

Appendix 1.

Selection of a basis set

Models of a chemical system generally consist of the combination of a theoretical method such as restricted HF and B3LYP with a basis set. Each such unique pairing of method with basis set represents a different approximation to the Schrödinger equation. *Choosing a model of a chemical system, almost always involves a trade-off between accuracy and computational cost.* More accurate methods and larger basis sets make jobs run longer.

A basis set is a mathematical description of the orbitals within a system used to perform the theoretical calculation. Larger basis sets approximate more accurately the orbitals by imposing fewer restrictions on the locations of the electrons in space.

Standard basis sets for electronic structure calculations use linear combinations of gaussian functions to form the orbitals. Basis sets assign a group of basis functions to each atom within a molecule to approximate its orbitals. These basis functions themselves are composed of a linear combination of gaussian functions; such basis functions are referred to as *contracted functions*, and the component gaussian functions are referred to as *primitives*. A basis function consisting of a single gaussian function is termed *uncontracted*.

Cartesian Gaussian functions are of the form

$$g_{ijk} = Nx^i y^j z^k e^{-\alpha r^2}$$

where i, j and k are nonnegative integers, α is a positive orbital exponent and N is a normalization constant.

- **Minimal basis sets**

Minimal basis sets contain the minimum number of basis functions that are needed for each atom. For example,

H: 1s

C: 1s, 2s, 2p_x, 2p_y, 2p_z

Minimal basis sets use fixed size atomic type orbitals. The **STO-3G** basis set is a minimal basis set (through it is not the smallest possible basis set). It uses three

gaussian primitives per basis function (“3G”). “STO” stands for “Slater type orbitals”, and the STO-3G basis set approximates Slater orbitals with gaussian functions. STO has the form

$$Nr^{n-1}e^{-\xi r}Y_l^m(\theta, \phi)$$

where $Y_l^m(\theta, \phi)$ are the spherical harmonic.

- **Split valence basis sets**

One way to increase the size of a basis set is to take more basis functions per atom. *Split valence basis sets*, such as 3-21G and 6-31G basis sets, have two (or more) sizes of basis function for each valence orbital. For example, in the above 3-21G and 6-31G basis sets we have:

H: 1s, 1s'

C: 1s, 2s, 2s', 2p_x, 2p_y, 2p_z, 2p_x', 2p_y', 2p_z'

Here the primed and unprimed orbitals differ in size. The *double zeta valence basis sets* form molecular orbitals from the linear combinations of two sets of functions for each atomic valence orbital. Similarly, *triple split valence basis sets* such as 6-311G, use three sets of contracted functions for each valence orbital type.

- **Polarized basis sets**

Split valence basis sets could be improved by adding orbitals with different shapes. *Polarized basis sets* add orbitals with angular momentums going beyond of requirement for the proper description of the ground state of each atom at the HF level. For example, polarized basis sets add *d*-functions to carbon atoms and some of them add *p*-functions to hydrogen atoms. Examples for polarized basis sets are the 6-31G(d) and the 6-311G(d, p) basis sets.

- **Diffused functions**

Basis sets with additional *diffuse functions* are large by size versions of *s*- and *p*-type split valence basis sets. Diffuse orbitals occupy a larger region of space. Basis sets with diffuse functions are important for systems where electrons may be far from the nucleus. One example for diffuse basis function is the 6-311+G(d, p) basis set (see table).

The table below summarizes the number of basis functions of each and the number of primitive gaussian functions they are constructed of.

Atomic number	1 - 2	3-10
$n_1 - n_2 n_3 n_4 G$ Basis functions	1s constructed of n_2 gaussians 1s' constructed of n_3 gaussians 1s'' constructed of n_4 gaussians	1s constructed of n_1 gaussians 2s, 2p each constructed of n_2 gaussians 2s', 2p' each constructed of n_3 gaussians 2s'', 2p'' each constructed of n_4 gaussians
$n_1 - n_2 n_3 n_4 G(d)$	$n_1 - n_2 n_3 n_4 G$ basis functions plus nothing	$n_1 - n_2 n_3 n_4 G$ basis functions plus 3d each constructed of 1 gaussian
$n_1 - n_2 n_3 n_4 G(d, p)$	$n_1 - n_2 n_3 n_4 G$ basis functions plus 2p each constructed of 1 gaussian	$n_1 - n_2 n_3 n_4 G(d)$ basis functions plus nothing
$n_1 - n_2 n_3 n_4 + G$	$n_1 - n_2 n_3 n_4 G$ basis functions plus nothing	$n_1 - n_2 n_3 n_4 G$ basis functions plus 3s, 3p each constructed of 1 gaussian
$n_1 - n_2 n_3 n_4 + + G$	$n_1 - n_2 n_3 n_4 G$ basis functions plus 2s constructed of 1 gaussian	$n_1 - n_2 n_3 n_4 + G$ functions plus nothing

Examples:

STO-3G – each occupied atomic orbital is constructed from three gaussian functions.

3-21G

Atomic number	1-2		3-10	
	Basis functions	Constructed of	Basis functions	Constructed of
	1s	2 gaussians	1s	3 gaussians
	1s'	1 gaussian	2s, 2p _x , 2p _y , 2p _z	2 gaussians
			2s', 2p _x ', 2p _y ', 2p _z '	1 gaussian
Total number	2	3	9	15

6-311G(d)

Atomic number	1-2		3-10	
	Basis functions	Constructed of	Basis functions	Constructed of
	1s	3 gaussians	1s	6 gaussians
	1s'	1 gaussian	2s, 2p _x , 2p _y , 2p _z	3 gaussians
	1s''	1 gaussian	2s', 2p _x ', 2p _y ', 2p _z '	1 gaussian
			2s'', 2p _x '', 2p _y '', 2p _z ''	1 gaussian
			3d _{xx} , 3d _{yy} , 3d _{zz} , 3d _{xy} , 3d _{xz} , 3d _{yz}	1 gaussian
Total number	3	5	19	32

6-311G(d,p)

Atomic number	1-2		3-10	
Basis functions	Basis functions	Constructed of	Basis functions	Constructed of
	1s	3 gaussians	1s	6 gaussians
	1s'	1 gaussian	2s, 2p _x , 2p _y , 2p _z	3 gaussians
	1s''	1 gaussian	2s', 2p _x ', 2p _y ', 2p _z '	1 gaussian
	2p _x , 2p _y , 2p _z	1 gaussians	2s'', 2p _x '', 2p _y '', 2p _z ''	1 gaussian
			3d _{xx} , 3d _{yy} , 3d _{zz} , 3d _{xy} , 3d _{xz} , 3d _{yz}	1 gaussian
Total number	6	8	19	32

6-311++G(d,p)

Atomic number	1-2		3-10	
Basis functions	Basis functions	Constructed of	Basis functions	Constructed of
	1s	3 gaussians	1s	6 gaussians
	1s'	1 gaussian	2s, 2p _x , 2p _y , 2p _z	3 gaussians
	1s''	1 gaussian	2s', 2p _x ', 2p _y ', 2p _z '	1 gaussian
	2p _x , 2p _y , 2p _z	1 gaussian	2s'', 2p _x '', 2p _y '', 2p _z ''	1 gaussian
	2s	1 gaussian	3d _{xx} , 3d _{yy} , 3d _{zz} , 3d _{xy} , 3d _{xz} , 3d _{yz}	1 gaussian
			3s, 3p _x , 3p _y , 3p _z	1 gaussian
Total number	7	9	23	36

The table below describes some common basis sets

Basis set [Applicable atoms]	Description	# basis functions for atomic number 3-10	# basis functions for atomic number 1-2
STO-3G [H – Xe]	Minimal basis set: is used for more qualitative results on very large systems (when one cannot afford even 3-21G).	5	1
3-21G [H – Xe]	Split valence: two sets of functions in the valence region provide a more accurate representation of orbitals. The basis is used for very large molecules for which 6-31G(d) is too expensive.	9	2
6-31G(d) [H – Cl]	Adds polarization functions to heavy atoms. The basis is used in many calculations of medium/large sized systems. (This basis set uses the 6-component type <i>d</i> functions)	15	2
6-311+G(d, p) [H – Br]	Triple zeta: adds extra valence functions (three sizes of <i>s</i> - and <i>p</i> -functions).	23	7

For additional details about basis sets please see

http://www.gaussian.com/g_ur/m_basis_sets.htm