Chemistry 460
Spring 2013
Dr. Jean M. Standard
April 25, 2013

Basis Set Notation

Using the LCAO-MO approximation, molecular orbitals can be represented as linear combinations of atomic orbitals,

$$\phi_i = \sum_{\mu=1}^{K} c_{\mu i} f_\mu \ . \tag{1}$$

In this equation, the functions ϕ_i are molecular orbitals, f_μ are atomic orbitals, $c_{\mu i}$ are numerical coefficients, and K is the total number of atomic orbital functions (or basis functions).

General Basis Set Types

1. Minimal Basis Set

A minimal basis set is constructed by using one basis function of each type occupied in the separated atoms that comprise a molecule. If at least one p-type or d-type orbital is occupied in the atom, then the complete set (3 p-type or 5 d-type) of functions must be included in the basis set.

For example, consider the carbon atom. In the atom, the electron configuration is $1s^2 2s^2 2p^2$. Thus, a minimal basis set for carbon atom would consist of 1s, 2s, and $2p_x$, $2p_y$, and $2p_z$ orbitals, a total of five basis functions. Rather than writing the three 2p orbitals explicitly each time, this set is often denoted 2p(3).

As a second example, consider the methane molecule, CH_4. For methane, a minimal basis set consists of

Atom/Basis Functions	No. Basis Functions	No. Atoms	TOTAL
C -- 1s 2s 2p(3)	5	1	5
H -- 1s	1	4	4

TOTAL NO. BASIS FUNCTIONS = 9

2. Double Zeta Basis Set

Since the electronic energy of a molecule decreases and approaches the exact value more closely as the number of basis functions is increased in a calculation (due to the Variation Principle), it is useful to construct basis sets that employ more than the minimal number of basis functions. The Double Zeta (DZ) basis set does this by using two basis functions of each type found in the separated atoms.

For the carbon atom, a double zeta basis set would consist of two 1s, two 2s, and two each of $2p_x$, $2p_y$, and $2p_z$ orbitals, for a total of 10 basis functions. The two 1s orbitals employed in the basis set are not identical. Rather, they have somewhat different orbital exponents . The same is true for the 2s and 2p orbitals. The basis functions in the double zeta basis set are denoted 1s, 1s′, 2s, 2s′, 2p(3), and 2p′(3).

To see why it might be useful to have basis functions of the same type with different orbital exponents, consider the general form of a 1s-type gaussian type orbital,

$$f_{1s}(r) = N_{1s} e^{-\zeta r^2} \ . \tag{2}$$

In Eq. (2), the factor N_{1s} is a constant, r is the distance from the origin (where the nucleus is located, and ζ is the orbital exponent.

For a given atom, the orbital exponent for the 1s orbital is known. However, in a molecule, the electron density around the atom will not be exactly the same as it is in the separated atom due to interaction with the other atoms in the molecule. Therefore, the orbital exponent for the 1s orbital for the atom in the molecule is not known exactly. By using two different 1s-type orbitals in the basis set, it allows for variations in the electron density of the atom due to bonding and other interactions.

The orbital exponent controls the size of the basis function. For example, two 1s-type orbitals are plotted in Fig. 1. The first, denoted $f_1(r)$, has an orbital exponent of 3.0 $Å^{-2}$, while the second, denoted $f_2(r)$, has an orbital exponent of 1.0 $Å^{-2}$.

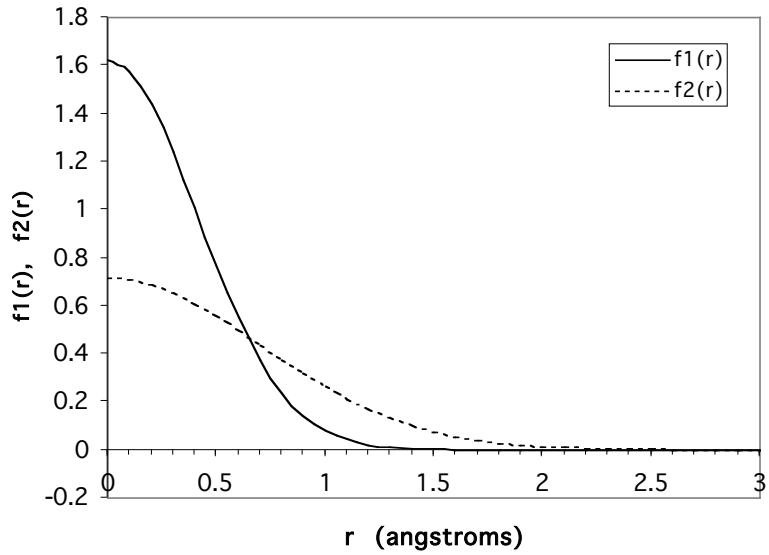

Figure 1. The shapes of two 1s-type orbitals with different orbital exponents. The orbital exponent for function 1 is 3.0 $Å^{-2}$, while the orbital exponent for function 2 is 1.0 $Å^{-2}$.

Notice in Fig. 1 that the basis function with a larger orbital exponent (f_1) is more compact (it dies off more rapidly). The basis function with the smaller orbital exponent (f_2) is broader and more diffuse. The pair of orbitals can be used to represent better the true molecular electron density than either alone could. The same procedure is used in representing the 2s and 2p orbitals with two basis functions with different orbital exponents.

Consider a double zeta basis set for the methane molecule, CH_4. This basis set would consist of the following basis functions.

Atom/Basis Functions	No. Basis Functions	No. Atoms	TOTAL
C -- 1s 1s' 2s 2s' 2p(3) 2p'(3)	10	1	10
H -- 1s 1s'	2	4	8

TOTAL NO. BASIS FUNCTIONS = 18

3. Triple Zeta Basis Set

The Triple Zeta (TZ) basis set goes even further in increasing the size of the basis set (in an effort to get even closer to the exact electronic energy). In this case, the triple zeta basis set employs three basis function of each type occupied in the separated atoms.

For the carbon atom, a triple zeta basis set would consist of three 1s, three 2s, and three each of $2p_x$, $2p_y$, and $2p_z$ orbitals, for a total of 15 basis functions. The three basis functions of each type employed in the basis set all have different orbital exponents . The basis functions in the triple zeta basis set are denoted 1s, 1s′, 1s″, 2s, 2s′, 2s″, 2p(3), 2p′(3), and 2p″ (3).

For the methane molecule, CH_4, the triple zeta basis set consists of the following basis functions.

Atom/Basis Functions	No. Basis Functions	No. Atoms	TOTAL
C -- 1s 1s' 1s″ 2s 2s' 2s″ 2p(3) 2p'(3) 2p″ (3)	15	1	15
H -- 1s 1s' 1s″	3	4	12

TOTAL NO. BASIS FUNCTIONS = 27

Note that there are also quadruple zeta and higher basis sets.

4. Split Valence Basis Set

Because the core electrons of an atom are less affected by the chemical environment than the valence electrons, they are sometimes treated with a minimal basis set while the valence electrons are treated with a larger basis set (of double zeta or higher quality). This is known as a split valence basis set. A basis consisting of a minimal basis set for the core electrons and a double zeta basis set for the valence electrons would be called a split valence double zeta basis set. Split valence triple zeta basis sets are also commonly used. Split valence basis sets are also used for larger molecules because they reduce the CPU time required for the calculation.

For the carbon atom, a split valence double zeta basis set would consist of a single 1s orbital, along with two 2s and two each of $2p_x$, $2p_y$, and $2p_z$ orbitals, for a total of 9 basis functions. The basis functions in a split valence double zeta basis set are denoted 1s, 2s, 2s′, 2p(3), and 2p′(3).

For the methane molecule, CH_4, the split valence double zeta basis set consists of the following basis functions. Note that for the H atom, because the 1s electron is considered the valence shell a double zeta basis set is used.

Atom/Basis Functions	No. Basis Functions	No. Atoms	TOTAL
C -- 1s 2s 2s' 2p(3) 2p'(3)	9	1	9
H -- 1s 1s'	2	4	8

TOTAL NO. BASIS FUNCTIONS = 17

Polarization Basis Functions

Another way to increase the size of the basis set in order to get closer to the exact electronic energy and wavefunction is to include polarization functions in the basis set. A polarization function is any higher angular momentum orbital used in a basis set that is not normally occupied in the separated atom. For example, for the hydrogen atom, the only orbital type that is occupied is s-type. Therefore, if p-type or d-type basis functions were added to the hydrogen atom they would be known as polarization functions. For first row elements like carbon, d-type and f-type basis functions would be considered to be polarization functions. For transition metals with occupied d-type orbitals, only f-type or higher functions would be considered polarization basis functions.

Polarization functions are not thought to be formally fully occupied in molecules. They are included solely to improve the flexibility of the basis set, particularly to better represent electron density in bonding regions. For example , consider the addition of a p-type orbital to the existing s-type orbital in a hydrogen atom. While the electron density of the isolated hydrogen atom is spherical like the s-type orbital, when it is involved in a bond in a molecule (like the C-H bond in methane), its electron density is no longer spherical. The electron density is shifted (or polarized) so that it lies more along the C-H bond.

The use of polarization basis functions allows for the atomic electron densities to be polarized in order to better represent the electron density of the molecule. Consider the sum of an s-type orbital and a p-type orbital illustrated in Fig. 2. One lobe of the p-type orbital is enhanced by the addition of the s-type orbital, while the lobe of opposite sign cancels and little is left. This has the effect of polarizing the electron density from its original spherical shape into a form that is polarized in one direction.

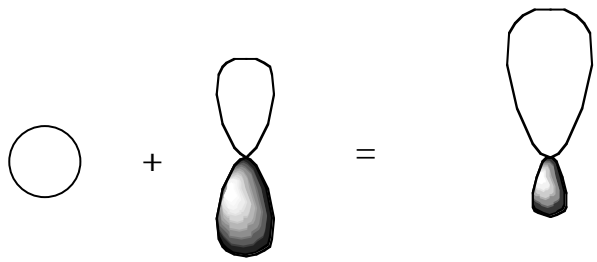

Figure 2. The addition of s-type and p-type basis functions. The p-type basis function allows for polarization of the electron density in a particular direction.

Diffuse Basis Functions

Diffuse basis functions are extra basis functions (usually of s-type or p-type) that are added to the basis set to represent very broad electron distributions. The orbital exponents of these diffuse basis functions are very small. They are especially important in representing the electron density in anions or in intermolecular complexes (where there may be particularly long bonds with electron density spread over a large region).

Fig. 3 illustrates a typical diffuse s-type function compared to an ordinary s-type basis function. The ordinary s-type basis function, denoted $f_1(r)$, has an orbital exponent of 3.0 Å^{-2}, while the diffuse basis function, denoted $f_2(r)$, has an orbital exponent of 0.2 Å^{-2}. Notice that the extent of the diffuse basis function is much larger. The ordinary s-type basis function has died off by a distance of about 1 Å, while the diffuse function extends out beyond 3 Å.

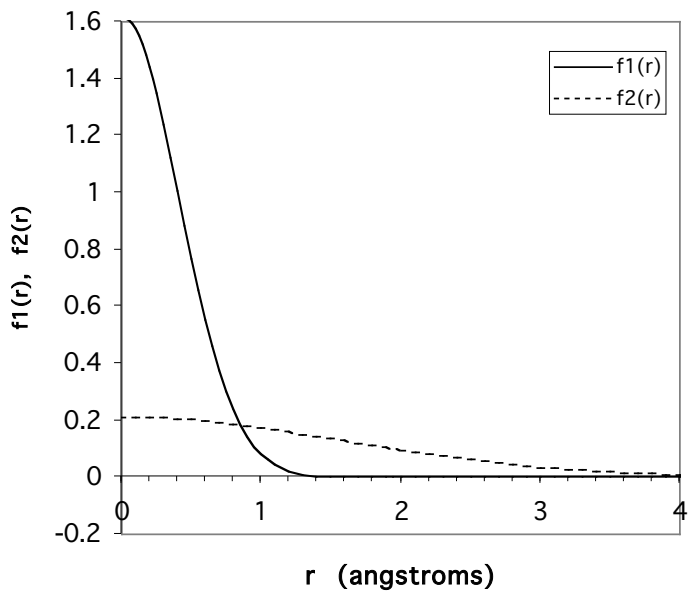

Figure 3. Comparison of an ordinary s-type basis function (function 1) to a diffuse s-type basis function (function 2). The orbital exponent for function 1 is 3.0 Å^{-2}, while the orbital exponent for function 2 is 0.2 Å^{-2}.

Pople Basis Sets

General Notation and Gaussian Primitives

Basis sets denoted by the general nomenclature N-M1G or N-M11G, where N and M are integers, are called Pople basis sets. The first, N-M1G, is a split valence double zeta basis set while the second, N-M11G is a split valence triple zeta basis set.

In Pople basis sets, the basis functions are made to look more like Slater Type Orbitals by representing each basis function as a sum of gaussian primitives. A gaussian primitive has the form

$$g_\ell(x, y, z) = N_\ell \; x^a y^b z^c \, e^{-\alpha_\ell r^2} \; . \tag{3}$$

In Eq. (3), N_ℓ and α_ℓ are constants and a, b, and c are integers. Each basis function f_μ in Eq. (1) consists of a linear combination of a small number of these gaussian primitives,

$$f_\mu = \sum_{\ell=1}^{L_\mu} g_\ell(x,y,z)\, d_{\ell\mu} \; . \tag{4}$$

In Eq. (4), $d_{\ell\mu}$ is a fixed coefficient and L_μ is the number of gaussian primitives used to represent the basis function.

For Pople basis sets, the ″G″ in the name simply indicates that gaussian primitives are used. The integers N and M in the basis set name give the values of L_μ used. For example, in the split valence double zeta basis set 6-31G for a carbon atom, the first number (N=6) represents the number of gaussian primitives used to construct the core orbital basis function (the 1s function). The second two numbers (M=3 and 1) represent the valence orbitals, 2s, 2s′, 2p(3), and 2p′(3). The first number after the dash in the basis set name (3 in this case) indicates the number of gaussian primitives used to construct the 2s and 2p(3) basis functions. The second number after the dash (1 in this case) gives the number of gaussian primitives used to construct the 2s′ and 2p′(3) basis functions.

Linear combinations of gaussian primitives are used to represent a basis function in order to improve the representation of the electron density close to the nucleus. Recall that a Slater Type Orbital has a cusp at the nucleus, while a Gaussian Type Orbital does not. By taking linear combinations of gaussian primitives, the cusp-like behavior is better reproduced.

To illustrate this, consider a Slater Type Orbital representing an s-type orbital with the form

$$s(r) = N \, e^{-\zeta r} , \tag{5}$$

where N and ζ are constants. A plot of this function with an orbital exponent of 1.0 is shown in Fig. 4 along with a Gaussian Type Orbital with the same orbital exponent and the form

$$g(r) = N \, e^{-\zeta r^2} . \tag{6}$$

Note the cusp at the origin for the Slater Type Orbital while the Gaussian Type Orbital has no cusp at the origin. In addition, the shape of the Gaussian Type Orbital is different due to the squared distance that appears in its exponent. All in all, the two orbitals do not look that similar.

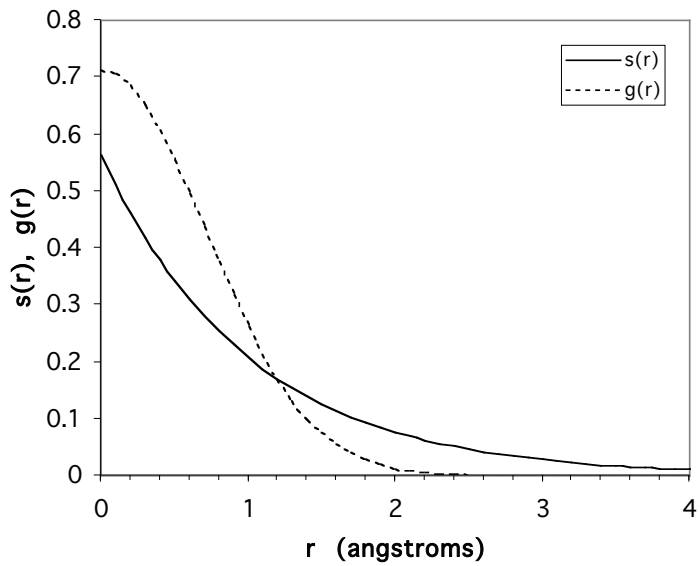

Figure 4. Comparison of Slater and Gaussian Type Orbitals. The orbital exponent for both orbitals is 1.0.

Now consider a basis function comprised of three gaussian primitives as given in Eq. (4) such that $L_\mu = 3$,

$$f(r) = \sum_{\ell=1}^{3} g_\ell(x,y,z)\, d_\ell \ . \tag{7}$$

The gaussian primitives for an s-type orbital have $a = b = c = 0$ in Eq. (3). Consider the specific case where $\alpha_1 = 1.0982 \times 10^{-1}$, $\alpha_2 = 4.0577 \times 10^{-1}$, and $\alpha_3 = 2.22766$. The three gaussian primitive functions $g_1(r)$, $g_2(r)$, and $g_3(r)$ are shown plotted individually in Fig. 5. Notice that the functions range from quite broad to fairly narrow in extent.

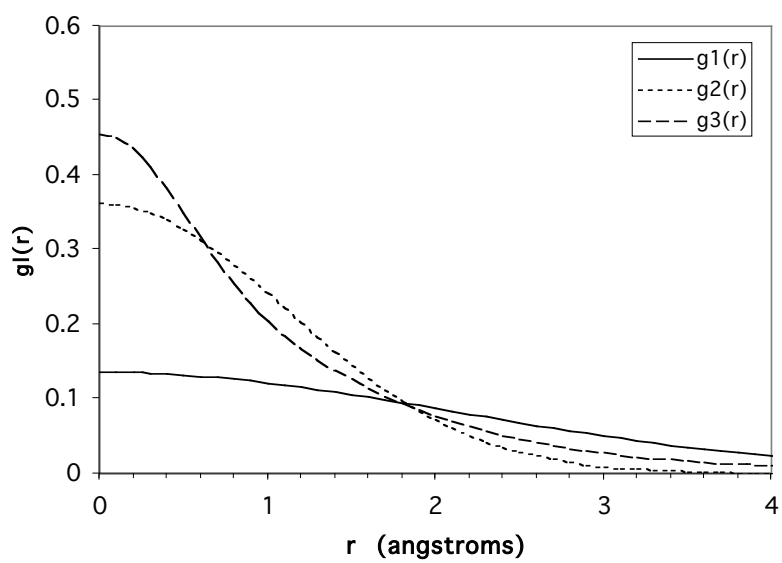

Figure 5. Plots of three gaussian primitive orbitals.

When the gaussian primitives shown in Fig. 5 are combined using the coefficients $d_1 = 4.446 \times 10^{-1}$, $d_2 = 5.353 \times 10^{-1}$, and $d_3 = 1.543 \times 10^{-1}$, a new basis function is produced according to Eq. (7). This basis function is shown in Fig. 6 compared to the Slater Type Orbital given in Fig. 4.

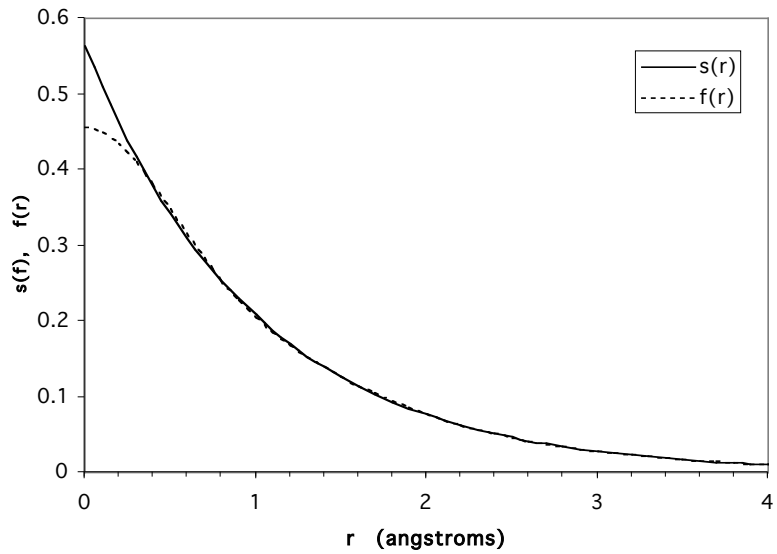

Figure 6. Construction of a basis function from the three gaussian primitive orbitals shown in Fig. 5 and comparison with the Slater Type Orbital of Fig. 4.

Note in Fig. 6 that the new basis function resembles the Slater Type Orbital much more closely than the single Gaussian Type Orbital shown in Fig. 4. The only slight variation is observed near the cusp. The new basis function, because it consists of gaussian primitive functions, does not have a cusp. However, the overall shape even close to the nucleus is very much like the Slater Type Orbital.

Notation for Polarization Functions

Polarization functions can also be added to Pople basis sets. There are two common methods for designating that polarization functions are included in a basis set. The first method is to use * or ** after the Pople basis set name; for example, 6-31G* or 6-31G**. The single * means that one set of d-type polarization functions is added to each non-hydrogen atom in the molecule. The double ** means that one set of d-type polarization functions is added to non-hydrogens and one set of p-type polarization functions is added to hydrogens.

The second method for including polarization functions in the basis set designation is more general. It is indicated by the notation (L1,L2) following the Pople basis set name; for example, 6-31G(d) or 6-31G(d,p). The first label indicates the polarization functions added to non-hydrogen atoms in the molecule. The notations 6-31G(d) and 6-31G(d,p) mean that one set of d-type polarization functions is added to all non-hydrogens. The notation 6-311(2df) means that two sets of d-type and one set of f-type polarization functions are added to non-hydrogens. The second label in the notation (L1,L2) indicates the polarization functions added to hydrogen atoms. The basis set 6-31G(d) has no polarization functions added to hydrogen, while the basis 6-31G(d,p) has one set of p-type polarization functions added to hydrogen atoms. The basis set 6-311G(2df,2pd) has two sets of p-type and one set of d-type polarization functions added to hydrogen atoms.

Notation for Diffuse Functions

The use of diffuse functions in a Pople basis set is indicated by the notation + or ++. The + notation, as in 6-31+G(d), means that one set of sp-type diffuse basis functions is added to non-hydrogen atoms (4 diffuse basis functions per atom). The ++ notation, as in 6-31++G(d), means that one set of sp-type diffuse functions is added to each non-hydrogen atom and one s-type diffuse function is added to hydrogen atoms.