Ab Initio and Density-Functional Treatments of Molecules

AB INITIO, DENSITY-FUNCTIONAL, SEMIEMPIRICAL,

Diatomic molecules: electronic wave function is a function of only one parameter (the internuclear distance)

Polyatomic molecule: electronic wave function depends on several parameters (bond distances, bond angles, and dihedral angles)

The equilibrium bond distances and angles are found as those values that minimize ($U = E_{el} + V_{NN}$)

Four **main methods**
(to calculating molecular properties):

- ab initio
- semiempirical
- density-functional
- Molecular mechanics

Semiempirical methods:
- ✓ use a simpler Hamiltonian than the correct $\hat{H}$
- ✓ use parameters (are adjusted to fit experimental data or the results of ab initio calculations)
- ✓ Example: is the Hückel MO treatment (uses a one-electron Hamiltonian and takes the bond integrals as adjustable parameters rather than quantities to be calculated theoretically)

ab initio (or **first principles**) **method**:
- ✓ uses the correct Hamiltonian
- ✓ does not use experimental data (only fundamental physical constants)
- ✓ HF SCF calculation is an ab initio calculation (seeks the antisymmetrized product Φ of one-electron functions that minimizes $\int \Phi^* \hat{H} \Phi d\tau$; $\hat{H}$ ≡ true Hamiltonian)
- ✓ Ab initio ≡ from the beginning (not 100% correct)
- ✓ An ab initio SCF MO calculation: taking ψ as an antisymmetrized product of spin-orbitals and uses a finite (and hence incomplete) basis set.

Density-functional method:

- ✓ does not attempt to calculate the molecular wave function
- ✓ calculates $\rho \rightarrow$ molecular electronic energy from ρ

Molecular-mechanics method:

- ✓ is not a QM method
- ✓ does not use a molecular Hamiltonian operator or wave function
- ✓ molecule $\equiv$ a collection of atoms held together by bonds
- ✓ molecular energy is expressed in terms of force constants for bond bending and stretching and other parameters.

The Hartree–Fock Self-Consistent-Field Method

- we must use spin-orbitals and must take an antisymmetric linear combination of products of spin-orbitals. This was pointed out by Fock (and by Slater) in 1930, and an SCF calculation that uses antisymmetrized spin-orbitals is called a **Hartree–Fock calculation**. We have seen that a Slater determinant of spin-orbitals provides the proper antisymmetry. For example, to carry out a Hartree–Fock calculation for the lithium ground state, we start with the function, where f and g are guesses for the $1s$ and $2s$ orbitals.

$$\phi = \frac{1}{\sqrt{6}} \begin{vmatrix} f(1)\alpha(1) & f(1)\beta(1) & g(1)\alpha(1) \\ f(2)\alpha(2) & f(2)\beta(2) & g(2)\alpha(2) \\ f(3)\alpha(3) & f(3)\beta(3) & g(3)\alpha(3) \end{vmatrix}$$

- We then carry out the SCF iterative process until we get no further improvement in f and g.
- This gives the lithium ground-state Hartree–Fock wave function.
- The differential equations for finding the Hartree–Fock orbitals have the same general form as

$$\hat{F}u_i = \varepsilon_i u_i, \qquad i = 1, 2, \ldots, n$$

- where u_i is the ith spin-orbital, the operator F, called the **Fock** (or **Hartree–Fock**) operator, is the effective HF Hamiltonian, and the eigenvalue ε_i is the orbital energy of spin-orbital i.
- Actually, above eq. applies only when the HF wave function can be written as a single Slater determinant, as it can for closed-subshell atoms and atoms with only one electron outside closed subshells. When the Hartree–Fock wave function contains more than one Slater determinant, the Hartree–Fock equations are more complicated than that.

- The orbital energy ε_i in the HF equations can be shown to be a good approximation to the negative of the energy needed to ionize a closed-subshell atom by removing an electron from spin-orbital i (Koopmans' theorem.

- In 1951, Roothaan proposed representing the Hartree–Fock orbitals as linear combinations of a complete set of known functions, called **basis functions**. Thus for lithium we would write the Hartree–Fock 1s and 2s spatial orbitals as

$$f = \sum_i b_i \chi_i, \quad g = \sum_i c_i \chi_i$$

- where the χ_i functions are some complete set of functions, and where the b_is and c_is are expansion coefficients that are found by the SCF iterative procedure. Since the χ_i (chi i) functions form a complete set, these expansions are valid. The Roothaan expansion procedure allows one to find the HF wave function using matrix algebra. The Roothaan procedure is readily implemented on a computer and is often used to find atomic HF wave functions and nearly always used to find molecular HF wave functions.

- A commonly used set of basis functions for atomic Hartree–Fock calculations is the set of Slater-type orbitals (STOs) whose normalized form is

$$\frac{(2\zeta/a_0)^{n+1/2}}{[(2n)!]^{1/2}} r^{n-1} e^{-\zeta r/a_0} Y_l^m(\theta, \phi)$$

- The set of all such functions with n, l, and m being integers obeying rules but with ζ having all possible positive values forms a complete set. The parameter ζ is called the orbital exponent. To get a truly accurate representation of the HF orbitals, we would have to include an infinite number of Slater orbitals in the expansions. In practice, one can get very accurate results by using only a few judiciously chosen Slater orbitals.

- Another possibility is to use Gaussian-type basis functions

THE HARTREE-FOCK METHOD FOR MOLECULES

A key development in quantum chemistry has been the computation of accurate self-consistent-field wave functions for many diatomic and polyatomic molecules.

For closed shells:

The molecular HF wave function ≡ an antisymmetrized product (Slater determinant) of spin-orbitals.

By the variation theorem: $E_{\text{HF}} = \langle D|\hat{H}_{\text{el}} + V_{NN}|D\rangle$

D: Slater-determinant HF wave function (normalized)

V_{NN}: does not involve electronic coordinates

the closed-shell orthogonal Hartree-Fock MOs satisfy:

$$\hat{F}(1)\phi_i(1) = \varepsilon_i\phi_i(1)$$

ε_i : the orbital energy

$$\hat{F}(1) = \hat{H}^{core}(1) + \sum_{j=1}^{n/2}[2\hat{J}_j(1) - \hat{K}_j(1)]$$ in atomic units

there are two electrons in each spatial orbital

$$\hat{H}^{core}(1) \equiv -\tfrac{1}{2}\nabla_1^2 - \sum_{\alpha}\frac{Z_\alpha}{r_{1\alpha}}$$

kinetic energy operator of one electron

potential-energy operators for the attractions between one electron and the nuclei

$$\hat{J}_j(1)f(1) = f(1) \int |\phi_j(2)|^2 \frac{1}{r_{12}} dv_2$$

Coulomb operator:
the potential energy of interaction between electron 1 and a smeared-out electron with electronic density - $|\phi_j(2)|^2$

an arbitrary function

$$\hat{K}_j(1)f(1) = \phi_j(1) \int \frac{\phi_j^*(2)f(2)}{r_{12}} dv_2$$

exchange operator
✓ has no simple physical interpretation
✓ arises from the requirement that the wave function be antisymmetric with respect to electron exchange.

over all space

✓ The orthogonality of the MOs greatly simplifies calculations (simpler)
✓ The VB method uses atomic orbitals, and AOs centered on different atoms are not orthogonal.

The true Hamiltonian operator and wave function involve the coordinates of all n electrons.

The Hartree-Fock Hamiltonian operator $\hat{F}$ is a one-electron operator

The operator $\hat{F}$ depends on its own eigenfunctions, which are not known initially. Hence the Hartree-Fock equations must be solved by an iterative process.

$$\hat{F}(1)\phi_i(1) = \varepsilon_i\phi_i(1)$$

a one-electron differential equation

multiply by $\phi_i^*(1)$
integrate over all space

$$\varepsilon_i = \int \phi_i^*(1)\hat{F}(1)\phi_i(1) \, dv_1$$

$$\hat{F}(1) = \hat{H}^{core}(1) + \sum_{j=1}^{n/2}[2\hat{J}_j(1) - \hat{K}_j(1)]$$

$$\varepsilon_i = \langle\phi_i(1)|\hat{H}^{core}(1)|\phi_i(1)\rangle + \sum_j[2\langle\phi_i(1)|\hat{J}_j(1)|\phi_i(1)\rangle - \langle\phi_i(1)|\hat{K}_j(1)|\phi_i(1)\rangle]$$

$$\varepsilon_i = H_{ii}^{core} + \sum_{j=1}^{n/2}(2J_{ij} - K_{ij})$$

$$\sum_{i=1}^{n/2} \varepsilon_i = \sum_{i=1}^{n/2} H_{ii}^{\text{core}} + \sum_{i=1}^{n/2} \sum_{j=1}^{n/2} (2J_{ij} - K_{ij})$$

over the n/2 occupied orbitals

$$E_{\text{HF}} = 2\sum_{i=1}^{n/2} H_{ii}^{\text{core}} + \sum_{i=1}^{n/2} \sum_{j=1}^{n/2} (2J_{ij} - K_{ij}) + V_{NN}$$

$$\sum_{i=1}^{n/2} \varepsilon_i = \sum_{i=1}^{n/2} H_{ii}^{\text{core}} + \sum_{i=1}^{n/2} \sum_{j=1}^{n/2} (2J_{ij} - K_{ij})$$

$$E_{\text{HF}} = 2\sum_{i=1}^{n/2} \varepsilon_i - \sum_{i=1}^{n/2} \sum_{j=1}^{n/2} (2J_{ij} - K_{ij}) + V_{NN}$$

sum of the orbital energies

avoids counting each interelectronic repulsion twice

Roothaan's proposal : expand the spatial orbitals ϕ_i as linear combinations of a set of one-electron basis functions χ_s (make feasible the calculation)

$$\phi_i = \sum_{s=1}^{b} c_{si}\chi_s$$

basis functions should form a complete set (exact representation, infinite number of basis functions)

In practice : a finite number b of basis functions

negligible error in MOs representation: b be large enough and the functions χ_s well chosen

$$\hat{F}(1)\phi_i(1) = \varepsilon_i\phi_i(1)$$

$$\phi_i = \sum_{s=1}^{b} c_{si}\chi_s$$

$$\sum_s c_{si}\hat{F}\chi_s = \varepsilon_i \sum_s c_{si}\chi_s$$

$\chi_r^* \times$ and integration

$$\sum_{s=1}^{b} c_{si}(F_{rs} - \varepsilon_i S_{rs}) = 0 , \qquad r = 1, 2, \dots, b$$

$$F_{rs} \equiv \langle \chi_r | \hat{F} | \chi_s \rangle \qquad S_{rs} \equiv \langle \chi_r | \chi_s \rangle$$
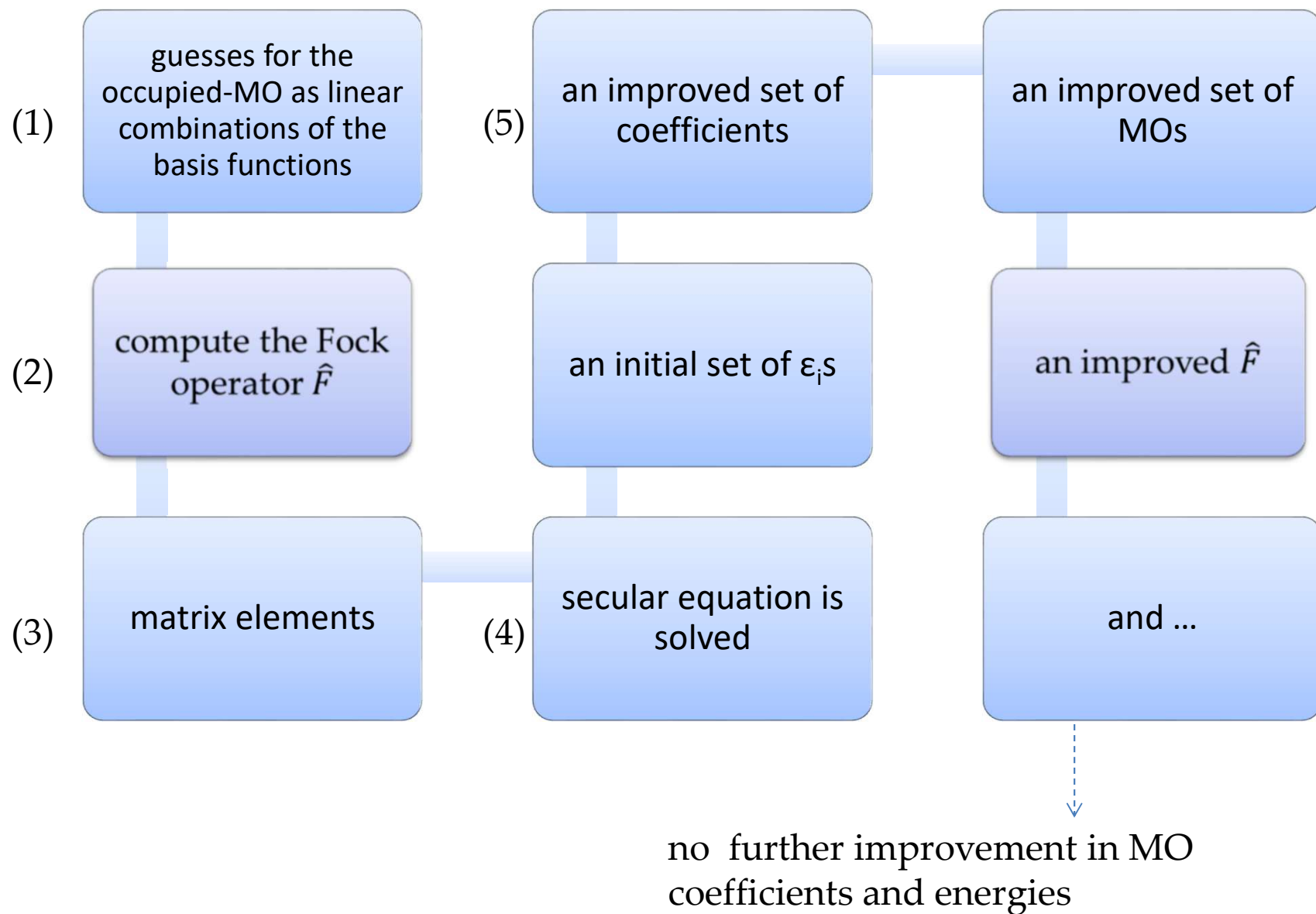
$$\sum_{s=1}^{b} c_{si}(F_{rs} - \varepsilon_i S_{rs}) = 0 , \qquad r = 1, 2, \ldots , b$$

a set of b simultaneous linear homogeneous equations in the b unknowns c_{si}, s = 1, 2, ..., b, that describe the MO .

For a nontrivial solution:

$$\det(F_{rs} - \varepsilon_i S_{rs}) = 0 \qquad \text{a secular equation}$$

Roothaan equations must be solved by an iterative process, since the F_{rs} depend on the orbitals ϕ_i (through the dependence of $\hat{F}$ on the ϕ_i's), which depend on c_{si}.

(1) guesses for the occupied-MO as linear combinations of the basis functions

(5) an improved set of coefficients

an improved set of MOs

(2) compute the Fock operator $\hat{F}$

an initial set of ε_is

an improved $\hat{F}$

(3) matrix elements

(4) secular equation is solved

and ...

no further improvement in MO coefficients and energies

$$\phi_i = \sum_{s=1}^{b} c_{si}\chi_s \tag{1}$$

$$\hat{F}(1) = \hat{H}^{\text{core}}(1) + \sum_{j=1}^{n/2} [2\hat{J}_j(1) - \hat{K}_j(1)] \tag{2}$$

$$\left.\begin{aligned}
\hat{H}^{\text{core}}(1) &\equiv -\tfrac{1}{2}\nabla_1^2 - \sum_{\alpha} \frac{Z_\alpha}{r_{1\alpha}} \\[2ex]
\hat{J}_j(1)f(1) &= f(1) \int |\phi_j(2)|^2 \frac{1}{r_{12}} dv_2 \\[2ex]
\hat{K}_j(1)f(1) &= \phi_j(1) \int \frac{\phi_j^*(2)f(2)}{r_{12}} dv_2
\end{aligned}\right\} \tag{2}$$

$$F_{rs} \equiv \langle \chi_r | \hat{F} | \chi_s \rangle, \qquad S_{rs} \equiv \langle \chi_r | \chi_s \rangle \tag{3}$$

$$\det(F_{rs} - \varepsilon_i S_{rs}) = 0 \tag{4}$$

$$\sum_{s=1}^{b} c_{si}(F_{rs} - \varepsilon_i S_{rs}) = 0, \qquad r = 1, 2, \ldots, b \tag{5}$$

Basis functions:

✓ Generally, each MO is written as a linear combination of one-electron functions (orbitals) centered on each atom.

✓ one can use Slater functions for the AOs.

✓ a complete set of AO BFs: an infinite number of SOs are needed,

✓ the true molecular HF wave function can be closely approximated with a reasonably small number of carefully chosen SOs.

✓ A minimal BS for a molecular SCF calculation consists of a single basis function for each inner-shell AO and each valence-shell AO of each atom.

✓ An extended basis set (EBS) is a set that is larger than a minimal set.

✓ Minimal-BS SCF calculations are easier than EBS calculations, but the latter are considerably more accurate.

Matrix Form of the Roothaan Equations

The Roothaan equations are most efficiently solved using matrix methods.

$$\sum_{s=1}^{b} F_{rs} c_{si} = \sum_{s=1}^{b} S_{rs} c_{si} \varepsilon_i, \qquad r = 1, 2, \ldots, b$$

c_{si} : relate the MOs ϕ_i to χ_s

$$\phi_i = \sum_s c_{si} \chi_s$$

$$F_{rs} = \langle \chi_r | \hat{F} | \chi_s \rangle$$

$$S_{rs} = \langle \chi_r | \chi_s \rangle$$

ε_i : orbital energies

Use of the matrix multiplication rule:

$$(\mathbf{C}\boldsymbol{\varepsilon})_{si} = \sum_m c_{sm} \varepsilon_{mi} = \sum_m c_{sm} \delta_{mi} \varepsilon_i = c_{si} \varepsilon_i$$

$$\sum_{s=1}^{b} F_{rs} c_{si} = \sum_{s=1}^{b} S_{rs} (\mathbf{C}\boldsymbol{\varepsilon})_{si}$$

C : be the square matrix of order b whose elements are the coefficients c_{si}

F : be the square matrix of order b whose elements are $F_{rs} = \langle \chi_r | \hat{F} | \chi_s \rangle$

S : be the square matrix whose elements are $S_{rs} = \langle \chi_r | \chi_s \rangle$

ε : be the diagonal square matrix whose diagonal elements are the $\varepsilon_1, \varepsilon_2, \varepsilon_3, \ldots$ orbital energies $\varepsilon_{mi} = \delta_{mi} \varepsilon_i$

← Kronecker delta

$$\sum_{s=1}^{b} F_{rs} c_{si} = \sum_{s=1}^{b} S_{rs} (\mathbf{C\varepsilon})_{si}$$

(r, i)th element of **FC**   (r, i)th element of **S(Cε)**

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

matrix form of the Roothaan equations.

$\{\chi_s\}$ is not an orthogonal set

Schmidt or some other procedure: linear combinations of the basis functions

$\{\chi'_s\}$: an orthonormal set

$$\chi'_s = \sum_t a_{ts}\chi_t \qquad S'_{rs} = \langle \chi'_r | \chi'_s \rangle = \delta_{rs}$$

See Szabo and Ostlund, Section 3.4.5, for details of the orthogonalization

in $\{\chi'_s\}$ the overlap matrix is a unit matrix: $\qquad S'_{rs} = \langle \chi'_r | \chi'_s \rangle = \delta_{rs}$

$$\mathbf{FC} = \mathbf{SC}\boldsymbol{\varepsilon} \quad \longrightarrow \quad \mathbf{F'C'} = \mathbf{C'}\boldsymbol{\varepsilon}$$

$$F'_{rs} = \langle \chi'_r | \hat{F} | \chi'_s \rangle$$

C' : relate the MOs ϕ_i to the orthonormal BFs: $\quad \phi_i = \sum_s c'_{si}\chi'_s$

F' and **C'** are readily calculated from **F** and **C**.

F and **F'** matrices and the **C** and **C'** matrices are related by (Problem 8.49c):

$$\mathbf{F'} = \mathbf{A}^{\dagger}\mathbf{F}\mathbf{A} \quad \text{and} \quad \mathbf{C} = \mathbf{A}\mathbf{C'}$$

where **A** is the matrix of coefficients a_{ts} in $\quad \chi'_s = \sum_t a_{ts}\chi_t$

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

The orbital energies ε_i, are the eigenvalues of the Fock matrix **F'** and each column of **C'** is an eigenvector of **F'**

F is Hermitian $\Rightarrow$ F' is a Hermitian matrix $\Rightarrow$ the eigenvector matrix **C'** can be chosen to be unitary

$\Rightarrow$ its inverse equals its conjugate transpose $\mathbf{C'}^{-1} = \mathbf{C'}^{\dagger}$ $\Rightarrow$ the MOs ϕ_i are orthonormal

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

Multiplication on the left by $\quad \mathbf{C'}^{-1} = \mathbf{C'}^{\dagger}$

$$\mathbf{C'}^{\dagger}\mathbf{F'}\mathbf{C'} = \boldsymbol{\varepsilon}$$

For an SCF calculation at a specified molecular geometry:

1. Choose a basis set χ_s
2. Evaluate the H_{rs}^{core}, S_{rs}, and $(rs|tu)$ integrals.
3. Use the overlap integrals S_{rs} and an orthogonalization procedure to calculate the **A** matrix of coefficients a_{ts} that will produce orthonormal basis functions $\chi_s' = \sum_t a_{ts}\chi_t$
4. Make an initial guess for the coefficients c_{si} in the MOs $\phi_i = \sum_s c_{si}\chi_s$ From the initial guess of coefficients, calculate the density matrix **P** in

$$P_{tu} \equiv 2\sum_{j=1}^{n/2} c_{tj}^* c_{uj}, \qquad t = 1, 2, \ldots, b, \qquad u = 1, 2, \ldots, b$$

5. Use

$$F_{rs} = H_{rs}^{core} + \sum_{t=1}^{b}\sum_{u=1}^{b} P_{tu}[(rs|tu) - \tfrac{1}{2}(ru|ts)]$$

to calculate an estimate of the Fock matrix elements F_{rs} from **P** and the $(rs|tu)$ and H_{rs}^{core} integrals.
6. Calculate the matrix **F'** using **F' = A†FA**.

7. Use a matrix-diagonalization method to find the eigenvalue and eigenvector matrices **ε** and **C'** of **F'**.

8. Calculate the coefficient matrix **C = AC'**.

9. Calculate an improved estimate of the density matrix from C using **P = 2CC†**, which is the matrix form of

$$P_{tu} \equiv 2 \sum_{j=1}^{n/2} c_{tj}^* c_{uj}, \qquad t = 1, 2, \ldots, b, \qquad u = 1, 2, \ldots, b$$

10. Compare the improved **P** with the preceding estimate of **P**. If the differences are negligible, the calculation has converged and one uses the converged SCF wave function to calculate molecular properties. If the calculation has not converged, one goes back to step (5)

1) One way to begin an SCF calculation:

$$F_{rs} = H_{rs}^{core} + \boxed{\sum_{t=1}^{b} \sum_{u=1}^{b} P_{tu}[(rs|tu) - \tfrac{1}{2}(ru|ts)]}_{\text{Is neglected}} \longrightarrow \qquad F_{rs} \approx H_{rs}^{core}$$

a very crude estimate

2) ab initio SCF calculations get the initial estimate of the density matrix by doing a semiempirical calculation on the molecule

3) a guess for the **P** matrix by using the density matrices of the atoms composing the molecule.

To find the equilibrium geometry of a molecule, one does a series of SCF calculations at many successive geometries.
For the second and later SCF calculations of the series, one takes the initial guess of **P** as **P** for the SCF wave function of a nearby geometry.

SCF wave function and Hartree-Fock wave function

the term SCF wave function is applied to any wave function obtained by iterative solution of the Roothaan equations

the basis set is large enough to give a really accurate approximation to the HF SCF wave function.

There is only one true HF SCF wave function, which is the best possible wave function that can be written as a Slater determinant of spin-orbitals.

By the extended-basis-set calculations $\rightarrow$ "near HF wave functions" or, less cautiously, "HF wave functions."

Hartree-Fock wave functions are **only approximations to the true wave functions**.

✓ a HF wave function gives a **very good approximation** to the **electron probability density** $\rho(x,y, z)$ for equilibrium configuration.

✓ A molecular property that **involves only one-electron operators** can be expressed **as an integral involving ρ**. Consequently, such properties are **accurately calculated** using HF wave functions (For example, the molecular **dipole moment**).

LiH :

with a near HF ψ: dipole moment = 6.00 D (experimental value = 5.83 D)

NaCl:

the calculated value = 9.18 D (experimental value 9.02 D)

✓ An error of about 0.2 D is typical in such calculations, but **where the dipole moment is small, the percent error can be large**.

CO:

(experimental moment = 0.11 D with the polarity C^-O^+,

the near-HF moment = 0.27 D with the wrong polarity C^+O^-.

a CI wave function gives 0.12 D with the correct polarity

A major weakness of the Hartree-Fock method is its failure to give accurate molecular dissociation energies.

N_2:
HF D_e = 5.3 eV   by an extended-bs   (true value = 9.9 eV)

F_2:
HF D_e = -1.4 eV   (true De = 1.66 eV)

can prove, for a many-electron MO wave function:

$$\rho = \sum \rho_{\text{each MO}} \times n_e(\text{occupying it})$$

$$\rho(x, y, z) = \sum_j n_j |\phi_j|^2$$

over different
orthogonal spatial
MOs

number of electrons in
the MO ϕ_j (0,1, or 2)

$B(\mathbf{r}_i)$: a function of the spatial coordinates x_i, y_i, z_i of electron i

the average value

$$\left\langle \psi \left| \sum_{i=1}^{n} B(\mathbf{r}_i) \right| \psi \right\rangle = \int \psi^* \sum_{i=1}^{n} B(\mathbf{r}_i)\psi \, d\tau = \sum_{i=1}^{n} \int |\psi|^2 B(\mathbf{r}_i) \, d\tau$$

electronic wave function.

electrons are indistinguishable:
terms are identical

$$\left\langle \psi \left| \sum_{i=1}^{n} B(\mathbf{r}_i) \right| \psi \right\rangle = \int n|\psi|^2 B(\mathbf{r}_1) \, d\tau$$

$$\langle \psi | \sum_{i=1}^{n} B(\mathbf{r}_i) | \psi \rangle = \int n|\psi|^2 B(\mathbf{r}_1) \, d\tau$$

B($\mathbf{r}_1$) depends only on x_1, y_1, z_1 before we integrate over x_1, y_1, z_1, we can integrate $n|\psi|^2$ over the spatial coordinates of electrons 2 to n and sum over all the spin coordinates.

$$\rho(\mathbf{r}) = n \sum_{\text{all } m_s} \int \cdots \int |\psi(\mathbf{r}, \mathbf{r}_2, \ldots, \mathbf{r}_n, m_{s1}, \ldots, m_{sn})|^2 \, d\mathbf{r}_2 \cdots d\mathbf{r}_n$$

$$\langle \psi | \sum_{i=1}^{n} B(\mathbf{r}_i) | \psi \rangle = \int \rho(\mathbf{r}_1) B(\mathbf{r}_1) \, d\mathbf{r}_1$$

$$\int \psi^* \sum_{i=1}^{n} B(\mathbf{r}_i) \psi \, d\tau = \int \rho(\mathbf{r}) B(\mathbf{r}) \, d\mathbf{r}$$

over the three spatial coordinates x, y, z.

DIPOLE MOMENTS

wave functions $\rightarrow$ molecular dipole moments

classical expression:

for a set of discrete charges Q_i

$$\boldsymbol{\mu}_{cl} = \sum_i Q_i \mathbf{r}_i$$

position vector

quantum-mechanical expression (by perturbation theory):

$$\boldsymbol{\mu} = -e \iiint \rho(x, y, z)\mathbf{r}\, dx\, dy\, dz + e \sum_\alpha Z_\alpha \mathbf{r}_\alpha$$

electron probability density.

THE VALENCE-ELECTRON APPROXIMATION

Cs_2, which has 110 electrons

In the MO method:

110 X 110 Slater determinant of molecular orbitals

MOs ~ functions containing variational parameters
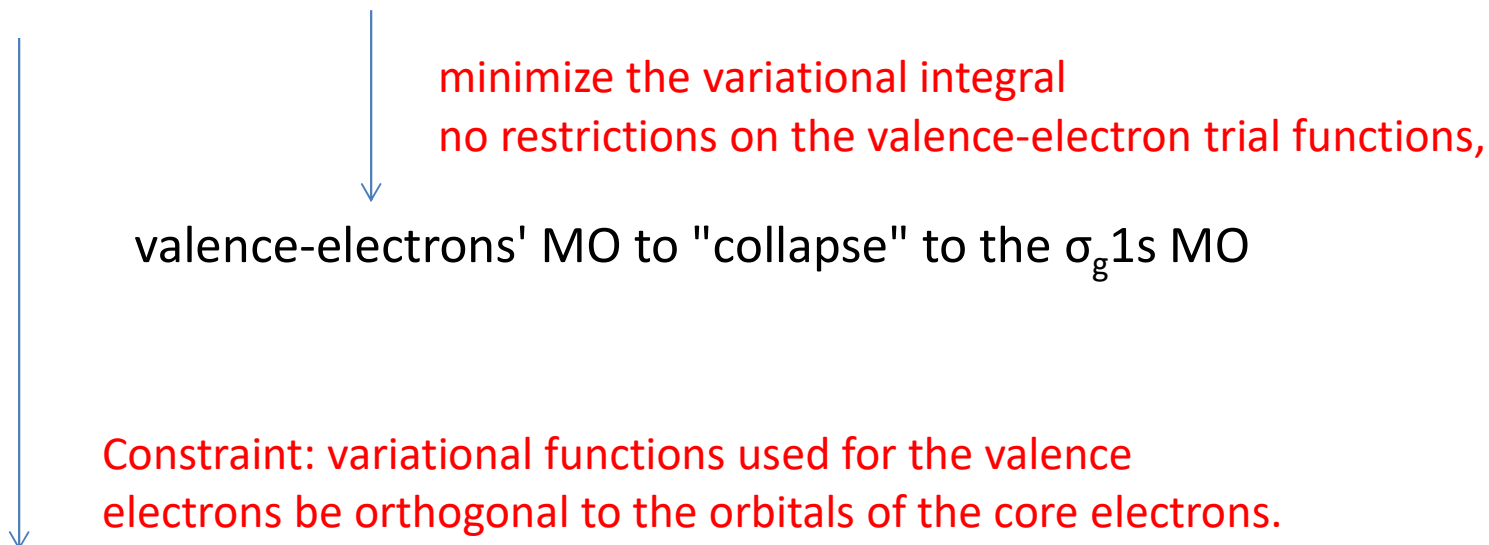
minimize the variational integral

the valence-electron approximation:

108 core electrons + two 6s valence electrons

molecular energy = core- and valence-electron energies.

1) core electrons ≡ point charges coinciding with the nucleus.

Hamiltonian (for Cs_2) = Hamiltonian for H_2

minimize the variational integral
no restrictions on the valence-electron trial functions,

valence-electrons' MO to "collapse" to the $\sigma_g 1s$ MO

Constraint: variational functions used for the valence
electrons be orthogonal to the orbitals of the core electrons.

more work

2) core electrons are treated as a charge distribution (effective repulsive
potential for the motion of the valence electrons).
effective Hamiltonian for the valence electrons

The valence-electron approximation is widely used in approximate
treatments of polyatomic molecules

SCF MO Wave Functions for Open-Shell States

A HF wave function in which electrons whose spins are paired occupy the
same spatial orbital is called a restricted HF (RHF) wave function.

- ✓ **RHF wave function** is generally used for closed-shell states
- ✓ **unrestricted HF** (UHF) and **restricted open-shell HF** (ROHF) are widely
  used for open-shell states.
- ✓ In ROHF method, paired electrons are given the same spatial orbital
  function
- ✓ $|1s\overline{1s}2s|$ for the ROHF wave function of the Li ground state
- ✓ $|1s\overline{1s'}2s|$ for UHF wave function, where 1s ≠ 1s'.
- ✓ the interaction between the 2sα and 1sα electrons differs from the
  interaction between the 2sα and 1sβ electrons (spatial orbitals, 1s and
  1s').

- ✓ UHF wf gives a slightly lower E than the ROHF wf
- ✓ UHF wf is more useful in predicting ESR spectra
- ✓ main problem with the UHF wf: is not an eigenfunction of $\hat{S}^2$ (nor a linear combination of a few UHF functions)
- ✓ true wf and the ROHF wf are eigenfunctions of $\hat{S}^2$.
- ✓ for the UHF wf, one have to check the deviation of (S^2) from $S(S + I)\hbar^2$

BASIS FUNCTIONS

Most QM methods begin with the choice of a set of BFs.

For diatomic molecules, the BFs are usually taken as AOs, (centered on <u>a</u> or <u>b</u>) each AO ≡ a linear combination of one or more STOs.
has the form

An STO centered on atom <u>a</u>
$$N r_a^{n-1} e^{-\zeta r_a} Y_l^m(\theta_a, \phi_a)$$

For nonlinear molecules, the real form of the STOs is used (Y_l^m is replaced by $(Y_l^{m*} \pm Y_l^m)/2^{1/2}$)

$$\phi_i = \Sigma_r c_{ri} \chi_r$$

STO basis function

LC-STO MOs

LC-STO method uses STOs centered on each of the atoms

The presence of more than two atoms causes difficulties in evaluating the needed integrals.

For a triatomic molecule: three-center, two-center, and one-center integrals.

For a molecule with four or more atoms: one also has four-center integrals

Solution of the Roothaan equations:

(rs|tu): is four-center If the BFs centered on different nucleus

H_{rs}^{core} integrals: involve either one or two centers

Number of BFs = b

b different possibilities for each BF in (rs|tu)
(rs|tu) = (sr|tu) = ...

number of different electron-repulsion integrals ≈ $b^4/8$

b = 20 to 400 → $b^4/8$ = 20000 to 3 X 10^9

Computer evaluation of three- and four-center integrals over STO basis functions is very time consuming.

To speed up integral evaluation, Boys proposed the use of Gaussian-type functions (GTFs) instead of STOs for the AOs in an LCAO wf.

$$g_{ijk} = N x_b^i y_b^j z_b^k e^{-\alpha r_b^2}$$

A Cartesian Gaussian centered on b

i, j, and k are nonnegative integers,
α is a positive orbital exponent,
x_b, y_b, z_b are Cartesian coordinates with the origin at nucleus b.

$$N = \left(\frac{2\alpha}{\pi}\right)^{3/4} \left[\frac{(8\alpha)^{i+j+k} i! j! k!}{(2i)!(2j)!(2k)!}\right]^{1/2}$$

normalization constant

i + j + k = 0 (that is, i = 0, j = 0, k = 0) : s-type Gaussian

i + j + k = 1 : p-type Gaussian

i + j + k = 2 : d-type Gaussian

Six d-type Gaussians: x_b^2, y_b^2, z_b^2, $x_b y_b$, $x_b z_b$, and $y_b z_b$.

five linear combinations ($x_b y_b$, $x_b z_b$, and $y_b z_b$, $x_b^2 - y_b^2$, $3z_b^2 - r_b^2$) have the same angular behavior as the five real 3d AOs; sixth ($x_b^2 + y_b^2 + z_b^2 = r_b^2$) is like a 3s function; sometimes omitted from the basis set.

there are ten f-type Gaussians, and these could be combined to have the angular behavior of the seven real 4f AOs.

The form of linear combinations $\quad N r_b^l e^{-\alpha r_b^2}(Y_l^{m*} \pm Y_l^m)/2^{1/2}$

the principal quantum number n is absent:

s AO (1s or 2s or ...) ≡ a LC of several Gaussians with different α, each Gaussian having the form exp $(-\alpha r_b^2)$

p_x orbital ≡ a LC of Gaussians, each of the form x_b exp $(-\alpha r_b^2)$ and so on.

The Cartesian Gaussians form a complete set.

spherical Gaussians:

$$N r_b^{n-1} e^{-\alpha r_b^2} (Y_l^{m*} \pm Y_l^m)/2^{1/2}$$

A GF does not have the desired cusp at the nucleus (a poor representation of AO for small r_b)
represent an AO as a LC of several GFs.
an LC-GTF SCF MO calculation involves many more integrals than the corresponding LC-STO SCF MO calculation
the number of two-electron integrals $\propto b^4$.
integral evaluation : much less computer time (by G)
Why? the product of two GFs centered at two different points is equal to a single GF centered at a third point (three- and four-center two-electron repulsion integrals are reduced to two-center integrals).

STO basis sets:

minimal (or **minimum**) **basis set** : one STO for each inner-shell and valence-shell AO of each atom

C_2H_2 : 1s, 2s,2p$_x$, 2p$_y$, and 2p$_z$ AOs on each C and a 1s STO on each H; five STOs on each C and one on each H = a total of 12 basis functions. two s-type STOs and one set of p-type STOs on each carbon and one s-type STO on each hydrogen; is denoted (2s1p) for the C functions and (1s) for the H functions; notation (2s1p/1s),

minimal STO set for first part of the periodic table:

	H, He	Li–Ne	Na–Ar	K, Ca	Sc–Kr
numbers of basis functions	1	5	9	13	18

A **double-zeta (DZ) basis set** : replaces each STO of a minimal basis set by two STOs that differ in ζ

C2H2 : 2 1s STOs on each H, 2 1s STOs, 2 2s STOs, 2 $2p_x$, 2 $2p_y$, and 2 $2p_z$ STOs on each carbon (24 basis functions; (4s2p/2s))

the number of variational parameters c_{ri} in a DZ-BS wf is twice that in a M-BS wf.

A **triple-zeta (TZ) basis set** : replaces each STO of a M-BS by three STOs that differ in ζ.

A **split-valence (SV) basis set** : two (or more) STOs for each valence AO but only one STO for each inner-shell (core) AO.

An SV-BS is minimal for inner-shell AOs and DZ (or TZ or ...) for the valence AOs (VDZ, VTZ,...)

AOs are distorted in shape upon molecule formation (polarization)

Thus, one adds BF STOs whose $\ell > \ell_{max}$ of the valence shell of the GS atom (polarized (P) basis set)

double-zeta plus polarization set (DZ + P or DZP) :

DZ set + five 3d functions on each "first- and second-row" atoms + three 2p functions ($2p_x$, $2p_y$, $2p_z$) on each H atom. (In QC, Li-Ne are called the first-row)

$C_2H_5OSiH_3$:is designated as (6s4p1d/4s2p1d/2s1p)

$\qquad\qquad\qquad\qquad\qquad\qquad$ Si $\qquad$ C, O $\qquad$ H

To increase accuracy, higher-ℓ polarization functions can be added.

Gaussian-basis-sets

Instead of using the individual GFs as BFs, take each BF as a normalized LC
of a few Gaussians:

$$\chi_r = \sum_u d_{ur} \, g_u$$

g_u's : normalized Cartesian Gaussians centered on the same atom and
having the same i, j, k , but different α's.
d_{ur} : contraction coefficients (constants)
χ_r : contracted CGTF
g_u's : primitive Gaussians.
using CG $\Rightarrow$ number of variational coefficients $\downarrow \Rightarrow$ computational time $\downarrow$
(little loss in accuracy if d_{ur}s are well chosen)
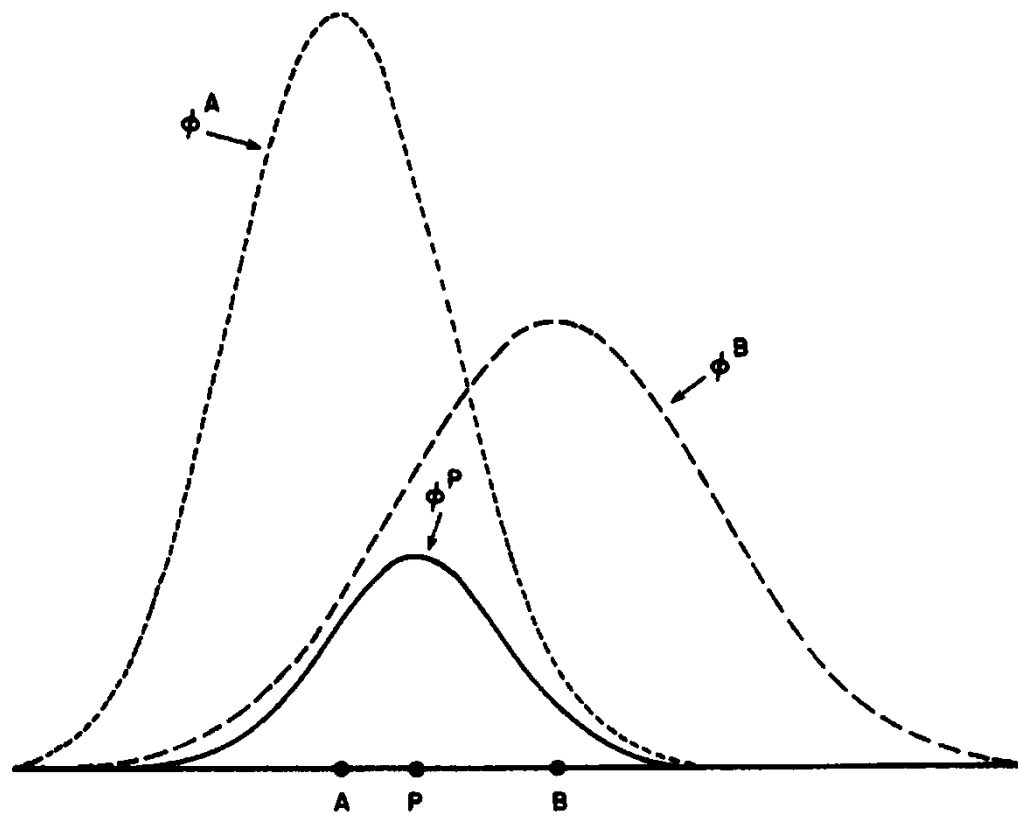
The product of two 1s Gaussians is a third 1s Gaussian.

The classifications given for STO BSs also apply to CGTF BSs

A minimal basis set : consists of one CGF for each inner-shell AO and for each valence-shell AO

A DZ basis set : two CGTFs for each such AO

A DZP set : adds CGs with higher ℓ to the DZ set, where $\ell = i + j + k$

αs and d_{ur}s of the BFs are kept fixed at the predetermined values

CGTF M-BS : no way for the BFs to adjust their sizes to differing molecular environments.

DZ BS : allow the AO sizes to vary from one molecule to another.

Example:

1s' and 1s" be two 1s CGTFs centered on a certain H atom

1s' (and 1s") be a linear combination of a few s-type PGs

Let for α in the primitives 1s' > Is". 1s" is spread out over a much larger region of space than Is'.

expression for a given MO contains $c_1$1s' + c_2Is"; c_1 and c_2 are found by the SCF process. The size of the function will increase as the ratio c_2/c_1 increases.

polarization functions vary the AO shapes, shifting charge density away from the nuclei and into the bonding regions in the molecule.

Example
adding 2p functions to a 1s on H
$c_1 1s + c_2 2p_x + c_3 2p_y + c_4 2p_z$
AO will be polarized in a direction determined by c_2, c_3, and c_4.
$c_3 = 0$, $c_4 = 0$, $c_1 > 0$, $c_2 > 0$. $c_2 2p_x$ will cancel some of the probability density of the $c_1 1s$ term on one side of the H atom and will augment it on the other side, thereby polarizing the 1s function in the positive x direction.

p-type AOs can be polarized by mixing in d-type AOs.

methods to form CG sets:
1) Minimal CGTF sets are often formed by fitting STOs.
a) one STO per AO, b) Each STO is approximated as a linear combination of N GFs, (coefficients and exponents)
N = 3 → STO-3G (H through Xe) is not quite as good as a MBS STO
    (6s3p/3s) contracted to [2s1p/1s]   (2s1p/1s)
parentheses for primitive Gaussians
brackets for contracted Gaussians.

1s STO having $\zeta = 1$ $S(r; 1) = \pi^{-1/2}e^{-r}$

normalized s-type Gaussian $(2\alpha/\pi)^{3/4}e^{-\alpha r^2}$

$$G_{3N}(r; 1) = c_1(2\alpha_1/\pi)^{3/4}e^{-\alpha_1 r^2} + c_2(2\alpha_2/\pi)^{3/4}e^{-\alpha_2 r^2} + c_3(2\alpha_3/\pi)^{3/4}e^{-\alpha_3 r^2}$$

$c_1 = 0.444615, c_2 = 0.535336, c_3 = 0.154340, \alpha_1 = 0.109814, \alpha_2 = 0.40575, \alpha_3 = 2.22746$

to fit a 1s STO with orbital exponent ζ

For $G_{3N}(r; 1)$, replacing each orbital exponent α_i by $\zeta^2\alpha_i$

ζ is called a scale factor.
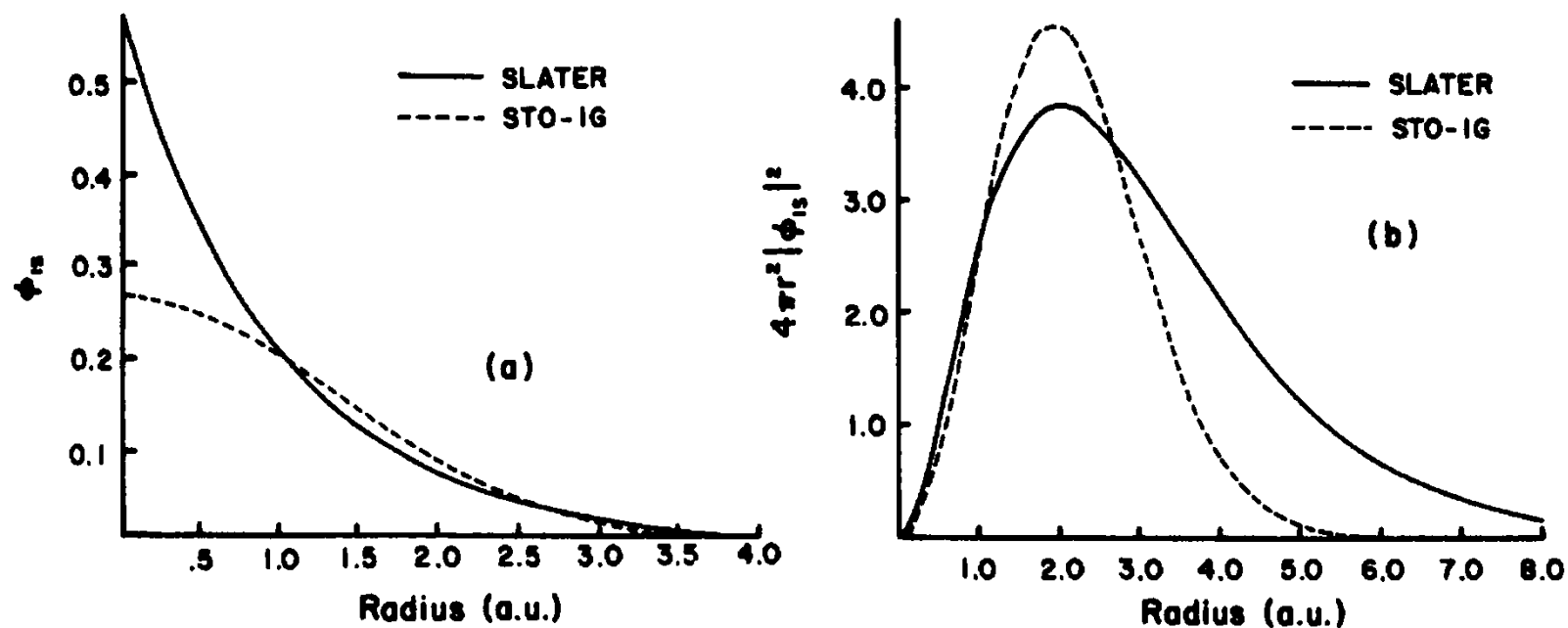
$$\phi_{1s}^{CGF}(\zeta = 1.0, \text{STO-1G}) = \phi_{1s}^{GF}(0.270950)$$

$$\phi_{1s}^{CGF}(\zeta = 1.0, \text{STO-2G})$$
$$= 0.678914\phi_{1s}^{GF}(0.151623) + 0.430129\phi_{1s}^{GF}(0.851819)$$

$$\phi_{1s}^{CGF}(\zeta = 1.0, \text{STO-3G}) = 0.444635\phi_{1s}^{GF}(0.109818) + 0.535328\ \phi_{1s}^{GF}(0.405771)$$

$$+ 0.154329\phi_{1s}^{GF}(2.22766)$$
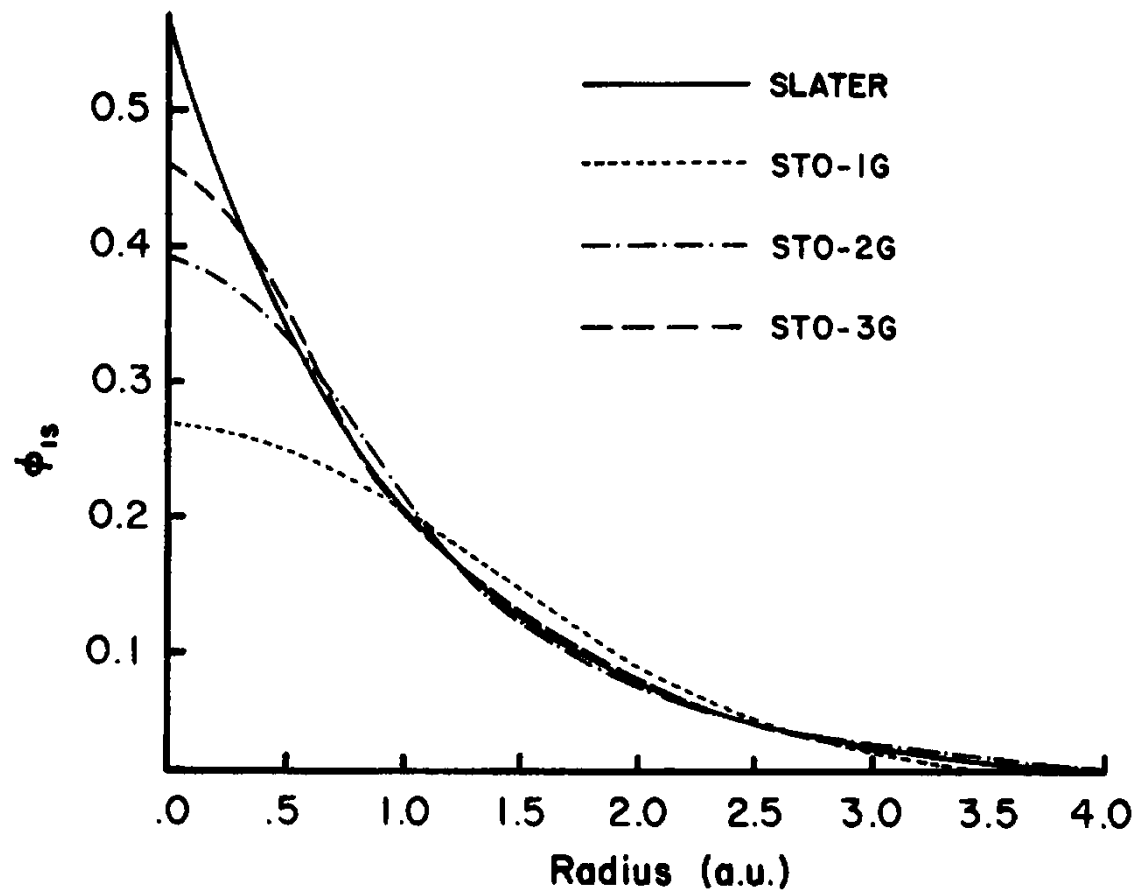
Comparison of the quality of the least-squares fit of a 1s Slater function ($\zeta = 1.0$) obtained at the STO-1G, STO-2G, and STO-3G levels.

2) start with atomic GTF SCF calculations
Huzinaga used a (9s5p) basis set of uncontracted Gaussians to do SCF calculations
on the atoms Li-Ne.

Example: the ground state of the O atom for the nine s-type basis GTFs:

orbital exponents:

g_1	g_2	g_3	g_4	g_5	g_6	g_7	g_8	g_9
7817	1176	273.2	81.2	27.2	9.53	3.41	0.940	0.285

diffuse function

expansion coefficients for the 1s SCF AO

g_1	g_2	g_3	g_4	g_5	g_6	g_7	g_8	g_9
0.0012	0.009	0.043	0.144	0.356	0.461	0.140	−0.0006	0.001

expansion coefficients for the 2s SCF AO

g_1	g_2	g_3	g_4	g_5	g_6	g_7	g_8	g_9
−0.0003	−0.002	−0.010	−0.036	−0.095	−0.196	−0.037	0.596	0.526

a split-valence [3s2p] set of contracted GTFs for O:

1s:

the g_1, g_2, g_3, g_4, g_5, and g_7 coefficients:1s AO >> 2s AO,

g_8 and g_9 coefficients 2s AO >> 1s AO,

g_6 : substantial contributions to both 1s and 2s.

$$1s = N(0.0012g_1 + 0.009g_2 + 0.043g_3 + 0.144g_4 + 0.356g_5 + 0.461g_6 + 0.140g_7)$$

$\uparrow$

because g_8 and g_9 have been omitted

2s:

we need two BFs for the 2s AO (from g_6, g_8, and g_9)

g_9 is called a diffuse function

The outer region of an AO changes the most upon molecule formation

(take g_9 as one of the basis)

$$2s = N'(-0.196g_6 + 0.596g_8), \qquad 2s' = g_9$$

The 2p and 2p' CGTFs can be formed similarly

3-21G set (H - Xe) and 6-31G set (H - Zn) :
SV basis sets of CGTFs.
3-21G:
inner-shell AO: (1s for Li-Ne; 1s, 2s, $2p_x$, $2p_y$, $2p_z$ for Na-Ar; and so on) $\rightarrow$ single CGTF (a LC of three PGs)
valence-shell AO: (1s for H; 2s and the 2p's for Li-Ne;...; 4s and the 4p's for K, Ca, Ga-Kr; 4s, the 4p's, and the five 3d's for Sc-Zn) $\rightarrow$ two basis functions (one CGTF that is a LC of two GP and one which is a single diffuse Gaussian.
6-31G:
inner-shell CGTF $\rightarrow$ six primitives
valence-shell AO $\rightarrow$ one CGTF (with 3 primitives) + one Gaussian (with one primitive)

α and d_{ur} were determined to minimize the SCF energies of atoms.
in 3-21G, α for H in an atomic calculation are increased using a scale factor
in 6-31G, valence orbital exponents of H and of Li through O are scaled to be more appropriate for molecular calculations.

6-31G* and 6-31G** [or 6-31G(d) and 6-31G(d,p)] (H - Zn)
VDZ polarized basis sets:

6-31G* = 6-31G + six d-type Cartesian-Gaussian polarization functions
on Li - Ca and ten f-type Car-GPFs on Sc - Zn.
6-31G** = 6-31G* + three p-type GPFs on H & He.
α of PFs were determined as the average of the optimum values found in
calculations on small molecules.

P atom:
6-31G* : 19 BFs centered on it (Is, 2s, $2p_x$, $2p_y$, $2p_z$, 3s, 3s', $3p_x$, $3p_y$, $3p_z$,
$3p'_x$, $3p'_y$, $3p'_z$, and six d's) and is [4s3p1d] for P.

3-21G$^{(*)}$ (H - Ar)
For second-row atoms, d orbitals contribute significantly to the bonding.
3-21G$^{(*)}$ (H - Ar) = 3-21G + six d-type GFs on each second-row atom.
For H-Ne, 3-21G$^{(*)}$ = 3-21G .

3-21+G, 6-31+G*, 3-21++G and 6-31++G*

Anions, compounds with lone pairs, and hydrogen-bonded dimers have significant electron density at large distances from the nuclei.
3-21+G (and 6-31+G*) = 3-21G (and 6-31G*) + four highly diffuse functions (s, p_x, p_y, p_z) on each non-hydrogen atom; very small α (0.01 to 0.1)
3-21++G (and 6-31++G*) = 3-21+G (and 6-31+G*) + a highly diffuse s function on each H atom.

cc-pVXZ basis sets (a CGTF BS; Dunning and co-workers):
(cc-pVDZ, cc-pVTZ, cc-pVQZ, and cc-pV5Z)
for use in methods (such as CI) that include ec.
cc-pVDZ (correlation consistent, polarized VDZ).
first-row atoms: cc-pVDZ is [3s2p1d]; cc-pVTZ is [4s3p2d1f].

cc-pCVXZ = cc-pVXZ + certain primitive Gaussians (CV stands for core/valence).
for calculations that include correlation effects involving the core electrons.

augmented sets aug-cc-pVDZ = cc-pVXZ + diffuse nonpolarization
for correlation calculations on anions and H-bonded species.

Example:

On the basis-set order form, choose the 3-21G basis set and enter O in the Elements box (Gaussian 94)

oxygen atom

can be ignored

a scale factor

O

0

s-type CGTF consisting of 3 PGs   S   3   1.00
                                                                                              large [inner-shell
                                   322.03700000   0.05923940   (core) 1s AO]
                                   48.43080000   0.35150000
                                   10.42060000   0.70765800

s-type and p-type CGTFs
valence 2s and 2p AOs   SP   2   1.00
                                   7.40294000   −0.40445300   0.24458600
                                   1.57620000   1.22156000   0.85395500
                                   SP   1   1.00
                                   0.37368400   1.00000000   1.00000000

α d_{ur} d_{ur}

assures us that the primitive Gaussians are normalized

1s CGTO

$$1s = 0.0592394g_s(322.037) + 0.3515g_s(48.4308) + 0.707658g_s(10.4206)$$

a normalized primitive s-type GTF with α = 322.037

The 3-21G set uses the same orbital exponents for the 2s and 2p AOs, so as to speed up calculations.

the valence CGTFs

$$2s' = -0.404453g_s(7.40294) + 1.22156g_s(1.5762)$$

$$2p'_x = 0.244586g_{p_x}(7.40294) + 0.853955g_{p_x}(1.5762), \ldots,$$

$$2s'' = g_s(0.373684)$$

$$2p''_x = g_{p_x}(0.373684), \ldots,$$

```
!  6-311++G(2d,2p)   EMSL  Basis Set Exchange Library   11/11/27 0:56
! Elements                                 References
! --------                                 ----------
! H, Li - Ne: R. Krishnan, J.S. Binkley, R. Seeger and J.A. Pople,
!                J. Chem. Phys. 72, 650 (1980)
! Na - Ar:    A.D. McLean and G.S. Chandler J. Chem. Phys. 72, 5639, (1980).
! K  - Ca:    J-P. Blaudeau, M. P. McGrath, L.A. Curtiss and L. Radom,
!                J. Chem. Phys. 107, 5016 (1997).
! Ga - Kr:    L. A. Curtiss, M. P. McGrath, J-P. Blandeau, N. E. Davis,
!                R. C. Binning, Jr. L. Radom, J. Chem. Phys. 103, 6104 (1995).
! I      :    M.N. Glukhovstev, A. pross, M.P. McGrath, L. Radom, J. Chem. Phys
!                103, 1878 (1995)
!
!
! Elements                                 References
! --------                                 ----------
! H-Ne: M.J. Frisch, J.A. Pople and J.S. Binkley, J. Chem. Phys. 80, 3265 (1984
!
!
! Elements                                 Reference
! --------                                 ----------
! H, Li-Cl: T. Clark, J. Chandrasekhar,  G.W. Spitznagel, P.V.R. Schleyer,
!                J. Comp. Chem. 4, 294 (1983).
!
!


****
O     0
S   6    1.00
   8588.5000000              0.00189515
   1297.2300000              0.0143859
    299.2960000              0.0707320
     87.3771000              0.2400010
     25.6789000              0.5947970
      3.7400400              0.2808020
SP    3    1.00
     42.1175000              0.1138890           0.0365114
      9.6283700              0.9208110           0.2371530
      2.8533200             -0.00327447          0.8197020
SP    1    1.00
      0.9056610              1.0000000           1.0000000
SP    1    1.00
      0.2556110              1.0000000           1.0000000
SP    1    1.00
```

6-311++G(2d,2p)

O   0
S 6 1.00
   8588.5000000   0.00189515
   1297.2300000   0.0143859
    299.2960000   0.0707320
     87.3771000   0.2400010
     25.6789000   0.5947970
      3.7400400   0.2808020
SP 3 1.00
     42.1175000   0.1138890   0.0365114
      9.6283700   0.9208110   0.2371530
      2.8533200   -0.00327447   0.8197020
SP 1 1.00
      0.9056610   1.0000000   1.0000000
SP 1 1.00
      0.2556110   1.0000000   1.0000000
SP 1 1.00
      0.0845000   1.0000000   1.0000000
D 1 1.00
      2.5840000   1.0000000
D 1 1.00
      0.6460000   1.0000000

POPULATION ANALYSIS

SCF wave functions can be analysied by population analysis, introduced by
Muliken [**Mulliken population analysis (MPA)**]
n-electrons $\rightarrow$ n_r + $n_{r\text{-}s}$
n_r : net populations in the basis functions χ_r
$n_{r\text{-}s}$: overlap populations for all possible pairs of basis functions

basis set $\{\chi_i\}$

$$\phi_i = \sum_s c_{si}\chi_s = c_{1i}\chi_1 + c_{2i}\chi_2 + \cdots + c_{bi}\chi_b$$

For simplicity: c_{si}'s and χ_{si}'s are real

The probability density associated with one electron in ϕ_i is

$$|\phi_i|^2 = c_{1i}^2\chi_1^2 + c_{2i}^2\chi_2^2 + \cdots + 2c_{1i}c_{2i}\chi_1\chi_2 + 2c_{1i}c_{3i}\chi_1\chi_3 + 2c_{2i}c_{3i}\chi_2\chi_3 + \cdots$$

Integrating (ϕ_i and χ_s's are normalized)

$$1 = c_{1i}^2 + c_{2i}^2 + \cdots + 2c_{1i}c_{2i}S_{12} + 2c_{1i}c_{3i}S_{13} + 2c_{2i}c_{3i}S_{23} + \cdots$$

overlap integrals

Muliken proposed:

$c_{1,i}^2$: contribution of one electron (in ϕ_i) to the net population in χ_1

$c_{2,i}^2$: contribution of one electron (in ϕ_i) to the net population in χ_2

...

$2c_{1i}c_{2i}S_{12}$: contribution of one electron (in ϕ_i) to the overlap population between χ_1 and χ_2

$2c_{1i}c_{3i}S_{13}$: contribution of one electron (in ϕ_i) to the overlap population between χ_1 and χ_3

For n_i electrons in the MO ϕ_i $\{n_i = 0,1,2)$:
Let,

$n_{ri} \equiv$ contribution of electrons (in ϕ_i) to the net population in χ_r

$n_{r\text{-}si} \equiv$ contribution of electrons (in ϕ_i) to the overlap population between χ_r and χ_s

$$n_{r,i} = n_i c_{ri}^2, \qquad n_{r\text{-}s,i} = n_i(2c_{ri}c_{si}S_{rs})$$

$\downarrow$ summing over the occupied MOs,

$$n_r = \sum_i n_{r,i} \quad \text{and} \quad n_{r\text{-}s} = \sum_i n_{r\text{-}s,i}$$

$$\Sigma_r n_r + \Sigma_{r>s} \Sigma_s n_{r\text{-}s} = n \quad \text{total number of electrons}$$

Example:

calculate the net and overlap population contributions for the H_2O

$$2a_1 = -0.027(O1s) + 0.820(O2s_\perp) + 0.132(O2p_z) + 0.152(H_11s + H_21s)$$

find n_r for each basis function. Use H_11s and H_21s as basis functions, rather than the symmetry-adapted basis functions.

Overlap integrals between basis STOs centered on different atoms can be found by interpolation in the tables of R. S. Muliken et al., J. Chem. Phys.,11,1248 (1949). (see Problem 15.21)

$$\langle H_11s|O1s \rangle = \langle H_21s|O1s \rangle = 0.054, \quad \langle H_11s|O2s_\perp \rangle = \langle H_21s|O2s_\perp \rangle = 0.471$$

$$\langle H_11s|O2p_y \rangle = -\langle H_21s|O2p_y \rangle = 0.319, \quad \langle H_11s|O2p_z \rangle = \langle H_21s|O2p_z \rangle = 0.247$$

$$\langle H_11s|H_21s \rangle = 0.238$$

net populations:

$$n_{O1s,2a_1} = 2(-0.027)^2 = 0.0015, \, n_{O2s_1,2a_1} = 2(0.820)^2 =$$

$$1.345, \, n_{O2p_z,2a_1} = 0.035, \, n_{H_11s,2a_1} = 2(0.152)^2 = 0.046, \, n_{H_21s,2a_1} = 0.046$$

The five lowest SCF MOs found by Pitzer and Merrifield at the experimental geometry are

$$1a_1 = 1.000(O1s) + 0.015(O2s_\perp) + 0.003(O2p_z) - 0.004(H_11s + H_21s)$$

$$2a_1 = -0.027(O1s) + 0.820(O2s_\perp) + 0.132(O2p_z) + 0.152(H_11s + H_21s)$$

$$1b_2 = 0.624(O2p_y) + 0.424(H_11s - H_21s)$$

$$3a_1 = -0.026(O1s) - 0.502(O2s_\perp) + 0.787(O2p_z) + 0.264(H_11s + H_21s)$$

$$1b_1 = O2p_x$$

The $O_{2s\perp}$ orbital is an orthogonalized orbital

$$O2s_\perp = 1.028[O2s - 0.2313(O1s)]$$

overlap populations

$$n_{O1s-H_1 1s, 2a_1} = 2(2)(-0.027)(0.152)(0.054) = -0.0009 = n_{O1s-H_2 1s, 2a_1}$$

$$n_{O2s_- - H_1 1s, 2a_1} = 0.235 = n_{O2s_1 - H_2 1s, 2a_1}$$

$$n_{O2p_z - H_1 1s, 2a_1} = 0.020 = n_{O2p_z - H_2 1s, 2a_1}, \quad n_{H_1 1s - H_2 1s, 2a_1} = 0.022$$

net population of O1s:

$$n_{O1s} = 2(1.000)^2 + 2(-0.027)^2 + 2(-0.026)^2 = 2.00$$

$$n_{O2s_\perp} = 1.85$$

$$n_{O2p_x} = 2.00$$

$$n_{O2p_y} = 0.78$$

$$n_{O2p_z} = 1.27, \quad n_{H_1 1s} = 0.54_5, \quad n_{H_2 1s} = 0.54_5.$$

The MO ϕ_i in a covalent molecule:

$S = \sum n_{r\text{-}s,i}$ (χ_r and χ_s lie on different atoms)

$S > 0$ (substantially), the MO is bonding

$S < 0$ (substantially), the MO is antibonding

$S \approx 0$, the MO is nonbonding

for the $3a_1$ MO
overlap of,

O1s with $H_1$1s : $2(2)(-0.026)(0.264)(0.054) = -0.0015$

O2s $_\perp$ with H_2 1s $2(2)(-0.502)$ X $(0.264)(0.471) = -0.250$,

O2p_z with $H_1$1s : 0.205

O2p_z with $H_2$1s : 0.205

$H_1$1s with $H_2$1s : $2(2)(0.264)(0.238) = 0.066$

$\sum = -0.03$ for the $3a_1$ MO $\approx$ zero $\rightarrow$ a nonbonding (lone-pair) MO

$\sum = 0.53$ for the $2a_1 \rightarrow$ bonding

$\sum = 0.50$ for the $1b_2 \rightarrow$ bonding

$\sum = 0.00$ for $1a_1$ (inner-shell)

1) apportioning the electrons into net populations in BFs and overlap populations for pairs of BFs,
2) apportioning the electrons among the basis functions only, with no overlap populations. Muliken proposed : splitting n_{r-s} equally between the BFs x_r and X_c

$$N_r = n_r + \frac{1}{2} \sum_{s \neq r} n_{r-s}$$ **gross population**

$$\sum_{r=1}^{b} N_r = n$$ number of electrons in the molecule

Example:
the contribution to the gross population of $O2s_\perp$ from the $2a_1$ MO

$$N_{O2s_\perp, 2a_1} = 2[(0.820)^2 + (0.820)(0.152)(0.471) + (0.820)(0.152)(0.471)] = 1.58$$

$$N_{O2s2s_\perp, 1a_1} = 0.00 \quad N_{O2s2s_\perp, 3a_1} = 0.25 \quad N_{O2s2s_\perp, 1b_1} = 0.00$$

$$N_{O2s2s_\perp, 1b_2} = 0.00$$

$$N_{O2s2s_\perp} = 1.83$$

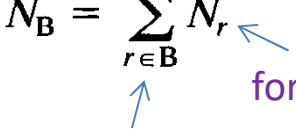$$N_{O1s} = 2.00 + 0.00 + 0 + 0.00 + 0 = 2.00 \qquad N_{O2s_\perp} = 1.83; N_{O2p_x} = 0 + 0 + 0 + 0 + 2 = 2$$

$$N_{O2p_y} = 0 + 0 + 1.12 + 0 + 0 = 1.12; N_{O2p_z} = 0 + 0.05_5 + 0 + 1.44_5 + 0 = 1.50$$

$$N_{H_1 1s} = 0.00 + 0.18_4 + 0.44_2 + 0.150 + 0 = 0.77_6; N_{H_2 1s} = 0.77_6$$

gross atomic population N_B for atom B:

$$N_B = \sum_{r \in B} N_r$$

all BFs centered on atom B

for basis function χ_r

number of electrons in the molecule

$$\sum N_B = n$$

net atomic charge q_B on atom B :

$$q_B \equiv Z_B - N_B$$

One should not put too much reliance on numbers calculated by population analysis.

1) assignment of ½ n_{rs} to each basis function is arbitrary and sometimes leads to unphysical results

2) a small change in BS can produce a large change in the calculated net charges.

For example:
net atomic charges on each H atom:

	CH_4	NH_3	H_2O	HF	
STO-3G	0.06	0.16	0.18	0.21	Comparison ?
3-21G	0.20	0.28	0.36	0.45	

Comparison

An improvement on MPA is **natural population analysis (NPA)**

uses ideas related to natural orbitals

1) one calculates a set of orthonormal natural atomic orbitals (NAOs) from $\{\chi_r\}$
2) NAOs $\rightarrow$ a set of orthonormal natural bond orbitals (NBOs), (core, lone pair, or bond orbital)
3) population analysis by NBOs. q_{NPA}s show less BS dependence than q_{MPA}s.

Other methods of assigning net atomic charges are discussed in the next section.

in view of the existence of improved methods, MPA should no longer be used [S. M. Bachrach in K. Lipkowitz and D. B. Boyd (eds.), Reviews in Computational Chemistry, vol. 5, VCH (1994), Chapter 3.]