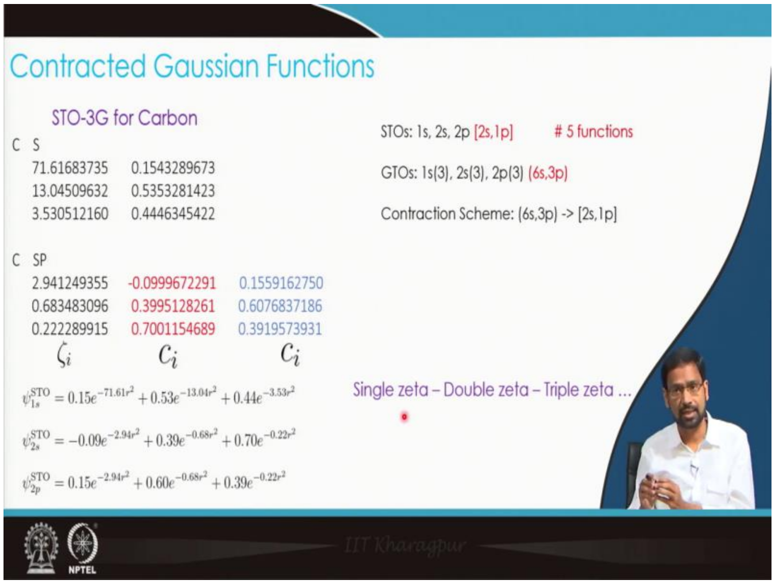
Approximate Methods in Quantum Chemistry
Professor Sabyashachi Mishra
Department of Chemistry
Indian Institute of Technology, Kharagpur
Lecture 36: Basis Sets - II

Hello students! Welcome to this lecture. In the last lecture we discussed about the functional forms of different basis functions and their strengths and weaknesses. We discussed the Slater type orbitals, the Gaussian type functions and the concept of contracted Gaussian type functions, where more than one Gaussian type functions are combined to reproduce a Slater type orbital.

(Refer Slide Time: 01:55)



Contracted Gaussian Functions

STO-3G for Carbon

```
C  S
   71.61683735    0.1543289673
   13.04509632    0.5353281423
   3.530512160    0.4446345422

C  SP
   2.941249355   -0.0999672291    0.1559162750
   0.683483096    0.3995128261    0.6076837186
   0.222289915    0.7001154689    0.3919573931
```

ζ_i $\quad c_i$ $\quad c_i$

$$\psi_{1s}^{\text{STO}} = 0.15e^{-71.61r^2} + 0.53e^{-13.04r^2} + 0.44e^{-3.53r^2}$$

$$\psi_{2s}^{\text{STO}} = -0.09e^{-2.94r^2} + 0.39e^{-0.68r^2} + 0.70e^{-0.22r^2}$$

$$\psi_{2p}^{\text{STO}} = 0.15e^{-2.94r^2} + 0.60e^{-0.68r^2} + 0.39e^{-0.22r^2}$$

STOs: 1s, 2s, 2p [2s,1p] # 5 functions

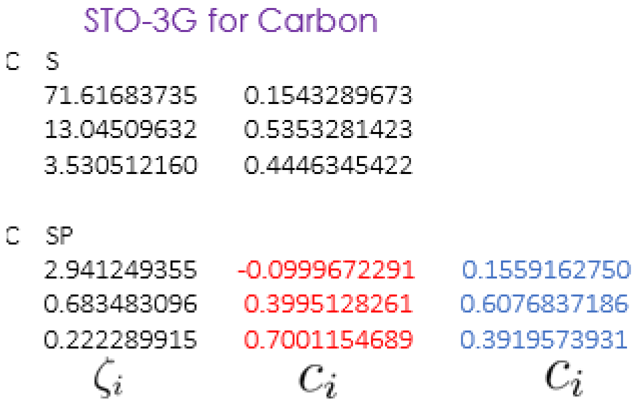
GTOs: 1s(3), 2s(3), 2p(3) (6s,3p)

Contraction Scheme: (6s,3p) -> [2s,1p]

Single zeta – Double zeta – Triple zeta ...

Let us discuss the STO-3G basis set for carbon atom. The electron configuration of C is $1s^2 2s^2 2p^2$. For each of these 5 orbitals (note, 2p has three degenerate orbitals) we require one STO. In total, we need 5 STOs of which 2 are s-type (1s and 2s) and 1 is p-type (with three different forms for x, y, and z direction). Together, this is represented by [2s,1p] notation. Now, each of these STOs are going to be represented by 3 GTFs within STO-3G basis. Hence, the number GTFs (3 for 1s, 3 for 2s and 3 for 2p) becomes (6s,3p). Again, each p-type function has three components. This makes 15 number of functions in the basis set (6 from 6s and 3 from each of the 3p). The contraction scheme for this basis becomes: (6s,3p) → [2s,1p]. Note the use of the round brackets for the GTFs and the square brackets for the STOs. Even if we are using 15 GTFs instead of 5 STOs, the computational cost is still less in the former due to the nature of the Gaussian functions.

Now let us look at the functional forms of the STO-3G basis. From the basis set library, we obtain the numbers for the orbital exponents and the coefficients.



STO-3G for Carbon

```
C   S
    71.61683735     0.1543289673
    13.04509632     0.5353281423
    3.530512160     0.4446345422

C   SP
    2.941249355    -0.0999672291    0.1559162750
    0.683483096     0.3995128261    0.6076837186
    0.222289915     0.7001154689    0.3919573931
```

ζ_i c_i c_i

The first block is for 1s orbital, where the first column represents the orbital exponents and the second column represents the coefficients. Using the above numbers, we can write that the STO for 1s orbital is given by,

$$\psi_{1s}^{\text{STO}} = 0.15e^{-71.61r^2} + 0.53e^{-13.04r^2} + 0.44e^{-3.53r^2}$$

The second block is for 2s and 2p orbitals (denoted by SP), where the 1st column represents the orbital exponents, while the 2nd and 3rd columns represent the coefficients for the 2s and 2p orbitals, respectively. Using the numbers in the basis set, we can write the corresponding STOs for 2s and 2p orbitals as,

$$\psi_{2s}^{\text{STO}} = -0.09e^{-2.94r^2} + 0.39e^{-0.68r^2} + 0.70e^{-0.22r^2}$$

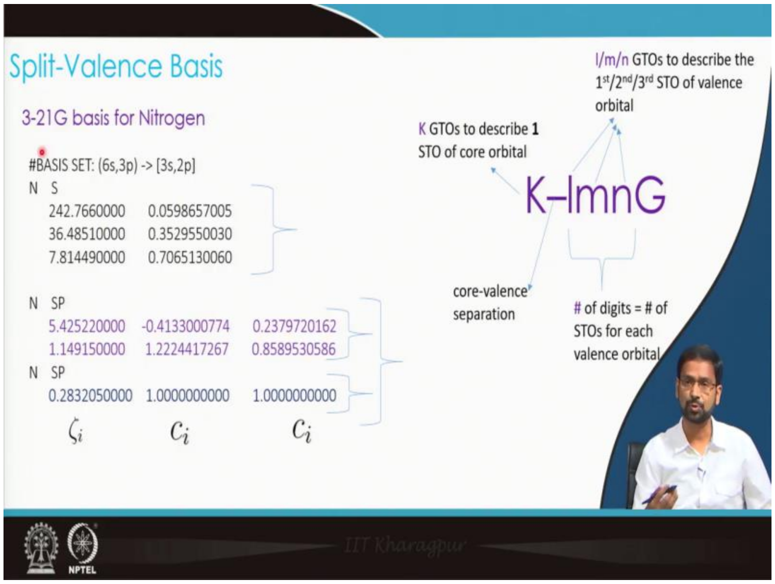
$$\psi_{2p}^{\text{STO}} = 0.15e^{-2.94r^2} + 0.60e^{-0.68r^2} + 0.39e^{-0.22r^2}$$

If you compare the values of the orbital exponents for 1s to those in 2s/2p, you would see large numbers for 1s orbital exponents. This describes the small size of the 1s orbital. On the other hand, smaller orbital exponents in 2s/2p allow these orbitals to participate in chemical bonding.

In the above discussion, we considered *one* STO for each orbital (which was then fitted by 3 GTFs). Since the STO itself is an approximate form of the atomic orbitals, it is often practiced to take more than 1 STOs to describe an orbital. Since STOs are characterized by their orbital exponents (zeta), we call those basis sets that take one STO for each orbital as single-zeta functions. Instead if we take 2 or 3 STOs for each orbitals, such basis set will be called as double-zeta or triple-zeta basis.

Since the core orbitals do not particiapate actively in chemical bonding, it is a common practice to express the core orbitals as a single-zeta function, while the valence orbitals are described by more number of STOs. Such a treatment, where core and valence are split, is known as split-valence basis.

(Refer Slide Time: 14:14)



Split-Valence Basis

3-21G basis for Nitrogen

```
#BASIS SET: (6s,3p) -> [3s,2p]
N   S
    242.7660000     0.0598657005
    36.48510000     0.3529550030
    7.814490000     0.7065130060

N   SP
    5.425220000    -0.4133000774     0.2379720162
    1.149150000     1.2224417267     0.8589530586
N   SP
    0.2832050000    1.0000000000     1.0000000000
```

ζ_i c_i c_i

l/m/n GTOs to describe the $1^{st}/2^{nd}/3^{rd}$ STO of valence orbital

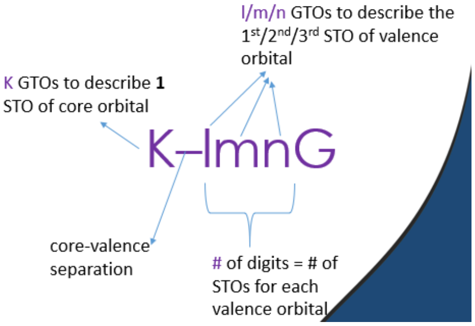
K GTOs to describe **1** STO of core orbital

K–lmnG

core-valence separation

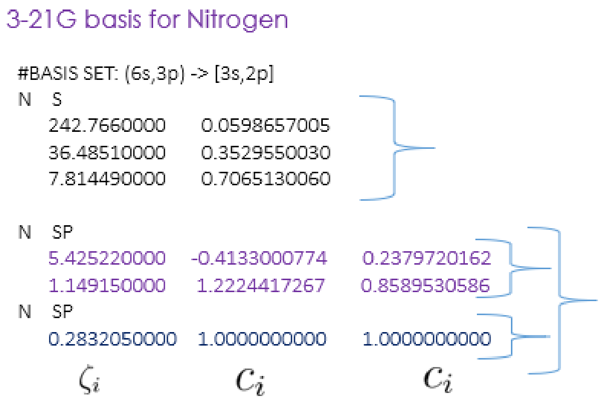
# of digits = # of STOs for each valence orbital

There are different approaches of obtaining the split-valence basis. We will discuss Pople's approach here. A general definition of split-valence basis *K*-*lmn*G is given below



The 'dash' represents the core-valence separation with core to the left and valence to the right of the 'dash'. *K* represents the number of GTOs to describe *one* STO. The number of digits after the 'dash' indicates the number of STOs to describe an orbitals (double-zeta if there are 2 digits and triple-zeta if there are 3 digits, and so on). The digits themselves (i.e., *l* or *m* or *n*) indicate the number of GTOs for that STO.

For example a basis set 3-21G suggests that the core orbital is described by 3 GTFs, the valence orbitals are described by two STOs, of which the 1st STO is described by 2 and the 2nd STO is described by 1 GTFs.



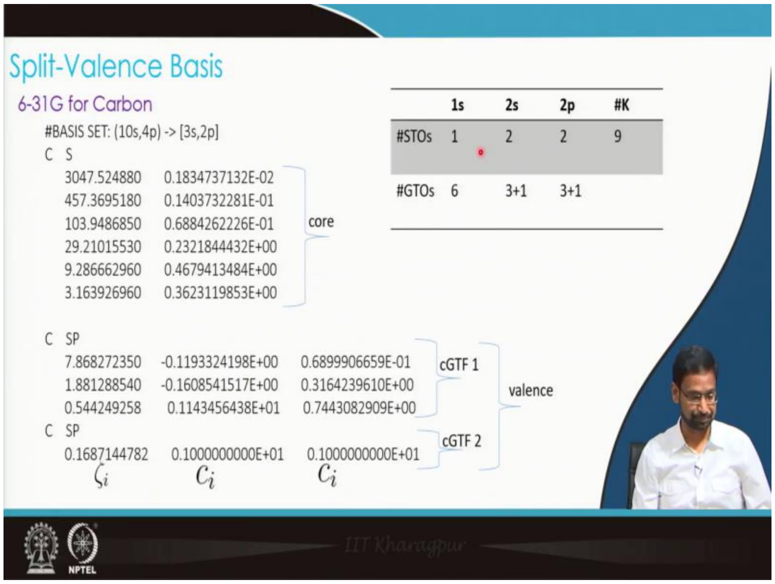
3-21G basis for Nitrogen

```
#BASIS SET: (6s,3p) -> [3s,2p]
N   S
    242.7660000     0.0598657005
    36.48510000     0.3529550030
    7.814490000     0.7065130060

N   SP
    5.425220000    -0.4133000774     0.2379720162
    1.149150000     1.2224417267     0.8589530586
N   SP
    0.2832050000    1.0000000000     1.0000000000
```

ζ_i c_i c_i

The first block (of three rows) correspond to the core 1s orbital represented by one STO (fitted to 3 GTFs), the second block (of two rows) is for 2s and 2p orbitals, where one STO is fitted to 2 GTFs and the third block (of one row) represents the second STO of the 2s and 2p orbitals, represented by a single GTF.

(Refer Slide Time: 18:42)



Split-Valence Basis

6-31G for Carbon

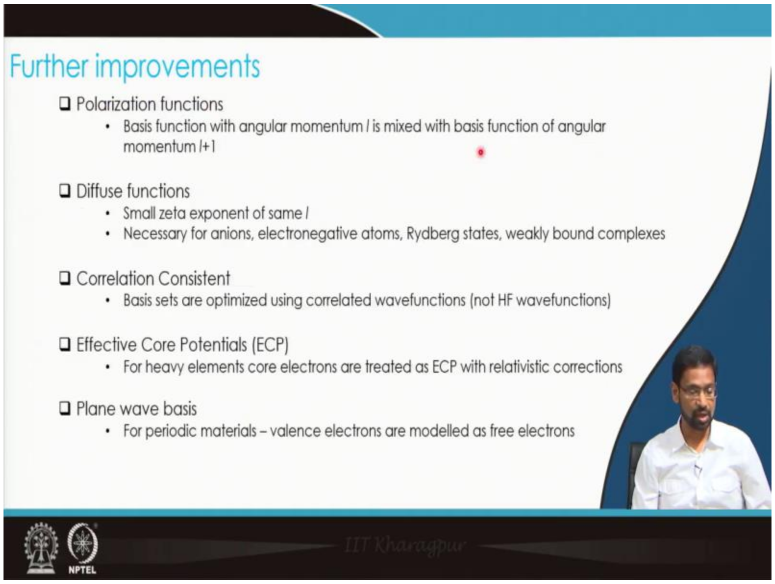
```
#BASIS SET: (10s,4p) -> [3s,2p]
C   S
    3047.524880     0.1834737132E-02
    457.3695180     0.1403732281E-01
    103.9486850     0.6884262226E-01
    29.21015530     0.2321844432E+00
    9.286662960     0.4679413484E+00
    3.163926960     0.3623119853E+00

C   SP
    7.868272350    -0.1193324198E+00    0.6899906659E-01
    1.881288540    -0.1608541517E+00    0.3164239610E+00
    0.544249258     0.1143456438E+01    0.7443082909E+00
C   SP
    0.1687144782    0.1000000000E+01    0.1000000000E+01
```

ζ_i c_i c_i

	1s	2s	2p	#K
#STOs	1	2	2	9
#GTOs	6	3+1	3+1	

Next, let us consider 6-31G basis for C atom. This is a split-valence double-zeta basis. Number of STOs for core orbital 1s is one, and the number of STOs for each valence orbital (2s and $2p_{x/y/z}$) is two. This makes total 3 s-type STOs and 2 p-type STOs which can be represented as [3s,2p], with total number of basis functions $K = 9$. The core STO is described by 6 GTFs and the two STOs for the valence orbitals are described by 3 and 1 GTFs. Taken together, we have 6+3+1=10 s-type GTFs and 3+1=4 p-type GTFs. This can be represented as (10s,4p). The overall contraction scheme is (10s,4p)→[3s,2p]. The values of the orbital exponents and the coefficients (obtained from basis set library) are given in the slide above.

(Refer Slide Time: 22:13)



Further improvements

- ☐ Polarization functions
  - Basis function with angular momentum l is mixed with basis function of angular momentum l+1
- ☐ Diffuse functions
  - Small zeta exponent of same l
  - Necessary for anions, electronegative atoms, Rydberg states, weakly bound complexes
- ☐ Correlation Consistent
  - Basis sets are optimized using correlated wavefunctions (not HF wavefunctions)
- ☐ Effective Core Potentials (ECP)
  - For heavy elements core electrons are treated as ECP with relativistic corrections
- ☐ Plane wave basis
  - For periodic materials – valence electrons are modelled as free electrons

What we discussed so far, is only the basic form of the basis set. Over the years, several improvements have been made on these basis sets. While we can not discuss all of them in detail, we will try to highlight some of the essential features of improved basis sets that are commonly used.

Upon chemical bond formation, the valence orbitals often get polarized, for example, the bonding orbital of H_2 develops a p_z character to the spherically symmetric 1s orbitals of H-atom. To allow this effect, the basis sets of angular momentum l are augmented with l+1 functions. In Pople's basis set notation, '*' at the end of K-lmnG indicates the polarized split-valence basis set.

Anions, electronegative atoms, and Rydberg states are often found to have electron density diffused over large space. For accurate treatment of such systems, basis sets include a function with very small orbital exponent (thus allowing diffused nature of the orbital). Such basis sets are called diffused basis set and the '+' sign indicates diffused basis set in Pople's notation (K-lmn+G).

All the basis functions we have discussed so far have been obtained from atomic Hartree-Fock calculations. There exists an alternate class of basis sets that add additional energy-optimized functions that make these basis sets suitable for correlated calculations. This class of functions, proposed by Dunning, are called correlation consistent (cc) basis set. There exists different variants of cc basis, for example: cc-pVDZ indicates a correlation consistent basis set of double-zeta quality with polarization functions added.

For heavy elements (e.g., transition metals, lanthanides and actinides) there are more number core electrons that do not take part in chemical bonding, but contribute to screening and relativistic effects. Explicit treatment of these core electrons increases computational cost. Hence, effective core potentials are used for these atoms, where the core electrons are treated as an effective potential while the valence electrons are treated explicitly. An example of the ECP is the LANL2DZ which indicate the ECP for the core electrons (along the line of Los Alamos National Laboratory class of ECPs) and double zeta basis for the valence electrons.

Plane wave basis functions are used for periodic systems. These are popular in calculations for solid-state periodic systems.

Thank you for your attention.