# Atomic Orbital Basis Set in Quantum Chemistry

Said Jalife

Duarte Group Seminar

Oxford University

# The Schrödinger Equation

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

DFT methods

Ab initio methods

$$\left(-\frac{1}{2}\nabla^2 + V_{ks}(r)\right)\psi_j(\boldsymbol{r}) = \mathcal{E}_j\psi_j(\boldsymbol{r})$$

$$F_j\psi_j(\boldsymbol{r}) = \mathcal{E}_j\psi_j(\boldsymbol{r})$$

*We need a basis set expansion to express the unkonwn molecular orbitals (or electron density) in terms of known functions*

Frank Jensen. Introduction to Computational Chemistry, 3rd Editon. 2017 by John Wiley & Sons, Ltd

2

# Linear Combination of Atomic Orbitals (LCAO)

Express MO's in terms of basis function conventionally called atomic orbitals.

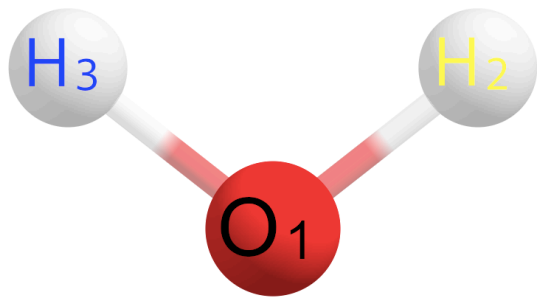$$\psi(\boldsymbol{r}) = \sum_{i=1}^{M} C_i\, \chi_i(\boldsymbol{r})$$

$\psi(\boldsymbol{r})$ = Molecular Orbital

$C_i$ = Coefficients

$\chi_i(\boldsymbol{r})$ = atomic orbitals

Example:

$\psi(\boldsymbol{r}) = c_1\chi_1(r) + c_2\chi_2(r) + c_3\chi_3(r) + c_4\chi_4(r) + c_5\chi_5(r) + c_6\chi_6(r) + c_7\chi_7(r)$

$(H_3)$        $(H_2)$        $(O_3)$

# Basis functions: Slater-type and Gaussian-type orbitals
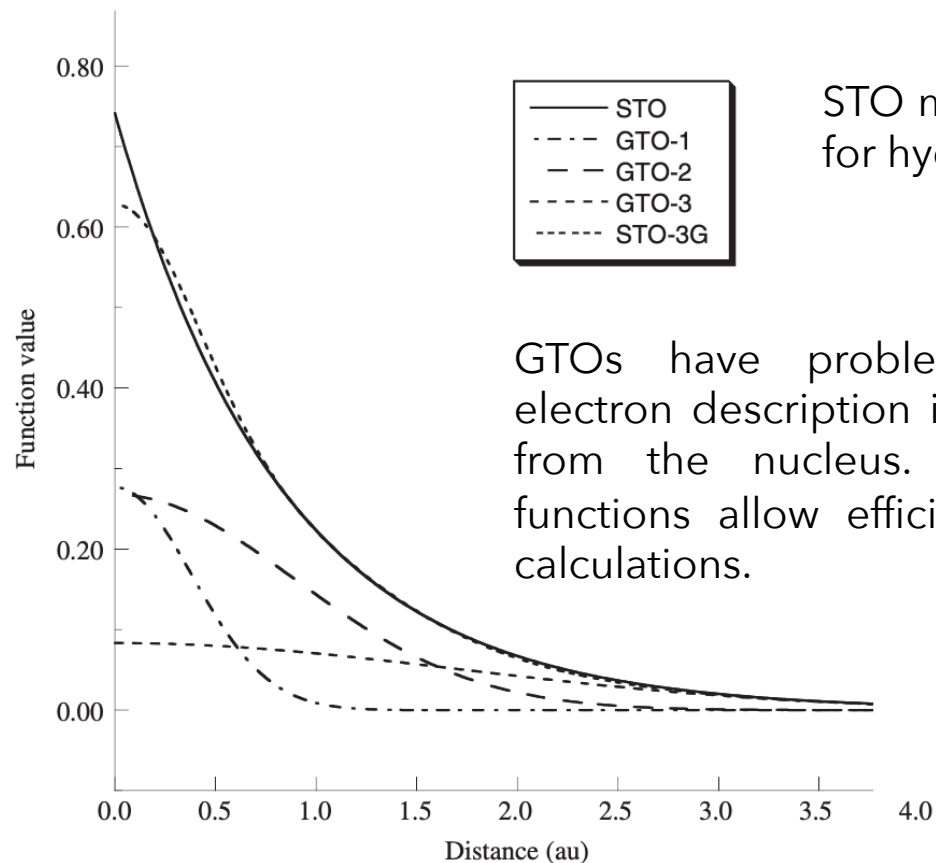
Spherical Slater-type Orbital

$$\chi_{\zeta,n,l,m}(r,\theta,\varphi) = N Y_{l,m}(\theta,\varphi) r^{n-1} e^{-\zeta r}$$

Spherical Gaussian-type Orbital

$$\chi_{\zeta,n,l,m}(r,\theta,\varphi) = N Y_{l,m}(\theta,\varphi) r^{2n-2-l} e^{-\zeta r^2}$$

Cartesian Gaussian-type Orbital

$$\chi_{\zeta,l_x,l_y,l_z}(x,y,z) = N x^{l_x} y^{l_y} z^{l_z} e^{-\zeta r^2}$$



STO mirros the exact orbials for hydrogen-like atoms.

GTOs have problems representing the electron description in regions near and far from the nucleus. However, Gaussian functions allow efficient molecular integral calculations.

# Primitive Gaussian-type orbitals

Spherical Gaussian-type Orbital

$$\chi_{\zeta,n,l,m}(r,\theta,\varphi) = NY_{l,m}(\theta,\varphi)r^{2n-2-l}e^{-\zeta r^2}$$

Cartesian Gaussian-type Orbital

$$\chi_{\zeta,l_x,l_y,l_z}(x,y,z) = Nx^{l_x}y^{l_y}z^{l_z}e^{-\zeta r^2}$$

The center is typically in the nuclear position, the sum of $l_x$, $l_y$ and $l_z$ defines the angular momentum (e.g., $lx + ly + lz = 1$ is a p-function), $\zeta$i is the exponent providing the radial extent of the function, and $N$ is a normalization constant.

Each individual GTO is called a **primitive function**

s-funtions $\qquad e^{-\zeta r^2} \qquad\qquad e^{-\zeta r^2}$

p-funtions
$$xe^{-\zeta r^2} \qquad\qquad Y_{1,-1}(\theta,\emptyset)e^{-\zeta r^2}$$
$$ye^{-\zeta r^2} \qquad\qquad Y_{1,0}(\theta,\emptyset)e^{-\zeta r^2}$$
$$ze^{-\zeta r^2} \qquad\qquad Y_{1,1}(\theta,\emptyset)e^{-\zeta r^2}$$

d-funtions
$$x^2e^{-\zeta r^2} \qquad\qquad Y_{2,-2}(\theta,\emptyset)r^2e^{-\zeta r^2}$$
$$y^2e^{-\zeta r^2} \qquad\qquad Y_{2,-1}(\theta,\emptyset)r^2e^{-\zeta r^2}$$
$$z^2e^{-\zeta r^2} \qquad\qquad Y_{2,0}(\theta,\emptyset)r^2e^{-\zeta r^2}$$
$$xye^{-\zeta r^2} \qquad\qquad Y_{2,1}(\theta,\emptyset)r^2e^{-\zeta r^2}$$
$$xze^{-\zeta r^2} \qquad\qquad Y_{2,2}(\theta,\emptyset)r^2e^{-\zeta r^2}$$
$$zye^{-\zeta r^2}$$

# Contraction of Gaussian-type Basis Sets

**A fixed linear combination of GTOs is called a contracted function**

The fact that many basis functions focus on describing the energetically important, but chemically unimportant, core electrons is the foundation for *contracted* basis sets. Avoid the computational effort on describing the chemically uninteresting.

$$\chi(\text{CGTO}) = \sum_{i}^{k} a_i \chi_i(\text{PGTO})$$

**Notation to indicate contraction of basis set**

(9s4p) → [3s2p]. Nine primitive s-functions, and four primitive p-functions are contracted to three s-functions, and two p-functions). (10s4p1d/4s1p) → [3s2p1d/2s1p]. Second row elements before the slash and hydrogen after.

Explicit contraction are represented in brackets {621/31}. The first, second, and third contracted s-functions are constructed of six, two and one primitive functions, respectively, whereas two contracted p-functions contain three and one primitive functions.

# Segmented vs General Contraction of Basis Sets

Segmented contraction

General contraction

$$\chi_1(\text{CGTO}) = \sum_{i=1}^{6} a_i \chi_i(\text{PGTO})$$

$$\chi_1(\text{CGTO}) = \sum_{i=1}^{10} a_i \chi_i(\text{PGTO})$$

$$\chi_2(\text{CGTO}) = \sum_{i=7}^{9} a_i \chi_i(\text{PGTO})$$

$$\chi_2(\text{CGTO}) = \sum_{i=1}^{10} b_i \chi_i(\text{PGTO})$$

$$\chi_3(\text{CGTO}) = \chi_{10}(\text{PGTO})$$

$$\chi_3(\text{CGTO}) = \sum_{i=1}^{10} c_i \chi_i(\text{PGTO})$$

Each primitive function is used only in one contracted function

All primitive functions are used in all the contracted functions

# Classification of Basis Sets

*n*-Zeta (*n*Z) notation

**Single Zeta** (SZ)

A *minimum* or *single zeta* (SZ) basis set employs only enough functions for a minimum description of the occupied orbitals of the isolated atom.
- H has a single *s*-function
- Second row atoms have two *s-functions* (1s and 2s) and one set of *p*-functions ($2p_x$, $2p_y$ and $2p_z$).

**Double Zeta** (DZ)

Doubling of all basis functions
- H has two *s*-function (1s and 1s')
- Second row atoms have four *s*-functions (1s, 1s', 2s and 2s') and two set of *p*-functions (2p and 2p').

Split valence basis

**Double Zeta Valance** (DZV)

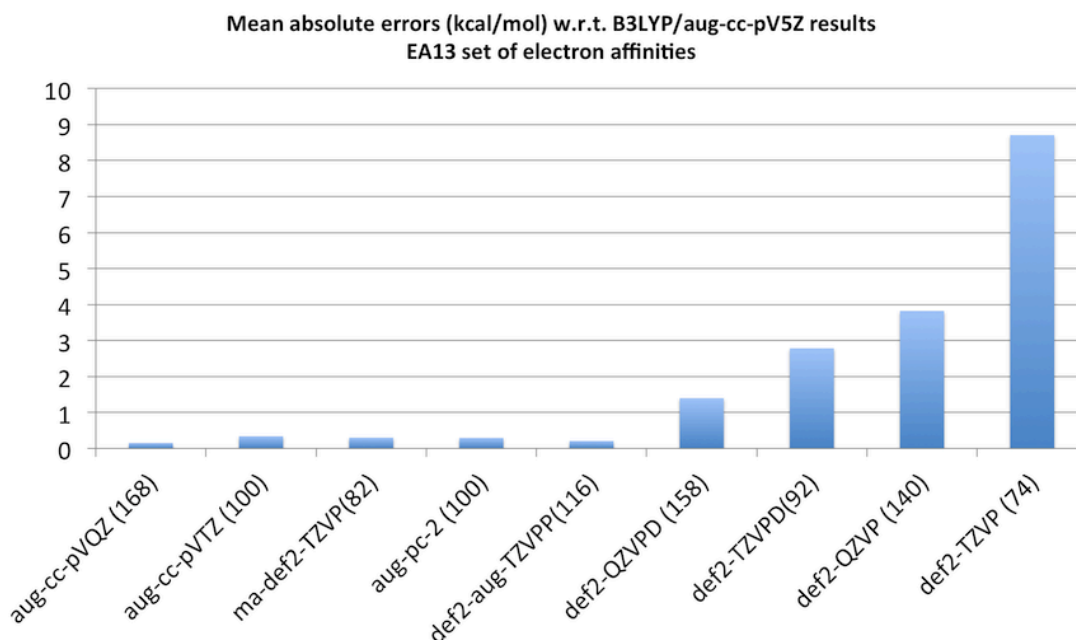Doubling only the valance basis functions
- Second row atoms have three *s*-functions (1s, 2s and 2s') and two set of *p*-functions (2p and 2p').

**Triple Zeta (TZ), (TZV), Quaduple Zeta (QZ), (QZV) Quintuple Zeta (5Z), …..**

# Polarization and augmentation with diffuse functions

**Double Zeta** (DZP)

Adding a single set of polarization function (higher angular momentum functions). For instance, a set of p-functions on hydrogens and a set of d-functions on non-hydrogen p-block atoms.

Mean absolute errors (kcal/mol) w.r.t. B3LYP/aug-cc-pV5Z results
EA13 set of electron affinities



If anions are considered or a property of interest dependes on the most loosely bound electrons, diffuse functions are needed, i.e, basis functions with small exponents.

# Pople Basis Set

**Example 6-31G Basis set**. Split double valence basis set. For C atom, 6-31G basis set is a contraction of (10s4p) → [3s2p]. It means, 6 PGTOs for the core orbitals. The valance orbitals split into two orbitals (2s and 2s', and 2p and 2p'), represented by 3 PGTOs and 1 PGTO, respectively.

The exponents and contraction coefficients for the s- and p-functions have been optimized by variational procedures at the HF level for atoms. $e^{-\zeta r^2}$

| Basis | Hydrogen | | Second row elements | | Third row elements | |
|---|---|---|---|---|---|---|
| | Contracted | Primitive | Contracted | Primitive | Contracted | Primitive |
| STO-3G | 1s | 3s | 2s1p | 6s3p | 3s2p | 9s6p |
| 3-21G | 2s | 3s | 3s2p | 6s3p | 4s3p | 9s6p |
| 6-31G(d,p) | 2s1p | 4s | 3s2p1d | 10s4p | 4s3p1d | 16s10p |
| 6-311G(2df,2pd) | 3s2p1d | 5s | 4s3p2d1f | 11s5p | 6s4p2d1f[a] | 13s9p[a] |

# Some basis sets in QM

Basis set composition for second row s- and p-block atoms

| | | | s-Block Elements | | | | | | | | | | | | p-Block Elements | | | | | | | | | | | |
| | | | Primitives | | | | | | Contracted | | | | | | Primitives | | | | | | Contracted | | | | | |
| Quality | Label | Name | s | p | d | f | g | h | s | p | d | f | g | h | s | p | d | f | g | h | s | p | d | f | g | h |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| SZ | Pople | STO-3G | 9 | 6 | | | | | 3 | 2 | | | | | 9 | 6 | | | | | 3 | 2 | | | | |
| DZ | Pople | 3-21G | 9 | 6 | | | | | 4 | 3 | | | | | 9 | 6 | | | | | 4 | 3 | | | | |
| | Jensen | pc-0 | 8 | 5 | | | | | 4 | 2 | | | | | 8 | 6 | | | | | 4 | 3 | | | | |
| DZP | Pople | 6-31G* | 16 | 10 | 1 | | | | 4 | 3 | 1 | | | | 16 | 10 | 1 | | | | 4 | 3 | 1 | | | |
| | Ahlrichs | Def2-SVP | 10 | 7 | 1 | | | | 4 | 3 | 1 | | | | 10 | 7 | 1 | | | | 4 | 3 | 1 | | | |
| | Jorge | DZP | 12 | 9 | 1 | | | | 7 | 4 | 1 | | | | 12 | 9 | 1 | | | | 7 | 4 | 1 | | | |
| | Sapporo | DZP | 11 | 7 | 2 | | | | 5 | 4 | 1 | | | | 12 | 9 | 3 | | | | 4 | 3 | 1 | | | |
| | Roos | ANO-DZP | 17 | 12 | 6 | | | | 5 | 2 | 2 | | | | 17 | 12 | 5 | | | | 4 | 3 | 1 | | | |
| | Dunning | cc-pVDZ | 12 | 8 | 1 | | | | 4 | 3 | 1 | | | | 12 | 8 | 2 | | | | 4 | 3 | 2 | | | |
| | Petersson | 2ZaP | 15 | 9 | 1 | | | | 6 | 4 | 1 | | | | 14 | 10 | 2 | | | | 5 | 4 | 2 | | | |
| | Jensen | pc-1 | 11 | 7 | | | | | 4 | 2 | | | | | 11 | 8 | 1 | | | | 4 | 3 | 1 | | | |
| TZP | Pople | 6-311G(2df) | 13 | 9 | 2 | 1 | | | 6 | 5 | 2 | 1 | | | 12 | 9 | 2 | 1 | | | 6 | 5 | 2 | 1 | | |
| | Ahlrichs | Def2-TZVP | 14 | 8 | 3 | | | | 5 | 4 | 3 | | | | 14 | 9 | 3 | 1 | | | 5 | 5 | 3 | 1 | | |
| | Jorge | TZP | 13 | 10 | 2 | 1 | | | 8 | 5 | 2 | 1 | | | 13 | 10 | 2 | 1 | | | 8 | 5 | 2 | 1 | | |
| | Sapporo | TZP | 15 | 9 | 4 | 1 | | | 7 | 6 | 3 | 1 | | | 13 | 9 | 4 | 2 | | | 5 | 4 | 2 | 1 | | |
| | Roos | ANO-TZP | 17 | 12 | 6 | | | | 6 | 3 | 2 | | | | 17 | 12 | 5 | 4 | | | 5 | 4 | 2 | 1 | | |
| | Dunning | cc-pVTZ | 15 | 10 | 2 | 1 | | | 5 | 4 | 2 | 1 | | | 15 | 9 | 3 | 1 | | | 5 | 4 | 3 | 1 | | |
| | Petersson | 3Zap | 17 | 12 | 2 | 1 | | | 6 | 5 | 2 | 1 | | | 17 | 13 | 3 | 1 | | | 6 | 5 | 3 | 1 | | |
| | Jensen | pc-2 | 13 | 9 | 1 | | | | 5 | 3 | 1 | | | | 13 | 10 | 2 | 1 | | | 5 | 4 | 2 | 1 | | |
| QZP | Ahlrichs | Def2-QZVP | 20 | 12 | 4 | 1 | | | 9 | 5 | 4 | 1 | | | 20 | 14 | 4 | 2 | 1 | | 9 | 6 | 4 | 2 | 1 | |
| | Jorge | QZP | 14 | 11 | 3 | 2 | 1 | | 9 | 6 | 3 | 2 | 1 | | 14 | 11 | 3 | 2 | 1 | | 9 | 6 | 3 | 2 | 1 | |
| | Sapporo | QZP | 16 | 10 | 5 | 3 | 1 | | 8 | 7 | 5 | 3 | 1 | | 16 | 11 | 5 | 2 | 2 | | 6 | 5 | 3 | 2 | 1 | |
| | Roos | ANO-QZP | 17 | 12 | 6 | 2 | | | 7 | 4 | 4 | 2 | | | 17 | 12 | 5 | 4 | 2 | | 6 | 5 | 3 | 2 | 1 | |
| | Dunning | cc-pVQZ | 16 | 12 | 3 | 2 | 1 | | 6 | 5 | 3 | 2 | 1 | | 16 | 11 | 4 | 2 | 1 | | 6 | 5 | 4 | 2 | 1 | |
| | Petersson | 4ZaP | | | | | | | | | | | | | 20 | 16 | 4 | 2 | 1 | | 7 | 6 | 4 | 2 | 1 | |
| | Jensen | pc-3 | 17 | 12 | 2 | 1 | | | 6 | 4 | 2 | 1 | | | 17 | 13 | 4 | 2 | 1 | | 6 | 5 | 4 | 2 | 1 | |
| 5ZP | Jorge | 5ZP | 18 | 12 | 4 | 3 | 2 | 1 | 10 | 7 | 4 | 3 | 2 | 1 | 18 | 12 | 4 | 3 | 2 | 1 | 10 | 7 | 4 | 3 | 2 | 1 |
| | Dunning | cc-pV5Z | 20 | 14 | 4 | 3 | 2 | 1 | 7 | 6 | 4 | 3 | 2 | 1 | 20 | 12 | 5 | 3 | 2 | 1 | 7 | 6 | 5 | 3 | 2 | 1 |
| | Petersson | 5ZaP | | | | | | | | | | | | | 23 | 19 | 5 | 3 | 2 | 1 | 8 | 7 | 5 | 3 | 2 | 1 |
| | Jensen | pc-4 | 21 | 15 | 3 | 2 | 1 | | 7 | 5 | 3 | 2 | 1 | | 21 | 16 | 6 | 3 | 2 | 1 | 7 | 6 | 6 | 3 | 2 | 1 |

# Optimization Criteria for some basis set

| Label | Name | Atomic sp(d) Exponents | | Polarization Exponents | | Contraction | |
|---|---|---|---|---|---|---|---|
| | | Method | Target | Method | Target | Method[a] | Coefficients |
| Pople | STO-3G | Fitted | Atoms | None | | Segmented | Fitted |
| | 6-31G | HF optimized<br>Identical s- and p-valence exponents | Atoms | HF optimized<br>Averaged | Molecules | Segmented | HF optimized |
| | 6-311G | MP2 optimized<br>Identical s- and p-valence exponents | Atoms | MP2 optimized | Atoms | Segmented | MP2 optimized |
| Karlsruhe | Def2-(S,TZ,QZ)VP | HF optimized | Atoms | From cc-pV$XZ$ | | Segmented | HF optimized |
| Jorge | $XZP$ | HF optimized for core<br>MP2 optimized for valence | Atoms | MP2 optimized | Atoms | Segmented | HF optimized |
| Sapporo | $XZP$ | HF optimized | Atoms | ANO fitted | Atom | Segmented | HF optimized |
| Roos | ANO-$XZP$ | HF optimized | Atoms | CISD optimized<br>Even-tempered | Atoms | General | ANO coefficients |
| Dunning | cc-pV$XZ$ | HF optimized | Atoms | CISD optimized<br>Even-tempered | Atoms | General | HF coefficients |
| Peterson | cc-pV$XZ$-F12 | HF optimized | Atoms | MP2-F12 optimized<br>Even-tempered | Molecules | General | HF coefficients |
| Petersson | $n$ZaP | HF L6 optimized for core[b]<br>MP2 even-tempered optimized for valence | Atoms | MP2 optimized<br>Even-tempered | Atoms | General | HF coefficients |
| Jensen | pc-$n$ | BLYP optimized | Atoms | BLYP optimized<br>Averaged | Molecules | General | BLYP coefficients |
| | pcseg-$n$ | BLYP optimized | Atoms | BLYP optimized<br>Averaged | Molecules | Segmented | BLYP optimized |

# Basis set superposition error

The basis superposition error (BSSE) in polyatomic systems arises due to the localised nature of the LCAO basis set. In the dimer, basis functions from one molecule can help compensate for the basis set incompleteness on the other molecule
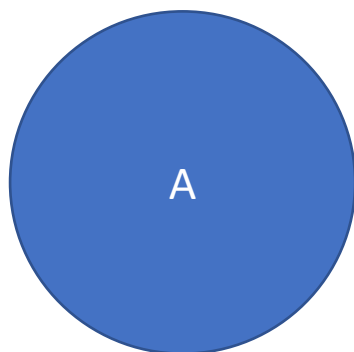
$$\Delta E_{complex} = E(AB)^*_{ab} - E(A)_a - E(A)_b$$
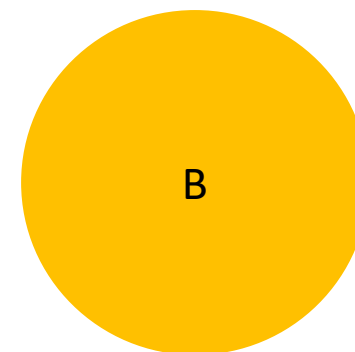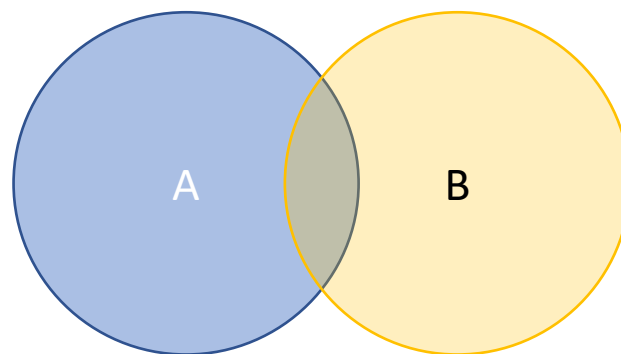


$$\Delta E_{CP} = E(A)^*_{ab} - E(A)^*_a - E(B)^*_{ab} - E(B)^*_b$$

# References

Frank Jensen. Introduction to Computational Chemistry, 3rd Editon. 2017 by John Wiley & Sons, Ltd

Christopher J. Cramer. Essentials of Computational Chemistry, 3rd Editon. 2002 by John Wiley & Sons, Ltd

Frank Jensen. Atomic orbitals basis set. WIRES. 2013, 3, 273-295.

AlexandreL.Magalhaẽs. Gaussian-Type Orbitals versus Slater-Type Orbitals: A Comparison. *J. Chem. Educ.* 2014, 91, 12, 2124-2127.

Abby L. Parrill , and Kenny B. Lipkowitz. Reviews of Computational Chemistry. Chapter 3. 2017 by John Wiley & Sons, Ltd