Chapter 3

Basis Sets for Computational Chemistry

J.M. García de la Vega and B. Miguel

Departamento de Química Física Aplicada, Facultad de Ciencias,

Universidad Autónoma de Madrid, 28049 Madrid, Spain [1]

3.1   Introduction

The widespread availability of programs for electronic structure and the advances in the technology of computers have improved the theoretical calculations which support the experimental research. Ab initio electronic structure methods are used to explain molecular properties. However, the quality of the computed properties depends on the quality of the data and/or the parameters supplied to the programs. The parameters in the *ab initio* calculations are those defining the functions of the basis set required for expanding the molecular orbitals or electron density. In order to achieve acceptable accuracy the choice of both parameters is essential. In practice, most popular computer programs for *ab initio* calculations contain internally defined basis sets from which the user must select an appropriate one. In principle this should facilitate the choice of the basis set.

During the past years there has been an important increase in the variety of functions used in atomic and molecular quantum mechanics. It is clearly not possible

to give a completely comprehensive account of all developments which have recently been made in the field of basis set construction. This field forms the foundation upon which the vast majority of current atomic and molecular electronic structure studies are based. In this review, we shall be concerned with the use of one-electron basis sets in quantum chemistry at introductory level, trying to show the background material needed to choose suitable basis sets for quantum chemical calculations. We also attempt to provide the nonexpert with rough guidelines for selecting among the many basis sets in the chemistry literature. Although the topic of basis sets is often considered highly technical, at least a quick understanding is essential to those who want to apply *ab initio* methods to questions of chemical interest. We shall proceed by reviewing some qualitative analytical features of wavefunctions. A certain amount of historical material about basis sets development will be presented. However, we make no attempt to present a complete overview, since these topics have been iteratively reviewed previously with different focus and extension [1-9].

The first step in any theoretical calculation of atoms and molecules is the determination of atomic and molecular wavefunctions, which are usually built up from a set of one-electron one-center functions. The major requirement is that the set has a size size that the functions were either a good representation or able to be adjusted to be a good representation of the final function, at least in the most significant regions. The basis set is fixed in a semiempirical way by calibrating calculations on a variety of molecules. A sequence of calculations with improving basis sets is applied to one molecule until the convergence is reached. The error in the calculation is estimated from the sensitivity of the results to further refinements in the basis sets. Clearly, the calibrated method is the only one that is practical for routine use in computational chemistry. The existence of such a vast multitude of basis sets is attributable, at least in part, to the difficulty of finding a single set of functions. The basis set should be flexible enough to produce good results over a wide range of molecular geometries and sufficiently small so that the problem is computationally easy and economically reasonable.

The truncation of the basis set is an important source of error in molecular electronic structure calculations. Most recent work has recognized the need for a more systematic approach to the problem of constructing basis sets with the aim of reducing the basis set truncation error in molecular electronic structure calculations. The great number of basis sets makes their choice of them a non trivial problem. Since the choice of a basis set must be guided by considerations of the desired accuracy of the results and the computational cost, it becomes imperative that the available information were sufficient so that an intelligent choice is possible. Some of the new basis sets appearing in the literature are used for a single project and then disappear into oblivion. Other are proposed as serious contenders to replace basis sets in common use. Even experts in the field find difficult to keep up with the literature as the proliferation of new basis sets continues.

The computer time required for some parts of an *ab initio* calculation depends strongly on the number of basis functions. In one of those parts, the integral evaluation goes as the fourth power of the number of functions. As stated above, even a small reduction in the size of the basis set can result in a huge overall speedup. The amount of computer time can often be measured in days. Therefore, there is great interest in finding small basis sets that produce results equivalent to those obtained with much larger sets. Since then several new basis sets have appeared in the theoretical literature aimed to further minimizing the computational expense of *ab initio* work while still maintaining an acceptable level of agreement with experimental results. For example, introducing minor changes in a determined basis set can aid to carry out a large calculation which would be easily controlable.

Since the first review, published more than 25 years ago[1], several authors have surveyed basis set construction. Wilson[10] examined methods for constructing basis sets of Gaussian functions and contracted Gaussian functions. Ahlrichs and Taylor [2] also discussed the choice of Gaussian basis sets. Feller and Davidson[6] have updated with their reviews in the 90's some practical aspects of Gaussian basis sets, and one of these authors [11] has selected the most useful database with information of computational chemistry and specially of Gaussian basis sets. The recent review of basis sets by Helgaker and Taylor[9] is other useful tool for theoretical chemists who use Gaussian basis sets. This review gives ample information about Gaussian basis set and their molecular integrals. Čársky and Urban's book[12] provides a particularly useful survey of both exponential and Gaussian basis sets. Hehre et al[13] provide some discussion of the basis set problem in their book. Huzinaga and coworkers [14]have published a monograph dealing with the construction of Gaussian basis sets. Poirier et al. [15]have collected in their Handbook the most common Gaussian basis sets.

The present work is divided into six sections. In section 2, the nature of orbital approximation as an expansion is considered, but taking into account that extensive details are usually included in quantum chemistry books. Section 3 is devoted to an overview of the exponential type orbitals which are less frequently employed in the electronic structure calculations. However, its use has recently increased and thus this kind of functions has received less attention in previous reviews. Section 4 deals with the primitive Gaussian type orbitals and section 5 with the contraction of those primitives. The Gaussian basis sets optimized with electronic correlation effects are considered in section 6. The important topic of the basis set superposition error is discussed in section 7. And finally in section 8, the state of the art of the basis set problem is overviewed and possible future directions are conjectured.

3.2 Orbital approximation

For a system with N electrons and ν nuclei the independent time Schrödinger equation is:

$$\left\{ \sum_{i=1}^{N} \left(-\frac{1}{2}\nabla_i^2 - \sum_{\alpha=1}^{\nu} \frac{Z_{i\alpha}}{r_{i\alpha}} \right) + \sum_{i<j}^{N} \frac{1}{r_{ij}} + \sum_{\alpha<\beta}^{\nu} \frac{Z_\alpha Z_\beta}{r_{\alpha\beta}} \right\} \Psi = E\Psi \qquad (3.1)$$

∇_i^2 being the Laplacian operator in the i electron coordinates, Z_α being the nuclear charge of α nucleus, $r_{i\alpha}$ being the distance between electron i and nucleus α, r_{ij} being the distance between electrons i and j and $R_{\alpha\beta}$ being the internuclear distance between α and β nuclei. In the *ab initio* methods, the exact electrostatic Hamiltonian of the system is considered, taking into account explicitly all the electrostatic interactions between nuclei and electrons in the system.

In equation (3.1) the Hamiltonian of a system composed by pointwise positive and negative charges appears so that the finite size of the atomic nuclei is neglected. Also it is a non-relativistic Hamiltonian and eq. (3.1) is the electronic Schrödinger equation under the Born-Oppenheimer approximation. Even with these approximations eq. (3.1) has only an analytical solution for extremely simple systems such as the hydrogen atom or the hydrogen ion molecule, due to the r_{ij}^{-1} term in the Hamiltonian.

The first approximation used in the resolution of eq. (3.1) is the so-called independent particles model or orbital approximation. It is a simplified model in which the repulsion term is neglected. Under this approximation the electronic wavefunction can be written as a product of monoelectronic functions. If this product is antisymmetrized, as the Pauli's principle requires, the resulting simplest wavefunction for a closed shell system is the Slater determinant:

$$\Psi = \frac{1}{\sqrt{N!}} |\phi_1\alpha(1) \ \ \phi_1\beta(2) \ \ .... \ \phi_{N/2}\alpha(N-1) \ \ \phi_{N/2}\beta(N) \ | \qquad (3.2)$$

$\{\phi_i\}_{i=1}^{N}$ being the monoelectronic function set or orbital set. When the variational principle is applied to eq. (3.1) and to this function (eq. (3.2)) the obtained integer-differential equations are known as Hartree-Fock[16-18] or HF equations. In atoms in which there is only one nucleus, it is possible to perform the integration over the angular coordinates of the electrons giving integer-differential equations in the r coordinate that can be solved numerically[19]. Some numerical resolutions to the HF equations have also been carried out for molecules[20].

In 1951 Roothaan[21] and Hall[22, 23] proposed the development of the orbitals as a linear combination of a known basis set. For example, in a molecule the molecular orbitals are written as a linear combination of atomic orbitals or LCAO. When including this development of the orbitals in the HF equations, a set of algebraic equations is obtained which is known as Hartree-Fock-Roothaan-Hall equations. They

can be solved using matrix techniques. If the set in which the orbitals are expanded were a complete basis set, the LCAO approximation would be exact independently of the chosen set. Unfortunately, the sets actually used must have a limited number of functions due to practical reasons of calculation. The non-completeness of the basis set introduces the so called truncation error of the basis set. The expansion of the wavefunction as a linear combination of basis functions will affect the accuracy of the Hartree-Fock-Roothaan-Hall equations. Since the basis set is not complete its choice is a source of error in the *ab initio* molecular calculations.

The HF equations are obtained with the wavefunction introduced in eq. (3.2). That is, the HF method is monoreferential, only one reference configuration associated to the HF wavefunction is considered. This HF wavefunction is not a complete representation of the electronic interactions, due to the average potential used to evaluate the electron interactions. It is clear that the electronic motion should be correlated, that is, the allowed regions in the space for one electron will be influenced by the presence of the other electrons. Actually, the electron motion should be more correlated than the description given by the HF method due to the strong repulsion between two electrons when r_{ij} goes to zero.

Löwdin[24] introduced the term *correlation energy* as the difference between the exact energy of a system and the limit HF energy. This correlation energy gives an idea of the quality of the HF approximation for a given system. The first proposed method to evaluate the correlation energy was the *configuration interaction* (CI). From the HF calculation the occupied and virtual orbitals are obtained. With these orbitals it is possible to construct excited wavefunctions and taking them as polielectronic basis set the Hamiltonian matrix is obtained. From its diagonalization a part of the correlation energy is recovered. This contribution would be dependent on the truncation in the polielectronic basis. Posteriorly a larger number of correlation methods have been proposed. For a review, the references in [25] are very useful.

3.3   Exponential Type Orbitals

The first intuition in the choice of a mathematical function for performing an atomic calculation is to use similar functions to those of a hydrogen atomic orbital. For this reason Zener[26] and Slater[27] proposed, for the description of the radial part of the atomic orbitals, the functions called Slater orbitals which are defined by,

$$r^{n^*-1} \; e^{-[(Z-s)/n^*]r} \tag{3.3}$$

n^* being the effective quantum number, s being the screening constant and Z being the atomic number. n^* and s parameters were defined by the Slater's rules. That is, they were not determined variationally, and their values were determined before

performing the energy calculation. Afterwards, Roothaan and Bagus[28] wrote an SCF code for atoms under the LCAO approximation. They introduced the functions

$$r^{n-1}e^{-\zeta r}\,Y_{lm}\left(\theta,\phi\right) \tag{3.4}$$

where n, l y m are the principal, azimuthal and magnetic quantum numbers, respectively; $Y_{lm}\left(\theta,\phi\right)$ is the spherical harmonic corresponding to l and m, and ζ is the orbital exponent. The ζ parameters were optimized variationally with respect to each total atomic energy. In general, these functions are named *Slater type orbitals* (STO) or *Slater type functions* (STF).

It is well-known the excellent behavior shown by the STO functions in the near and far regions of the atomic nucleus. They satisfy the nuclear cusp condition[29] because of their exponential relationship with the nucleus-electron distance. This fact allows the STO basis sets to reproduce correctly the regions near the nucleus without increasing the angular quantum numbers. On the other hand, they decay properly in the long r distances, in a similar way to that of the atomic orbitals in the regions far away from the nucleus.

For the atoms from He to Kr, basis sets of STO functions were optimized with respect to the atomic energy by Clementi[30]. This work was posteriorly reviewed, improved and extended until the Xe atom by Clementi and Roetti[31], and until the Rn atom by McLean and McLean[32]. The atomic orbitals thus obtained are collected in extended tables which include the optimal exponents ζ and quantum numbers n and l of the employed STOs, as well as the coefficients of the atomic orbitals expanded in the given STO basis sets. In those papers three sizes of basis sets depending on the number of employed STOs for every occupied atomic orbital were proposed. Clementi and Roetti named them single-zeta (SZ or single-ζ), double-zeta (DZ) and limit Hartree-Fock (LHF) basis sets. In the SZ basis sets the number of STOs coincides with the number of occupied atomic subshells. That is, there is only one ζ parameter for each subshell. If there are two STOs by subshell, so they have identical n and l but different ζ, a DZ basis set is defined. If the size of the basis set is augmented until the basis set completeness is achieved, the defined basis set is called LHF. This nomenclature has become standard in the quantum chemistry field, irrespectively of the type of functions defining the basis set. If the basis set is a DZ only for the valence subshells, that is, it is a SZ for the core shells, the basis set is named split-valence or SV. Similarly, when there are three functions for each occupied atomic orbital, we refer to triple-zeta basis sets (TZ) basis sets, whereas if there are three functions for each valence atomic orbitals, then we have a triple-zeta-valence (TZV), and so on.

Unfortunately, the initial and successful work developed in the atomic calculations with STO functions was almost stopped. The exponential functions on r^2 (named *Gaussian type orbitals* or GTO) started to become the standard basis sets employed in the calculation of the molecular electronic structure. This regres-

sion in the use of the STOs was due to the complexity in the evaluation of multicenter integrals required in the molecular calculations. Nowadays, some authors have tried to perform molecular calculations with the STO basis sets. McLean and Yoshimine[33] gave the first compilation of these calculations optimizing exponents at fixed bond lengths for several diatomic and linear polyatomic molecules. Cade and coworkers[34-36] accomplished a systematic study of diatomic molecules, and some cations and anions of the considered species. First they computed homonuclear molecules[34], then hydrides [35], and finally the oxides and several heteronuclear molecules[36].

Knowing that STO functions behave properly in the conflictive regions, that is $r \rightarrow 0$ and $r \rightarrow \infty$, a big effort has been invested in the development of the multicenter integrals, thus renewing the interest in STOs, *Exponential type orbitals* or ETOs in general[37]. At present, some programs include these STO basis sets in their calculations, HF and CI[38-40] and DFT [41, 42], solving the multicenter integrals with different levels of accuracy and methods (for example, numerical integration, analytical resolutions, development in accurate GTO expansions...).

This renewed interest has been reflected on the published papers concerning STO basis sets. Some authors have focussed their attention on the initial basis sets of Clementi and Roetti[31] and have improved them taking advantage of the present computational resources and new optimization algorithms. Following this idea, Koga et al.[43, 44] with a reoptimization of the exponents for atoms, cations and anions, have obtained an improvement for the total energy of 0.1 mhartrees and 0.053 hartrees in the He to Kr atoms and Cd atom, respectively. Likewise, Bunge et al.[45, 46] have reoptimized basis sets whose sizes are slightly larger than those of Clementi's basis sets carrying out a partial optimization of the principal quantum numbers n. They have obtained energies with 8 significant figures, comparing with the results from the numerical resolution of the HF's equations, which implies a decrease in the energy error between 21 and 2770 %. The most elaborated optimizations of Clementi's basis sets have been carried out by Koga et al.[43, 47]. They performed a full double optimization of exponents and principal quantum numbers. They found that the optimized n are slightly lower than the n employed by Bunge et al., achieving ten significant figures in the total atomic energy. This optimization of exponents and n quantum numbers have also been carried out in SZ and DZ basis sets by the same authors[48, 49].

Optimization criteria different from the energy minimization have also been proposed and used for obtaining basis sets. For example, the maximum overlap between the functions of a large reference basis set and the functions of the optimized basis set[50], or the minimal geometrical distance between the subspace associated to the reference basis set and the optimized one[51-53] or the minimization of the mono- and bielectronic contributions to the total atomic energy [54, 55]. GTO and STO universal basis sets has been obtained by Trsic et al.[56-58] using the Generator

Coordinate HF method[59]. These types of basis sets have been applied to the calculations of molecular systems.

The first step after the choice of the mathematical function to be employed as basis set is the preparation for its use in the molecular calculations. This requirement implies that the basis set must provide an accurate description of the atoms forming the molecule and the proper estimation of their deformation in the molecular electronic structure. The first condition is satisfied with the optimization of all the exponents with respect to the atomic energy. The second one will be determined by the flexibility of the basis set in the description of the atomic orbital relaxation when the molecule is formed. That is, the proper description of the valence region is influenced by the primitive contraction scheme and by the employed diffuse functions. The basis set optimization demands a high computational cost which increases with the size of the considered basis set. This bottleneck has suggested some authors to design alternative ways for generating basis sets with a limited lower number of variational parameters. Usually, these techniques have been used for GTO basis sets and they mainly could be grouped in *even tempered*[60-62] or *well tempered*[63-67] basis sets, which will be presented in detail in the next section.

Under the idea of decreasing the number of variational parameters to be optimized with respect to the energy, STO basis sets with only one exponent for each subshell (single exponent Slater functions or SESF) have recently been proposed[68, 69]. Their radial part is given by:

$$R_s(r) = e^{-\zeta_{1s} r} \sum_{i=1}^{k_{1s}} a_i^s r^{i-1} + e^{-\zeta_{2s} r} \sum_{j=1}^{k_{2s}} b_j^s r^{j-1} + ...$$

$$R_p(r) = r\, e^{-\zeta_{2p} r} \sum_{i=1}^{k_{2p}} a_i^p r^{i-1} + ... \tag{3.5}$$

that is, every SESF representing an orbital subshell is described by a linear combination of STOs with the same exponent ζ_{1s}, ζ_{2s}, ζ_{2p} ..., and the same l and m quantum numbers, but different principal quantum numbers. The SESF energies are comparable to the LHF ones, either for atoms or for molecules[70], but have the advantage of reducing the number of exponents to be optimized. In Table 3.1 some results of SESF basis sets for the C atom are collected. The SESF size is given by the $\{k_{1s}\, k_{2s}\}\{k_{2p}\}$ numbers which indicate the upper limit in the sum of eq. (3.5) for the $1s$, $2s$ and $2p$ SESFs.

Recently, Koga and coworkers[71-73] have proposed the use of STOs with non-integer quantum numbers, which was firstly suggested by Parr[74, 75]. The non-integer quantum numbers and exponents are considered variational parameters. The non-integer SZ (NSZ) basis sets give energies lower than those with SZ-STO[31]. The most favorable case corresponds to the heavy atoms, for instance, the Lr atom in which the improvement in the energy is 56.8 hartrees. The NSZ basis sets not

Table 3.1: SESF basis set for the C atom (^{3}P) state, optimized with respect to the energy

Basis Set	ζ_{1s}	ζ_{2s}	ζ_{2p}	E(a.u.)	Virial Ratio
{33} {4}	6.122253	1.458034	1.753768	-37.685803358	-2.000000000
{33} {5}	6.115960	1.458734	1.942717	-37.687891980	-2.000000001
{34} {4}	6.779258	1.728607	1.752483	-37.686225152	-2.000000000
{34} {5}	6.771966	1.714643	1.942551	-37.688333474	-2.000000000
{35} {5}	6.819651	1.716263	1.942562	-37.688333874	-2.000000000
{44} {4}	6.715725	1.643143	1.752494	-37.686234034	-2.000000000
{44} {5}	6.717072	1.627785	1.942140	-37.688342713	-2.000000000
{45} {5}	6.635614	1.662539	1.942092	-37.688344748	-2.000000000
{55} {5}	5.925759	1.568993	1.942018	-37.688346241	-2.000000000

only give better total energies, but also produce an improvement of the orbital energies. Therefore, the unphysical positive energies of d and f orbitals given by the conventional SZ disappear. Additionally, NDZ basis sets have been prepared improving the quality of conventional DZ basis sets and obtaining energies closer to the corresponding numerical HF solutions.

Considering the improvement in the total energy produced by the non-integer quantum numbers and the decreasing of variational parameters of SESF basis sets, a mixed strategy has been proposed[76]. Results show that non-integer SESFs (NS-ESF) give energies which are comparable or even better for some atoms, than those obtained with the conventional SZ-STO basis sets with the same number of primitive functions.

The use of *exponential type orbitals* (ETO) other than STOs has been proposed under the assumption that the former are similar to the latter in the quality for describing the electronic density. In fact their multicenter integrals cannot be solved by simpler algorithms. Filter and Steinborn[77] proposed the use, as ETO basis sets, of modified Bessel functions (currently named *B functions* or BTO), which have a very simple Fourier transform. They are defined by:

$$\hat{K}_{n-1/2}\left(\zeta r\right) r^l\, Y_{lm}\left(\theta, \phi\right) \tag{3.6}$$

where

$$\hat{K}_{n-1/2}(\zeta r) = \sqrt{\frac{\pi}{2}}\, e^{-\zeta r} \sum_{i=0}^{n-1} \frac{(n-1+i)!\,(\zeta r)^{n-i-1}}{(n-1-i)!\, i!\, 2^i} \tag{3.7}$$

As eq.(3.7) shows the BTO functions are a linear combination of STO functions, so their good behavior at short and long distances is ensured. The main properties of these BTO functions are related with the computation of the integrals. A two-center distribution can be written as a multiple one-center distribution, whereas a four-center integral can be considered the sum of two-center Coulomb integrals. These two features together with the simple form of their Fourier transform allow to evaluate all the required molecular integrals. Recently, BTO basis sets of SZ and DZ size for atoms from He to Kr have been obtained. At DZ level, BTOs yield almost the same accuracy as STOs, while at SZ level, STO basis sets are significantly better than BTOs. Molecular calculations for small polyatomic molecules at HF level show that the accuracy in energy is similar for both ETO basis sets[78].

3.4 Gaussian Type Orbitals

The Gaussian functions, generally named *Gaussian type orbitals* (GTO) or *Gaussian type functions* (GTF), are the usual alternative functions to the STOs in the molecular calculations. The spherical GTOs were proposed for the first time by Boys[79] and McWeeny[80]. These functions are defined by

$$r^{n-1}e^{-\alpha r^2}\,Y_{lm}\left(\theta,\phi\right) \tag{3.8}$$

α being the GTO exponent, l and m being the azimuthal and magnetic quantum numbers, respectively, whereas n is given by

$$n = l+1, l+3, l+5... \tag{3.9}$$

Generally it is only used $n = l+1$ in this relation.

Cartesian GTOs are used in molecular calculations because the multicenter integrals are easily evaluated due to the Gaussian theorem that allows to express the product of two Gaussian functions centered in two different points of the space as another Gaussian function centered in a third point located on the line that joins the two initial points. These Cartesian GTOs are defined by:

$$x^i y^j z^k e^{-\alpha r^2} \tag{3.10}$$

in which i, j and k are integer positive numbers. The radial part of Cartesian GTOs is related with the lowest r power of the spherical GTO of each symmetry. The above mentioned restriction, $n = l+1$, imposed to spherical GTOs, comes from the posterior use of their exponents in Cartesian GTOs. That is, the GTO basis set optimized at atomic level only has $1s$ functions to represent the s-type atomic orbitals, $2p$ functions for the p-type, $3d$ functions for the d-type... However, Huzinaga[81] has illustrated that it is not necessary to increase n in the GTOs when the $n = l+1$ restriction is satisfied.

The first derivative of a GTO with respect to r, when r tends to zero, is null, in contrast to the non null value achieved by some STO. It is well-known the bad description of the GTO in the $r \to 0$ region, which is a smooth function with no cusp in contrast with what is obtained when STO are used. This behavior is reflected in the non-fulfillment of the nuclear cusp condition. In the other limit region, when $r \to \infty$, they are functions that decay too fast compared to STOs. These features imply that usually GTO with a high unphysical exponent have to be included to improve the behavior at $r \to 0$ and that l quantum numbers higher than the l of the atomic orbitals are required to obtain accurate atomic energies. In spite of this behavior they are widely used in the computational chemistry because they allow a more simple computation of multicenter integrals.

The initial works in the optimization of GTOs are due to Huzinaga[81] who carried out a study of (10s 6p) basis sets for the atoms with Z=1 to 10. Later, van Duijneveldt[82] extended Huzinaga's work increasing the sizes of the basis set analyzed up to (14s 9p). The methodology used by both authors was based on the variational determination of the exponents with respect to the HF atomic energies, which is the same procedure to that used by Clementi[83] for the STO basis sets. The (10s 6p) and (14s 9p) basis imply a space of 16 and 23 dimensions, respectively, in which a local energy minimum is looked for. The difficulties found in the optimization of individual exponents of large basis sets are associated to the smoothness or non existing minima of the energy surface. This effect can be illustrated with van Duijneveldt's basis set[82] with the same size of Faegri's basis set[84], (14s 9p) for the Ne atom, in which the difference between the energies obtained by both authors is 0.08 mhartrees, due to the employment of a more sophisticated optimization by the second author. Along the years a great amount of papers obtaining primitive GTO functions have appeared. Those which are most outstanding are referenced in Table 3.2 (first and second row and first and second transition metal series).

Even though the individual optimizations of the exponents provide the best basis set description they involve a high computational cost. Therefore, some authors have designed alternative ways for obtainig primitive GTOs using a limited number of parameters. The most fruitful proposal defines the GTO exponents with a geometrical series:

$$\zeta_k = \alpha \, \beta^k \qquad k = 1, 2 ... N \tag{3.11}$$

for every l symmetry, N being the number of desired primitives of an l class, and α and β being two optimizable parameters. They were named *even tempered* basis sets and were proposed by Reeves and Harrison[60] and widely employed by Ruedenberg's group[61, 62].

Conventionally the basis sets employed in electronic calculations have a small number of functions in order to restrict the computational requirements. However, for achieving high accuracy the basis set size should be augmented. Since

Table 3.2: Primitive GTO basis sets.

Atoms	Basis set	Authors	Reference
H- He	(10s)	Huzinaga	[81]
Li-Ne	(10s 6p)	Huzinaga	[81]
	(13s 8p), (14s 9p)	van Duijneveldt	[82]
	(11s 7p)	Huzinaga and Sakai	[85]
	(18s 13p)	Partridge	[86]
	(8s 4p)	Whitman and Hornback	[87]
	(7s 3p)	Roos and Siegbahn	[88]
	(6s 3p), (9s 5p)	Huzinaga	[89, 90]
Na-Ar	(12s 9p)	Veillard	[91]
	(10s 6p)	Roos and Siegbahn	[88]
	(9s 5p), (12s 3p)	Huzinaga	[89, 90]
	(24s 15p)	Schmidt and Ruedenberg	[62]
	(20s 15p)	Partridge	[92]
Sc-Zn	(12s 6p 4d)	Ross et al.	[93]
	(13s 7p 5d)	Hyla-Krispin et al.	[94]
	(20s 12p 9d), (20s 12p 10d),	Huzinaga et al.	[3,95-97]
	(26s, 17p, 13d)	Huzinaga et al.	[3,95-97]
	(16s 11p 8d)	Faegri and Speis	[98]
	(20s, 12p, 9d)	Partridge	[86]
K-Zn	(14s 9p 4d)	Wachters	[99]
Y-Cd	(17s 11p 8d)	Huzinaga	[95]
	(14s 8p 7d)	Hyla-Krispin et al.	[94]
	(22s 14p 12d), (22s 14p 13d),	Huzinaga et al.	[3,95-97]
	(26s, 19p, 16d)	Huzinaga et al.	[3,95-97]
	(17s 12p 8d)	Gropen	[100]
	(20s 14p 11d)	Faegri and Biran	[101]
	(24s 16p 10d), (26s 16p 10d)	Partridge and Faegri	[102, 103]

the flexibility of the basis set is increased with the increasing in size, the basis set optimization becomes a less important step in the obtention of larger basis sets. Silver and Nieuwpoort[104] introduced the concept of *universal basis set* as the basis set that might satisfactorily describe several atoms. Following this idea Wilson and coworkers[105-107] suggested the use of even-tempered STO and GTO basis sets large enough to reproduce the atomic energies of the elements belonging to the first row. For example, in their initial work[105] they generated a universal STO basis set, nine s and six p functions, for the B to Ne atoms. The atomic energies obtained thus have an accuracy of 0.5–1.5 mhartrees. The great advantage of these universal basis sets is the transferability of the integrals. As a consequence the use of the same basis set for each atom in the molecular system and the computation and storage of integrals is simplified and reduced. A more sophisticated universal and even-tempered basis sets were proposed by Clementi and Corongiu[108, 109]. They imposed six constraints to the basis set, related with its size, and the number of function for an $l-$type. They are called *geometrical basis sets*. In a generator coordinate version of the HF equations, the obtained equations are numerically integrated. From the discretization points used to perform this integration Trisic et al.and [56-59] defined the exponents of the Gaussian basis sets. Since there is no need for a variational optimization of the basis functions, the same discretization points are valid for different atoms, and from them universal Gaussian basis sets are defined. The accuracy obtained for the atomic energies of Li to Ne atoms varies between 24 and 100 mhartrees, respect to the LHF basis set of Clementi and Roetti[31].

Another series to reduce the number of parameters was introduced by Huzinaga and coworkers[63-67]. In this case, more sophisticated relationships arise and the number of optimizable parameters increase from two to four. These expressions receive the name of *well tempered* and they are an extension of the formula(3.11), by adding two more parameters in the form:

$$\zeta_k = \alpha\,\beta^{k-1}\left[1 + \gamma\left(\frac{k-1}{N}\right)^{\delta}\right] \qquad k = 1, 2...N \qquad (3.12)$$

The *well tempered* series have been used to generate basis sets for atoms from H to Rn[110].

3.5 Basis Set Contraction

Since in the chemical bond formation the inner atomic shells do not experiment great changes, it is generally assumed that their description in the molecular environment will be similar to their description in the atom. This is the main idea behind the contraction of the basis set functions in quantum chemistry. Usually basis set

contraction has been carried out for GTO basis sets. For this reason, in this section we will refer to GTOs, but all the methodology presented here can be extended to any kind of basis set functions.

A contracted GTO basis set is defined by the so-called contracted GTO (CGTO), each of them given by a fixed linear combination of primitive GTOs. The molecular orbitals will be a linear combination of the CGTOs whose coefficients will be obtained in the electronic molecular calculation. However, neither the exponents of the primitive GTOs nor the coefficients involved in the CGTO definition are modified along the calculation.

Usually, once the primitive GTO basis set is chosen, an atomic SCF calculation in this basis set is performed[111]. In this calculation the exponents are optimized and from the final set of atomic orbitals the coefficients that will define the CGTOs are extracted. There are mainly three reasons which justify the use of CGTOs as basis sets:

- the saving in the computational time of each SCF iteration which is achieved when some primitive basis sets are contracted. Also, this saving affects the post-HF calculations in which the time depends on the fifth power of the basis set size,

- the saving in storage and computational time in the molecular integral calculations, reflected in the size of the matrices in which the integrals are kept in the computer memory. m and n being the number of primitive and contracted GTOs, respectively, the contraction process reduces in $(m/n)^4$ the number of integrals to be computed and stored,

- and the complexity in the optimization of a wavefunction that depends on several non-linear parameters[60].

Two main schemes of contraction have been proposed: *the segmented contraction* [112, 113] and *the general contraction* [114]. In the first one every primitive GTO only contributes to one CGTO, which implies that each primitive GTO has only a significant weight in the description of only one atomic orbital. Several authors [113,115-119] expand one STO function with exponent equal to 1.00 as a linear combination of primitive GTOs. The so obtained CGTOs are known as STO-NG basis sets. The exponents and coefficients of the primitive GTO were obtained in the minimization of the integral of the square difference between the radial part of the corresponding STO and the CGTO. A posterior evaluation of the atomic exponents for the scaling of the CGTO is required.

The change of a primitive GTO basis set $\{\chi_i\}_{i=1}^n$ for a CGTO basis set $\{\varphi_i\}_{i=1}^m$

Table 3.3: Segmented scheme of contraction with the (33/3) basis set of Huzinaga[14] for Carbon atom(3P)

.

$$\chi^t_{GTO} = \begin{vmatrix} e^{-153.17\,r^2} \\ e^{-23.073\,r^2} \\ e^{-4.9233\,r^2} \\ e^{-5.7256\,r^2} \\ e^{-.45504\,r^2} \\ e^{-.14707\,r^2} \\ r\,e^{-4.2513\,r^2} \\ r\,e^{-.86327\,r^2} \\ r\,e^{-.20135\,r^2} \end{vmatrix} \qquad U = \begin{vmatrix} .0707 & .0000 & .0000 \\ .3954 & .0000 & .0000 \\ .6633 & .0000 & .0000 \\ .0000 & -.0814 & .0000 \\ .0000 & .5748 & .0000 \\ .0000 & .5024 & .0000 \\ .0000 & .0000 & .1099 \\ .0000 & .0000 & .4627 \\ .0000 & .0000 & .6275 \end{vmatrix}$$

could be written in terms of matrices

$$\varphi_{CGTO} = \chi_{GTO}\,U \qquad\qquad (3.13)$$

φ_{CGTO} being the vector of $1 \times m$ dimension containing the CGTO basis set, χ_{GTO} being the vector of $1 \times n$ dimensions with the primitive GTO basis set and U being the matrix of the basis change with $n \times m$ dimensions. Since each primitive GTO only contributes to one CGTO, the U matrix is blocked in the segmentation contraction scheme. In Table 3.3 the transposed χ_{GTO} vector and U matrix are given for a simple example, i.e.: the (33/3) Huzinaga's basis set[14]. This segmented scheme has been widely used in the computational chemistry calculations. As examples, we could mention the STO-NG (N=3,4,5 or 6) [113,115-119], the 3-21G[120-122], the 4-31G[123-126], the 6-31G[123-126] etc. However, there are remarkable defaults in the segmented contraction scheme. Usually, the CGTO basis set constructed under this scheme gives energies higher than those obtained with the primitive GTO set. This discrepancy in the energy grows as the size of the primitive set does. This occurs in the basis sets for heavy elements. Often, there are primitive functions that contribute significantly to more than one atomic orbital[127], so their assignation to one or another CGTO becomes unclear. This problem could be avoided including the primitive twice in the CGTO which leads to an undesirable increasing of the primitive basis set size.

Each CGTO is defined as an atomic orbital in the general contraction scheme. That is, all the given primitive GTOs in the basis set of a given l symmetry contribute to all the CGTO of this symmetry. As it is shown in the example of Table 3.4 the U matrix is not blocked for an l symmetry. This scheme has several advantages:

- the better description of the atomic orbitals, that is the CGTO basis set gives

Table 3.4: General scheme of contraction in the cc-pVDZ basis set of Dunning et al.[128] for Boron atom(2P).

$$\chi_{GTO}^t = \begin{vmatrix} e^{-4570.\ r^2} \\ e^{-685.9\ r^2} \\ e^{-156.5\ r^2} \\ e^{-44.47\ r^2} \\ e^{-14.48\ r^2} \\ e^{-5.131\ r^2} \\ e^{-1.898\ r^2} \\ e^{-.3329\ r^2} \\ e^{-.1043\ r^2} \\ r\ e^{-6.001\ r^2} \\ r\ e^{-1.241\ r^2} \\ r\ e^{-.3364\ r^2} \\ r\ e^{-.9538\ r^2} \\ r^2\ e^{-.3430\ r^2} \end{vmatrix} \qquad U = \begin{vmatrix} .0007 & -.0001 & .0000 & .0000 \\ .0053 & -.0011 & .0000 & .0000 \\ .0271 & -.0054 & .0000 & .0000 \\ .1014 & -.0219 & .0000 & .0000 \\ .2720 & -.0597 & .0000 & .0000 \\ .4484 & -.1387 & .0000 & .0000 \\ .2901 & -.1314 & .0000 & .0000 \\ .0143 & .5395 & .0000 & .0000 \\ -.0035 & .5808 & .0000 & .0000 \\ .0000 & .0000 & .0355 & .0000 \\ .0000 & .0000 & .1981 & .0000 \\ .0000 & .0000 & .5052 & .0000 \\ .0000 & .0000 & .4795 & .0000 \\ .0000 & .0000 & .0000 & 1.000 \end{vmatrix}$$

the same energy as the uncontracted primitive set,

- the easier analysis of the molecular wavefunction due to the physical meaning of the CGTO

- and the reduction of the basis set superposition error [129], which will be discussed in section 7.

As mentioned above, the CGTOs generated with the general scheme reproduce exactly the atomic orbitals, but the proper molecular description also demands a good description of valence molecular orbitals. Then a modified general scheme is introduced. For each valence atomic orbital, besides one CGTO another uncontracted GTO is added. Preferably these new GTOs will have the same exponent as the more diffuse primitive GTOs included in the initial basis set in order to do not increase the basis set size. An example is presented in Table 3.5.

In spite of the advantages of the general scheme, it is still more generally used the segmented scheme. This is due to the fact that at computational level the segmented scheme is easier to implement than the general scheme. This is strengthened by the widely diffusion of some computational codes in other chemistry fields. Fortunately, in the past years some codes have implemented the general contraction scheme.

Table 3.5: Modified general scheme of contraction with the cc-pVDZ basis set of Dunning et al.[128] for Boron atom(2P).

$$
\chi^t_{GTO} =
\begin{vmatrix}
e^{-4570.\ r^2} \\
e^{-685.9\ r^2} \\
e^{-156.5\ r^2} \\
e^{-44.47\ r^2} \\
e^{-14.48\ r^2} \\
e^{-5.131\ r^2} \\
e^{-1.898\ r^2} \\
e^{-.3329\ r^2} \\
e^{-.1043\ r^2} \\
r\ e^{-6.001\ r^2} \\
r\ e^{-1.241\ r^2} \\
r\ e^{-.3364\ r^2} \\
r\ e^{-.9538\ r^2} \\
r^2\ e^{-.3430\ r^2}
\end{vmatrix}
\qquad
U =
\begin{vmatrix}
.0007 & -.0001 & .0000 & .0000 & .0000 & .0000 \\
.0053 & -.0011 & .0000 & .0000 & .0000 & .0000 \\
.0271 & -.0054 & .0000 & .0000 & .0000 & .0000 \\
.1014 & -.0219 & .0000 & .0000 & .0000 & .0000 \\
.2721 & -.0597 & .0000 & .0000 & .0000 & .0000 \\
.4484 & -.1387 & .0000 & .0000 & .0000 & .0000 \\
.2901 & -.1315 & .0000 & .0000 & .0000 & .0000 \\
.0143 & .5395 & .0000 & .0000 & .0000 & .0000 \\
-.0034 & .5808 & 1.000 & .0000 & .0000 & .0000 \\
.0000 & .0000 & .0000 & .0355 & .0000 & .0000 \\
.0000 & .0000 & .0000 & .1981 & .0000 & .0000 \\
.0000 & .0000 & .0000 & .5052 & .0000 & .0000 \\
.0000 & .0000 & .0000 & .4795 & 1.000 & .0000 \\
.0000 & .0000 & .0000 & .0000 & .0000 & 1.000
\end{vmatrix}
$$

3.6   Correlated Basis Set

Nowadays, the GTO optimizations are directed to the preparation of basis sets flexible enough to describe the electronic correlation effects. It is generally accepted the useful division of the electronic correlation into two contributions to the energy of the system[130]. The so-called nondynamical correlation is due to the existence of some excited configurations with energies close to the HF configuration energy that are significantly coupled with it. This electron correlation is corrected using an MC-SCF (or *multiconfigurational SCF*)[130-133] calculation. For closed shell molecules at distances close to the equilibrium of the bond length, this contribution is not very important. Nonetheless, it increases in the molecular dissociation or in general when some excited states of the same symmetry are almost degenerated with the state under study. The second contribution is the dynamical correlation, which is tied to the culombic cusp, and requires the inclusion of more terms in the development of the wavefunction that will describe better its singularity. The most extended method to include the dynamical correlation is the *Configuration Interaction* or CI. This is a variational method in which the wavefunction is expanded as a linear combination of a reference configuration and all the possible multiexcitations over it. Also, the *Multireference Configuration Interaction* or MRCI[134] method treats this lack in the wavefunction by taking several configurations and performing excitations over them. Therefore, two opposite requirements of the basis set size play an important role in the post-HF calculations: on the one hand, the basis sets should not be very

large due to the matrices size in the calculation, but on the other hand the basis sets should include correlated primitives. For these reasons, the basis sets used in correlation calculations are normally either DZ or SV with polarization.

Thus, the general idea is to optimize the exponents of the primitive GTO set with respect to the atomic correlation energy. These correlated functions have shown to treat correctly the polarized effects in molecular calculations. In 1980 Pople and coworkers[135] considered the possibility of optimizing basis sets with respect to the energy of the atomic ground at second order Moller-Plesset (MP) perturbation level. As a consequence, they published the 6-311G** basis set for the first and second groups of elements, whose exponents and coefficients were optimized. Calculations for some molecules showed that the new basis set did not improve geometrical parameters at second order MP but the bond lengths between heavy atoms were improved at high MP levels. Also they found a better convergence in the perturbative series. Later, Ahlrichs et. al.[136-138] carried out several studies on the influence of the diffuse functions in correlated calculations. They optimized polarization d, f and g-type functions for different correlation methods showing that the results mainly depend on the incompleteness of the basis set. That is, their results have empirically shown that for a given correlation method the ultimate accuracy is determined by the initial set expansion[7].

However, great advances in this scheme of optimization have been achieved in the last years. Nowadays, there are new codes capable to manage the correlation calculations at a reasonable computer time and computer memory requirements. There are two groups of correlated basis sets, the *Atomic Natural Orbitals* (ANO)[139] and the *correlation-consistent* (cc)[128] basis sets. Almlöf and Taylor[139] have suggested the application of the general contraction scheme to the Natural Orbitals (NO). These orbitals are obtained from an atomic SD-CI (simple and Double Configuration Interaction), they are the eigenvectors of the first order density matrix and their eigenvalues are called occupation numbers. The occupation numbers give the relative importance of the corresponding NO in the total wave function. The occupation numbers are used as a threshold for the choice of the NO that will be included in the contraction scheme.

This method for generating ANO basis sets have several advantages: the polarized effects are included, the basis set superposition error is minimized (see next section), it is easy to improve the basis set in a systematic way etc.[7]. However, some modifications have been added to the ANO generation, depending on the atoms. For example, for atoms with only one valence electron the NO are obtained from their homonuclear molecules[139]. Likewise, in atoms where orbitals exhibit a near degeneration (s-p atoms or transitional metals), an averaged density matrix is used, instead of the density matrix for the ground state. It represents an averaged density between the ground state and several excited states which are close in energy[140-142]. This idea of defining an average density matrix has been widely

employed by Widmark et al.[143, 144]. They defined an averaged density with the ground state of the atom, anion and cation and also with the atom perturbed in an external homogeneous electric field. For H to Kr atoms there are ANO basis sets averaged for the most significant electronic states of the atom, the ground state of the anion and cation and even the diatomic molecule at its equilibrium geometry[145].

The main disadvantage of the above mentioned ANO basis set is the large number of primitives required, because the number of computer integrals is proportional to the fourth-power of the number of primitives. For this reason, Dunning suggested smaller basis sets which take into account correlation effects[128]. With a (12s 6p) primitive basis set contracted under the Raffenetti scheme[114] and optimized at HF level, a CI calculation with single and double excitations from all the degenerate HF configurations is performed. Subsequently, this basis set is systematically augmented in order to quantify the contribution of the added polarization function to the correlation energy. For example, adding a first 3d function the correlation energy achieved for that orbital reaches 78 %; when a second 3d function is added we obtain up to 96 % of that energy, and when a third 3d function is added calculations show up to 99 % of that energy. Similarly, adding two 4f, that is, (2d 1f) and (3d 2f), and one 5g (3d 2f 1g), their contributions are studied. Results showed that the first polarized function of a given l symmetry reduces significantly the correlation energy and that the subsequent functions of the same symmetry give a lower convergence to the limit of correlation energy for a given symmetry. Once the contributions of the different functions have been evaluated, the construction of a *correlated consistent* basis set is made with all the correlation functions which decrease the correlation energy by similar amounts, as well as with all the correlating functions which reduce the energy in larger amounts. For atoms belonging to the second row of elements Dunning found a basis set with DZV polarized size or polarized valence double zeta (cc-pVDZ).

These basis sets show an excellent behavior in the calculations of atoms and molecules. In fact, cc-pVDZ achieves 99 % of the correlation energy obtained with ANO basis sets, although the latter have almost double number of primitives. They do not give a good description of the anions, but this default has been solved by enlarging the basis set with more additional diffuse functions[146]. These basis sets have also been published for atoms in the third row[147]. They have been included in the Gaussian 94 package[148] that is synonymous of becoming a widely used basis set for molecular correlated calculations.

3.7   Basis set superposition error

The non-completeness of the basis sets employed in molecular calculations has a significative importance in the computation of molecular dissociation energies. This non-completeness of the basis set arises from the truncation of the basis set functions introduced by the LCAO approximation. When two atoms, A and B, approach to form the AB molecule, the total energy decreases by two factors; the stabilization of the system due to the bond formation and the improvement in the individual atomic description. This second effect is an artifact, because the basis set centered in A will be influenced by the basis set centered in B and viceversa, causing an improvement in the description of both atoms. This effect is called *basis set superposition error* or BSSE, which was introduced in the preceding sections. For more details see two reviews[149, 150].

The interaction energy between A and B at a given R distance is given by:

$$\Delta E = E_{AB} - (E_A + E_B)$$

E_{AB}, E_A and E_B being the energies of AB molecule, A atom and B atom, respectively. If the basis set in the atomic calculation of A or B is large enough, the atomic energies will be near HF and the improvement of the basis set with another functions will not modify the total energy. In this case the basis set is named *saturated basis set*. Therefore, the use of saturated basis set prevents the BSSE. However, the increase of basis set size causes the increase of the computational cost. Moreover, the saturation of the atomic basis set does not imply the saturation of the molecular basis set, because polarization functions are usually required for the proper bond description.

The basis sets commonly employed in molecular calculations are not saturated for atomic systems, because of the computational and storage requirements of the bielectronic integrals. Moreover, the description of A and B improves because basis sets centered in A and B are not saturated when the two atoms approach. If polarization functions have also been added that effect grows. Thus, $\Delta E(R)$ for a non saturated basis set is overestimated due to the better description of A in the molecular system AB at finite distance R.

Boys and Bernardi[151] proposed a method to evaluate the BSSE. First, the $\Delta E(R)$ is computed with the energies of A, B and AB obtained with two saturated basis sets for A and B, respectively. Then, the energy of A is computed with the two basis sets used in the calculation of AB system, including polarization functions and with all the functions centered in the same points as they were centered in AB at a given R. The energy of B is computed in a similar way. In both computations of atomic energies the molecular integrals of the AB system appear. $\Delta'E(R)$ being the difference between E_{AB} while the two new atomic energies are defined by those

authors as:

$$\epsilon(R) = \Delta'E(R) - \Delta E(R)$$

which in fact, is a positive amount, since $|\Delta'E(R)| \leq |\Delta E(R)|$. This $\epsilon(R)$ allows to quantify the completeness or saturation of the chosen basis set. Boys and Bernardi's method is a useful evaluation and correction of the BSSE; so $\epsilon(R)$ is taken as the BSSE mathematical definition.

For small molecules the above methodology corrects the energy giving results which are very similar to those obtained with larger basis sets. However, following this procedure this estimation of the BSSE is overestimated [152] because in the atomic basis sets there are functions which are shared between several atoms in the molecular environment while they describe only one in the atom. Additionally, this method is not easily generalizable for molecular systems.

3.8   Concluding Remarks

The theoretical treatment of the electronic structure of atoms and molecules requires the approximate solution of the corresponding electronic Schrödinger equation. The choice of the basis sets is one of the most important factors in *ab initio* quantum chemical calculations, since it ultimately determines the accuracy of the calculation. A wrong choice can render meaningless results, even for large and time-consuming calculations. With the use of parallel architectures in modern computers and the growing awareness that basis set truncation errors are frequently the dominant source of error in current studies of electronic structure, significant improvements in the accuracy of such studies can be expected in the forthcoming years. Our intention has been to provide a detailed background and update of this topic so that the interested reader can explore it in more depth in the literature. In this concluding section, we would like to remark some of future perspectives related with the topic of basis sets.

Numerical and seminumerical integration techniques have been employed in studies of diatomic molecules for which one angular variable may be separated and treated analytically[20]. Such calculations have provided very useful benchmark results within the Hartree-Fock approximations which can be employed in assessing the accuracy of calculations performed by basis sets expansion methods. The numerical solutions of the molecular Schrödinger equation for arbitrary polyatomic molecules would involve a three-dimensional numerical integration and would be very demanding computationally[153]. Although research continues along these lines, no general numerical method for polyatomic molecules has appeared yet. If numerical methods increase, an abandonment of the basis set approach to building wavefunctions could arrive.

The use of two-electron basis functions in addition to traditional one-electron basis sets in the calculation of correlation energies have also been considered[154]. It is well known that the two-electron correlation cusp is treated poorly by expansions in products of one-electron functions. The limits of what can be accomplished using one-electron basis functions might already be approaching, given the slow convergence of the correlation energy with angular quantum numbers. The alternative to traditional one-electron basis sets is to consider, explicitly in the wavefunction, two-electron basis functions. The two-electron functions have a Gaussian dependence on the interelectronic distance (Gaussian-type geminals). Persson and Taylor[155, 156] have shown that only a few of such Gaussian correlation factors are necessary to dramatically reduce the error in the correlation energy.

Unfortunately, there are not efficient algorithms to solve all the multicenter integrals with STO functions. However, some solutions are possible employing part of STO multicenter integrals with Gaussian expansions of STO[157] and they are producing excellent results[158, 159]. With the use of STO we are including in our calculation the dynamic correlation, which is bonded to the culombic cusp condition. As a consequence, the number of functions used is lower than the corresponding number of Gaussian functions, in which it is necessary to include correlated orbitals with superior numbers of radial and angular nodes[39].

3.9   Acknowledgments

Bibliography

[1] F.R. Burden and R.M. Wilson, Adv. Phys. **21**, 825 (1972).

[2] R. Ahlrichs and P.R. Taylor, J. Chim. Phys. **78**, 315 (1981).

[3] S. Huzinaga, Comput. Phys. Rep. **2**, 279 (1985).

[4] E.R. Davidson and D. Feller, Chem. Rev. **86**, 681 (1986).

[5] S. Wilson, Adv. Chem. Phys. **67**, 439 (1987).

[6] D. Feller and E.R. Davidson, in *Reviews in Computational Chemistry*, edited by K.B. Lipkowitz and D.B. Boyd (VHC Publishers, New York, 1990).

[7] J. Almlöf and P. R. Taylor, Adv. Quantum Chem. **22**, 301 (1991).

[8] T.L. Sordo and J.A. Sordo, in *Computational Chemistry: Structure, Interactions and Reactivity*, Vol. 77 of *Studies in Physical and Theoretical Chemistry*, edited by S. Fraga (Elsevier Science Publishers, Amsterdam, 1992), p. 223.

[9] T. Helgaker and P.R. Taylor, in *Modern Electronic Structure Theory: Part II*, edited by D.R. Yarkony (World Scientific, Singapore, 1995), p. 727.

[10] S. Wilson, in *Methods in Computational Molecular Physics*, edited by G.H. Diercksen and S. Wilson (Dordrecht, Reidel, 1983).

[11] D. Feller, J. Comput. Chem. **17**, 1571 (1992).

[12] P. Čársky and M. Urban, *Ab Initio Calculations* (Springer-Verlag, Heidelberg, 1980).

[13] W.J. Hehre, L. Radom, P. v.R.Schleyer, and J.A. Pople, *Ab Initio Molecular Orbital Theory* (Wiley-Interscience, New York, 1986).

[14] S. Huzinaga, *Gaussian Basis Sets for Molecular Calculations* (Elsevier, Amsterdam, 1984).

[15] R.A. Poirier, R.E. Kari, and I.G. Csizmadia, *Handbook of Gaussian Basis Sets* (Elsevier, Amsterdam, 1985).

[16] D.R. Hartree, Proc. Cambridge Phil. Soc. **24**, 89 (1928).

[17] V. Fock, Z. Phys. **61**, 126 (1930).

[18] J.C. Slater, Phys. Rev. **35**, 210 (1930).

[19] C. Froese Fischer, *The Hartree-Fock Method for Atoms. A Numerical Approach* (Willey-Interscience, New York, 1977).

[20] J. Kobus and L. Laaksonen ans D. Sundholm, Comp. Phys. Comun. **98**, 346 (1996).

[21] C.C.J. Roothaan, Rev. Mod. Phys. **23**, 69 (1951).

[22] G.G. Hall, Proc. Roy. Soc. London Sr. A **205**, 541 (1951).

[23] G.G. Hall, Proc. Roy. Soc. London Sr. A **208**, 328 (1951).

[24] P.O. Löwdin, Adv. Chem. Phys. **2**, 207 (1959).

[25] S. Wilson, *Electron Correlation in Molecules* (Clarendon Press, Oxford, 1984).

[26] C. Zener, Phys. Rev. **36**, 51 (1930).

[27] J.C. Slater, Phys. Rev. **36**, 57 (1930).

[28] C.C.J. Roothaan and P.S. Bagus, *Methods in Computational Physics* (Academic Press, New York, 1963), Vol. II.

[29] T. Kato, Comm. Pure Appl. Math. **10**, 151 (1957).

[30] E. Clementi, IBM J. Res. Develop. Suppl. (1965).

[31] E. Clementi and C. Roetti, At. Data Nucl. Data Tables **14**, 177 (1974).

[32] A.D. McLean and R.S. Mclean, At. Data Nucl. Data Tables **26**, 197 (1981).

[33] A.D. McLean and M. Yoshimine, *Tables of Linear Molecule Function* (IBM J. Res. Dev., San José, California, 1968), Vol. 12, p. 206.

[34] P.E. Cade and W. Huo, At. Data Nucl. Data Tables **12**, 415 (1973).

[35] P.E. Cade and A.C. Wahl, At. Data Nucl. Data Tables **13**, 339 (1974).

[36] P.E. Cade and W. Huo, At. Data Nucl. Data Tables **15**, 1 (1975).

[37] C.A. Weatherford and H.W. Jones, *ETO Multicenter Molecular Integrals* (D. Reidel Publ. Co., Dordrecht, 1982).

[38] A.D. McLean, M. Yoshimine, B.H. Lengsfield, P.S. Bagus, and B. Liu, in *Modern Techniques in Computational Chemistry, MOTECC*, edited by E. Clementi (Escom, Leiden, 1989), p. 593.

[39] J. Fernández Rico, R. López, A. Aguado, I. Ema, and G. Ramírez, J. Comput. Chem. in press (1998).

[40] A. Bouferguene, M. Fares, and P.E. Hogan, Int. J. Quantum Chem. **56**, 801 (1996).

[41] D. Ellis and G.S. Painter, Phys. Rev. B **2**, 2887 (1970).

[42] E.J. Baerends, D. Ellis, and P. Ros, Chem. Phys **2**, 41 (1973).

[43] T. Koga, H. Tatewaki, and A.J. Thakkar, Phys. Rev. A **47**, 4510 (1993).

[44] T. Koga, Y. Seki, A.J. Thakkar, and H. Tatewaki, J. Phys. B: At. Mol. Opt. Phys. **26**, 2529 (1993).

[45] C.F. Bunge, J.A. Barrientos, A.V. Bunge, and J.A. Cogordan, Phys. Rev. A **46**, 3691 (1992).

[46] C.F. Bunge, J.A. Barrientos, and A.V. Bunge, At. Data Nucl. Data Tables **53**, 113 (1993).

[47] T. Koga, S. Watanabe, K. Kanayama, R. Yasuda, and A.J. Thakkar, J. Chem. Phys. **103**, 3000 (1995).

[48] T. Koga and A.J. Thakkar, Theor. Chim. Acta **85**, 363 (1993).

[49] T. Koga, Y. Seki, and A.J. Thakkar, Bull. Chem. Soc. Jpn. **66**, 3135 (1993).

[50] E. Francisco, L. Seijo, and L. Pueyo, J. Solid State Chem. **63**, 393 (1986).

[51] J.M. García de la Vega, J. Fernández Rico, and J.I. Fernández-Alonso, J. Mol. Struct (Theochem) **184**, 1 (1989).

[52] J.M. García de la Vega and B. Miguel, J. Comput. Chem. **12**, 1172 (1991).

[53] J.M. García de la Vega and B. Miguel, J. Solid State Chem. **116**, 275 (1995).

[54] J.A. Sordo and L. Pueyo, Int. J. Quantum Chem. **28**, 687 (1985).

[55] J.A. Sordo and L. Pueyo, J. Mol. Struct. (Theochem) **120**, 9 (1985).

[56] J.R. Mohallem and M. Trsic, J. Chem. Phys. **86**, 5043 (1987).

[57] H.F.M. da Costa, M. Trsic, and J.R. Mohallem, Mol. Phys. **62**, 91 (1987).

[58] A.B.F. da Silva, H.F.M. da Costa, and M. Trsic, Mol. Phys. **68**, 433 (1989).

[59] J.R. Mohallem, R.M. Dreizler, and M. Trsic, Int. J. Quantum Chem. Sym. **20**, 45 (1986).

[60] C.M. Reeves and M.C. Harrison, J. Chem. Phys. **39**, 11 (1963).

[61] in *Energy, Structure and Reactivity. Proceedings of the 1972 Boulder Conference on Theoretical Chemistry*, edited by K. Ruedenberg, R.C. Raffenetti, and R.D. Bardo (Wiley, New York, 1973).

[62] M.W. Schmidt and K. Ruedenberg, J. Chem. Phys. **71**, 3951 (1979).

[63] S. Huzinaga, M. Klobukowski, and H. Tatewaki, Can. J. Chem. **63**, 1812 (1985).

[64] S. Huzinaga, M. Klobukowski, and H. Tatewaki, Chem. Phys. Lett. **120**, 509 (1985).

[65] S. Huzinaga and B. Miguel, Chem. Phys. Lett. **175**, 289 (1990).

[66] S. Huzinaga and M. Klobukowski, Chem. Phys. Lett. **212**, 260 (1993).

[67] M. Klobukowski, Can. J. Chem. **72**, 1741 (1994).

[68] J.M. García de la Vega and B. Miguel, Chem. Phys. Lett. **207**, 270 (1993).

[69] J.M. García de la Vega and B. Miguel, Int. J. Quantum Chem. **51**, 397 (1994).

[70] J.M. García de la Vega, B. Miguel, and G. Ramírez, J. Phys. B **29**, 5273 (1996).

[71] T. Koga, K. Kanayana, and A.J. Thakkar, Int. J. Quantum Chem. **62**, 1 (1997).

[72] T. Koga, Y. Seki, A.J. Thakkar, and H. Tatewaki, J. Phys. B: At. Mol. Opt. Phys. **30**, 1623 (1997).

[73] T. Koga and K. Kanayana, Chem. Phys. Lett. **266**, 123 (1997).

[74] R.G. Parr and H.W. Joy, J. Chem. Phys. **26**, 424 (1957).

[75] A.F. Saturno and R.G. Parr, J. Chem. Phys. **29**, 490 (1958).

[76] T. Koga, J.M. García de la Vega, and B. Miguel, Chem. Phys. Lett. **283**, 97 (1998).

[77] E. Filter and E.O. Steinborn, Phys. Rev. A **18**, 1 (1978).

[78] J. Dotterweich, I. Ema, J.M. García de la Vega, B. Miguel, and E.O. Steinborn, (in preparation).

[79] S.F. Boys, Proc. Roy. Soc. (London) A **200**, 542 (1950).

[80] R. McWeeny, Nature **166**, 21 (1950).

[81] S. Huzinaga, J. Chem. Phys. **42**, 1293 (1965).

[82] F.B. van Duijneveldt, IBM Technical Research Report, RS 945 (1971).

[83] E. Clementi, C.C.J. Roothaan, and M. Yoshimine, Phys. Rev. **127**, 1618 (1962).

[84] K. Faegri Jr. and J. Almlöf, J. Comput. Chem. **7**, 396 (1986).

[85] S. Huzinaga and Y. Sakai, J. Chem. Phys. **50**, 1371 (1969).

[86] H. Partridge, J. Chem. Phys. (1989).

[87] D.R. Whitman and C.J. Hornback, J. Chem. Phys. **51**, 398 (1969).

[88] B. Roos and P. Siegbahn, Theor. Chim. Acta **17**, 209 (1970).

[89] S. Huzinaga, Approximate Atomic Wave Functions I, University of Alberta (1971).

[90] S. Huzinaga, Approximate Atomic Wave Functions II, University of Alberta (1971).

[91] A. Veillard, Theor. Chim. Acta **12**, 405 (1968).

[92] H. Partridge, J. Chem. Phys. (1987).

[93] B. Roos, A. Veillard, and G. Vinot, Theor. Chim. Acta **20**, 1 (1971).

[94] I. Hyla-Krispin, J. Demuynck, A. Strich, and M. Bénard, J. Chem. Phys. **75**, 3954 (1981).

[95] S. Huzinaga, J. Chem. Phys. **66**, 4245 (1977).

[96] M. Klobukowski and S. Huzinaga, J. Chem. Phys. **84**, 323 (1986).

[97] S. Huzinaga, M. Klobukowski, Z. Barandiaran, and L. Seijo, J. Chem. Phys. **84**, 6315 (1986).

[98] K. Faegri Jr. and H.J. Speis, J. Chem. Phys. **86**, 7035 (1987).

[99] A.J.H. Wachters, J. Chem. Phys. **52**, 1033 (1970).

[100] O. Gropen, J. Comput. Chem. **8**, 982 (1987).

[101] K. Faegri Jr. and G. Biran, J. Comput. Chem. **10**, 495 (1989).

[102] H. Partridge and K. Faegri Jr., Theor. Chim. Acta **82**, 207 (1992).

[103] H. Partridge, NASA Technical Report (1992).

[104] D.M. Silver and W.C. Nieuwpoort, Chem. Phys. Lett. **57**, 421 (1978).

[105] D.M. Silver, S. Wilson, and W.C. Nieuwpoort, Int. J. Quantum Chem. **14**, 635 (1978).

[106] S. Wilson, Theor. Chim. Acta **57**, 53 (1980).

68

[107] S. Wilson, Theor. Chim. Acta **58**, 31 (1980).

[108] E. Clementi and G. Corongiu, Chem. Phys. Lett. **90**, 359 (1982).

[109] E. Clementi and G. Corongiu, IBM Tech. Rep. POK-11 (1982).

[110] S. Huzinaga, B. Miguel, and M. Klobukowski, Well-Tempered Gaussian Basis Sets, Department of Chemistry, Univeristy of Alberta, Canada (1993).

[111] T.H. Dunning Jr. and P.J. Hay, in *Modern Theoretical Chemistry*, edited by H.F. Schaefer (Plenum Press, New York, 1977), Vol. III.

[112] T.H. Dunning Jr., J. Chem. Phys. **53**, 2823 (1970).

[113] W.J. Hehre, R.F. Stewart, and J.A. Pople, J. Chem. Phys. **51**, 2657 (1969).

[114] R.C. Raffenetti, J. Chem. Phys. **58**, 4452 (1973).

[115] R.F. Stewart, J. Chem. Phys. **50**, 2485 (1969).

[116] R.F. Stewart, J. Chem. Phys. **52**, 431 (1970).

[117] J.A. Pople and W.J. Hehre, J. Comput. Phys. **27**, 161 (1978).

[118] W.J. Pietro, B.A. Levi, W.J. Hehre, and R.F. Stewart, Inorg. Chem. **19**, 2225 (1980).

[119] W.J. Pietro and W.J. Hehre, J. Comput. Chem. **4**, 241 (1983).

[120] J.S. Binkley, J.A. Pople, and W.J. Hehre, J. Am. Chem. Soc. **102**, 939 (1980).

[121] M.S. Gordon, J.S. Binkley, J.A. Pople, W.J. Pietro, and W.J. Hehre, J. Am. Chem. Soc. **104**, 2797 (1982).

[122] W.J. Pietro, M.M. Francl, W.J. Hehre, D.J. Defrees, J.A. Pople, and J.S. Binkley, J. Am. Chem. Soc. **104**, 5039 (1982).

[123] R. Ditchfield, W.J. Hehre, and J.A. Pople, J. Chem. Phys. **54**, 724 (1971).

[124] W.J. Hehre, R. Ditchfield, and J.A. Pople, J. Chem. Phys. **56**, 2257 (1972).

[125] P.C. Hariharan and J.A. Pople, Theor. Chim. Acta **28**, 213 (1973).

[126] M.S. Gordon, Chem. Phys. Lett. **76**, 163 (1980).

[127] T.H. Dunning Jr., J. Chem. Phys. **66**, 1382 (1977).

[128] T.H. Dunning Jr., J. Chem. Phys. **90**, 1007 (1989).

[129] R.D. Bardo and K. Ruedenberg, J. Chem. Phys. **59**, 5956 (1973).

[130] B.O. Roos, in *Lectures in Quantum Chemistry*, edited by B.O. Roos (Springer-Verlag, Berlin, 1992), p. 177.

[131] H.J. Werner, Adv. Chem. Phys. **69**, 1 (1987).

[132] R. Shepard, Adv. Chem. Phys. **69**, 63 (1987).

[133] B.O. Roos, Adv. Chem. Phys. **69**, 399 (1987).

[134] R.J. Buenker, in *Current aspects of Quantum Chemistry*, edited by R. Carbó (Elsevier, Amsterdam, 1982), Vol. 21.

[135] R. Krishnan, J.S. Binkley, R. Seeger, and J.A. Pople, J. Chem. Phys **72**, 650 (1980).

[136] R. Ahlrichs, P. Scharf, and K. Jankowski, Chem. Phys. Lett. **98**, 381 (1985).

[137] K. Jankowski, R. Becherer, P. Scharf, H. Schiffer, and R. Ahlrichs, J. Chem. Phys. **82**, 1413 (1985).

[138] R. Becherer and R. Ahlrichs, Chem. Phys. Lett. **99**, 389 (1985).

[139] J. Almlöf and P.R. Taylor, J. Chem. Phys. **86**, 4070 (1987).

[140] C.W. Bauschlicher, Chem. Phys. Lett. **142**, 71 (1987).

[141] T.J. Lee, A.P. Rendell, and P.R. Taylor, J. Chem. Phys. **92**, 489 (1990).

[142] H. Partridge, C.W. Bauschlicher, L.G.M. Pettersson, A.D. McLean, B. Liu, M. Yoshimine, and A. Komornicki, J. Chem. Phys. **92**, 5377 (1990).

[143] P.O. Widmark, P.A. Malmqvist, and B.O. Ross, Theor. Chim. Acta **77**, 291 (1990).

[144] P.O. Widmark, B.J. Persson, and B.O. Ross, Theor. Chim. Acta **79**, 419 (1991).

[145] K. Pierloot, B. Dumez, P.O. Widmark, and B.O. Roos, Theor. Chim. Acta **90**, 87 (1995).

[146] R.A. Kendall, T.H. Dunning Jr., and R.J. Harrison, J. Chem. Phys. **96**, 6796 (1992).

[147] R.A. Woon and T.H. Dunning Jr., J. Chem. Phys. **98**, 1358 (1993).

[148] M. J. Frisch *et al.*, *Gaussian 94, Revision C.2* (Gaussian, Inc., Pittsburgh PA, 1995).

[149] G. Chalasinski and M.M. Szczesniak, Chem. Rev. **94**, 1723 (1994).

[150] F.B. van Duijneveldt, J.G.C.M. van Duijneveldt-van de Rijdt, and J.H. van Lenthe, Chem. Rev. **94**, 1873 (1994).

[151] S.F. Boys and F. Bernardi, Mol. Phys. **19**, 553 (1970).

[152] K. Morokuma and K. Kitaura, in *Chemical Applications of Atomic and Molecular Electrostatic Potentials*, edited by P. Politzer (Plenum Press, New York, 1981).

[153] R.M. Dickson and A.D. Becke, J. Chem. Phys. **99**, 3898 (1993).

[154] W. Kultzelnigg and W. Kopper, J. Chem. Phys **94**, 1985 (1991).

[155] B.J. Persson and P.R. Taylor, J. Chem. Phys. **105**, 5915 (1996).

[156] B.J. Persson and P.R. Taylor, Theoret. Chem. Acc. **97**, 240 (1997).

[157] J. Fernández Rico, R. López, I. Ema, and G. Ramírez, Comput. Phys. Commun. **105**, 216 (1997).

[158] A. Aguado, J.M. García de la Vega, and B. Miguel, J. Chem. Soc. Faraday Trans. **93**, 29 (1997).

[159] J.M. García de la Vega and B. Miguel, unpublished results .