pp) = (pp|pp)

4. One-electron integrals for diagonal matrix elements

$$\left\langle \mu \left| -\frac{1}{2}\nabla^2 - \sum_k \frac{Z_k}{r_k} \right| \mu \right\rangle = -\mathrm{IP}_\mu - \sum_k (Z_k - \delta_{Z_A Z_k})\gamma_{Ak}$$

5. One-electron integrals for off-diagonal matrix elements

$$\left\langle \mu \left| -\frac{1}{2}\nabla^2 - \sum_k \frac{Z_k}{r_k} \right| \nu \right\rangle = \frac{(\beta_A + \beta_B)S_{\mu\nu}}{2} \tag{5.10}$$

β: empirical parameter

- Also gives poor geometry

- When only π-type orbitals are included in the secular equation, it's called the PPP model (Pariser-Parr-Pople)

- INDO (Intermediate Neglect of Differential Overlap) Formalism

For one center two-electron integrals:

$$(ss|ss) = G_{ss}$$

$$(ss|pp) = G_{sp}$$

$$(pp|pp) = G_{pp}$$

$$(pp|p'p') = G_{pp'}$$

$$(sp|sp) = L_{sp}$$

Different parameters estimated from spectroscopic data!

- Improvement observed in the accuracy of bond angle, but overall molecular geometry still poor

- INDO/S by Ridley and Zerner was parameterized specifically for spectroscopic problems (use experimental geometry)
  good for d→d transition because transition localized on a single center; less robust for spectroscopic transitions not well localized to a single center

- MINDO/3 (third generation of modified INDO model) by Dewar was parameterized for both energy (heat of formation) and structure. STO orbital of s and p can have different ζ exponents. β takes into account of orbital types (not just atom types), γ adopts different form, 138 compounds with C, H, O, and H are used for parameterization…

- SINDO1 (symmetric orthogonalized INDO) by Jug and Schulz includes d functions for atoms of the second row in the periodic table. Doing better for hypervalent molecules (ex: phsphorus-containing compounds) than previous models

- NDDO (Neglect of Diatomic Differential Overlap) Formalism

- In INDO, the two-center two-electron integral takes on the value γ_{AB} irrespective of which orbitals on atoms A and B are considered. NDDO relaxes this constraints. All integrals $(\mu\nu|\lambda\sigma)$ are retained provided μ and ν are on the same atom and λ and σ are on the same atom. Most modern semiempirical models are NDDO.

- MNDO (modified neglect of differential overlap) by Dewar and Thiel for molecules containing elements H, He, Li, Be, B, C, N, O, F, Mg, Al, Si, P, S, Cl, Zn, Ge, Br, Sn, I, Hg, Pb. Assume a valence orbital set comprised only of s and p orbitals. Nuclear repulsion energy is added to the SCF energy. Some integrals are parameterized, some integrals are calculated. Very poor for hydrogen bonding.

- AM1 (Austin Model 1) by Dewar was set out to improve the nuclear repulsion term for better structure. Parameterized for B, F, Mg, Al, Si, P, S, Cl, Zn, Ge, Br, Sn, I, and Hg. (Very often employed in the literature)

- PM3 (Parameterized Model 3) by Stewart was parameterized simultaneously for H, C, N, O, F, Al, Si, P, S, Cl, Br, and I. Parameters are optimized in a mathematical approach, rather than considering chemical reasonableness. Parameters also available for Li, Be, Na, Mg, Ca, Zn, Ga, Ge, As, Se, Cd, In, Sn, Sb, Te, Hg, Tl, Pb, and Bi. (Very often employed)

- MNDO/d by Thiel and Voityuk has d orbitals for seond-row and heavier elements. Handles hypervalent molecules well. Good heat of formation. Describes hydrogen bonding poorly as MNDO.

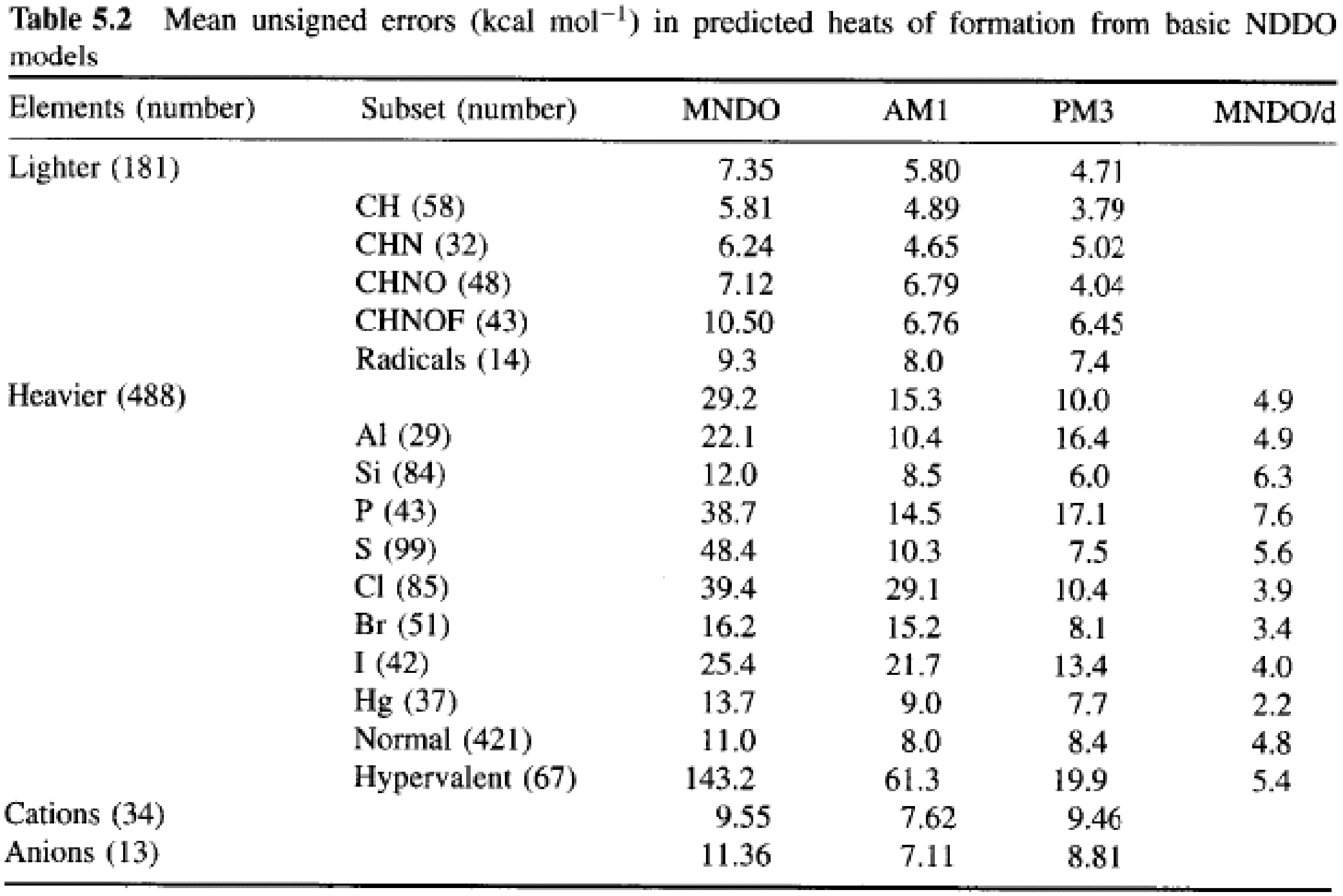


Table 5.2 Mean unsigned errors (kcal mol^{-1}) in predicted heats of formation from basic NDDO models

Elements (number)	Subset (number)	MNDO	AM1	PM3	MNDO/d
Lighter (181)		7.35	5.80	4.71	
	CH (58)	5.81	4.89	3.79	
	CHN (32)	6.24	4.65	5.02	
	CHNO (48)	7.12	6.79	4.04	
	CHNOF (43)	10.50	6.76	6.45	
	Radicals (14)	9.3	8.0	7.4	
Heavier (488)		29.2	15.3	10.0	4.9
	Al (29)	22.1	10.4	16.4	4.9
	Si (84)	12.0	8.5	6.0	6.3
	P (43)	38.7	14.5	17.1	7.6
	S (99)	48.4	10.3	7.5	5.6
	Cl (85)	39.4	29.1	10.4	3.9
	Br (51)	16.2	15.2	8.1	3.4
	I (42)	25.4	21.7	13.4	4.0
	Hg (37)	13.7	9.0	7.7	2.2
	Normal (421)	11.0	8.0	8.4	4.8
	Hypervalent (67)	143.2	61.3	19.9	5.4
Cations (34)		9.55	7.62	9.46	
Anions (13)		11.36	7.11	8.81	

- Weak interactions normally not well modeled by semiempirical methods because electron correlation not well treated

- Relative conformer energies not very good in general.

- For MNDO, steric crowding tends to be too strongly disfavored and small ring compounds are predicted to be too stable.

- For dipole moment, mean unsigned errors fo 125 molecules are 0.45, 0.35 and 0.38 D for MNDO, AM1, and PM3, respectively.