

# CHM 673

## Lecture 7: Basis sets

---

Suggested reading:

Chapter 3.5, 3.6 from S&O

Chapter 5, specifically 5.1-5.4 from Jensen

<http://vergil.chemistry.gatech.edu/courses/chem6485/pdf/basis-sets.pdf>

Database of basis sets: <https://www.basissetexchange.org/>

# Hartree-Fock equation: How to solve practically?

---

$$\hat{f}_i |\chi_i\rangle = \varepsilon_i |\chi_i\rangle$$

- One-electron wave functions (orbitals) can be expanded in the basis set
- HF equations become matrix equations: perfect for solving on a computer!
- More complete (bigger) basis set will result in a better solution (lower HF energy)

# Introducing a basis

The spin orbitals are expressed through a set of  
(known, spatial) basis functions:

$$\chi_i = \sum_{\mu=1}^K C_{\mu i} \phi_{\mu i} \quad \text{basis } \phi_{\mu i}, \mu=1 \dots K$$

$$f \sum_{\nu} C_{\nu i} \phi_{\nu} = \varepsilon_i \sum_{\nu} C_{\nu i} \phi_{\nu}$$

$$\sum_{\nu} C_{\nu i} \underbrace{\int dx_1 \phi_{\mu}^*(x_1) f(x_1) \phi_{\nu}(x_1)}_{F_{\mu\nu}} = \varepsilon_i \sum_{\nu} C_{\nu i} \underbrace{\int dx_1 \phi_{\mu}^*(x_1) \phi_{\nu}(x_1)}_{S_{\mu\nu}}$$

Fock matrix                              overlap matrix

$$\sum_{\nu} F_{\mu\nu} C_{\nu i} = \varepsilon_i \sum_{\nu} S_{\mu\nu} C_{\nu i}$$

Roothaan Equations (1951)

$$FC = SC\varepsilon$$

matrix equations!

$$\hat{f}_i |\chi_i\rangle = \varepsilon_i |\chi_i\rangle$$



Clemens Roothaan  
(1918 – 2019)

# Overlap matrix

$$\Sigma_\nu F_{\mu\nu} C_{\nu i} = \varepsilon_i \Sigma_\nu S_{\mu\nu} C_{\nu i}$$

K – total number of basis functions

$\phi_\mu$ ,  $\mu=1\dots K$  - atomic basis functions

Overlap matrix  $S_{\mu\nu} = \langle \phi_\mu | \phi_\nu \rangle$

K x K Hermitian matrix

different basis functions are orthonormal but in general not orthogonal

Computational cost  $\sim K^2$

$$\begin{pmatrix} 1 & & & \\ & 1 & & \\ & & 1 & \\ & & & 1 \end{pmatrix} \quad K \times K$$

# Fock matrix

$$\Sigma_{\nu} F_{\mu\nu} C_{\nu i} = \varepsilon_i \Sigma_{\nu} S_{\mu\nu} C_{\nu i}$$

K – total number of basis functions

$\phi_{\mu}, \mu=1\dots K$  - atomic basis functions

Fock matrix

$$F_{\mu\nu} = \langle \phi_{\mu} | \hat{f} | \phi_{\nu} \rangle$$

K x K Hermitian matrix

$$F_{\mu\nu} = \langle \phi_{\mu} | \hat{f} | \phi_{\nu} \rangle = \langle \phi_{\mu} | \hat{h} | \phi_{\nu} \rangle + \sum_j^{N_{elec}} \langle \phi_{\mu} | \hat{J}_j - \hat{K}_j | \phi_{\nu} \rangle$$

$$\langle \phi_{\mu} | \hat{h} | \phi_{\nu} \rangle \quad \text{computational cost} \sim K^2$$

$$\sum_j^{N_{elec}} \langle \phi_{\mu} | \hat{J}_j - \hat{K}_j | \phi_{\nu} \rangle = \sum_j^{N_{elec}} (\langle \phi_{\mu} \chi_j | \phi_{\nu} \chi_j \rangle - \langle \phi_{\mu} \chi_j | \chi_j \phi_{\nu} \rangle) =$$

$$\sum_j^{N_{elec}} \sum_{pq}^K C_{pj} C_{qj} (\langle \phi_{\mu} \phi_p | \phi_{\nu} \phi_q \rangle - \langle \phi_{\mu} \phi_p | \phi_q \phi_{\nu} \rangle) =$$

$$\sum_{pq}^K D_{pq} (\langle \phi_{\mu} \phi_p | \phi_{\nu} \phi_q \rangle - \langle \phi_{\mu} \phi_p | \phi_q \phi_{\nu} \rangle) \quad \text{computational cost} \sim K^4$$

$$D_{pq} = \sum_j^{N_{elec}} C_{pj} C_{qj} \quad \text{density matrix}$$

$$F_{\mu\nu} = h_{\mu\nu} + \sum_{pq}^K G_{\mu\nu pq} D_{pq}$$

# HF self-consistent field procedure

1. Specify molecular geometry, basis functions, and electronic state of interest (i.e., singlet, triplet, etc – stay tuned!)
2. Form overlap matrix  $S$
3. Guess initial molecular orbital density matrix  $D$
4. Form Fock matrix  $F$
5. Solve  $FC = SC\epsilon$
6. Use new molecular orbital coefficients  $C$  to compute new density matrix  $D$
7. Go to step 4; repeat until density matrix  $D$  no longer changes from one iteration to the next

In other words, HF method boils down to **matrix diagonalization problem**.

But the Fock matrix  $F$  depends on its eigenvectors  $C$ , that's why we need iterations to bring  $F$  and  $C$  into consistency

# Hartree-Fock model

---

Solving Roothaan's equations result in  $K_{\text{basis}}$  spatial orbitals:

- $N_{\text{elec}}/2$  orbitals are occupied
  - $K - N_{\text{elec}}/2$  orbitals are virtual
- 
- The Fock matrix and the total HF energy depend only on the occupied orbitals
  - Virtual orbitals have no direct physical interpretation, except as electron affinities (in Koopmans' theorem)

# Computational cost of the HF model

$K$  is the number of basis functions

Computing the Fock matrix involves the following integrals:

$\sim K^2$  of overlap  $S_{\mu\nu} = \langle \phi_\mu | \phi_\nu \rangle$  and one-electron integrals  $\langle \phi_\mu | \hat{h} | \phi_\nu \rangle$

$\sim K^4$  of two-electron integrals  $\langle \phi_\mu \phi_p | \phi_\nu \phi_q \rangle$

- computed and stored on hard drive (older codes)
- recomputed on-the-fly as needed (newer codes)

Diagonalization of the Fock matrix of dimensionality  $K \times K$  is proportional to  $K^3$

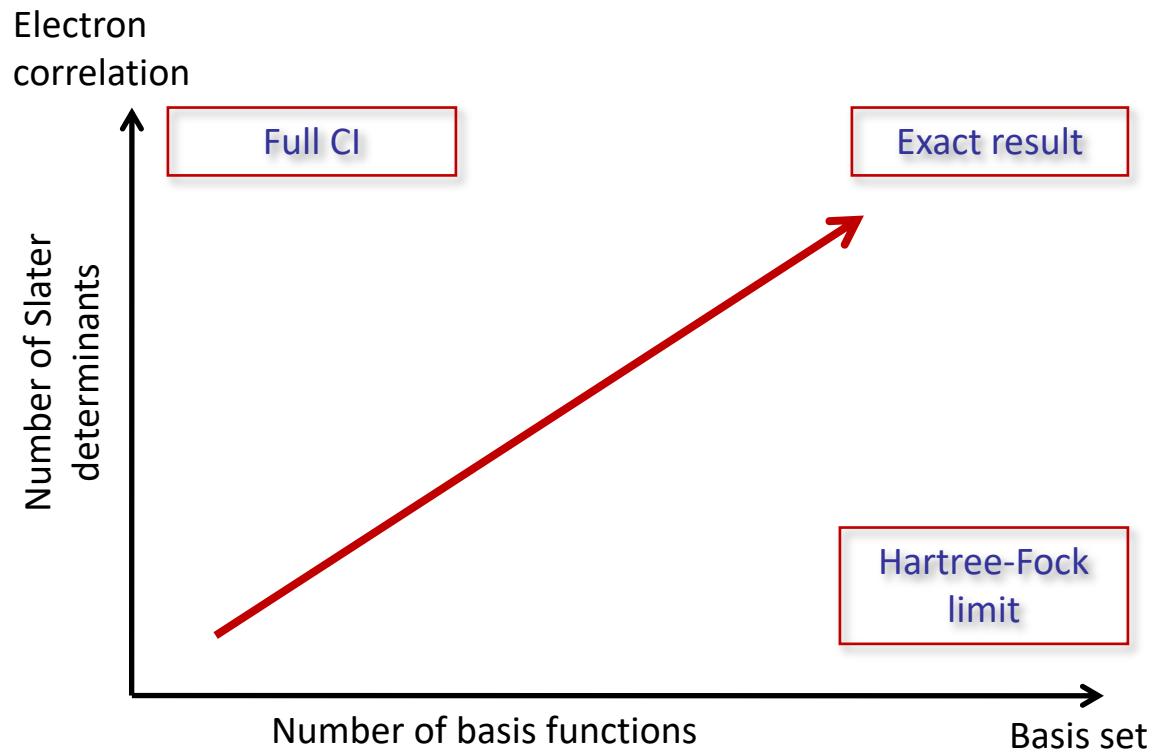
The overall scaling of the HF model is  $\sim K^3 - K^4$  depending on the implementation

# Size of the basis set

If the basis (number of basis functions) is increased → variational space increases →  
The energy becomes lower and closer to the exact HF energy

In the limit of the infinite number of basis functions (complete basis set, CBS) →  
the Hartree-Fock limit

But this is not the exact solution of  
the SE, only the best energy with a  
single-determinantal wave function



# Choosing the basis set

---

How to pick the basis set?

- Basis functions should have a behavior that agrees with physics of the problem
- Basis functions should make it easy to calculate all the required integrals

## Molecules

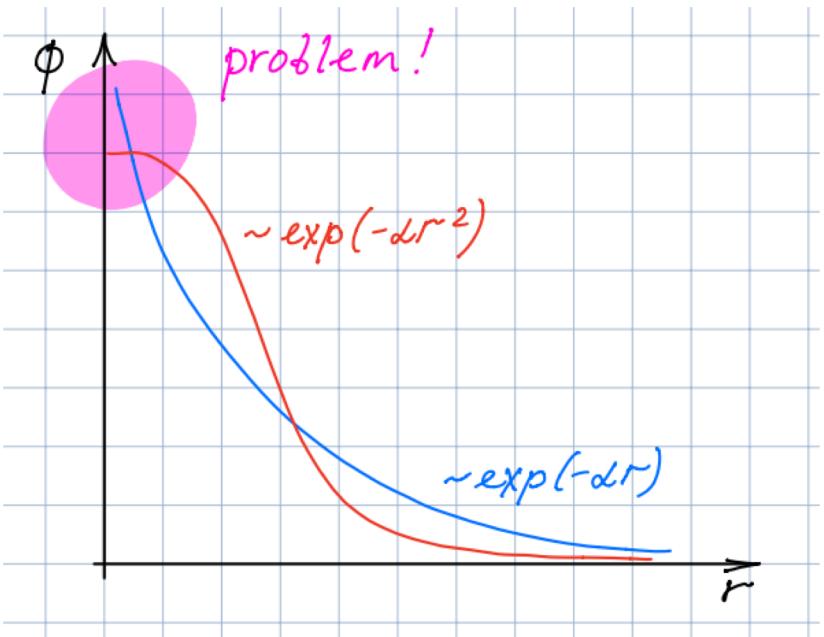
- Gaussian functions (integrals are semi-analytical)
- exponential functions (close resemblance to H atom eigenfunctions)

## Crystals/solids

- plane waves (exact solutions of a free electron)

More complete (larger) basis set will result in a *variationally* better solution (lower HF energy) – but not necessarily practically better solution!

# Exponential versus Gaussian basis sets



Gaussian functions are not perfect

## Slater-type orbitals (STO)

$$\phi_{abc}^{STO}(x, y, z) = N x^a y^b z^c e^{-\zeta r}$$

N – normalization

$L = a+b+c$  - angular momentum

$\zeta$  determines the width of the orbitals

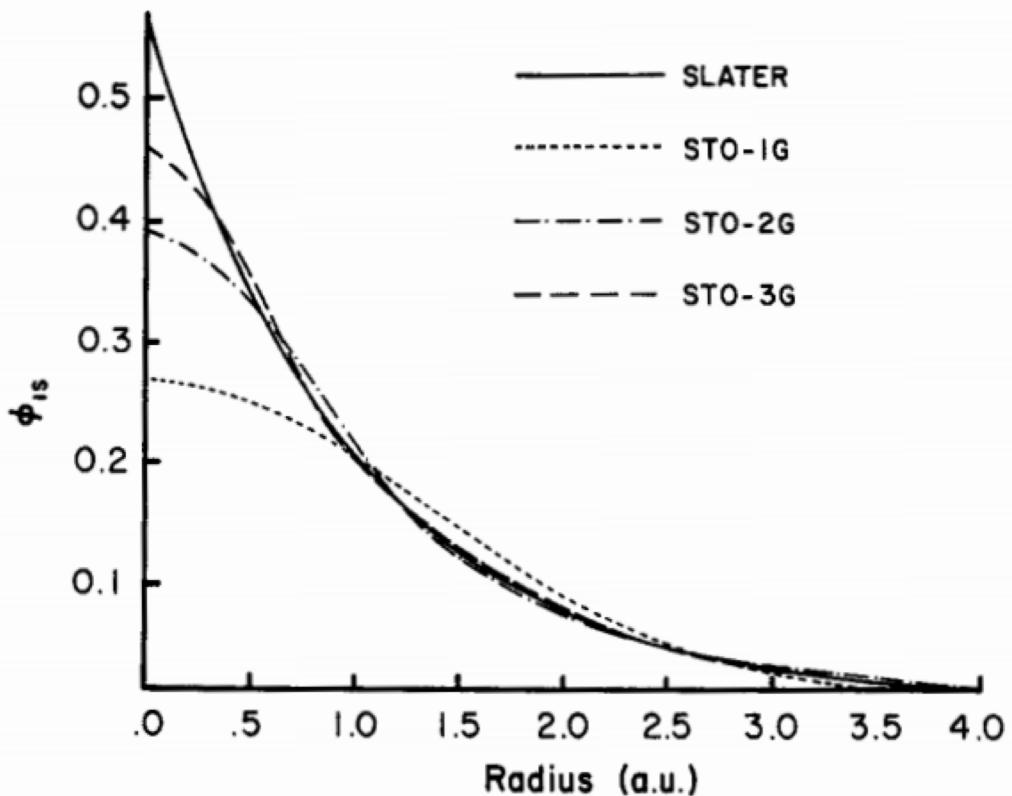
- Exact for one electron in H atom
- Lack radial nodes and not pure spherical harmonics (how to get 2s, 2p etc?)
- Correct short-range and long-range behavior

## Gaussian-type orbitals (GTO)

$$\phi_{abc}^{GTO}(x, y, z) = N x^a y^b z^c e^{-\zeta r^2}$$

- No longer H atom like
- Easier to compute (integrals are semi-analytical)

# Gaussian basis sets: contractions



**Figure 3.3** 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.

Exponential functions are more accurate but Gaussian functions make easier integral calculations

Solution: **Contracted Gaussian functions (CGF)**

Choose in advance the number of Gaussians  $n$ , their widths  $\zeta_i$  and weights  $c_i \rightarrow$  this composition becomes a single predefined basis function

$$\phi_{abc}^{CGF}(x, y, z) = N \sum_{i=1}^n c_i x^a y^b z^c e^{-\zeta_i r^2}$$

Note: each integral over a pair of basis functions will be a sum of integrals over participating Gaussians

# Minimal basis set

---

Minimal basis set:

one basis function for each atomic orbital in the atom

H, He: 1 (1s)

Li ... Ne: 5 (1s 2s 2p<sub>x</sub> 2p<sub>y</sub> 2p<sub>z</sub>)

Na ... Ar: 9 (1s 2 s 2p<sub>x</sub> 2p<sub>y</sub> 2p<sub>z</sub> 3s 3p<sub>x</sub> 3p<sub>y</sub> 3p<sub>z</sub>)

represent each basis function with 3 Gaussians → STO-3G

Limited variational flexibility

This basis is too small to provide chemical accuracy

# Double and triple zeta basis sets

Let's allow some expansion, contraction and anisotropy of orbitals

**double zeta** bases (2 b.f. for each atomic orbital)

H: 2 (1s, 1s')

C: 10 (1s, 1s', 2s, 2s', 2p, 2p')

check for **DZ** in the basis name , e.g. **DZP**, cc-pV**DZ**

**triple zeta** bases (3 b.f. for each atomic orbital)

H: 3 (1s, 1s', 1s'')

C: 15 (1s, 1s', 1s'', 2s, 2s', 2s'', 2p, 2p', 2p'')

check for **TZ**, e.g. **TZP**, aug-cc-pV**TZ**

**quadruple zeta** (4 b.f. for each atomic orbital): **QZ**

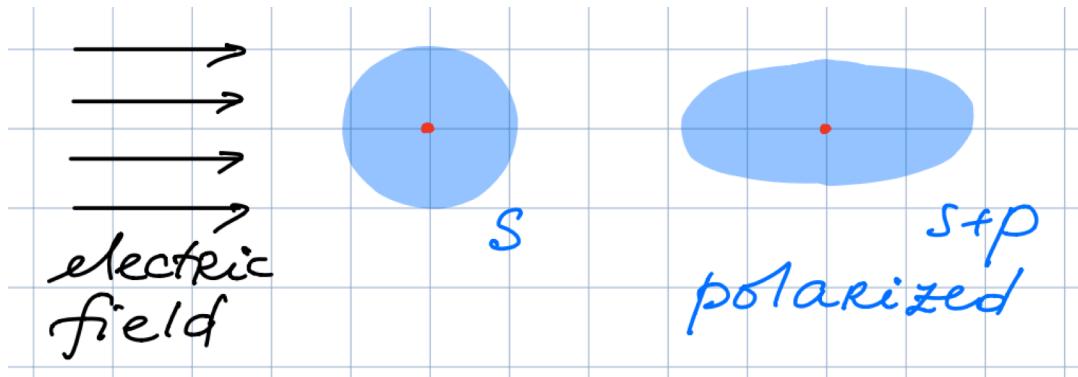
**split-valence bases**: inert core (1 b.f.), valence orbitals (2 or 3 b.f) e.g. C: 1s, 2s, 2s', 2p, 2p'

6-31G ←

1 b.f. for core electrons  
contracted from  
6 Gaussians

2 b.f. for valence electrons (double-zeta)  
contracted from  
3 and 1 Gaussians

# Polarization functions



Polarization functions: higher angular momentum functions that allow anisotropy

**p** b.f. to polarize **s** orbital, **d** b.f. to polarize **p** orbital etc.

- important for ions, polar molecules, molecules with electronegative atoms (F, O, etc)
- more important for heavy elements than for Hydrogen

Check for letter “**p**” in the basis name: DZP, cc-pVDZ

Pople bases: \* or explicit names of added functions: 6-31G\*\* = 6-31G(**d,p**)

non-Hydrogens      Hydrogens

# Diffuse functions

---

**Diffuse functions:** functions with very shallow Gaussians, can describe electron density far from nuclei

s b.f. for Hydrogen, s + p b.f. for C,N,O etc

- important for **anions**, electronic excited states, non-covalent complexes
- more important for heavy elements than for Hydrogen

Check for letter “+” or “aug” in the basis name: DZP+, aug-cc-pVDZ, 6-31+G\*

# Common Gaussian basis sets

description	Pople	Dunning	Correlation-consistent
Double-zeta polarized	6-31G* 6-31G**	DZP	cc-pVDZ
Double-zeta polarized with diffuse functions	6-31+G* 6-31++G**	DZP+	aug-cc-pVDZ
Triple-zeta polarized	6-311G* 6-311G**	TZP	cc-pVTZ
Triple-zeta polarized with diffuse functions	6-311+G* 6-311++G**	TZP+	aug-cc-pVTZ
Quadruple-zeta polarized	-	-	cc-pVQZ
...			

# Spherical versus Cartesian bases

---

d functions: 5 components ( $d_{x^2-y^2}$ ,  $d_{z^2}$ ,  $d_{xy}$ ,  $d_{xz}$ ,  $d_{yz}$ ) – “spherical” or “pure angular momentum” functions

Computers prefer to work with 6 d functions ( $d_{x^2}$ ,  $d_{y^2}$ ,  $d_{z^2}$ ,  $d_{xy}$ ,  $d_{xz}$ ,  $d_{yz}$ ) – “Cartesian functions”

f functions: 7 Spherical and 10 Cartesian components

Variational energy is slightly lower for 6 d function basis, but generally the results will be very similar

Different basis sets (and different programs!) have different defaults for the spherical vs cartesian components

This is the most common discrepancy between results computed in different packages

# Basis sets: examples

H atom, STO-3G

H	0	
S	3	1.00
		0.3425250914D+01
		0.6239137298D+00
		0.1688554040D+00
****		



exponents  $\zeta$



contraction coefficients  $c_i$

# Basis sets: examples

C atom, STO-3G

C	0				
S	3	1.00			
		0.7161683735D+02	0.1543289673D+00		1s
		0.1304509632D+02	0.5353281423D+00		
		0.3530512160D+01	0.4446345422D+00		
SP	3	1.00			
		0.2941249355D+01	-0.9996722919D-01	0.1559162750D+00	2s, 2p
		0.6834830964D+00	0.3995128261D+00	0.6076837186D+00	
		0.2222899159D+00	0.7001154689D+00	0.3919573931D+00	
****					
exponents $\zeta$					
contraction coefficients $c_i$ for 2s functions					
contraction coefficients $c_i$ for 2p functions					

# Basis sets: examples

H atom, 6-31G\*\*

H	0	
S	3	1.00
		0.1873113696D+02
		0.2825394365D+01
		0.6401216923D+00
S	1	1.00
		0.1612777588D+00
P	1	1.00
		0.1100000000D+01
****		

exponents  $\zeta$

contraction coefficients  $c_i$

1s

1s'

Polarization functions

# Basis sets: examples

C atom, 6-31G\*\*

C	0		
S	6	1.00	
		0.3047524880D+04	0.1834737132D-02
		0.4573695180D+03	0.1403732281D-01
		0.1039486850D+03	0.6884262226D-01
		0.2921015530D+02	0.2321844432D+00
		0.9286662960D+01	0.4679413484D+00
		0.3163926960D+01	0.3623119853D+00
SP	3	1.00	
		0.7868272350D+01	-0.1193324198D+00
		0.1881288540D+01	-0.1608541517D+00
		0.5442492580D+00	0.1143456438D+01
SP	1	1.00	
		0.1687144782D+00	0.1000000000D+01
D	1	1.00	
		0.8000000000D+00	1.0000000

1s (core)

2s, 2p

2s', 2p'

Polarization functions

# Basis sets: examples

C atom, 6-311++G\*\*

C	0		
S	6 1.00		
	4563.240	0.00196665	
	682.0240	0.0152306	
	154.9730	0.0761269	
	44.45530	0.2608010	
	13.02900	0.6164620	
	1.827730	0.2210060	
SP	3 1.00		
	20.96420	0.114660	0.0402487
	4.803310	0.919999	0.237594
	1.459330	-0.00303068	0.815854
SP	1 1.00		
	0.4834560	1.000000	1.000000
SP	1 1.00		
	0.1455850	1.000000	1.000000
SP	1 1.00		
	0.0438000	1.0000000	1.0000000
D	1 1.00		
	0.626	1.000000	
****			

1s (core)

2s, 2p

2s', 2p'

2s'', 2p''

s, p diffuse functions

Polarization functions