

Basis Set and their correlations with quantum chemical computations

Ram Kumar

*Department of Physics, BSNV PG College, Lucknow, UP, India
e-mail:- rktshri@gmail.com*

Abstract

A basis set is a set of functions combined linearly to model molecular orbitals. Basis functions can be considered as representing the atomic orbital of the atoms and are introduced in quantum chemical calculations because the equations defining the molecular orbitals are otherwise very difficult to solve. The selection of a basis set for quantum chemical calculations is very important. It is sometimes possible to use small basis sets to obtain good chemical accuracy, but calculations can often be significantly improved by the addition of diffuse and polarization functions. The purpose of the current study is to find basis set correlations with quantum chemical computations.

Keywords: Basis Set, Quantum chemistry, DFT, MP2, NLO materials, SHG, STO, GTO.

Introduction

Computational chemistry has become a key to investigate materials which are very far away from our reach. It also helps the researcher to know about the chemical systems before performing the actual experiments. Over the past few decades, the ab-initio quantum chemistry has turned into a fundamental tool in the investigation of atoms and molecules. The approximations made are generally mathematical approximations using a simple functional form for a function or wave function. Generally, ab-initio calculations produce very good qualitative results and the advantage of this method is that they ultimately converge to the near accurate solution, once all the approximations are included properly [1].

A wide variety of methods are implemented in the computer programs to approximate wave functions, which are the solutions to complex wave equations. These molecular wave functions carry the information about all the electrons, spatial as well as spin coordinates. The wave function is calculated quantum mechanically by solving the non-relativistic Schrodinger wave equation. Most of these methods are developed on the basis of one-electron model in which electrons move independently under the influence of average potential of other electrons, nuclei, and external fields (Hartree Fock method). These one-electron wave functions constitute atomic orbitals. Since the molecule of any substance is made up of

atoms, so a very effective method has been developed to expand the molecular orbitals in the form of centered atomic orbitals [2-4].

These methods can be further classified into two classes. The first one is post-Hartree Fock schemes in the framework of configuration interaction or coupled cluster methods which are based on wave-function and can be used for molecules containing up to a few tens of atoms and are likely to provide precise results. The second one is electron density based density functional theory (DFT) which may be employed over the systems containing a few tens up to a few hundreds of atoms [5-7]. DFT provides the wide-range of applicability and reasonably accurate results in many cases.

Basis Set

A basis set is a set of functions combined linearly to model molecular orbitals. Basis functions can be considered as representing the atomic orbitals of the atoms and are introduced in quantum chemical calculations because the equations defining the molecular orbitals are otherwise very difficult to solve.

Ab-initio quantum chemical methods try to derive information by solving the Schrödinger equation without fitting to experimental data. The mathematical sketch of the orbital, which is used to perform the theoretical calculation, is called the basis set. The basis set can be understood as expressing function for each electron in particular region of space [8]. The molecular orbitals are defined as a linear combination of a set of one-electron functions recognized as basis functions. An individual molecular orbital is expressed as,

$$\Phi_i = \sum_{\mu=1}^N c_{\mu} \chi_{\mu} \dots \dots \dots \dots \dots \dots (1)$$

Where the coefficients c_{μ} and χ_{μ} are known as orbital expansion coefficient and arbitrary basis function, respectively. The use of an appropriate basis set is a fundamental requirement for the accomplishment of molecular quantum mechanical calculation. Larger basis sets impose fewer restrictions on electrons and accurately approximates the exact molecular orbital. However, a complete basis means that an infinitely large number of functions must be used which is impossible in actual calculations. The smaller the basis used the poorer the representation. Thus, the type of basis functions used in the calculation influences the accuracy of the results. There are two types of atomic functions existing to construct basis functions, commonly used in electronic structure calculations: *Slater Type Orbitals* (STO) and *Gaussian Type Orbitals* (GTO).

Slater and Gaussian type orbitals

Strictly speaking atomic orbitals are solutions of the Hartree-Fock equations for the atom, i.e., wave functions of a single electron in the atom. Later on the term atomic orbital was replaced by basis function

or contraction when appropriate. Early, the Slater type orbitals were used as basis functions due to their similarity to atomic orbitals of the hydrogen atom. Slater type orbitals (STOs) are defined as (in spherical polar co-ordinates)

$$\Phi_i(\xi, n, l, m; r, \theta, \phi) = N r^{n-1} e^{-\xi r} Y_{lm}(\theta, \phi) \quad \dots (2)$$

Where N is normalization constant, ξ is called orbital exponent, Y_{lm} is the angular momentum part (function describing shape). The n, l and m are principal, angular momentum and magnetic quantum number. Slater orbitals (in Cartesian co-ordinates)

$$\Phi_{klm} = N x^k y^l z^m e^{-\xi r} \quad \dots (3)$$

The exponential term depends on the nucleus-electron distance, which provides the exact orbitals for the hydrogen atom. This exponential dependence gives a reasonably fast convergence with increasing numbers of functions. STOs can be used also with DFT methods that do not contain exact exchange and in which the Coulomb energy is obtained by fitting the density into a set of auxiliary functions.

The calculation of three and four center two electron integrals in STO takes much time and provides no analytic form. In order to conquer this problem, Gaussian type orbitals (GTOs) were introduced. In 1950, Boys suggested GTOs instead of STOs which permits the analytical evaluation of integrals. Modern programs, that are used to calculate two-electron integrals, are prepared to work on Cartesian coordinates and they produce pure spherical functions by transforming the Cartesian components to spherical. Gaussian type orbitals in Cartesian coordinates are defined as

$$\Phi(\alpha, l, m, n, x, y, z) = N x^k y^m z^n \exp(-\xi r^2) \quad \dots (4)$$

Where N is normalization constant α is called exponent and l, m and n are not quantum numbers but simply integral exponents and $r^2 = x^2 + y^2 + z^2$. Unlike the GTO functions [8], STOs exhibits the correct cusp behavior at nucleus ($r \rightarrow 0$) with a discontinuous derivatives while GTO has zero slope at the nucleus and the desired exponential decay in the tail regions (from nucleus, as $r \rightarrow \infty$), GTO falls off too rapidly.

Classification of basis sets

The number of functions to be used is an important factor to decide the types of function (STO/GTO) and the nucleus location. The smallest possible number of functions is called minimal basis set. Commonly used minimal basis sets of this type are STO-2G, STO-3G, STO-6G, STO-3G* and polarized version of STO-3G. Subsequently improved basis set is produced by doubling of all basis functions, known as a

Double Zeta (DZ) type basis set. The term zeta is taken due to the exponent of STO basis functions is often represented by the Greek letter ζ . The variation of the DZ type basis only doubles the valence orbitals number to produce a split valence basis. The term DZ basis also stands for split valence basis sets or sometimes valence double zeta (VDZ) [43]. The commonly used split-valence basis sets are 3-21G, 3-21G*-polarized, 6-31G, 6-31+G*, 6-311+G*, 6-311++G*.

The next advancement in basis set size is a Triple Zeta (TZ) in which the number of functions is three times of minimal basis set. Some of the core orbitals again may be saved by splitting the valence orbitals only to produce a triple split valence basis set [10-11]. The next levels of basis sets are also used as quadruple zeta (QZ) and pentuple zeta (PZ), but large basis sets are generally given clearly in terms of the number of basis functions of each type. In most of the calculations, higher angular momentum functions are also significant, and these are represented by polarization functions. These polarization functions are added to the selected *sp*-basis. The inclusion of a single set of polarization functions to the DZ basis produces a DZ plus polarization (DZP) type basis set. Similarly, the multiple sets of polarization functions with different exponents can be added to the *sp*-basis sets. If two sets of polarization functions are included to a TZ *sp*-basis, then this type of basis set is called a TZ plus double polarization (TZ2P) type basis. Examples of these basis sets are cc-pVDZ as double-zeta, cc-pVTZ as triple-zeta, cc-pVQZ as quadruple-zeta, ccpV5Z as pentuple-zeta, etc. Besides zeta basis sets, there are basis sets as suggested by Pople called Pople style basis sets.

Pople style basis sets.

- **STO-*n*G basis sets:** these Pople style basis sets are STOs, which contains *n* primitive GTOs (PGTO). This is a minimum type basis in which the exponents of the PGTO are established by fitting to the STO, instead of optimizing them by a variational method. The basis sets for *n* = 2–6 have been established which shows that if more than three PGTOs are using for denoting the STO provides the little improvement. So the STO-3G basis is minimum basis set, which is widely used.
- ***k-nlmG basis sets*** These basis sets are of the split valence type basis set and proposed by Pople and co-workers. Here *k* indicating, the numbers of PGTOs and are used to represent the core orbitals and *nlm* indicates both number of functions, the valence orbitals are split into, and number of PGTOs are used for the representation of these split functions. These basis sets have the restriction of having the same exponent for both the *s* and *p*-functions. This raises the computational efficiency and decreases the flexibility of the basis set.
- **3-21G** This basis set is a split valence basis set, in which the core orbitals are made by the contraction of three PGTOs and the valence orbitals inner part is made by the contraction of two PGTOs

and the outer part of the valence is represented by one PGTO. This basis consists of same number of PGTOs as the STO-3G. However, 3-21G basis is more flexible.

- **6-31G** This basis set is also a type split valence basis in which the core orbitals are a contraction of six PGTOs, the inner part of the valence orbitals is a contraction of three PGTOs and the outer part of the valence is represented by one PGTO. This basis set consists of same number of functions as 3-21G, but the representation of each function is better due to the use of more PGTOs.
- **6-311G** This is a type of triple split valence basis in which the core orbitals are a contraction of six PGTOs and the valence split into three functions, represented by three, one and one PGTOs, respectively.

The addition of polarization and diffuse functions significantly improves the quality of basis set. Inclusion of polarization in 6-311G basis set is symbolized as 6-311G(d,p) or 6-311G**.

In cases, where the electron distribution differs much from the ground state, it is necessary to add diffuse functions in the basis, which have small exponent ranging from 0.01 to 0.1 allowing the electrons to spread farther away. That is, the diffuse functions improve the description of basis at large distances from the nuclei. The diffuse functions on heavy atoms are indicated with "+", if they are also included on hydrogen atoms, it is indicated with "++". Diffuse functions are important for anions, excited states and intermolecular complexes.

Conclusion

A basis set is a set of functions combined linearly to model molecular orbitals. Single Gaussian functions are not well suited to describe the spatial extend and nodal characteristics of atomic orbitals. For this, basis functions are described as a sum of several Gaussian functions. The suitability of the basis set depends upon the chemical problem in hand. Generally largest basis set available is applied in order to model molecular orbitals but due to high computational cost associated with higher basis sets some compromise is made. Basis sets have been constructed from Slater, Gaussian, plane wave and delta functions. Slater functions have the correct behavior at the origin and in the asymptotic regions. Delta functions are used in several quantum chemistry programs. Delta functions are simple but require thousands of functions to achieve accurate results, even for small molecules. The most important basis sets are contracted sets of atom-centered Gaussian functions. The quantum Chemistry programs are generally optimized to exploit basis sets of the contracted Gaussian function type and has a large number of built-in standard basis sets which the user can access quickly and easily.

References

- [1] S. Hazarika, R. L. Hebb & Syed S. H. Rizvi (2018) Prediction of Ripening Stage of Cameo Apple Using Fourier-Transform Infrared Spectroscopy, *International Journal of Fruit Science*, 18:2, 188-198, DOI: 10.1080/15538362.2017.1377671
- [2] Dunning T.H., Hay P.J. (1977) Gaussian Basis Sets for Molecular Calculations. In: Schaefer H.F. (eds) *Methods of Electronic Structure Theory. Modern Theoretical Chemistry*, vol 3. Springer, Boston, MA. https://doi.org/10.1007/978-1-4757-0887-5_1
- [3] Sathyanarayana, D.N. (2004). *Vibrational spectroscopy : theory and applications* / D.N. Sathyanarayana.
- [4] Gerald Knizia, *Journal of Chemical Theory and Computation* 2013 9 (11), 4834-4843, DOI: 10.1021/ct400687b
- [5] van Mourik T, Bühl M, Gaigeot MP. Density functional theory across chemistry, physics and biology. *Philos Trans A Math Phys Eng Sci*. 2014;372(2011):20120488. Published 2014 Feb 10. doi:10.1098/rsta.2012.0488
- [6] Hasnip Philip J., Refson Keith, Probert Matt I. J., Yates Jonathan R., Clark Stewart J. and Pickard Chris J. 2014 Density functional theory in the solid state *Phil. Trans. R. Soc. A*. 3722013027020130270
- [7] Al-Mahayni, H, Wang, X, Harvey, J-P, Patience, GS, Seifitokaldani, A. Experimental methods in chemical engineering: Density functional theory. *Can J Chem Eng*. 2021; 99: 1885–1911. <https://doi.org/10.1002/cjce.24127>
- [8] Nagy, B. and Jensen, F. (2017). Basis Sets in Quantum Chemistry. In *Reviews in Computational Chemistry* (eds A.L. Parrill and K.B. Lipkowitz). <https://doi.org/10.1002/9781119356059.ch3>
- [9] Brian C. Ferrari and Chris J. Bennett 2019 *J. Phys.: Conf. Ser.* 1290 012013
- [10] Boys, S.F., *Proc. Roy. Soc. (London)*, 1950, 200A, 542.
- [11] Kjær H, Sauer SP. Pople Style Basis Sets for the Calculation of NMR Spin-Spin Coupling Constants: the 6-31G-J and 6-311G-J Basis Sets. *J Chem Theory Comput*. 2011 Dec 13;7(12):4070-6. doi: 10.1021/ct200546q. Epub 2011 Nov 9. PMID: 26598352.