

Session 2

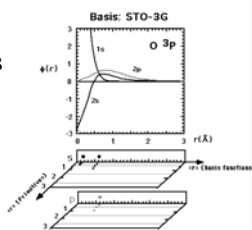
Basis Sets

CCCE-2008

1

Session 2: Basis Sets

- Two of the major methods (*ab initio* and DFT) require some understanding of basis sets and basis functions
- This session describes the essentials of basis sets:
 - What they are
 - How they are constructed
 - How they are used
 - Significance in choice



The r_i for each basis function is indicated by the position of a colored dot along a horizontal axis. Basis functions may consist of individual Gaussian primitives or a linear combination of Gaussians. In the latter case, the r_i for the underlying Gaussians will be indicated by a row of dots which appear to come out of the page.

CCCE-2008

2

Running a Calculation

- In performing *ab initio* and DFT computational chemistry calculations, the chemist has to make several decisions of input to the code:
 - The molecular geometry (and spin state)
 - The basis set used to determine the wavefunction
 - The properties to be calculated
 - The type(s) of calculations and any accompanying assumptions

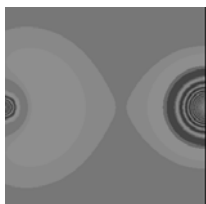
CCCE-2008

3

Running a calculation

- For *ab initio* or DFT calculations, many programs require a basis set choice to be made

- The basis set is an approximate representation of the atomic orbitals (AOs)
- The program then calculates molecular orbitals (MOs) using the Linear Combination of Atomic Orbitals (LCAO) approximation



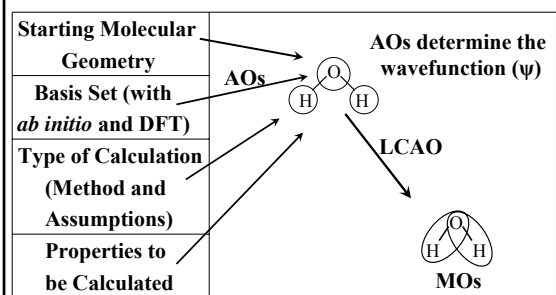
CCCF-2008

4

Computational Chemistry Map

Chemist Decides:

Computer calculates:



CCCF-2008

5

Critical Choices

- Choice of the method (and basis set) used is critical
 - Which method?
 - Molecular Mechanics, *Ab initio*, Semiempirical, or DFT
 - Which approximation?
 - MM2, MM3, HF, AM1, PM3, or B3LYP, etc.
 - Which basis set (if applicable)?
 - Minimal basis set
 - Split-valence
 - Polarized, Diffuse, High Angular Momentum,

CCCF-2008

6

Why is Basis Set Choice Critical?

- The basis set needs to be able to approximate the actual wave function sufficiently well to give chemically meaningful results
 - Also needs a reasonable computational “cost”
 - Integrals should be evaluated quickly and accurately
- Trade-offs
 - Choice always involves a balance between accuracy and computational cost
 - More accurate methods using larger basis sets will take more computer time

CCCE 2008

7

Theoretical Models

- Goal of computational chemistry is to mathematically represent chemical reality
 - Improving the basis set and the degree of electron correlation improves the ability of the computational model to approach reality
- Ultimate goal is an exact solution of the Schrödinger equation

CCCE 2008

8

Comparison of Some Methods for Accuracy

		DFT Location??						
		HF	MP2	MP3	MP4	QCISD(T)	...	Full CI
BASIS SET ↓	Minimal STO-3G		Electron Correlation →					
	Split Valence 3-21G							
	Polarized 6-31G(d)							
	6-311G(d,p)							
	Diffuse 6-311+G(d,p)							
	High ang. Momentum 6-311+G(2d,p)							
	6-311++G(3df,3dp)							
.....								
∞	HF Limit							Schrödinger Equation

CCCE 2008

9

Possible Basis Functions

1. Hydrogen-like Orbitals

Derived for a *one-electron* atom

- Not truly accurate for a *many-electron* atom

Form: $\Psi(r, \theta, \phi) = R(r)Y_l^m(\theta, \phi)$

$R(r)$ = radial function

Y_l^m = spherical harmonic

- + Advantages: Mutually orthogonal
- Disadvantages: Complex form is awkward for calculations; Most atoms of interest > one electron

CCCF-2008

10

Possible Basis Functions

- True wavefunction should be antisymmetric to electron interchange – use spin orbitals
 - Antisymmetric linear combination of products of spin orbitals used in an SCF calculation
 - **HF-SCF calculation**
 - Numerical methods were originally used to solve and find the Hartree-Fock orbitals
- Roothaan: Represent the HF orbitals as linear combinations of a set of known (basis) functions
 - Commonly used set of basis functions for atomic HF calculations is the set of Slater-type orbitals (STOs)

CCCF-2008

11

Possible Basis Functions

2. Slater Type Orbitals (STOs)

- Normalized form:

$$\mu(r, \theta, \phi) = \frac{(2\zeta/a_0)^{n+0.5}}{[(2n)!]^{0.5}} r^{(n-1)} e^{(-\zeta r/a_0)} Y_l^m(\theta, \phi)$$

where n, m, and l are integers and ζ (orbital exponent) is a variational parameter

- Improve results by using a linear combination of several STOs to represent each HF orbital
- HF-SCF atomic calculations require lots of computation
 - Hartree did this numerically in the 1930's

CCCF-2008

12

Slater Type Orbitals

- + Advantages: Have a complete set
 - Radial behavior closely matches hydrogenic orbitals
 - Disadvantages:
 - No nodes, as with H-like orbitals
 - Not mutually orthogonal
 - For larger molecules, computer evaluation of the many integrals involved is quite time consuming
- Need to reduce the computational cost

CCCE 2008

13

Possible Basis Functions

3. Gaussian Type Orbitals (GTOs)

Proposed by S.F. Boys in 1950

GTO form:

$$g(r, \theta, \phi) = \left[\frac{2^{(2n+1.5)}}{(2n-1)! \sqrt{\pi}} \right]^{0.5} \zeta^{(2n+1)/4} r^{(n-1)} e^{-\alpha r^2} Y_l^m(\theta, \phi)$$

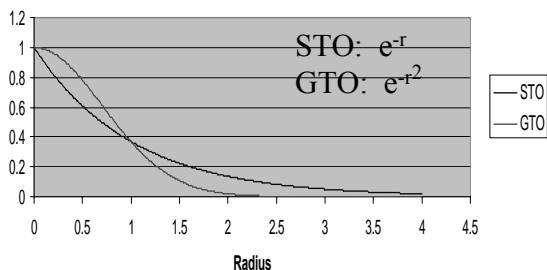
- + Advantages: Have a complete set
 - Computer evaluation of integrals much faster
 - Closed integrals; Integrated GTO gives a GTO
- Disadvantages: Not mutually orthogonal
 - Representation of e- probability is poor near and far away from the nucleus

CCCE 2008

14

Comparison

STO vs. GTO



CCCE 2008

15

GTOs

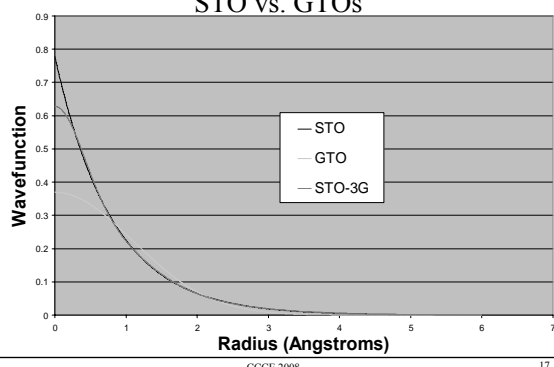
- Linear combinations of GTO's are used to approximate STOs (which are themselves approximations)
 - A single GTO basis function has significant errors when compared to a STO, especially near the nucleus (See previous slide)
 - If several GTOs are combined in a linear combination, the basis function is greatly improved
 - See next slide

CCCE 2008

16

Comparisons

STO vs. GTOs



CCCE 2008

17

Use of GTOs

- Individual GTOs not used as basis functions:
 - Use a normalized linear combination of a few GTOs (called *primitives*), each with different α values to give a “contracted” Gaussian function

$$g_c = \sum c_i g_p$$

where g_c is a contracted gaussian,
 g_p is a primitive gaussian, and c_i is
a contraction coefficient

- A linear combination of these primitives (typically 1-7) is used to approximate the STO

CCCE 2008

18

Use of GTOs

- Using contracted GTOs instead of primitive GTOs as the basis set has advantages:
 - Number of variational coefficients to be determined is reduced, which saves a lot of computational time
 - Accuracy is NOT reduced significantly, as long as the contraction coefficients (c_i 's) are well chosen
- Increasing the number of primitive GTOs used in each contracted Gaussian improves the accuracy
- Different types of basis sets use different numbers and types of GTOs

CCCE 2008

19

Minimal Basis Sets

- Jargon Used → **STO-NG** (Single ζ)
 - N is the number of primitive GTOs used
 - Example: STO-3G
 - Three primitive GTOs used per AO
 - Popular starting point for calculations
 - STO-3G basis functions have been developed for most of the elements in the Periodic table
- Minimal basis sets do not adequately describe non-spherical (anisotropic) electron distribution in molecules (as in polar covalent bonds)

CCCE 2008

20

Minimal Basis Sets

- GTO representation of a STO for $1s$ AO:

$$\Phi_{STO}^{1s}(r) = \pi^{-0.5} \zeta^{1.5} e^{-\zeta r} \quad \text{where } \zeta = 1$$

$$\Phi_{STO-3G}^{1s}(r) = c_1 \left(\frac{2\alpha_1}{\pi} \right)^{3/4} e^{-\alpha_1 r^2} + c_2 \left(\frac{2\alpha_2}{\pi} \right)^{3/4} e^{-\alpha_2 r^2} + c_3 \left(\frac{2\alpha_3}{\pi} \right)^{3/4} e^{-\alpha_3 r^2}$$

where $c_1 = 0.444615, c_2 = 0.535336, c_3 = 0.154340$
and $\alpha_1 = 0.109818, \alpha_2 = 0.405771, \alpha_3 = 0.222766$

- “ c ” values are called the **contraction coefficients**
- The exponents are the alpha (α) values

CCCE 2008

21

Split Valence Basis Sets

- Jargon Used → **K-LMG** (Double ζ)
 - Differentiate between core and valence electrons
 - Developed to overcome problems of inadequate description of anisotropic electron distributions using minimal basis sets (Size is adjusted)
- K = number of sp-type inner shell primitive GTOs
L = number of inner valence s- and p-type primitive GTOs
M = number of outer valence s- and p-type primitive GTOs

CCCF-2008

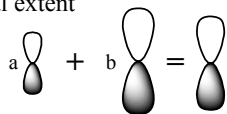
22

Split Valence Basis Sets

- Each split valence atomic orbital is composed of a variable proportion of two (or more) functions of different size or radial extent

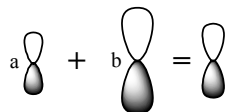
For a larger e^- cloud
(longer bond)

$$a < b$$



For a smaller e^- cloud
(shorter bond)

$$a > b$$



→ **a** and **b** are normalized and sum to 1

CCCF-2008

23

Split Valence Basis Sets

- Examples:
 - 3-21G** (Used as the semiempirical basis set)
 - Three** primitives for the inner shell (STO-3G); each valence orbital is constructed with **two** sizes of basis function (*Two* GTOs for contracted valence orbitals; *One* GTO for extended valence orbitals)
 - 6-311G**
 - STO-6G for inner shell; **Three** sizes of basis function for each valence orbital (*Three* GTOs for contracted valence orbitals, and *two different sizes* of GTO for extended valence orbitals)

CCCF-2008

24

Polarized Basis Sets

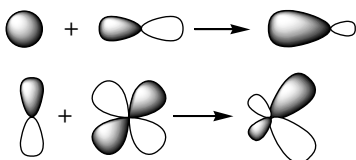
- Jargon Used → 6-31G(d) or 6-31G* (older)
 - Also have 6-31G(d,p) or 6-31G**
 - (d) or * type
 - d-type functions added to atoms with $Z > 2$
 - f-type functions added to transition metals
 - (d,p) or ** type
 - p-type functions added to H atoms
 - d-type functions added to atoms with $Z > 2$
 - f-type functions added to transition metals
- 6-31G(d) is another popular basis set choice

CCCE-2008

25

Polarized Basis Sets

- In molecule formation, AOs become distorted in shape (polarization)
 - Orbitals are influenced by other nuclei
 - Polarization accounts for these influences which distort the orbital shape



CCCE-2008

26

Diffuse Basis Sets

- Jargon Used → 6-31+G(d) or 6-31++G(d)
 - 6-31+G(d)**
 - 6-31G(d) basis set with an additional larger p-function for atoms with $Z > 2$
 - 6-31++G(d)**
 - 6-31+G(d) basis set with an additional larger s-function for H atoms
- Diffuse basis sets are useful for describing anions, molecules with lone pairs, excited states, and transition states (loosely held e^-)

CCCE-2008

27

Basis Set Progression

- Increasing number of GTOs Used:
 Minimal < Split Valence < Polarized < Diffuse
 →
- Get an increasingly good approximation to the actual wave function
 - The number of integrals increases as N^4 where N is the number of basis functions
 - During the minimization process, the orbital exponents are adjusted to define a new basis set to start another iteration
- **Computational cost has to be considered**

CCCE 2008

28

Comparison of Some Methods for Accuracy

		DFT Location??							
		HF	MP2	MP3	MP4	QCISD(T)	...	Full CI	
BASIS SET ↓	Minimal STO-3G		Electron Correlation →						
	Split Valence 3-21G								
	Polarized 6-31G(d) ~ DZVP								
	6-311G(d,p) ~ TZVP								
	Diffuse 6-311+G(d,p)								
	High ang. Momentum 6-311+G(2d,p)								
	6-311++G(3df,3dp)								
								
	∞	HF Limit						Schrödinger Equation	

CCCE 2008

29

Basis Set Choice and Expense

axial-methylcyclohexane on SGI Indigo2
 (Spartan cpu time in sec.)

Method/Basis Set	s.p.	opt.
AM1/STO-3G	~1	10
HF/STO-3G	72	983
HF/ 3-21G(d)	193	2214
HF/ 6-31G(d,p)	2632	34655 (9.6 h)

- As larger basis sets are used, the energy decreases (Variational Principle)

CCCE 2008

30

Common Basis Sets

Brief Description of Standard Basis Sets

Basis Set	Description
STO-3G	Minimal basis; qualitative results - large systems
3-21G	Double ζ ; more accurate results on large systems
6-31G(d)	Moderate set; Common use for medium systems
6-31G(d,p)	Used where H is site of interest; More accurate
6-31+G(d)	Used with anions, excited states, lone pairs, etc.
6-31+G(d,p)	Used with anions, etc., where H is site of interest
6-311++G(d,p)	Good for final, accurate energies, but expensive

- Many other sets are in use. Existing sets can be modified

CCCF-2008

31

Optimization of a Basis Set

- Variational Principle: Energy values are bounded from below
 - The lower the calculated energy, the better
 - Procedure:
 - Vary the constants and exponents that describe the Gaussian functions *sequentially* until the lowest energy is obtained
 - Such a basis set may only apply to that individual molecule, however

CCCF-2008

32

Lab: Gaussian Orbitals

- **Question:** How are Gaussian orbitals used to approximate a Slater Type Orbital?
- **Importance:** Basis sets are approximations based on mathematical use of two or more Gaussian functions
- **Goal:** Visualize an “STO-3G” basis set - what does the resultant function look like?
- **Computational Tool:** Spreadsheet
- Refer to Basis Set Case Study handout for detailed instructions

CCCF-2008

33
