

CHAPTER 12

GAUSSIAN BASIS SETS AND MOLECULAR INTEGRALS

Trygve Helgaker

*Department of Chemistry, University of Oslo
P.O. Box 1033, Blindern, N-0315 Oslo, Norway*

Peter R. Taylor

*San Diego Supercomputer Center
P.O. Box 85608, San Diego, California 92186-9784, USA*

Contents

1. Introduction	727
2. Qualitative Considerations	728
2.1. Quantum Chemistry of Some Model Systems	728
2.2. Molecular Orbitals and Electron Correlation	731
2.3. Slater Orbitals and Gaussian Orbitals	734
2.4. Correlating Orbitals	737
3. Gaussian Basis Sets	741
3.1. Basis Set Convergence	741
3.2. Contracted Gaussians from Fits to STO's	743
3.3. Contracted Gaussians from Atomic Calculations	744
3.4. Segmented Versus General Contractions	748
3.5. Even-Tempered Sequences	750
3.6. Polarization Functions	752
3.7. Correlated Calculations	753
4. Basis Sets for SCF Calculations	753
4.1. Small Basis Sets	755
4.2. Larger Basis Sets	757
4.3. Extended Basis Sets	759
5. Basis Sets for Correlated Calculations	760
5.1. Nondynamical Correlation	760

5.2. Small Basis Sets	762
5.3. Atomic Natural Orbitals	764
5.4. Correlation-Consistent Basis Sets	768
5.5. Basis Set Recommendations	770
6. Basis Sets and Molecular Properties	771
6.1. Basis Functions and Energy Derivatives	771
6.2. Spectroscopic Constants	774
6.3. Electric Properties	775
6.4. Magnetic Properties	779
6.5. Basis Set Recommendations	782
7. Basis Set Superposition Error	783
7.1. Counterpoise Correction	784
7.2. Basis Set Improvements and BSSE	785
8. Basis Sets: Miscellaneous Topics	787
8.1. Excited and Ionized States	787
8.2. Core Correlation	787
8.3. Bond Functions	789
9. Molecular Integrals — Introduction	790
10. Charge Distributions	792
10.1. Properties of GTO's	792
10.2. Hermite Gaussian Functions	793
10.3. Overlap Distributions	796
10.4. Overlap Distributions by Recursion	799
11. Simple One-Electron Integrals	800
11.1. Multipole Moments	801
11.2. Momentum and Kinetic Energy Integrals	804
12. Coulomb Integrals	805
12.1. Electrostatics for Gaussian Charge Distributions	806
12.2. The Incomplete Gamma Function	809
12.3. Hermite Integrals	815
12.4. Cartesian Integrals	819
12.5. The Obara-Saika Scheme	821
12.6. Integral Prescreening	828
13. Integrals over CGTO's	829
13.1. Cartesian Gaussian Integrals	829
13.2. Spherical Harmonic Gaussian Integrals	832
14. Property Integrals	833
14.1. Geometrical Derivative Integrals	833
14.2. Derivatives of London Orbitals	837
14.3. Spin-Orbit Integrals	839
15. Rys Quadrature	840
16. Molecular Symmetry and Integrals	842
16.1. Symmetry-Distinct Integrals over CGTO's	843
16.2. Symmetry-Adapted Basis Functions	845

16.3. Computational Implementations	847
16.4. Differentiated Integrals and Symmetry	849
17. Conclusions	849
Acknowledgments	850
References	850

1. Introduction

In this review, we shall be concerned with the use of one-electron basis sets in quantum chemistry, and with the evaluation of integrals over these basis sets. Our intentions are pedagogical — we hope to provide the reader with the background material needed to choose suitable basis sets for quantum chemical calculations and to understand how the integrals are evaluated. We also hope to provide enough fundamental details for the interested reader to pursue these topics in the current literature. We have not tried to review the fields of Gaussian basis sets and integral evaluation comprehensively, since this would require an entire volume (or more) to itself.

The choice of the basis set is one of the most important factors in designing reliable quantum chemical calculations, and deserves considerable attention. We shall proceed by reviewing some qualitative analytical features of wave functions, and how the analytical behavior is or is not reflected by different types of basis function. A certain amount of historical material about basis set development will be presented, but, as we have said, we make no attempt at completeness here, since these matters have been reviewed many times previously (see, for example, Refs. 1–3). We shall concentrate on the design of accurate calculations in our discussion, although recommendations for basis sets to be used at various levels of accuracy are included. The basis set requirements for calculating molecular properties are also discussed.

In the second part of this chapter we consider various methods for calculating integrals over the Gaussian basis functions that are invariably the choice for modern quantum chemical calculations. Again, historical material is presented without any claim to completeness — much of the earlier literature in this subject is referenced in other reviews, and those of Saunders^{4,5} are particularly recommended. We shall concentrate on devel-

opments of the last decade or so, and also on issues related to calculating derivatives of the energy as well as the energy itself. We shall also touch briefly on molecular property integrals and fine-structure integrals.

As a final introductory note, it is desirable to clarify some terminology. Throughout this work, we employ the term “*ab initio*” to refer only to nonempirical calculations based on wave functions. That is, our usage excludes nonempirical density-functional-based methods that would have a legitimate claim to be included under this rubric. This is in no sense a slight: the latter methods are beyond the scope of this chapter and are nowhere discussed, so it makes no sense to use a more restrictive construction than *ab initio* when no ambiguity or confusion can occur.

2. Qualitative Considerations

The goal of *ab initio* quantum chemical calculations is to predict the properties of atoms and molecules using only the principles of quantum mechanics. Some insight into the problems encountered in trying to achieve this goal can be obtained by considering several simple systems as illustrative examples.

2.1. Quantum Chemistry of Some Model Systems

Perhaps the simplest possible system is the hydrogen atom, for which the Hamiltonian in atomic units takes the form

$$H = -\frac{1}{2}\nabla^2 - \frac{1}{r}. \quad (1)$$

Here r is the electron–nucleus distance. The wave functions — solutions to the wave equation

$$H\Psi = E\Psi, \quad (2)$$

for different energy levels E of the hydrogen atom — are products of associated Laguerre functions⁶ in the radial coordinate and spherical harmonic angular functions. Hence they involve an exponential radial dependence $\exp(-r)$. A simpler set of functions displaying exponential radial dependence is given by

$$\mathcal{N}r^{n-1} \exp(-\zeta r), \quad (3)$$

where \mathcal{N} is a normalization constant, n the “principal quantum number,” and ζ is the orbital exponent. The radial function Eq. (3) is commonly

referred to as a *Slater-type orbital* (STO) or exponential-type orbital. The ground state of the hydrogen atom is an STO with $n = 1$ and $\zeta = 1$. The exponential form for the radial wave function is not surprising — the Hamiltonian of Eq. (1) displays singular behavior as $r \rightarrow 0$, while the exponential function shows a cusp behavior in this limit which cancels the singular term. We note that the hydrogenic orbitals (solutions to Eq. (1)) form a complete set only if continuum solutions are admitted, as well as bound-state solutions.⁷ STO's, on the other hand, form a complete set when only bound-state solutions are admitted.

Let us consider now a more complicated situation, the helium atom, in which there are two electrons. The Hamiltonian becomes

$$H = -\frac{1}{2}\nabla_1^2 - \frac{1}{2}\nabla_2^2 - \frac{2}{r_1} - \frac{2}{r_2} + \frac{1}{r_{12}}, \quad (4)$$

where subscripts 1 and 2 have been used to label the electrons. The wave equation involving this Hamiltonian can be regarded as intractable with respect to analytical solution, and thus an alternative approach is required. One strategy is to use a variational method, in which a trial wave function $\Psi(\alpha)$ involving adjustable parameters α is constructed, and the energy functional

$$E(\alpha) = \frac{\int \Psi^*(\alpha) H \Psi(\alpha) d\tau}{\int \Psi^*(\alpha) \Psi(\alpha) d\tau} \quad (5)$$

is made stationary with respect to variations of the parameters α . Assuming that the two electrons are spin-paired, we need to choose a trial wave function that is symmetric in the coordinates of electrons 1 and 2. Clearly, the hydrogen wave function suggests the desirability of exponential radial behavior. However, there is now also a singularity in the Hamiltonian as $r_{12} \rightarrow 0$, which would imply cusp behavior in the wave function in this limit. Hylleraas⁸ suggested a trial function of the form

$$\exp(-[r_1 + r_2]/2) \sum_{nlm} c_{nlm} (r_1 + r_2)^n (r_1 - r_2)^{2l} r_{12}^m, \quad (6)$$

where the coefficients c are to be determined variationally. The results from optimization of trial functions of this type are excellent. Formal analysis of the cusp as $r_{12} \rightarrow 0$ shows that the wave function behaves as $1 + \frac{1}{2}r_{12}$, supporting the use of factors in r_{12} like those in Eq. (6).

Somewhat more complication ensues if we introduce two nuclei into our two-electron problem. The Hamiltonian for the hydrogen molecule at a fixed internuclear separation is

$$H = -\frac{1}{2}\nabla_1^2 - \frac{1}{2}\nabla_2^2 - \frac{1}{r_{A1}} - \frac{1}{r_{A2}} - \frac{1}{r_{B1}} - \frac{1}{r_{B2}} + \frac{1}{r_{12}} + \frac{1}{R_{AB}}. \quad (7)$$

Here we have used R_{AB} to denote the fixed internuclear distance which appears in the (constant) nuclear repulsion energy. The reduction of the symmetry of the system from spherical in the helium case to cylindrical for H_2 introduces a number of complications. Nevertheless, by choosing to use confocal elliptical coordinates,

$$\xi_i = \frac{(r_{Ai} + r_{Bi})}{R_{AB}}; \quad \eta_i = \frac{(r_{Ai} - r_{Bi})}{R_{AB}}, \quad (8)$$

we can expect good results from a trial wave function of the form

$$\exp(-\alpha[\xi_1 + \xi_2]) \sum_{ijklm} \xi_1^i \xi_2^j \eta_1^k \eta_2^l r_{12}^m. \quad (9)$$

Again, this approach is very successful,⁹ and only a few terms are needed for high accuracy.

We should point out one subtlety concerning the exponential cusps at the nuclei. These are a direct consequence of the assumption that the nuclei are point particles like the electrons. This is an approximation, and one that in relativistic calculations, for instance, is better dispensed with. Once the nuclei are assumed to be finite in size, the electron–nuclear cusp disappears.

In addition to the formal behavior of the wave function in the region of the electron–nuclear and electron–electron cusps, much analysis has been devoted to the asymptotic behavior (see, for example, Ref. 10). This is again exponential in nature, for many-electron as well as one-electron systems.

In view of the foregoing discussion, the characteristics of a general approach to molecular electronic structure appear clear. The molecular nonrelativistic Coulomb Hamiltonian (again with fixed nuclei — the clamped-nucleus Born–Oppenheimer Hamiltonian) is

$$H = -\frac{1}{2} \sum_i \nabla_i^2 + \sum_{i>j} r_{ij}^{-1} - \sum_{i,A} Z_A r_{iA}^{-1} + \sum_{A>B} Z_A Z_B R_{AB}^{-1}. \quad (10)$$

To expand the unknown wave function, the foregoing analysis suggests that trial functions that involve one- and two-particle functions should be constructed: the former should display exponential behavior at the nuclei and at long distances and the latter should behave linearly near zero in r_{12} . Fermion statistics and spin symmetry must of course be properly incorporated, but the desired analytical properties of the trial wave function are those we have described. It might then come as a surprise to a reader with mathematical inclinations and no experience of quantum chemistry that such functions are in fact almost never used! Unfortunately, trial functions of this very desirable type lead to almost unmanageable complications in the wave function optimization. Such trial functions have been used rather seldom, and then only for small systems with rather few electrons, because of the very severe computational problems that arise from the many-electron integrals that appear in the variational energy expression Eq. (5); alternative strategies such as perturbation expansion of the wave function and energy suffer from exactly the same problems. A compromise seems necessary between mathematical desirability and computational feasibility: that compromise is driven by the feasibility of evaluating and manipulating the many-electron integrals that will appear when expressions such as Eq. (5) are expanded. In order to understand how this compromise will affect our computed results, we shall now broadly review the methodology of *ab initio* molecular electronic structure theory.

2.2. Molecular Orbitals and Electron Correlation

We can obtain some physical insight into the requirements for different trial functions by partitioning the problem into different levels of treatment. As a first approximation, we may assume that the detailed interaction between electrons can be replaced by an averaged interaction potential: one in which each electron interacts with a smeared-out charge distribution provided by the other electrons. This is the mean-field or *Hartree-Fock* approximation. It can be obtained by writing a trial wave function in which each electron occupies a one-electron function termed an *orbital*; the overall many-electron wave function then takes the form

$$\Psi = \mathcal{A}(\psi_1(\tau_1) \dots \psi_N(\tau_N)). \quad (11)$$

Here \mathcal{A} is an antisymmetrizing operator designed to ensure that the overall wave function obeys Fermi statistics (i.e., is antisymmetric with respect to

electron interchange). The functions ψ_i are the Hartree–Fock orbitals, τ_i denotes the space and spin coordinates of electron i .

The Hartree–Fock orbitals are unknown functions of the electron coordinates. Their form may be determined by substituting the wave function Eq. (11) into the variation principle and making the energy stationary. The solution to the resulting Hartree–Fock equations then defines the optimum orbitals. The form of the equations is given explicitly later in this section and is also discussed in detail in the chapter by Almlöf. For the present, we mention only that for atoms and diatomic molecules, numerical methods can be used to solve the Hartree–Fock equations to high accuracy, giving numerical Hartree–Fock orbitals. However, for polyatomic molecules no general convenient numerical method is available, and we proceed by expanding the unknown Hartree–Fock orbitals in a fixed *basis set*,

$$\psi_i = \sum_{\mu} \chi_{\mu} C_{\mu i}, \quad (12)$$

where the elements χ_{μ} of the basis set are chosen with an eye to computational tractability and efficiency, and with some attention to the physics of electronic motion.

It is obvious that the Hartree–Fock wave function is not a complete representation of the electronic structure, because the true interaction between electrons is not via an averaged potential, but is governed by the Coulomb repulsion term in the Hamiltonian. In effect, the true electronic motion will be correlated more than in the Hartree–Fock model, because of the strong Coulomb repulsion that is exerted at small interelectronic distances. Löwdin¹¹ coined the term “correlation energy” for the difference between the exact nonrelativistic energy of a system and the Hartree–Fock energy. At first sight, the correlation energy seems rather unimportant, because it is a small fraction of the total energy. Indeed, for first-row atoms and molecules the Hartree–Fock energy is typically more than 99% of the total energy. However, the energy differences that are of interest in chemistry, such as binding energies, are also only 1% or less of the total energy. Hence we must not be surprised if the Hartree–Fock approximation is not adequate for describing various phenomena of interest.

We can readily see that the explicit inclusion of interelectronic coordinates in the He and H₂ wave functions in Sec. 2.1 above is related directly to the phenomenon of electron correlation. These functions provide a very

rapidly convergent representation, but, as we have stated, their use in more general systems is accompanied by enormous computational complications. Thus, we must find an alternative method of accounting for electron correlation. Let us suppose that a basis set has been used to approximate the solution of the Hartree–Fock equations. In general, this set will include more elements than there are occupied Hartree–Fock orbitals, and thus there will be a complementary space of so-called *virtual orbitals* as well as the occupied orbitals produced. N -electron functions in which one or more occupied orbitals are replaced by these virtual orbitals can be formed, and a trial function can be constructed as a linear combination of these N -electron functions. This is the *configuration interaction* (CI) method. If a complete basis set were used for the Hartree–Fock calculation, and all the possible resulting N -electron configurations were employed in the CI calculation, our results would be exact. In practice, complete sets cannot be employed, and then the results may be affected by the truncation of both the one-electron basis set and, if additional truncation of the configuration space is performed, by truncation of the N -electron basis. We may expect that the convergence of the correlation energy with truncation of the one-electron basis will be slow. This is because the cusp behavior as an interelectronic distance tends to zero is quintessentially a two-electron phenomenon. Such a two-electron cusp can be described rather well by only a few functions that display cusp behavior themselves, as the results of Hylleraas⁸ and of James and Coolidge⁹ show. But if our two-electron functions can only be represented using products of (in this sense, smooth) one-electron functions, a large number of such functions will be required. There is a contrast here with representing the one-electron nuclear cusps, which as we saw can be accurately described with exponentials. The correlation cusp inherently will require long expansions in products of one-electron functions.

In practice, it is convenient to consider electron correlation as arising from two different sources. What we have discussed so far — the two-electron cusp behavior in the wave function — is termed *dynamical* correlation, since its origins are in the dynamics of the electron motion. In some circumstances, however, such as when chemical bonds are formed or broken, the Hartree–Fock model itself is qualitatively incorrect. This may be because other electron configurations are similar in energy to the Hartree–Fock configuration, and interact strongly with it. Such *non-dynamical* correlation is a failure of Hartree–Fock to provide appropriate

orbitals or an appropriate zeroth-order wave function. These defects can be remedied by multiconfigurational Hartree–Fock calculations that include all the important configurations. Since this requires only that the basis set be capable of describing all the important zeroth-order configurations, we can expect that the basis set requirements for such calculations will be similar to those for Hartree–Fock calculations. The difficulties of describing the correlation cusp thus will not affect treatments of nondynamical correlation.

2.3. Slater Orbitals and Gaussian Orbitals

As we have said above, the simplest *ab initio* approach to molecular electronic structure calculations is to invoke the independent-particle or Hartree–Fock approximation. While the Hartree–Fock equations can readily be solved numerically for atoms, their general application to molecules was only made feasible by the introduction of basis set expansion methods by Roothaan¹² and by Hall.¹³ Our aim is to obtain an optimum set of one-electron orbitals: *molecular orbitals* (MO's) for a molecule, by minimizing the energy of a single electron configuration like that of Eq. (11), subject to the orthonormality constraint

$$\int \psi_i^*(\tau) \psi_j(\tau) d\tau = \delta_{ij} . \quad (13)$$

We expand the unknown MO's in a fixed basis set (Eq. (12)) and obtain from the variation principle (and the orthonormality constraint) the following equation system for the closed-shell case:

$$\mathbf{FC} = \mathbf{SC}\boldsymbol{\varepsilon}, \quad (14)$$

where the “Fock matrix” elements are

$$F_{\mu\nu} = h_{\mu\nu} + \sum_{\lambda\sigma} D_{\lambda\sigma} \left(g_{\mu\nu\lambda\sigma} - \frac{1}{2} g_{\mu\lambda\nu\sigma} \right), \quad (15)$$

and **S** is the overlap matrix over the basis functions

$$S_{\mu\nu} = \int \chi_\mu(\tau) \chi_\nu(\tau) d\tau. \quad (16)$$

(We shall henceforth assume that the basis functions are real, except where we specifically indicate to the contrary. This involves no loss of generality in

the nonrelativistic case and in the absence of magnetic fields.) The orbital expansion coefficients (Eq. (12)) to be determined form the matrix \mathbf{C} , while the matrix $\boldsymbol{\epsilon}$ is a diagonal matrix of “orbital energies.” Finally, the Fock matrix elements have been expressed in terms of one-electron

$$h_{\mu\nu} = \int \chi_\mu(\tau) \left(-\frac{1}{2}\nabla^2 - \sum_A Z_A r_A^{-1} \right) \chi_\nu(\tau) d\tau, \quad (17)$$

and two-electron integrals

$$g_{\mu\nu\lambda\sigma} = \iint \chi_\mu(\tau_1) \chi_\nu(\tau_1) r_{12}^{-1} \chi_\lambda(\tau_2) \chi_\sigma(\tau_2) d\tau_1 d\tau_2, \quad (18)$$

and the “density matrix”

$$D_{\mu\nu} = 2 \sum_i C_{\mu i} C_{\nu i}, \quad (19)$$

where i runs over occupied MO’s.

The matrix form of the Hartree–Fock equations is usually referred to as the *self-consistent field* (SCF) equations, because the equations Eq. (14) are generally solved iteratively, beginning with a guess at \mathbf{C} that is used to form the Fock operator, which is in turn inserted in Eq. (14), whose solution determines a new \mathbf{C} . Iterations are repeated until the incoming and outgoing MO’s agree to within a given tolerance. The use of a fixed expansion basis set is referred to as the *linear combination of atomic orbitals* (LCAO) approximation. This terminology reflects the scheme’s origins in the traditional qualitative approach to constructing molecular orbitals as simple linear combinations of atomic orbitals. In fact, in the years immediately following the introduction of the SCF equations, considerable effort was invested in using basis functions that closely resemble atomic orbitals, namely, the Slater orbitals of Eq. (3). A review of this early work was given by Allen and Karo.¹⁴ The simplest STO basis sets contained one function for each orbital occupied in the atom, and were termed *minimal* or single-zeta basis sets. While such sets should provide a good description of the core electrons, we can expect that the atomic valence charge density will deform in the molecule, and therefore that a more flexible description of at least the valence atomic orbitals may be required, and to improve the molecular description larger STO sets were formed by using

two exponentials per occupied orbital in the atom — double-zeta basis sets. The desirability of including orbitals not occupied in the atom (such as *d*-type functions for first-row atoms) to allow for “polarization” of the atomic electron density on molecule formation was also recognized (see Sec. 3.6) and led, for example, to double-zeta plus polarization STO sets. In addition, very large sets were used in a number of calculations,^{15,16} providing much useful information on the convergence of STO expansions as well as “near-Hartree–Fock” quality results for some molecular properties.

The use of STO basis sets in LCAO-SCF calculations came at a price, however. In order to perform SCF calculations, it is necessary to compute the one- and two-electron integrals of Eqs. (17) and (18). Combinations of analytical and numerical methods were implemented for diatomic molecules (that is, for one- and two-center integrals), and then extended to linear polyatomic molecules, but the computer time required to evaluate the integrals was very large. Moreover, estimates of the computer time that would be required if the multicenter numerical methods were extended to treat nonlinear polyatomic molecules indicated that STO’s would simply not be feasible for this, the most general molecular application. It appeared that an alternative approach was required.

A different type of expansion basis function had actually been introduced before the SCF equations were derived: Boys¹⁷ had suggested that functions with a Gaussian rather than exponential dependence could be useful in electronic structure calculations. Such functions were first used by Boys and coworkers¹⁸ and by Preuss.¹⁹ Boys suggested the “Cartesian Gaussian” functions, given in unnormalized form as

$$x_A^l y_A^m z_A^n \exp(-ar_A^2), \quad (20)$$

centered at point *A* and involving “angular quantum numbers” *l*, *m* and *n*. Such a *Gaussian-type orbital* (GTO) can of course also be written as a linear combination of functions involving spherical harmonic angular factors, and vice versa, and we will return to this issue later. At first glance, the basis function of Eq. (20) has little to recommend it from the point of view of analytical behavior. It lacks the cusp behavior of the exponential function as *r* → 0, and at large distances it will die off much too quickly. These disadvantages are manifested in using GTO’s as a basis set to solve the hydrogen atom problem. Here the exact bound-state solutions require an infinite expansion in GTO’s, but a finite expansion in STO’s. We may

thus expect the convergence of GTO basis set expansions to be much worse than STO expansions. GTO's have one enormous advantage, however: all of the integrals required for an SCF calculation can be evaluated straightforwardly. While the analytical formulae can become tediously long, the computer time requirements are very modest. And in practice the slow convergence of GTO expansions can be alleviated somewhat by using fixed linear combinations of GTO's – the process of *basis set contraction* that we discuss in Sec. 3.

2.4. Correlating Orbitals

So far, we have discussed STO and GTO basis sets only in the context of SCF calculations. Of course, for most basis sets an SCF calculation will yield a number of unoccupied virtual MO's that can, as we discussed in Sec. 2.2, be used in some type of calculation of the correlation energy. We have already established that the convergence of the correlation energy with respect to any approach using (smooth) one-particle basis functions will be slow. What one-electron basis functions will be required for an adequate description of electron correlation? And what are the relative merits of STO's and GTO's in this context?

As a first step in answering these questions, we need to define a set of orbitals that will let us analyze the correlation energy conveniently. For two-electron systems, the *natural orbitals* (NO's) of the system (see, for example, Ref. 20) are ideal for this purpose. The NO's are eigenvectors of the one-particle density matrix, summed over spin. Specifically, for a real N -electron wave function Ψ we can expand the one-particle density

$$\gamma(1|1') = \int \Psi(1, 2, \dots, N)\Psi(1', 2, \dots, N)d\sigma_1 d\tau_2 \dots d\tau_N, \quad (21)$$

where the notation indicates integration over the spatial coordinates of electrons 2 through N and summation over the spin coordinates (σ) of all electrons, as

$$\gamma(1|1') = \sum_{pq} \gamma_{pq} \psi_p(1)\psi_q(1'). \quad (22)$$

Here ψ is a set of orthonormal orbitals (say, the Hartree–Fock orbitals, occupied and virtual) and γ is the density matrix. The natural orbitals φ have the property that

$$\gamma(1|1') = \sum_r n_r \varphi_r(1)\varphi_r(1'), \quad (23)$$

that is, the density matrix is diagonal in the NO basis. The diagonal elements are termed *occupation numbers*, and measure the importance of each NO in the expansion of the density. For the two-electron systems, the occupation numbers also measure the importance of each NO in the wave function.²¹ We can therefore examine the NO's for a two-electron system to try to understand how different correlation contributions arise.

We display here the radial part of the NO's for the ground state of He. Since the ground state is a closed shell, the NO's display full atomic symmetry. We consider first the *s* symmetry NO's, shown in Fig. 1. The most striking feature is the radial extent of all the NO's: the 1s, which has an occupation number near two and which closely resembles the Hartree–Fock 1s orbital, and the correlating orbitals, 2s, 3s, etc., is similar. The main difference between the NO's is that the number of radial nodes changes: the 1s NO is nodeless, as expected, and the correlating orbitals have successively more radial nodes. This is mandated by the orthonormality of the NO's. The effect is that these *s* symmetry NO's describe *radial correlation*: the tendency that if one is electron is to be found near the nucleus, the other will be found further away. In three dimensions, of course, these radial nodes become nodal surfaces. Thus the *s* symmetry NO's display the same nodal surfaces as typical atomic *s* orbitals, hence the notation 1s, 2s, etc. It is crucial to note, however, that the spatial form of these NO's is completely different from excited state atomic orbitals. For example,

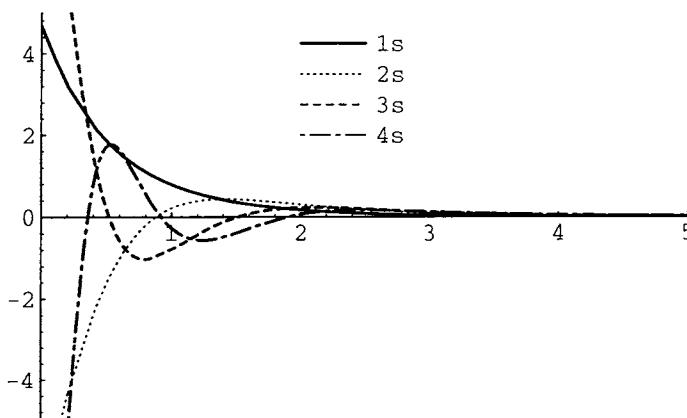


Fig. 1. He *s* symmetry natural orbitals.

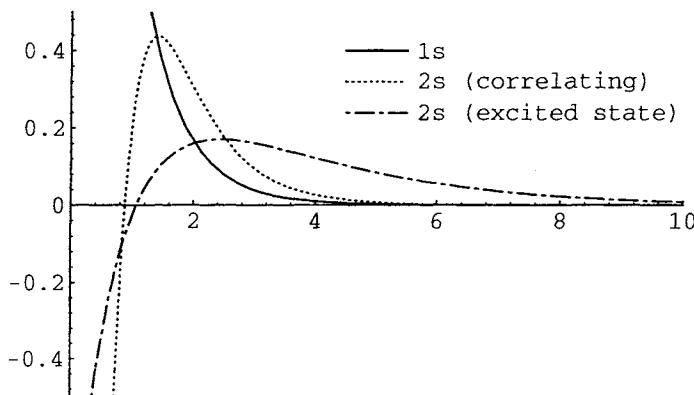
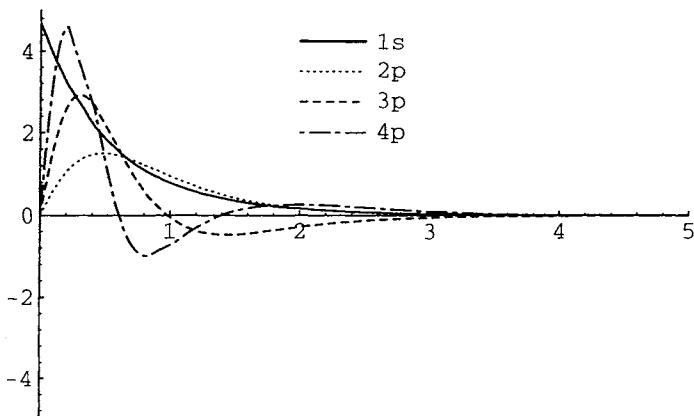


Fig. 2. He 1s and 2s orbitals.

Fig. 3. He p symmetry natural orbitals.

In Fig. 2 we have plotted the 1s and 2s NO's for He, together with the 2s orbital from the 1s2s lowest triplet state of He. As we noted above, the correlating NO has much the same radial extent as the strongly occupied NO, whereas the excited state 2s orbital is much more diffuse. This gives us an immediate clue to the basis set requirements for radial correlation: we need basis functions *with the same spatial extent as the strongly occupied NO's*, but with more radial nodes.

Of course, the exact wave function for He cannot be expanded using (one-center) s orbitals alone. The radial forms of the lowest p symmetry

NO's are shown in Fig. 3, together with the $1s$ NO. Once again we see that the spatial extent of the correlating NO's is similar to that of the strongly occupied NO. However, the nodal structure is obviously different from the s symmetry NO's. For the $2p$ orbital we have a nodal surface passing through the origin arises from the angular part of the wave function. For the $3p$ orbital, we have an additional nodal surface to maintain orthonormality to the $2p$ orbital, and so on. These p symmetry orbitals are said to describe *angular correlation*: the tendency that when one electron is on one side of the nucleus, the other electron will be found on the other side. Higher angular types such as d or f symmetry NO's contribute to the same effect, and these NO's also display the same radial extent as the strongly occupied orbitals, like the s and p symmetry NO's. Again, this indicates what is required for a good description of angular correlation: we need basis functions with more angular nodes than the strongly occupied orbitals, and with about the same spatial extent.

The above discussion used NO's for He as the example. However, the conclusions are quite general, even though the formal analysis is less rigorous (for instance NO occupation numbers for many-electron systems only provide a guide to the importance of NO's for the density, not to the wave function itself). Thus the most useful functions for describing correlation effects in given MO's have the same spatial extent as the occupied orbitals, but have additional nodal surfaces. For atoms, for example, we require functions with more radial and more angular nodes than the strongly occupied orbitals. This also allows us to draw some conclusions about the relative efficacy of GTO's and STO's in describing correlation effects. Since it is the nodal surfaces that are the most important feature, the difference between Gaussian and exponential behavior is of less consequence than it was for occupied orbitals, and there is less difference in convergence between STO and GTO expansions for the correlation energy than for the Hartree–Fock energy. We emphasize, however, that the overall rate of convergence of the correlation energy with any one-particle basis set expansion will be slow. Various studies show that the convergence of the correlation energy as a function of angular quantum number l behaves as a fairly low power of l : $(l + \frac{1}{2})^{-4}$ for He, for example (see Ref. 22 and references therein). The result is that while, say, 90% of the correlation energy for a small molecule can be recovered with moderate effort, increasing this fraction to 95% is difficult, and increasing it further is almost impossible. Very high angular

momentum functions will be required for accurate results, as exemplified by the first-row atom calculations of Sasaki and Yoshimine.²³ Overall, therefore, the basis set requirements for reliable correlations are much more stringent than for reliable SCF calculations, and we shall devote considerable attention to this point.

3. Gaussian Basis Sets

3.1. Basis Set Convergence

We have mentioned that Gaussian basis functions display neither exponential nuclear cusp behavior nor exponential decay. We can illustrate this by plotting a normalized STO and GTO,^a both with unit exponent, as in Fig. 4. The STO is, of course, the exact hydrogen $1s$ solution. The deviation of the GTO from the STO behavior at short and long range is very noticeable. The total energy of this single GTO is $-0.09 E_h$, which is in error by more than $0.4 E_h$, but of course the exponent is far from optimum. A GTO exponent of 0.285 gives a total energy of $-0.4244 E_h$, still in error by $0.075 E_h$. Figure 5 shows this GTO on the same plot as the exact solution. A comparison with Fig. 4 illustrates how the smaller GTO exponent produces much better agreement between the GTO and STO, at least at distances beyond $1 a_0$ from the nucleus. The remaining error in the energy arises from the short-range behavior of the GTO. There is no further remedy for this using a single Gaussian: the only recourse is to use more GTO's in the expansion. For example, in Fig. 6 an optimized four-term GTO expansion is shown together with the exact solution. The only significant deviation (at least to the eye) is in the region very close to the nucleus, and the total energy of this GTO expansion is in error by only $0.0007 E_h$.

It is evident that the convergence of total energies with respect to the number of basis functions used will be considerably slower for GTO expansions than for STO expansions. The practical question, of course, is by how much. Early studies of this question gave rather pessimistic results. For atoms, for example, at least three times the number of GTO's as STO's were required for a given error in the Hartree-Fock energy.²⁴ Since an error of about $0.001 E_h$ relative to the Hartree-Fock limit requires some four

^aWe should note that throughout this work we normalize according to spherical coordinates, with volume element $r^2 dr \sin \theta d\theta d\phi$. The quantities plotted include only the radial part of the normalization factor.

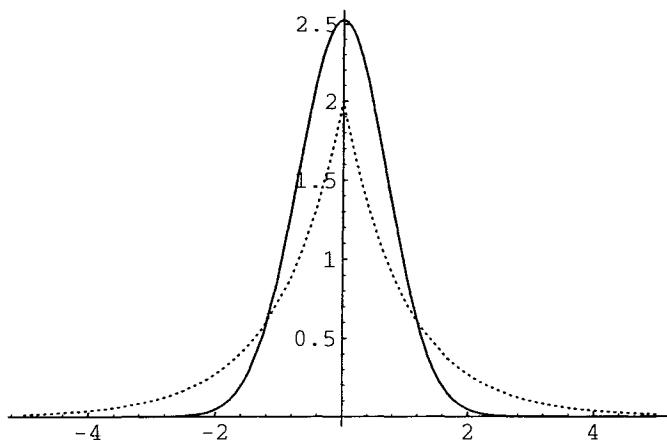


Fig. 4. Unit exponent normalized GTO and STO. Solid line: GTO; Dashed line: STO.

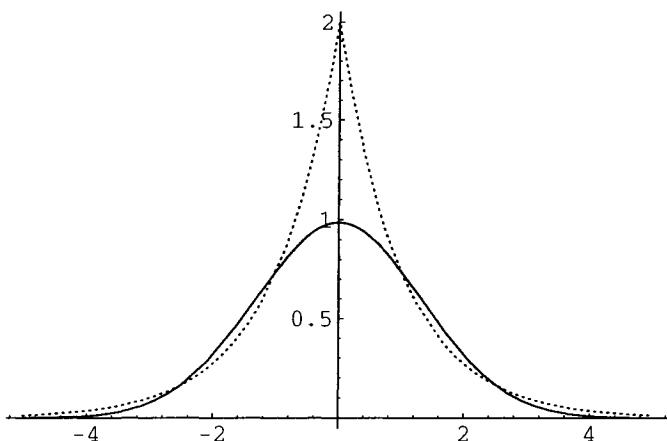


Fig. 5. Optimum GTO for H 1s. Solid line: GTO; Dashed line: STO.

s-type and three *p*-type STO's for a first-row atom, this suggests that at least twelve *s*- and nine *p*-type GTO's — a basis we shall denote by (12*s* 9*p*) — would be required. The number of two-electron integrals (Eq. (18)) depends on the fourth power of the basis set, and so would be roughly two orders of magnitude larger for the GTO set than for the STO set. The computational consequence of this is not so much one of computer time (GTO integrals can be computed much faster than STO integrals) but

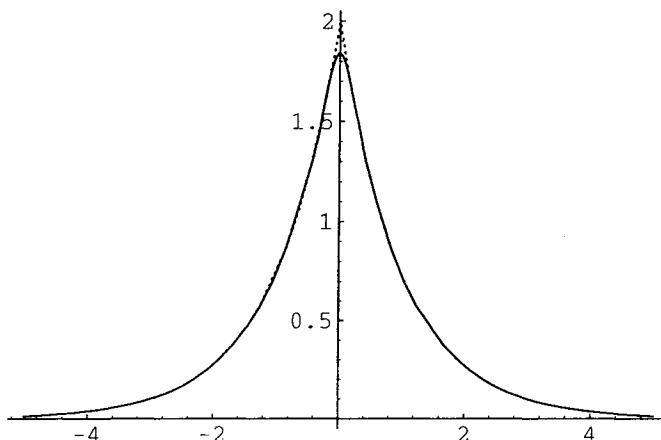


Fig. 6. Four-term GTO expansion for H 1s. Solid line: GTO; Dashed line: STO.

one of storage. The integrals will take up two orders of magnitude more space on disk; similarly, the Fock matrix itself will occupy almost ten times as much memory for the GTO set as for the STO set. This would create critical problems for calculations on many molecules.

One possible solution to the problem of storing and handling large data sets involves recomputing the data rather than storing it, and this approach to molecular electronic structure calculations, which has many ramifications beyond basis set issues, is described in detail in Chapter 3 by Almlöf. More conventionally, it is natural to seek ways in which the dimension of the basis can somehow be reduced. This can be achieved by using fixed linear combinations of GTO's in the calculation, rather than individual GTO's. Such fixed combinations are termed *contracted Gaussian-type orbitals* (CGTO's)^{25,26}; the individual GTO's contributing to these contractions are termed *primitive GTO's*. The size of the one- and two-electron integral matrices, etc., is now determined by the number of CGTO's, rather than the number of primitive GTO's in the original basis.

3.2. Contracted Gaussians from Fits to STO's

The construction of a CGTO basis can be approached in several different ways. In one approach, we retain STO's as the conceptual basis for our calculations, but sidestep the difficulties of STO integral evaluation by explicitly representing the STO's as fixed expansions of GTO's.^{27,28} The

contraction coefficients are obtained by least-squares fitting. Considerable effort was invested in this approach in the early days of computational quantum chemistry, and it still survives in some very small basis sets for molecular calculations. Perhaps the most popular use of fitting has been in representing a minimal STO basis by relatively few GTO's per STO, commonly three.²⁹ In practice, the additional requirement that the same GTO exponents are used for, e.g., the $2s$ and $2p$ STO's is imposed. This naturally will reduce the accuracy of the fit for these functions. Such an "STO-3G" basis, in the terminology of Pople and coworkers, is probably the smallest basis set, in terms of both primitive and contracted functions, that can be used for molecular SCF calculations. The advantages of size and computational economy have led to widespread use of this basis: its reliability and predictive accuracy are only fair, as discussed in Sec. 4.1. For greater accuracy, we could consider constructing GTO expansions of larger STO sets, but the most effort has been invested in using CGTO's directly, rather than as approximations to STO's.

3.3. Contracted Gaussians from Atomic Calculations

The more common approach to the use of CGTO's abandons STO's completely and uses basis functions determined in calculations on atomic systems. This approach is thus based on the observation that (energetically, at least) molecule formation is a relatively small perturbation on the constituent atoms. We may expect, for example, that inner shells, and even the inner regions of the atomic valence orbitals, will not change shape very much on molecule formation. GTO's describing these regions will then appear with the same relative weight in the molecular orbitals as in the atomic orbitals of the separated atoms, and there should be little loss in energy or flexibility if these GTO's are combined into one function using these relative weights.

The design of a CGTO basis requires the specification of a set of primitive GTO exponents, and their corresponding contraction coefficients. Traditionally, this is performed as a two-step procedure. First, the exponents are obtained, commonly by optimization of the atomic SCF energy, or perhaps by a fitting procedure. The contraction coefficients are then taken from the atomic orbitals obtained in this primitive GTO basis. In particular, the inner shells and possibly the inner part of the valence shell are represented as CGTO's using the atomic coefficients, consistent with our

discussion above about their relative insensitivity to the molecular environment. The primitive GTO's contributing mainly to the valence or outer valence region are left uncontracted. For light atoms, at least, as we discuss later in this section, each primitive GTO contributes predominantly to one atomic orbital, and it is reasonable to impose the restriction that each primitive GTO should contribute to only one CGTO. Such a contraction scheme is termed *segmented* and is discussed at greater length in Sec. 3.4, together with possible generalizations.

High accuracy necessitates the use of large primitive GTO basis sets, and as mentioned above the first issue that must be settled is the choice of GTO exponents. While it would be feasible to optimize these in each molecular calculation, this would involve a great deal of effort. Other disadvantages include the possibility of arriving at different stationary points of the exponent optimization at different geometries, creating "bumpy" potential energy surfaces. Almost universally, GTO exponents are optimized in atomic calculations, and the exponents are then kept fixed when the basis is used in molecular calculations. Whatever the source, it is also universal that *shells* of functions are used. That is, the same exponents are used for p_x , p_y , and p_z functions, defining a P shell, or for the five d angular components, etc. This is not at all what would result from molecular optimization of exponents, of course, where one would naturally expect to see some anisotropy among different angular components unless symmetry constrains the outcome. However, enormous computational advantages in the integral calculation result if shells are used, as we shall discuss in Sec. 9 and later. Indeed, such advantages can be compounded if the same exponents are used for *different* angular shells, for instance, a so-called L shell that contains both s and p functions, but such a constraint can compromise the quality of the basis unless many GTO's are used.

Even with the restriction to atomic calculations, optimization of the energy with respect to GTO exponents is not a trivial problem. This is clearly a highly nonlinear optimization, and multiple minima can easily arise. The most common cause is when a given GTO contributes dominantly to one atomic orbital with one value of the exponent, and dominantly to another AO when the exponent has a slightly different value. The ideal solution would obviously be to include more GTO's, but this increases the size of the final basis. Complete details of appropriate optimization strategies are beyond the scope of this review, but we should mention

the increasing use of analytical derivative methods, in which derivatives of the energy with respect to the GTO exponents are calculated as well as the energy itself.³⁰⁻³² Powerful optimization techniques can be employed when these data are available. Nevertheless, GTO exponent optimization remains a tricky business, especially for large basis sets. We shall not discuss this specific issue further: the interested reader is referred elsewhere.³⁰ Instead, we shall assume here that the numerical problems have been solved, and the major issues in GTO exponent optimization are those of providing the best possible description of the chemistry and physics of interest.

Self-consistent field calculations are an obvious tool to use for GTO optimization in atoms. Once the size of the desired set has been decided, it remains only to optimize the atom's SCF energy with respect to the GTO exponents. A variety of atom-optimized GTO sets are available: the occurrence of different "optimized" exponents in sets of the same size from different sources speaks to the difficulties associated with multiple minima in the optimizations, or inadequate convergence of the optimizations. A useful discussion is given by van Duijneveldt in his compendium³³ of optimum exponents for first-row atoms. Van Duijneveldt also devoted some effort to deciding what constitutes "s/p balance" in first-row basis sets, that is, how many shells of each angular type should be present. Some care is always needed in this area, since improving the description of the nuclear cusp often provides the maximum improvement in the SCF energy, but does not necessarily improve the basis set for use in molecules, for which flexibility in the valence region is usually the desire.

In view of the importance of organic chemistry, or more generally the chemistry of first-row systems, it is not surprising that the greatest variety of basis sets is available for first-row atoms, as a perusal of the compendium of Poirier *et al.*³ will show. And while heavier elements are less well served, there are still many optimized GTO sets available. However, for heavier elements a problem arises with the orbital coefficients of the GTO sets: we find that a given primitive GTO will often contribute significantly to more than one atomic orbital. We may compare the *s* AO's of oxygen and sulfur, expanded in double zeta quality GTO sets,^{34,35} in Tables 1 and 2. For oxygen, we see that most primitive GTO's contribute dominantly to either the *1s* or the *2s* AO, with the exception of the sixth, which contributes significantly to both. This would not cause difficulty if we used primitive GTO sets as basis functions, but, as we discussed in Sec. 3.1, it

Table 1. Atomic *s* orbitals for O (3P).

Exponent	1s	2s
7816.54	0.001177	0.000266
1175.82	0.008968	0.002065
273.188	0.042865	0.009786
81.1696	0.143894	0.035740
27.1836	0.355553	0.095078
9.5322	0.461367	0.195897
3.4136	0.140166	0.037412
0.9398	-0.000580	-0.595676
0.2846	0.001391	-0.525755

Table 2. Atomic *s* orbitals for S (3P).

Exponent	1s	2s	3s
35713.98	0.001016	0.000280	0.000081
5396.628	0.007690	0.002146	0.000621
1249.709	0.037859	0.010598	0.003090
359.9343	0.136973	0.041089	0.011925
119.1713	0.339271	0.113395	0.033811
43.98356	0.437323	0.216824	0.065143
17.62667	0.182799	0.087157	0.031091
5.420090	0.009894	-0.556102	-0.231433
2.073873	-0.000851	-0.565247	-0.373061
0.426425	0.000250	-0.031377	0.663394
0.151872	-0.000116	0.006635	0.532427

is necessary to use contracted basis sets. So far, at least, we have considered contraction as simply forming fixed linear combinations of particular GTO's and leaving others uncontracted. In the case of the oxygen ($9s\ 5p$) basis, the sixth GTO should clearly either be left uncontracted, or perhaps should be duplicated so that it can appear in two CGTO's. The situation becomes more difficult with sulfur, as we can see from Table 2. Here the GTO's with exponents 119.1713 and 43.98356 contribute substantially to both the 1s and 2s AO's (but with different relative weights). Similarly, the GTO's with exponents 5.420090 and 2.073873 contribute to both 2s and

$3s$. Hence we again appear to require duplication of primitive GTO's in the contracted functions, or more GTO's must be left uncontracted. A more effective strategy for dealing with this problem is to develop a more general contraction scheme, as we now discuss.

3.4. Segmented Versus General Contractions

In the foregoing discussion, we have spoken of contraction as involving combining primitive Gaussians into a single function. If we write this as a transformation

$$\chi_\mu = \sum_p T_{\mu p} \xi_p, \quad (24)$$

in which χ denotes a CGTO, and ξ a primitive GTO, we thus expect the transformation to have the structure of Fig. 7. Here each column of T has only one nonzero element: each primitive contributes to only one CGTO. This is a reasonable assumption where GTO's are being used to represent STO's, since different GTO's will be used to expand each STO. It is also reasonable if we are combining primitives only to try to represent the nuclear cusp region in a first-row atom. However, it is a restrictive procedure — a more general approach would involve a transformation of the form of Fig. 8. Here the transformation matrix might have no zero elements. Such an approach is termed *general contraction*, whereas the more restricted contraction of Fig. 7 is a segmented contraction. The latter is clearly a special case of the former. General contraction³⁶ is a more flexible approach, and is especially suited to situations in which one or more primitive GTO's contribute strongly to different atomic or molecular orbitals. As we have seen in Sec. 3.3, if only segmented contractions are used, such a primitive must either be duplicated, or excluded from contraction.

General contractions obviously become more attractive than segmented contractions in situations like the sulfur basis discussed in Sec. 3.3. Since the primitive GTO's can contribute to all CGTO's, there is no need to leave large numbers of functions uncontracted, as would have to be done with a segmented contraction. Hence more compact sets of CGTO's can be used, especially for heavier elements.³⁶

The difficulties with primitive GTO's that contribute to more than one AO increase down the periodic table. For heavy elements, especially with large, accurate basis sets, segmented contractions without duplication are hardly possible, unless only a few primitives are contracted. Thus the ad-

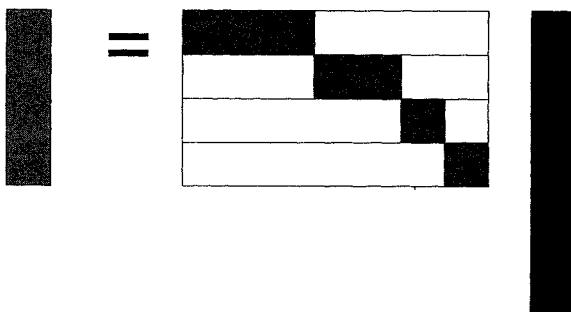


Fig. 7. Segmented contraction



Fig. 8. General contraction.

vantages of general contraction grow with increasing atomic number. It is of interest here to note that Dunning — if not the inventor then certainly the systematizer of segmented contraction — recommended general contraction for use with his GTO sets for the elements Ga–Kr and In–Xe (Ref. 37).

It should be obvious that the general contraction will be more computationally intensive than segmented contraction. The “transformation” of a set of two-electron integrals from a primitive GTO basis to a CGTO basis is an N^4 procedure for a segmented contraction, for N GTO’s, but (in general) an MN^4 procedure for a general contraction if there are M CGTO’s. However, we shall demonstrate in Sec. 13.1 that obtaining CGTO integrals from GTO integrals can be performed significantly more efficiently than these factors would suggest, so the issue of cost of general contraction versus segmented contraction should not be overemphasized.

Another area in which general contractions are important is related to the role of atomic orbitals in molecular calculations. The acronym LCAO used for calculations in which CGTO's are used as the basis clearly originates in the use of atomic orbitals as basis functions in quantum chemical calculations. The single-zeta/double-zeta terminology also refers to AO's, in the sense of how many basis functions represent them. These views were appropriate in the days of small STO basis sets, but most CGTO basis sets simply provide a flexible function space in which the AO's could be represented. In many cases, a program user who examines the coefficients of CGTO's to a specific MO will find it almost impossible to identify which AO's contribute substantially, obscuring any simple AO analysis of the wave function. It was partly this growing disparity between LCAO basis functions and atomic orbitals themselves that led Schmidt and Ruedenberg³⁸ to recommend the use of general contractions, since this allows the use of arbitrarily accurate SCF AO's as LCAO basis functions. A more complete discussion is given by Schmidt and Ruedenberg³⁸ for the SCF case, and by Almlöf and Taylor³⁹ for correlated calculations.

3.5. Even-Tempered Sequences

As we have discussed, one problem with exponent optimization is the occurrence of multiple minima, which results in part from the fact that many nonlinear parameters are being optimized. One way to sidestep this problem is to reduce the number of free parameters defining the GTO basis, by exploiting any formal or empirically observed relationship between GTO exponents. One such empirical relationship is displayed in Fig. 9 for the oxygen (9s 5p) primitive set.³⁴ To a good approximation, the exponents are seen to form a geometric sequence. We can thus envisage specifying an n -term set of GTO exponents as the *even-tempered sequence*⁴⁰

$$\alpha\beta^k, k = 0, 1, \dots n - 1. \quad (25)$$

In our oxygen example, the β value for the s space is about 0.32, and for the p space a little more. We note that such a basis approaches completeness as n increases, provided that α and β are allowed to vary with n . In other words, simply fixing α and β and increasing n will not lead to a complete basis. The obvious advantage of an even-tempered GTO set is that the exponent sequence is determined by only two parameters, α and β , for each angular momentum. Thus optimizing a (9s 5p) primitive

GTO set is a 14-parameter optimization (although the s and p exponent optimizations may be decoupled to a large extent), whereas optimizing an even-tempered sp set of *any* size is a four-parameter optimization (again, with considerable decoupling of the s and p optimizations). Obviously, the imposition of the geometric relationship between exponents is a constraint on the optimization, so an even-tempered set of a given size should give a higher atomic energy than a fully optimized set of the same size. The difference corresponds roughly to the effect of one s and one p function for first-row atoms, so that an even-tempered (10s 6p) set is roughly equivalent to an optimized (9s 5p) set.

A considerable effort to develop and investigate the properties of even-tempered basis sets was invested by Ruedenberg and coworkers in the 1970's. Among the valuable findings of those studies were simple formulae that provide good estimates of the parameters α and β for a given atom and a given expansion length n , and formulae that provide very high accuracy extrapolation of finite-basis atomic SCF energies to the Hartree-Fock limit.^{38,41} Other efforts in these areas include "well-tempered" sets,⁴² for which the same exponents are used for (some) s and p functions. It is also obvious that an even-tempered set should provide an excellent starting guess for a full optimization of the GTO exponents. Finally, extension of an existing basis for special purposes, like adding more diffuse functions to better describe some molecular properties, as in Sec. 6, can conveniently and simply be carried out by approximating the original set as an even-tempered sequence and adding more functions $\alpha\beta^n$, $\alpha\beta^{n+1}$ and so on.

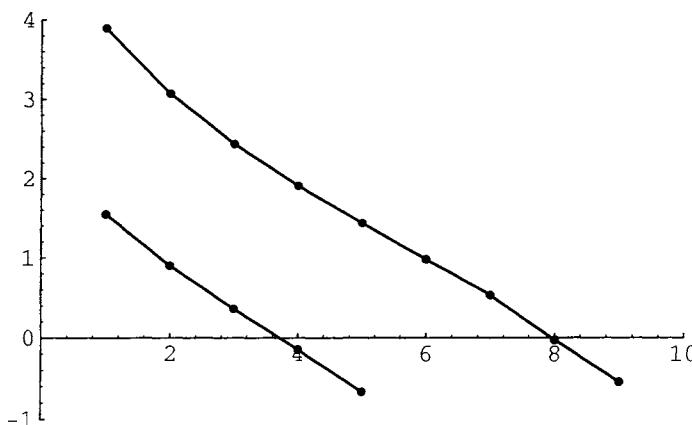


Fig. 9. $\log_{10}(\text{orbital exponent})$ for the O (9s 5p) basis.

3.6. Polarization Functions

One of the consequences of the lower symmetry of molecules, compared to that of their constituent atoms, is that functions of symmetries different from those of the occupied molecular orbitals can contribute to the SCF wave function. For instance, the σ component of the $2p$ orbital on H in H_2 will contribute to the occupied $1\sigma_g$ orbital. These higher angular functions describe the polarization of the charge distribution, that is, the distortion of the atomic charge clouds in the formation of the molecule. Hence basis functions with higher angular momentum than the occupied atomic orbitals are referred to as *polarization functions*. The first question one might ask is how to determine optimum exponents for these functions. If they are to be effective in describing distortion of the charge density, it is natural to expect that they should contribute most where the charge density of the occupied atomic orbital to be polarized is at its maximum. Thus a reasonable rule-of-thumb for choosing polarization exponents for molecular SCF calculations is to determine where the radial maximum of the atomic density occurs and to use a polarization function exponent that has a similar maximum. Optimization actually leads to somewhat more diffuse functions than this prescription, since it is the outer part of the electron density that is most affected by polarization effects. An alternative, therefore, would be to explore a range of molecules and adjust polarization function exponents on the basis of maximum energy lowering, say. In fact, very extensive experience indicates that SCF energies, at least, are not particularly sensitive to the choice of polarization exponents, and a wide variety of “recommended values” for such exponents can be found in the literature (see, for example, Ref. 3). However, higher angular momentum functions serve another purpose in our calculations — the description of electron correlation effects — and as we shall now see, it is often better to consider this aspect rather than SCF charge polarization when determining exponents.⁴³

We may note here another area in which GTO fits to STO's find application. Three term f -type GTO representations of STO's appear to be more effective as polarization functions than a single GTO when used in transition metal calculations. Only the integral evaluation time is affected by the use of more primitives in a contraction, since the size of the CGTO basis, which determines the computational cost of most of the remaining steps in the calculation, is the same.

3.7. Correlated Calculations

Up to this point, we have concentrated on calculations at the SCF level in GTO optimization. Even for atoms, these sets will not be well-suited to correlated calculations, because as we have seen an orbital-based description of the correlation cusp will necessitate the use of higher angular functions in the basis. This problem will, of course, arise in molecular calculations as well. Can we develop an optimization strategy *at the atomic level* for adding higher angular functions to our atomic basis sets for use in molecules? For a considerable period of time it was tacitly assumed that the answer to this question was "no," and higher angular function exponents were obtained from, for example, fits of correlated molecular calculations with (roughly) optimized exponents⁴³ (although see Ref. 44).

In 1987, however, a series of investigations using natural orbitals from atomic wave functions as basis sets showed that contraction coefficients, at least, for correlated molecular calculations, could be obtained from atomic calculations.⁴⁵ Dunning then showed that atom-optimized exponents for higher angular functions could also be determined very satisfactorily from atomic calculations, and developed a new series of basis sets for accurate molecular calculations.⁴⁶ Both these sets and the use of atomic natural orbital basis functions are treated in sections of their own in this review. It suffices at this point to assert that atomic calculations may be used successfully to optimize GTO exponents for both SCF and correlated calculations, although it is perhaps pertinent to note here that both methods rely on general contraction.

4. Basis Sets for SCF Calculations

We have already discussed the qualitative considerations that underlie the design of basis sets for SCF calculations on molecules. We expect that an adequate description of the nuclear cusps will be required: presumably basis sets that achieve this for each atom in the molecule will be satisfactory here. We expect that flexibility in the valence-shell description will be required, so that the atomic charge density can deform in the molecule. This will also require some higher angular momentum polarization functions. On the other hand, there is the all-important practical consideration of computational cost: the basis sets we design must be small enough to make the calculations we wish to perform computationally feasible.

Table 3. SCF results for H₂O.

Basis	<i>E</i> _{SCF} (<i>E</i> _h)	<i>r</i> _e (Å)	∠HOH (°)	<i>ω</i> ₁	<i>ω</i> ₂	<i>ω</i> ₃
<i>sp/s</i> basis sets						
STO-3G	-74.965901	0.989	100.0	4140	2170	4391
DZ (7s 3p)	-75.878683	0.974	107.2	3728	1765	3839
4-31G	-75.895492	0.966	109.4	3778	1683	3914
VDZ (7s 4p)	-75.916744	0.960	111.1	3875	1653	4017
6-31G	-75.972623	0.964	109.8	3808	1677	3950
DZ (9s 5p)	-76.002237	0.963	111.3	3914	1651	4083
<i>spd/sp</i> basis sets						
VDZP (7s 4p)	-75.961664	0.947	105.3	4142	1750	4243
6-31G*	-75.998968	0.956	103.9	3959	1798	4064
6-31G**	-76.015308	0.949	104.6	4080	1753	4185
cc-pVDZ	-76.027389	0.946	104.5	4112	1775	4209
DZP (9s 5p)	-76.042226	0.952	105.5	4119	1738	4224
TZ2P	-76.056174	0.941	106.5	4123	1740	4225
<i>spdf/sp</i> and <i>spdf/spd</i> basis sets						
TZ2Pf	-76.056971	0.941	106.5	4118	1734	4221
cc-pVTZ	-76.058329	0.941	106.0	4127	1753	4229
cc-pVQZ-g	-76.065664	0.940	106.2	4131	1751	4230
Experiment		0.957	104.5	3832	1649	3943

(See text for a complete description of basis sets.)

We shall classify basis sets by size, and discuss the smallest sets first. In order to illustrate some general features of different basis sets, we have performed a number of calculations on the molecule H₂O, using the SIRIUS/ABACUS program system.⁴⁷ In Table 3, we display results for the total SCF energy, equilibrium geometry and harmonic vibrational frequencies obtained with several basis sets of different sizes. Cartesian angular functions were used throughout. The basis sets are grouped by general type, and then ordered by decreasing total energy. It is reasonable to suppose that the largest basis sets listed produce values for the energy and properties that are close to the "Hartree-Fock limit," that is, the complete basis set solution to the Hartree-Fock equations. We have given experi-

mental results⁴⁸ in the table also, in order to provide a perspective on the errors attendant on the neglect of electron correlation. From comparisons with the largest basis results in the table we can see that the Hartree-Fock method underestimates the bond length by almost 0.02 Å, overestimates the angle by almost 2°, and overestimates all vibrational frequencies. The percentage error in the stretching frequencies ω_1 and ω_3 is similar to that in the bending mode, so the absolute error in the stretching modes is considerably larger than that in the bend. We shall discuss the results of Table 3 further in the following sections.

4.1. Small Basis Sets

The smallest sets in routine use for molecular calculations are *minimal* basis sets, in which one contracted function is used for each occupied orbital in the atom. Thus for a first-row atom we have a [2s 1p] set. (Techniques for using subminimal basis sets, like floating spherical Gaussian orbitals,⁴⁹ have been little used in recent years.) The overwhelming majority of minimal basis set calculations employ three-term GTO expansions of STO's: the STO-3G basis of Hehre and coworkers.²⁹ A vast number of quantum chemical calculations have been performed with STO-3G basis sets. The accuracy achievable with such sets is necessarily limited, since the description of the atomic cores is poor; the flexibility of the valence description is limited and is constrained by use of the same exponents for the valence-shell *s* and *p* GTO primitives. STO-3G sets can be useful for qualitative or preliminary investigations, at least for first-row molecules, although, as Table 3 shows, predicted geometries and frequencies are only fair. For example, the STO-3G set is the only basis in the table for which the error in the bending frequency, compared to near-Hartree-Fock results, is substantially greater than for the stretching frequencies.

The next step up from minimal basis sets is to employ a double zeta (DZ) description of at least the valence shell. This gives a [3s 2p] basis set for first-row atoms. Such sets are termed *split valence* or *valence double zeta* (VDZ). The most popular are probably the 3-21G (Ref. 50), 4-31G (Ref. 51), and 6-31G (Ref. 52) sets of Pople and coworkers. Again, the same exponents are used for the valence-shell *s* and *p* GTO primitives. The total energies obtained with the larger of these sets are much better than the STO-3G results, and geometry and vibrational frequency predictions, for example, are fairly good, as can be seen from our H₂O example. Indeed, in terms

of comparison with *experiment*, the 6-31G basis frequencies are the best in the table. Of course, since there is a large correlation contribution to the frequencies, this agreement is quite spurious, arising from a cancellation of errors: neglect of correlation and basis set incompleteness. The error compared to larger basis sets and to the Hartree–Fock limit values is still considerable. Since error cancellations cannot be expected to be universal, the agreement between computed values and experiment with these VDZ basis sets at the SCF level should not be overemphasized. In addition, these basis sets seem to be less satisfactory for heavier systems.

While some of the older small primitive GTO sets (e.g., van Duijn-eveldt's ($6s\ 3p$) set) could be used to construct VDZ basis sets, little work seems to have been done in this area. Recently, Schäfer and coworkers have reported optimized split-valence sets for atoms up to Kr (Ref. 32). These sets are unusual in that the GTO exponents *and* the contraction coefficients have been optimized simultaneously in the atom. Their VDZ set is based on a ($7s\ 4p$) primitive set, and thus should be more economical, in terms of integral evaluation, than the Pople sets. Results for H_2O are listed in Table 3: they indicate that these new VDZ sets produce results very similar to larger sets, and thus look promising for use in large molecules.

Using a double-zeta description of both core and valence shells leads to traditional DZ basis sets. For first-row atoms, Dunning's [$4s\ 2p$] contraction⁵³ of Huzinaga's ($9s\ 5p$) primitive set³⁴ has been popular for more than twenty years.^b A [$6s\ 4p$] contraction of Huzinaga's ($11s\ 7p$) primitive set is of similar quality for second-row atoms. We may note from Table 3 that the larger VDZ and DZ basis sets produce rather similar results for properties other than the energy. That is, the error compared to the Hartree–Fock limit is not greatly dependent on enlarging the *sp* space, at least for the geometry and frequencies.

It should not have escaped the reader's attention that up to this point there has been little mention of basis sets for transition metal atoms or heavier species. This is due in large part to the heavy basis set requirements for even a qualitatively useful description of transition metal *d* shells in molecules. A minimal basis set description is completely unsatisfactory.

^bThe reader should note that there are at least two contraction schemes for the *s* space — denoted ($6, 1, 1, 1$) and ($5, 2, 1, 1$) — in use and referred to as “Huzinaga–Dunning.” Also, in the former contraction pattern, the sixth, eighth and ninth primitives should be left uncontracted. A DZ contraction in which the first six primitives are contracted together is quite unsatisfactory (see Sec. 3.3).

There have been attempts to generate transition metal basis sets of VDZ quality, including the recent work by Schäfer and coworkers,³² but these authors conclude that a triple-zeta description of the *d* shell is desirable. Mention should also be made of the MINI and MIDI basis sets of Tatewaki and coworkers,^{54,55} which are designed to enable small contracted sets to be used in molecular calculations. Otherwise, the transition metal basis sets in wide use involve many GTO's and CGTO's, and thus fall into the regime of "larger basis sets" we now discuss.

4.2. *Larger Basis Sets*

For lighter atoms, the most significant improvement over DZ-quality basis sets is the inclusion of polarization functions. For the Huzinaga–Dunning DZ basis sets, the addition of a single *d* function on first- or second-row atoms yields a DZP set. The corresponding set for H would have a *p* set added to the original [2s]. For the Pople split-valence basis sets, the addition of a *d* set on first-row atoms is denoted by an asterisk, and the addition of a *d* set on first-row atoms and a *p* set on H by a double asterisk, thus, 6-31G* or 6-31G**. Arguments have been given both for and against the addition of polarization functions to non-hydrogen atoms only — see, for example, the discussions by Roos and Siegbahn,⁵⁶ Dunning,⁵⁷ and Hariharan and Pople.⁵⁸ For the case of H₂O, the results of Table 3 suggest that it is desirable to add polarization functions to all atoms — the 6-31G* basis geometry and frequencies are inferior, in comparison to near-Hartree–Fock results, to all the other sets that feature polarization functions on both atoms. Especially noteworthy is the performance of the VDZ set from Schäfer and coworkers,³² augmented with polarization functions (Dunning's exponent recommendations⁵⁷ in this case).

Exponents for the added polarization functions have been derived from a variety of sources.³ For SCF polarization, the exponents are not critical, and in fact using the same exponent for the atoms B through F does not seem to cause problems. However, if the basis set is ultimately to be used in a correlated calculation, the values of the exponents of higher angular momentum functions are much more critical. We defer discussion of this topic to Sec. 5, but it is obvious that it may be preferable to choose polarization function exponents with an eye to their correlation contributions, unless the basis set being constructed will only be used for SCF calculations.

The results for energy differences and structural parameters at the DZP SCF level are usually close to those from very large basis SCF calculations. Larger basis sets, such as “triple zeta” (TZ) or valence triple zeta can provide even better agreement when augmented with polarization functions. Basis sets like TZP, or TZ2P (two sets of d functions, or p functions on H), can be developed using a [5s 3p] contraction of Huzinaga’s (10s 6p) sets³⁴ for first-row atoms, or McLean and Chandler’s [6s 5p] contraction⁵⁹ of the (12s 9p) set.³⁵ Table 3 lists results for a TZ2P basis set (O d exponents of 1.479 and 0.491; H p exponents of 1.299 and 0.433): the results are very close to the Hartree–Fock limit.

We have also listed results obtained with Dunning’s “correlation-consistent” cc-pVDZ set⁴⁶ in Table 3: they are similar to those obtained with other VDZ or DZ sets augmented by a single set of polarization functions.

For transition metal atoms, as we stated above, there are really no “small sets.” Participation of the d electrons in the bonding creates very heavy demands on the basis set. This results from the contribution of different electronic states of the atom to the molecular wave function. Most transition metals have low-lying occupations of the general forms $d^n s^2$ and $d^{n+1} s^1$, and the optimum atomic d orbitals for these occupations are very different, with the d orbitals for the $d^{n+1} s^1$ occupation being considerably more diffuse. In some cases, such as Ni or Pd, the occupation $d^{n+2} s^0$ is also low-lying, and again has very different orbitals from the other occupations. If the different contributions of these atomic states to the molecular wave function are to be described correctly, the d basis must be very flexible.⁶⁰ A double-zeta description of the d orbitals may be inadequate for this purpose. Thus even the smallest useful transition metal basis sets are likely to include three d functions. The most commonly used sets are probably those based on the (14s 9p 5d) primitive sets of Wachters.⁶¹ Unfortunately, even primitive sets this large are not adequate for molecular calculations. Since they were optimized for the $d^n s^2$ occupations, they must be supplemented with a more diffuse d primitive to describe the $d^{n+1} s^1$ occupation.⁶⁰ Further, since the outermost metal p orbital (i.e., the 4p orbital for the first transition row) is not occupied in the atom, there are no primitive GTO’s suitable for describing this orbital in the original basis, and thus one or more additional p primitives are required. Typically, a primitive set like (14s 11p 6d) might be used, contracted to [8s 6p 4d] or [8s 6p 3d]. For smaller primitive

sets, such as those of Roos *et al.*,⁶² even more modification of the original basis may be required. In view of the excellent results obtained for first-row systems, the optimized sets of Schäfer *et al.*³² for transition metals deserve attention.

For consistency with the first-row basis sets, polarization functions should be added for transition metals. A single *f* primitive is usually observed to perform fairly well for this purpose at the SCF level but indifferently at the correlated level⁶³: a contracted *f* function (for example, a two- or three-term GTO fit to an STO) is therefore to be preferred.

4.3. Extended Basis Sets

The basis sets we have described in the previous section will usually give results for SCF energy differences, geometries, and vibrational frequencies (for example) that are close to the Hartree–Fock limit values. One-electron properties like multipole moments, or properties such as polarizabilities will still require special attention, but this is the province of Sec. 6. Nevertheless, sometimes it is desirable to extend basis sets to obtain even closer agreement with Hartree–Fock values. Adding an *f* polarization set to the TZ2P set for first- or second-row atoms is one possible extension, although, as can be seen from Table 3, the effects of an *f* function (exponent 0.9) are relatively small. We also list results not only for the correlation-consistent cc-pVTZ set,⁴⁶ which is a [4s 3p 2d 1f] on O and [3s 2p 1d] on H, but also for the cc-pVQZ set with the *g* function on O and the *f* function on H omitted, giving [5s 4p 3d 2f] on O and [4s 3p 2d] on H. This set should provide results for the properties listed that are close to the Hartree–Fock limit. The effects of the omitted *g* and *f* functions would be negligible at the SCF level.

It is also possible to use larger primitive *sp* sets: van Duijneveldt's (13s 8p) basis³³ for first-row atoms gives very accurate atomic SCF wave functions. Unfortunately, primitive sets of this size are not easy to contract using a segmented scheme, at least not without duplication of some functions. An [8s 5p] or [8s 6p] segmented contraction of the (13s 8p) set is common. On the other hand, if general contraction is used, any desired level of contraction is possible. [5s 4p] or even [4s 3p] sets are suitable for accurate SCF calculations.

For heavier atoms, primitive sets of the quality of the van Duijneveldt (13s 8p) first-row sets are available from the work of Partridge,^{64,65} and

Partridge and Fægri.⁶⁶ The size of these sets makes them difficult to handle using a segmented contraction, and general contraction is the recommended strategy.

5. Basis Sets for Correlated Calculations

From our earlier discussions, we can expect that the convergence of the correlation energy with respect to basis set size will be much slower than for the SCF energy, at least for describing the dynamical correlation energy that arises from the two-electron cusp behavior in the wave function. This places much more stringent requirements on the basis set, as we shall see. Before considering this, however, we shall discuss the basis set requirements for nondynamical correlation.

5.1. *Nondynamical Correlation*

From the discussion in Sec. 2.2, we can think of nondynamical correlation as arising within the valence shell orbitals. Since these orbitals will be described well by a basis that yields good Hartree–Fock results, we can expect that nondynamical correlation effects will also be well-described by such a basis. The only exception would be a system in which the valence shell should be expanded beyond the naive atomic estimate. Such situations include the anions O⁻ and F⁻, in which there are very strong nondynamical correlation effects involving orbitals with a 3p-like nodal structure, and certain systems containing second-row atoms, in which the valence shell should really include the 3d orbitals. In the former case, proper attention must be paid to augmenting the basis set with diffuse p functions, while in the latter an extended set of d functions may be required.

For our H₂O test case, we expect nondynamical correlation to be relatively unimportant. In Table 4 we list complete active space SCF (CASSCF) results for H₂O in various basis sets, again obtained with SIRIUS/ABACUS.⁴⁷ These calculations probably reflect dynamical correlation effects arising from the valence orbitals as much as nondynamical correlation, but they serve to illustrate the basis set convergence of CASSCF calculations. The active space used involves six electrons allocated among two a_1 , two b_2 , and one b_1 MO. More extended active spaces yield much better results for H₂O (Ref. 67), but would not be compatible with the smallest basis sets used here.

Table 4. CASSCF results for H₂O.

Basis	<i>E</i> _{CASSCF} (<i>E</i> _h)	<i>r</i> _e (Å)	∠HOH (°)	<i>ω</i> ₁	<i>ω</i> ₂	<i>ω</i> ₃ (cm ⁻¹)
<i>sp/s</i> basis sets						
STO-3G	-75.020290	1.028	96.8	3571	2034	3815
DZ (7s 3p)	-75.933471	1.006	102.9	3283	1687	3415
4-31G	-75.951901	0.997	104.7	3356	1627	3500
VDZ (7s 4p)	-75.972013	0.989	106.4	3468	1613	3612
6-31G	-76.028743	0.994	105.2	3385	1625	3533
DZ (9s 5p)	-76.057151	0.993	106.6	3489	1607	3664
<i>spd/sp</i> basis sets						
VDZP (7s 4p)	-76.012986	0.970	102.2	3773	1698	3885
6-31G*	-76.053453	0.981	101.1	3571	1740	3694
6-31G**	-76.067331	0.973	101.5	3698	1698	3815
cc-pVDZ	-76.079103	0.970	101.6	3738	1716	3848
DZP (9s 5p)	-76.093727	0.975	102.4	3746	1683	3864
TZ2P	-76.107675	0.962	103.6	3771	1690	3886
<i>spdf/sp</i> and <i>spdf/spd</i> basis sets						
TZ2Pf	-76.108499	0.963	103.6	3764	1683	3881
cc-pVTZ	-76.109802	0.963	103.1	3769	1701	3883
cc-pVQZ-g	-76.117359	0.961	103.4	3779	1699	3892
Experiment		0.957	104.5	3832	1649	3943

(See text for a complete description of basis sets.)

The convergence of the CASSCF energy and properties of H₂O is generally similar to the SCF values of Table 3, although a detailed examination will show that the convergence with basis set is actually slightly slower than for the SCF results. On the other hand, the convergence is generally more systematic than at the SCF level — the bond length is uniformly overestimated compared to experiment, but the discrepancy decreases as the basis is expanded, unlike the SCF results. Similarly, the stretching frequencies seem to converge fairly uniformly from below. In any event, Table 4 demonstrates rather convincingly that the basis set requirements for CASSCF calculations are the same as for the SCF calculations. We may note that, other than the bending frequency, the large basis set CASSCF results and SCF

results (Table 3) straddle experiment: the CASSCF stretching frequencies are smaller than experiment, as is the HOH angle, and the bond length is larger, exactly opposite to the situation for SCF. Further, the difference between experimental and CASSCF results is smaller than between experiment and SCF. This is all typical of CASSCF calculations. The remaining errors can only be removed by treating dynamical correlation, to which we now turn our attention.

5.2. Small Basis Sets

The sort of simple analysis of correlation effects in atoms we employed in Sec. 2.4 describes correlation in terms of radial and angular contributions. Radial and angular correlation are of similar importance, and thus any basis set with a pretension to acceptability for a correlated calculation should describe both of them. *At the very least*, this implies two functions for each strongly occupied orbital included in the correlation treatment, one to describe the orbital itself and one for radial correlation, and a set of higher angular functions (polarization functions, to extend the terminology used earlier) for describing angular correlation. Thus, for example, if we intend to correlate only the valence-shell electrons in a first-row atom, the minimum size of basis set that is useful would be split-valence plus polarization. It is absolutely crucial to realize that smaller basis sets, particularly those lacking polarization functions, will give a completely inadequate picture of dynamical correlation effects.

In view of this reasoning, the smallest acceptable basis sets for correlated calculations will be of split-valence plus polarization (VDZP) or DZP type. Sets of this size designed for use in molecular SCF calculations can be used without modification if desired, but in general the polarization exponents selected for use in SCF calculations are not so effective in describing angular correlation. Correlation-optimized exponents tend to be somewhat larger. Since the correlation energy is more sensitive to the polarization function exponents than is the SCF energy, several authors have suggested that correlation-optimized polarization functions should be used at all times. Optimized exponents have been given by Redmon and coworkers,⁶⁸ Krishnan and coworkers,⁴⁴ and by Ahlrichs and Taylor,⁴³ who give the simple formulae

$$a_d = 0.02Z^2, \quad 3 \leq Z \leq 10, \quad (26)$$

and

$$a_d = 0.077Z - 0.69, \quad 11 \leq Z \leq 18 \quad (27)$$

for d exponents for first- and second-row atoms. These were based on correlated calculations on closed-shell hydrides. A hydrogen p exponent of 1.0 is appropriate. Multiple sets of polarization functions can be obtained by using an even-tempered sequence whose geometric mean is the d exponent from these formulae. A factor of 1/3 between the exponents is suitable.

If higher angular functions are to be added, various optimization studies have confirmed a rough scaling of 1.2 for each step in l for first- and second-row atoms.⁶⁹ For example, Eq. (26) suggests 1.28 as a d exponent for use on O. A reasonable f exponent would be 1.54, with 1.84 for a g exponent, etc. Again, these can be split into even-tempered sequences for multiple polarization functions. We may note here that for both multiple polarization sets and higher angular quantum numbers, the use of Cartesian angular functions becomes less and less desirable. In addition to the true spherical harmonic functions of the corresponding l value, the Cartesians include "contaminant" functions of lower angular type (p functions in a Cartesian F set, for example). The contaminants create linear dependence problems in large sets, and substantially increase the size of the two-electron integral list while adding little or nothing to the basis set. The use of spherical harmonic basis functions is to be strongly preferred.

Table 5 lists H₂O energies, geometries and frequencies obtained using a variety of basis sets and the CCSD(T) method⁷⁰ for treating electron correlation. These calculations were performed with the MOLECULE-SWEDEN-TITAN programs.^{71,72} CCSD(T) is a coupled-cluster method that includes all single and double excitations, augmented with a noniterative estimate of the effect of connected triple excitations — coupled-cluster methods are described in detail in Chapter 16 by Bartlett. Calibration calculations on a number of systems have established that this is a very accurate correlation treatment for a molecule like H₂O, whose exact wave function is dominated by the Hartree-Fock configuration. We note here that only the eight valence electrons of H₂O were correlated in these calculations, and spherical harmonic basis functions were used. The effect of neglecting correlation effects that involve core electrons is discussed below and at greater length in Sec. 8.2, but it will be smaller than most of the basis set effects that we will discuss. In view of the expected accuracy of the correlation treatment, it is appropriate here to compare the computed results with experiment.

Table 5. CCSD(T) results for H₂O.

Basis	$E_{\text{CCSD}(\text{T})}$ (E_h)	r_e (Å)	$\angle \text{HOH}$ (°)	ω_1	ω_2	ω_3
<i>spd/sp</i> basis sets						
cc-pVDZ	-76.241305	0.966	101.9	3822	1690	3927
DZP	-76.249074	0.974	103.3	3819	1652	3937
<i>spdf/sp</i> and <i>spdf/spd</i> basis sets						
TZ2Pf	-76.320942	0.962	104.0	3799	1639	3906
cc-pVTZ	-76.332217	0.959	103.6	3842	1668	3945
ANO	-76.345617	0.962	104.0	3827	1653	3942
<i>spdfg/spdf</i> basis sets						
cc-pVQZ	-76.359798	0.958	104.1	3845	1659	3951
ANO	-76.363655	0.959	104.3	3836	1653	3947
Experiment		0.957	104.5	3832	1649	3943

(See text for a complete description of basis sets.)

Results for two fairly typical small and medium-sized basis sets for use in correlated calculations are given in Table 5: the DZP and TZ2Pf sets described in Sec. 4.2 above. Unlike the SCF case, additional polarization functions (beyond DZP) have a very significant effect on the harmonic vibrational frequencies, although the effect on the geometry is smaller and on a par with the SCF results. The greatest difference is for the total energy, which at the CCSD(T) level changes by more than 0.07 E_h on expanding the basis from DZP to TZ2Pf, five times the SCF energy lowering. The remaining results in Table 5 are for basis sets designed with two different approaches to the construction of correlation-optimized sets: the atomic natural orbital procedure and correlation-consistent optimized GTO exponents. We shall now discuss these approaches in detail.

5.3. Atomic Natural Orbitals

A quite different strategy for the design and construction of basis sets for correlated calculations was proposed by Almlöf and Taylor.⁴⁵ Their philosophy was an extension of the earliest and simplest LCAO treatments, in which the basis functions were chosen as atomic orbitals, or close approx-

imations thereto, as opposed to simply being some set of computationally convenient expansion functions. But while the choice of appropriate LCAO basis functions (SCF AO's) is obvious for molecular SCF calculations, it is not at all obvious what should be used for correlated molecular calculations. What choice of orbitals includes the effect of correlation? The choice made by Almlöf and Taylor was to use the *natural orbitals* (NO's) of correlated atomic wave functions (usually single-reference or multi-reference CISD). As discussed in Sec. 2.4, the natural orbitals are eigenfunctions of the one-particle density matrix:

$$\gamma(1|1') = \sum_r n_r \varphi_r(1) \varphi_r(1'). \quad (28)$$

The quantities n_r are *occupation numbers* of the natural orbitals φ_r , and represent one measure of the contribution a given NO makes to the density matrix. NO's with small occupation numbers can thus be presumed to be less important. The advantage of using NO's is that they provide an obvious strategy for basis set truncation — a correlated calculation in a given primitive GTO set is performed, a threshold is selected, and only those NO's whose occupation numbers are above the threshold are retained. For example, for the neon atom we observe the pattern for NO occupation numbers depicted in Fig. 10. These results were obtained using an eight-electron CISD wave function and a $(13s\ 8p\ 6d\ 4f\ 2g)$ primitive basis.

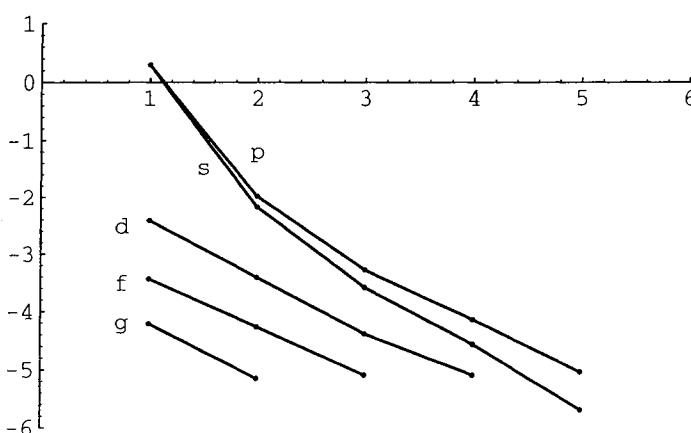


Fig. 10. $\log_{10}(\text{occupation number})$ for Ne natural orbitals.

For occupation numbers above 10^{-5} there is a rather well-defined pattern, and (for instance) selecting at threshold of 10^{-3} , 10^{-4} , and 10^{-5} yields the basis sets $[3s\ 2p\ 1d]$, $[4s\ 3p\ 2d\ 1f]$, and $[5s\ 4p\ 3d\ 2f\ 1g]$, respectively, when the $1s$ orbital is also included. (Below a threshold of 10^{-5} the occupation numbers form less of a pattern, and it becomes progressively more difficult to decide whether to include or omit different NO's, since invariably some will have occupation numbers close to any choice of threshold.) In this way, we have defined a set of contraction coefficients using the natural orbital coefficients.

Except for two-electron systems, there is no well-defined connection between the occupation number of a natural orbital and its contribution to the correlation energy. Nevertheless, considerable experience has shown that the larger the NO occupation number, the more important it is for the energy. For example, we can truncate the NO space and recompute the correlation energy, obtaining the results of Table 6 for the case of Ne. It is obvious that the convergence of the correlation energy with truncation of the NO space follows the occupation numbers rather well. We can therefore assume that using NO occupation numbers as a criterion in basis set truncation is appropriate for the energy, as well as the density matrix or wave function. The $[5s\ 4p\ 3d\ 2f\ 1g]$ ANO set shows a contraction error (loss in energy compared to the GTO set) of about 1 kcal/mol per l value in a correlated calculation, roughly independent of the size of the primitive GTO set. We therefore have a method for exhausting a primitive set in a modest number of contracted functions.

First- and second-row atoms all show a similar pattern of occupation numbers in CISD calculations with large primitive basis sets, and thus *atomic natural orbital* (ANO) contracted basis sets such as $[4s\ 3p\ 2d\ 1f]$ or $[5s\ 4p\ 3d\ 2f\ 1g]$ for the first row, or $[5s\ 4p\ 2d\ 1f]$, etc., for the second, were suggested by Almlöf and Taylor as appropriate for use in correlated molecular calculations.⁴⁵ Several points may be noted here. First, since the transformation from the primitive basis to the ANO basis is, for each symmetry block, a dense matrix, ANO sets necessitate the use of a general contraction. Second, some care is required to maintain "symmetry and equivalence restrictions" (Ref. 73), that is, to ensure that all angular components of the same shell have identical radial terms and that there is no mixing between atomic symmetries. These issues are discussed at greater length in a review article by Almlöf and Taylor.³⁹

Table 6. SCF and CCSD(T) energies for Ne.

Basis	$E_{\text{SCF}} (E_h)$	$E_{\text{CCSD(T)}} (E_h)$
ANO basis sets		
$(13s\ 8p\ 6d\ 4f\ 2g)^a$	-128.546574	-128.857376
$[5s\ 4p\ 3d\ 2f\ 1g]$	-128.546511	-128.850831
$[4s\ 3p\ 2d\ 1f]$	-128.546408	-128.827218
$[3s\ 2p\ 1d]$	-128.545770	-128.750646
Correlation-consistent basis sets		
cc-pVQZ	-128.543470	-128.843690
cc-pVTZ	-128.531862	-128.802454
cc-pVDZ	-128.488776	-128.680693

^aPrimitive GTO set used for ANO contracted sets.

Results of CCSD(T) calculations on H₂O using two ANO basis sets are given in Table 5. These are based on a $(13s\ 8p\ 6d\ 4f\ 2g)$ primitive set on O, and an $(8s\ 6p\ 4d\ 3f)$ set on H. The smaller ANO set is a $[4s\ 3p\ 2d\ 1f]$ contraction on O and $[3s\ 2p\ 1d]$ on H, while the larger is $[5s\ 4p\ 3d\ 2f\ 1g]$ on O and $[4s\ 3p\ 2d\ 1f]$ on H. Both sets produce the lowest CCSD(T) energies of their general type in the table, and both yield excellent spectroscopic constants. Comparison of the two sets suggests that the CCSD(T) stretching frequencies are still not converged with respect to the basis set, and that basis set expansion appears to increase the frequencies, which are already somewhat larger than experiment in the larger ANO set. If anything, core correlation would be expected to increase the frequencies also. However, the differences here are only of the order of a few cm⁻¹, and could arise from deficiencies in the correlation treatment, which is a subject beyond the scope of this work.

As a philosophy, the ANO approach is a rather general one. While it has usually been applied using CISD wave functions, there is no reason why other correlation treatments should not be used. Similarly, it has usually been applied using very large primitive basis sets, to try to ensure that the basis set was close to saturation in the angular functions used, although this is not a requirement. A particular advantage of the use of the same large primitive set for all contracted sets is that the smaller basis sets are proper subsets of the larger ones. Thus a well-defined hierarchy of contracted

basis sets exists, whose results should ultimately converge to those obtained using the primitive basis without contraction. There are a few inadequacies, however. One problem is in the calculation of some molecular properties,⁷⁴ although as we shall see in Sec. 6 almost any energy-optimized basis set will show deficiencies in this respect. Another concerns situations in which atomic states of different orbital occupations contribute significantly to the bonding. As we discussed earlier, transition metal chemistry is one such situation, and the chemistry of molecules containing very electronegative elements like O, F, and Cl, where ionic terms play an important role, is another. One solution is to use not the density matrix for a particular atomic state to define the NO's, but an average over atomic and/or ionic states of interest.^{75,76} It is also possible to take this further and include in the average a range of atomic and ionic states, and states perturbed by an electric field.^{77,78} Averaging over electronic states is also indispensable in the development of ANO basis sets for transition metal atoms.⁷⁹

5.4. Correlation-Consistent Basis Sets

One of the most significant conclusions from the success of ANO basis sets is that it is perfectly acceptable to use correlated atomic calculations to optimize basis sets for correlated molecular calculations. Dunning exploited this to simplify the process of obtaining basis sets for correlated calculations⁴⁶: returning to the optimization of exponents and developing a hierarchy of atomic basis sets of different sizes for use in molecules. His approach is to begin with an accurate (minimal contracted basis) SCF description, and then to add functions to describe correlation. The polarization functions are taken as even-tempered sequences: interestingly, the optimized values roughly reflect the empirical scaling of 1.2 between successive l values (Sec. 5.2). Again, a general contraction procedure is required to handle these *correlation-consistent (cc)* basis sets. As with more traditional basis sets, the number of primitive GTO's increases as the *cc* contracted sets are enlarged. The smallest set is of split-valence plus polarization type, [3s 2p 1d] for the first row, for example, and is denoted *cc-pVDZ*. Larger sets like *cc-pVTZ*, a [4s 3p 2d 1f] contracted set, and *cc-pVQZ*, [5s 4p 3d 2f 1g] were also suggested. The primitive *sp* sets used are (9s 4p), (10s 5p), and (12s 6p), respectively; the polarization sets are always uncontracted.

Energies obtained in calculations on Ne are included in Table 6. The largest *cc* sets produce total energies similar to those from ANO sets of the same size. The loss in energy in the smaller sets is due in part to the error in the SCF energy. The ANO sets reproduce the SCF energy almost exactly, since the strongly occupied NO's are very close to the occupied SCF orbitals, whereas the SCF energy in the *cc* sets varies considerably because of the quality of the primitive basis set used. However, the loss in correlation energy along the sequence *cc-pVQZ*, *cc-pVTZ*, *cc-pVDZ* is almost the same as along the corresponding sequence of ANO sets. The *cc* sets are somewhat cheaper to use, since the initial calculation of integrals over the basis set will require less computer time.

Results obtained with various *cc* basis sets are listed in Table 5. The *cc-pVTZ* and *cc-pVQZ* sets yield spectroscopic constants similar to those of the corresponding ANO sets: the geometries agree somewhat better with experiment and the frequencies somewhat worse.

Martin⁸⁰ has made a detailed comparison of the performance of ANO and *cc* basis sets in the calculation of molecular binding energies. Of particular interest are the differences between [4s 3p 2d 1f] and [5s 4p 3d 2f 1g] level binding energies from the two types of basis set. While these generally favor the ANO sets, the differences are always small and often negligible, especially for the largest basis sets.

Of some interest is the performance of small ANO sets, like [3s 2p 1d], and their *cc-pVDZ* counterparts. In the former case, the use of a large primitive set with a contraction effectively optimized for the atom results in a rather significant bias in the basis set towards describing the atom. Hence binding energies are significantly underestimated compared to the use of, say, a traditional DZP basis set, where the single polarization exponent is chosen as a compromise, favoring neither atom nor molecule. Given the atomic optimization performed in generating the *cc-pVDZ* set, the same atomic bias might be expected to be present. In the authors' experience, at least, this seems to be less of a problem than with the [3s 2p 1d] ANO set. Part of the explanation may lie in basis set superposition errors (see Sec. 7): the ANO set describes the atom sufficiently well that superposition errors are smaller than with the other sets. Cancellation of errors may thus play a role when the *cc-pVDZ* or DZP sets are used. The results in Table 5 suggest that the *cc-pVDZ* set is a reasonable alternative to the traditional DZP basis set (although both the bond angle and bending frequency are anomalously large).

5.5. Basis Set Recommendations

The H₂O calculations we have presented in this review are obviously not an adequate foundation to use in identifying preferred basis sets for use in correlated molecular calculations. There is, however, a very considerable body of literature that addresses this question (see, for example, Ref. 81 and references therein), and we have drawn on that information to formulate some recommendations here.

For small basis sets, we have already established that at least VDZ plus polarization quality is required. The traditional DZP set for the first and second rows, together with an [8s 6p 3d] or [8s 6p 4d] contraction of the Wachters set⁶¹ for transition metals augmented with a contracted *f* set,⁸² seem perfectly acceptable for this purpose. An alternative for the first and second rows is the cc-pVDZ set,⁴⁶ although this set has not been used and calibrated to nearly the extent of the traditional sets as yet. For higher accuracy, the choice for the first and second rows is probably the Dunning cc-pVTZ set,⁴⁶ which is cheaper to use than the corresponding ANO set given by Almlöf and Taylor.⁴⁵ An appropriate transition metal set would probably have to be designed from scratch, perhaps using one of Partridge's primitive sets.⁸³

For the highest accuracy, sets containing at least *g* functions for the first and second rows are required. It is not so clear here that the cc-pVQZ set⁴⁶ is necessarily preferable to an ANO set with the same number of CGTO's, especially for situations in which it is desirable to recover the maximum amount of correlation energy and the optimum description of atomic correlation that the ANO set affords. Where even larger basis sets are required, there is a cc-pV5Z set, [6s 5p 4d 3f 2g 1h] for first-row atoms (Woon and Dunning, unpublished work), and ANO sets of this size or even larger have also been developed and used.⁸⁴ However, such sets become impractically large for polyatomic molecules, and a hybrid approach, using cc primitive sets and ANO contractions, may be the best route to higher accuracy.

Irrespective of their ability to describe correlation effects, none of the basis sets listed here are capable of describing accurately all the molecular properties we may desire to calculate. The design of basis sets or extension of basis sets for the prediction of properties is discussed in the next section.

6. Basis Sets and Molecular Properties

Many time-independent properties can be defined conveniently as derivatives of the energy with respect to certain parameters, that is, the energy of the system is written as a Taylor expansion in the strengths of various applied perturbations

$$E(\lambda) = E_0 + \sum_i \lambda_i \frac{dE}{d\lambda_i} \Big|_{\lambda=0} + \frac{1}{2} \sum_{ij} \lambda_i \lambda_j \frac{d^2E}{d\lambda_i d\lambda_j} \Big|_{\lambda=0} \dots \quad (29)$$

First-order properties are first derivatives of the energy with respect to an applied perturbation (evaluated for a perturbation strength of zero). For example, the dipole moment corresponds to the first derivative of the energy with respect to an applied electric field F , evaluated at zero field strength:

$$\mu = \frac{dE}{dF} \Big|_{F=0}, \quad (30)$$

while harmonic force constants correspond to the second derivative of the energy with respect to displacements of the nuclei, such as

$$k_{A_x B_y} = \frac{d^2E}{dA_x dB_y} \Big|_{\mathbf{C}=\mathbf{C}_0}, \quad (31)$$

where the notation $\mathbf{C} = \mathbf{C}_0$ implies that all nuclear coordinates \mathbf{C} are evaluated at the equilibrium geometry \mathbf{C}_0 . Another example of a second derivative is nuclear magnetic shielding constants, which are mixed second derivatives of the energy with respect to an external magnetic field and nuclear magnetic moments. Basis sets that perform well in the calculation of the molecular energy cannot necessarily be expected to perform well in the calculation of molecular properties. Calculation of reliable properties will require that we describe the molecule equally well in the presence and absence of the perturbation, and we now discuss how this affects the choice of CGTO basis set.⁸⁵

6.1. Basis Functions and Energy Derivatives

One of the issues that appears immediately when we consider the basis set requirements for computing reliable properties is the behavior of the basis set itself under the perturbation. For example, when we compute

the gradient of the energy with respect to nuclear displacements by finite differences, the basis set clearly moves with the nuclei, and thus the basis set changes as the perturbation is applied. In other words, the basis set depends on the perturbation. What if the gradient is computed analytically as the first derivative of the energy? The expressions for the analytical energy derivatives with respect to nuclear displacements⁸⁶ include terms resulting from the dependence of the basis set on the nuclear coordinates, so again we see that the basis set is perturbation-dependent. The net effect is that the molecular basis set “readjusts” or responds satisfactorily when geometric perturbations are applied.

A quite different situation is encountered in the calculation of electric properties. Here the same basis set is almost always employed whether the field is present or not. Now, for electric field perturbations, the effect of applying the perturbation is to multiply the wave function by a factor in \mathbf{r} . For example, an s orbital will acquire some p character when the perturbation is applied. Whether this orbital response is described correctly or not will depend on the original choice of basis set (what p functions are present, in this case).⁸⁷ Here we have a potential problem — if the polarizing effect of the electric field is not represented in the original basis, the response of the molecule to an electric field perturbation will not be properly described.

These considerations are subtle, and are worth reiterating. If our basis is fixed and does not itself change under the perturbation, it must be chosen flexible enough that the response to the perturbation can be accommodated by the LCAO coefficients. If the basis set is perturbation-dependent, however, it should be more capable of describing the response to the perturbation adequately. Hence for geometric perturbations, we can expect that basis sets that yield good energies should also yield good properties (equilibrium geometries and harmonic frequencies, say), whereas for electric-field perturbations such as dipole moments or polarizabilities it may be necessary to augment the basis. For higher derivatives of the energy the same reasoning will apply, and we may expect that mixed derivatives (such as dipole moment gradients for infrared intensities) will suffer from the same problems as the electric field derivatives.

The situation for magnetic perturbations is more complicated. For gauge-invariant methods, the basis set (London orbitals, for example) depends explicitly on the applied perturbation.⁸⁸ Hence the computation of

the energy derivative automatically accounts for the basis set response term. In contrast, if the basis set does not depend on the perturbation, then again extension or manipulation of the basis set will be required to describe the response as well as to ensure approximate gauge invariance. The basis set requirements for calculation of mixed derivatives will be determined by the responses to each perturbation — mixed electric/magnetic properties will usually require extension of the basis set to describe the electric field response, for example.

We can summarize our reasoning so far as follows. For geometric perturbations, at least for atom-centered basis sets, we expect good results with energy-optimized basis sets. For electric-field perturbations we must expect to extend our energy-optimized basis sets, since we must ensure that the basis set is flexible enough to describe the response of the wave function to the perturbation. For magnetic properties, an overriding concern is usually to ensure gauge invariance. The use of London orbitals to achieve this results in a basis set that is again perturbation-dependent, and in a physically reasonable way; we may then expect that our usual energy-optimized basis sets will yield good magnetic properties.

We should mention here several alternatives in constructing basis sets suitable for the prediction of molecular properties, particularly electric field derivatives. As we have said, the obvious approach is to enlarge the basis set, and we shall devote considerable attention to this strategy. An alternative would be to make the basis set perturbation-dependent, as is already the case in calculations of geometric perturbations. Sadlej⁸⁹ has developed electric-field-dependent basis set methods for just this purpose, but these have received less use and attention than they might, especially given that modern analytical derivative methods can deal with these basis sets straightforwardly. Sadlej's method has been more widely used, in fact, as a way of identifying functions to be added to existing energy-optimized basis sets (see, for example, Ref. 90).

Finally, we can contemplate treating the basis function positions as wave function parameters, and optimizing them in a variational calculation. This is the approach based on so-called "floating functions," in which the energy is optimized with respect to the basis function positions. Floating functions can provide significantly improved property values,⁹¹ but are not always cost-effective. A hybrid approach is to augment the basis with the specific aim of duplicating the effect of floating the basis functions: we consider

this, together with more conventional aspects of basis set manipulation and augmentation, in the next sections.

6.2. Spectroscopic Constants

Under this heading we include those parameters that can be derived from a knowledge of the potential energy surface in the absence of external fields, including the equilibrium geometry and force constants. From our discussion of the basis set response contribution to energy derivatives, we know that this is included for nuclear motion perturbations, whether the derivatives are obtained analytically or numerically, for basis sets moving rigidly with the nuclear framework. We may thus expect that basis sets that give reliable energies will also give reliable energy derivatives with respect to the nuclear coordinates.⁹² The convergence of equilibrium geometries and force constants with respect to basis set should be similar to that of the energy. For example, SCF equilibrium geometries and force constants obtained using TZ2P or even DZP quality basis sets should be close to Hartree–Fock-limit values, as is well illustrated by the results of Table 3, and of Table 4 for CASSCF wave functions. The convergence of correlated values — that is, the contribution of dynamical correlation — is slower, because the convergence of the correlation energy itself is slower, but the basis set requirements are essentially the same as for the energy. Again, this is illustrated perfectly by the H₂O results of Table 5.

We may note that we can evaluate the gradient using the derivative of the Hamiltonian operator only (the Hellmann–Feynman force) *provided* the basis set is modified to properly describe the response of the original basis set to the perturbation in the nuclear positions. As discussed above, one approach is to allow the basis functions to float off the nuclei.^{91,93} The other is to augment the basis.^{94,95} The perturbation operator is, say, $\partial/\partial A_x$, whose effect on a Gaussian basis function $x_A^l y_A^m z_A^n \exp(-ar_A^2)$ is

$$\begin{aligned} \frac{\partial}{\partial A_x} x_A^l y_A^m z_A^n \exp(-ar_A^2) &= 2ax_A^{l+1} y_A^m z_A^n \exp(-ar_A^2) \\ &\quad - lx_A^{l-1} y_A^m z_A^n \exp(-ar_A^2). \end{aligned} \quad (32)$$

Thus basis functions corresponding to the derivatives of the original basis functions must be added to the basis. These will have the same exponents as the original functions, but different angular quantum numbers. In this way, the extended basis set will be able to describe the effects of the perturbation correctly.

6.3. Electric Properties

The simplest electric property is the dipole moment: the first derivative with respect to an applied electric field. In general, our basis sets contain no explicit dependence on an applied field, which has the consequence that the SCF or MCSCF (but not CI or most other correlated methods) dipole moment, for example, can be evaluated either as the energy derivative or, more economically, as the expectation value of the dipole operator. However, this lack of dependence of the basis set on an applied field means that the only way the response of the original basis set to a field can be described is by augmenting it. Otherwise, the response of the basis set to a field is likely to be described only poorly, and we might expect that the dipole moment results from our energy-optimized basis sets would be poor, and would not show very satisfactory convergence behavior with respect to the basis set. This is exactly what is observed in most calculations. Without suitable augmentation of the basis, dipole moment results are not satisfactory, and their “convergence” with respect to a sequence of larger and larger energy-optimized basis sets is erratic.

How should energy-optimized basis sets be augmented or modified for dipole moment calculations? We note first that the form of the operator is $F\mathbf{r}$, where F here is a scalar field strength. The basis set response thus corresponds to multiplying the basis functions by \mathbf{r} , which effectively raises the angular quantum number of each basis function. However, it is not necessary to extend a given basis by incrementing the angular quantum numbers of all basis functions. In general, external fields will be small compared to the internal fields generated by the nuclei and other electrons, and we can expect that the outer regions of the electron density will respond most to the external field. It will therefore be most important to include the response of those basis functions that describe this region: the most diffuse basis functions. For example, d functions with small exponents will be required to allow the outer part of the valence p orbitals to respond to the field.

This general philosophy can be applied to other properties. For instance, dipole polarizabilities are determined by response to an electric field (this time as the second derivative of the energy), while quadrupole polarizabilities are determined by response to an external field gradient. The latter operator results in incrementing the angular quantum numbers of the basis functions by two, introducing diffuse f functions to describe the response

of valence p functions, for example. The same arguments are relevant to properties related to *internal* fields, such as the electric field gradient at the nuclei. This operator has the general form $Y_{2m}r^{-3}$, that is, d -like behavior and a strong weighting towards the nuclear center. The inner shell contributions tend to average to zero because of the influence of the nuclear attraction itself, so the main effect is to increment the angular quantum numbers of the functions that describe the valence region (predominantly the inner valence region) by two. In fact, such functions will usually be similar to the higher angular functions that must be added to describe correlation effects, so basis sets that yield good correlation energies will be fairly well adapted to calculating internal electric field gradients.⁷⁴ It is the response to external fields that most requires extension of typical energy-optimized basis sets, and we illustrate this with calculations on H₂O, again performed using SIRIUS/ABACUS.⁴⁷

Before discussing the results of Table 7, we should make some general observations. First, there is good evidence that the extension of basis sets to improve electric properties can be carried out entirely at the SCF level, since the inclusion of correlation effects does not seem to introduce any new problems.^{96–98} We have therefore presented only SCF results here. Second, the calculations have been performed at the optimum (SCF) geometries for each individual basis. In fact, only for extension of the STO-3G set is there any significant difference between the optimum geometry in the underlying energy-optimized set and the extended set, so the effects of this can be neglected in our comparisons.

We may first observe that all basis sets considered in Table 3 appear also in Table 7. The crucial need to extend *all* of these basis sets to avoid significant errors in the polarizabilities, at least, is obvious in Table 7. (Once again, we may regard the final entry in the table, an “augmented cc-pVQZ” basis discussed below, as representing near Hartree–Fock quality results). Other than the original cc-pVQZ basis, all of the energy-optimized sets yield α_{yy} values^c in error by more than 25%, and even the cc-pVQZ set result is in error by 15%. The convergence of the other polarizability tensor components is better, but still poor: the α_{xx} component converges faster than α_{zz} . Basis set convergence of the dipole moment is better than for the polarizability, but at least one d function is required for acceptable values, in line with our reasoning above.

^cThe molecule lies in the xz plane with the z axis as the C_2 rotation axis.

Table 7. SCF dipole moments and polarizabilities for H₂O.

Basis	<i>E</i> _{SCF} (<i>E</i> _h)	μ (D)	α_{xx} (a_0^3)	α_{yy}	α_{zz}
<i>sp/s</i> basis sets					
STO-3G	-74.965901	1.71	5.51	0.04	2.57
DZ (7s 3p)	-75.878683	2.36	7.08	1.10	3.90
4-31G	-75.895492	2.39	6.81	1.58	4.01
VDZ (7s 4p)	-75.916744	2.44	6.61	1.79	3.89
6-31G	-75.972623	2.40	6.79	1.69	4.05
DZ (9s 5p)	-76.002237	2.46	6.80	2.03	4.29
<i>sp/s</i> basis sets + extensions					
STO-3G ^a	-75.072127	2.30	11.12	9.56	10.45
6-31G ^a	-75.975714	2.31	7.41	3.31	5.25
<i>spd/sp</i> basis sets					
VDZP (7s 4p)	-75.961664	2.07	6.74	3.35	5.04
6-31G*	-75.998968	2.11	6.71	2.97	4.91
6-31G**	-76.015308	2.06	6.73	3.12	4.95
cc-pVQZ	-76.027389	2.05	6.72	3.07	5.01
DZP (9s 5p)	-76.042226	2.09	6.96	3.62	5.40
TZ2P	-76.056174	2.03	7.84	5.74	6.64
<i>spd/sp</i> basis sets + extensions					
6-31G*** ^a	-76.027939	2.10	8.53	6.82	7.68
DZP (9s 5p) ^a	-76.046963	2.05	8.54	6.77	7.71
DZP (9s 5p) ^b	-76.047568	1.96	8.41	6.18	7.27
DZP (9s 5p) ^c	-76.050625	2.00	8.62	7.22	7.78
<i>spdf/sp</i> and <i>spdf/spd</i> basis sets					
TZ2Pf	-76.056971	2.00	7.85	5.76	6.67
cc-pVTZ	-76.058329	2.03	7.78	5.35	6.57
cc-pVQZ-g	-76.065664	1.97	8.22	6.40	7.29
<i>spdf/sp</i> and <i>spdf/spd</i> basis sets + extensions					
TZ2Pf ^c	-76.058307	1.95	8.80	7.73	8.15
aug-cc-pVTZ	-76.061740	1.94	8.89	7.71	8.20
aug-cc-pVQZ-g	-76.066748	1.94	8.87	7.80	8.26

^aAugmented with Spackman diffuse *d* on O and *p* on H (see text).^bAugmented with other diffuse *d* on O and *p* on H (see text).^cAugmented with diffuse *spd* on O and *sp* on H (see text).

Consider now those values obtained with basis sets augmented with additional diffuse functions. For the smaller sets, a d exponent of 0.07 on O and a p exponent of 0.08 on H was used. These correspond to averages of values optimized by Spackman⁹⁹ for use in augmenting the 6-31G set for calculating polarizabilities, and are about one-tenth or less of the values that would typically be used as SCF polarization functions. Even with these additional functions, the STO-3G basis is not satisfactory (the substantial overshooting of the Hartree–Fock limit values arises because both the bond length and angle in the extended STO-3G basis are much too large). The improvement in the 6-31G basis values on extending the basis is also only modest, and there is still a large error in α_{yy} . A much greater improvement is seen when the diffuse functions are added to a set that already contains polarization functions, as can be seen from the extended 6-31G** and extended DZP results. We have also investigated an alternative choice of exponent for the diffuse functions added to the DZP set, taking one-third of the usual polarization exponents. The results are not greatly different from those with Spackman’s exponents — previous work⁹⁶ has suggested that using one-third of the usual values is close to optimum for the dipole moment but is somewhat too large for the polarizability, and this is consistent with the results of Table 7.

In addition to adding diffuse polarization functions, the polarizability, at least, can be improved by adding diffuse sp functions on O and a diffuse s function on H, to improve the description of the outer part of the unperturbed and perturbed electron density. The exponents used were again chosen by “even-tempering,” taking one-third of the most diffuse exponents in the original DZP basis to define the diffuse functions added. These results are the best of any spd/sp basis in Table 7. Better results can be obtained with sets that contain higher angular functions. The extended TZ2Pf basis incorporates diffuse spd functions on O and diffuse sp functions on H, chosen according to the same even-tempered criterion as used to extend the DZP basis. Finally, the cc sets are extended to the augmented sets developed by Kendall *et al.*¹⁰⁰ for describing negative ions. The diffuse functions needed to describe the latter appear to do an excellent job for polarizabilities and dipole moments.

We have not listed results for dipole gradients here, but they parallel the results for polarizabilities, consistent with our analysis in Sec. 6.1.

6.4. Magnetic Properties

In general, there has been less experience with the computation of magnetic properties of molecules than with electric properties. This is partly because such properties are often more complicated — involving for example imaginary and triplet perturbations. Also, properties involving an external magnetic field and a finite basis set are plagued by the gauge-origin problem. Unless this problem is solved, the calculated results will depend on the choice of the gauge origin for the vector potential representing the external field. This dependence is unphysical and may lead to meaningless results. Several methods have been proposed to deal with this problem. The most satisfactory solution is to use London atomic orbitals, also referred to as *gauge-invariant atomic orbitals*¹⁰¹ (GIAO's) (see also Ref. 102). London orbitals are obtained from standard orbitals by attaching a complex phase factor that depends on the external vector potential, see Sec. 14.2. Calculations using London orbitals are strictly gauge origin independent. Furthermore, because of the attached phase factors, the London orbitals behave in a physically reasonable way in an external magnetic field. Hence the use of London orbitals should result in calculations that are fairly robust with respect to modifications of the basis set.

The use of London orbitals is not the only way to handle the gauge invariance problem. Popular alternatives include the individual gauge for localized orbitals (IGLO) approach of Kutzelnigg and coworkers,^{103,104} and the localized orbitals/local origin (LORG) approach of Hansen and Bouman.^{105,106} For vibrational circular dichroism, Stephens has proposed the distributed origins (DO) method.¹⁰⁷ It should be realized that the basis set requirements for all these methods are different. In general, the London orbital method converges faster than the above methods,^{102,108–110} and in the following we discuss results obtained using London orbitals only.

We have calculated several magnetic properties for water using the same basis sets as for electric properties discussed above: no attempt has been made to tailor these sets for magnetic properties. In Table 8 the SCF results obtained with SIRIUS/ABACUS⁴⁷ are listed together with the experimental values.

The magnetic susceptibility (χ) represents the second derivative of the electronic energy with respect to an external magnetic field. Since the response of the orbitals to this field is directly incorporated into the London orbitals, we expect that the calculated susceptibilities should be fairly

stable with respect to the basis set. This is confirmed by our calculations. The sp/s basis sets are clearly inadequate, giving errors of about 5% from the basis set limit. The energy-optimized spd/sp basis sets offer some

Table 8. SCF magnetic susceptibility, nuclear magnetic shieldings, and OH spin–spin coupling constant for H_2O .

Basis	$\chi (10^{-30} \text{JT}^{-2})$	$\sigma(\text{H}) (\text{ppm})$	$\sigma(\text{O}) (\text{ppm})$	$J(\text{OH}) (\text{Hz})$
<i>sp/s</i> basis sets				
STO-3G	-194.7	32.82	347.3	+7.7
DZ ($7s\ 3p$)	-212.6	32.07	341.3	-37.3
4-31G	-215.3	32.10	325.7	-67.2
VDZ ($7s\ 4p$)	-218.3	32.27	327.5	-119.1
6-31G	-216.9	32.18	325.2	-75.8
DZ ($9s\ 5p$)	-223.5	32.08	315.4	-72.4
<i>sp/s</i> basis sets + extensions				
STO-3G ^a	-236.0	30.83	365.8	-38.9
6-31G ^a	-218.9	32.28	327.8	-78.0
<i>spd/sp</i> basis sets				
VDZP ($7s\ 4p$)	-219.2	31.51	354.7	-111.9
6-31G*	-217.2	31.81	331.6	-71.6
6-31G**	-217.2	31.28	347.5	-74.4
cc-pVDZ	-218.1	31.80	353.4	-76.9
DZP ($9s\ 5p$)	-223.5	31.26	341.8	-78.0
TZ2P	-227.0	31.63	340.1	-111.5
<i>spd/sp</i> basis sets + extensions				
6-31G*** ^a	-230.9	31.02	338.2	-83.1
DZP ($9s\ 5p$) ^a	-230.7	31.19	339.8	-82.4
DZP ($9s\ 5p$) ^b	-226.0	31.65	342.0	-85.9
DZP ($9s\ 5p$) ^c	-231.1	31.61	341.5	-90.3
<i>spdf/sp</i> and <i>spdf/spd</i> basis sets				
TZ2Pf	-227.2	31.53	344.3	-110.8
cc-pVTZ	-224.5	31.44	343.3	-79.3
cc-pVQZ-g	-227.1	31.23	340.3	-103.0

^aAugmented with Spackman diffuse *d* on O and *p* on H.

^bAugmented with other diffuse *d* on O and *p* on H.

^cAugmented with diffuse *spd* on O and *sp* on H.

Table 8. (Continued)

Basis	χ (10^{-30} JT^{-2})	$\sigma(\text{H})$ (ppm)	$\sigma(\text{O})$ (ppm)	$J(\text{OH})$ (Hz)
<i>spd/spd and spd/sp basis sets + extensions</i>				
aug-cc-pVTZ	-229.9	31.32	336.9	-90.9
TZ2Pf ^c	-229.9	31.52	334.9	-108.6
aug-cc-pVQZ-g	-229.7	31.19	336.9	-103.5
Experiment ^d	-218 ± 30	30.052 ± 0.015	344.0 ± 17.2	$\pm 73.5 \pm 2.1$

^cAugmented with diffuse *spd* on O and *sp* on H.

^dRef. 111 for χ , Ref. 112 for $\sigma(\text{H})$, Ref. 113 for $\sigma(\text{O})$, Ref. 114 for $J(\text{OH})$.

improvement, but satisfactory convergence is only obtained by adding diffuse orbitals. The calculations at the extended *spd/sp* level are only 1% from the Hartree-Fock limit. The importance of diffuse basis functions is not surprising, since the outer valence region is most strongly affected by an external field, as already observed for electric properties. The need for diffuse functions notwithstanding, it should be noted that the magnetic susceptibility converges much faster than the polarizability (Table 7), since the orbitals depend explicitly on the magnetic field but are independent of the electric field.

The computed shieldings display much the same behavior. The calculations at the *sp/s* level are about 3% from the Hartree-Fock limit — the hydrogen shieldings are overestimated and the oxygen shielding underestimated. The hydrogen shieldings are much improved by adding polarization functions and are in error by less than 1% at the *spd/sp* level, but little improvement is seen for oxygen at this level. Again we find that the extended *spd/sp* basis sets yield reasonable numbers; both the hydrogen and oxygen shieldings at this level are less than 1% from the Hartree-Fock limit.

The indirect spin-spin coupling constant $J(\text{OH})$ is much harder to converge. This property does not involve an external magnetic field, and the calculations are therefore unaffected by the use of London orbitals. The convergence is clearly erratic. The *sp/s* results are completely unreliable and the STO-3G basis set gives the wrong sign. Some improvement is observed at the *spd/sp* level, but the numbers are still in error by about 20% (assuming a Hartree-Fock limit of about -100 Hz). The addition of diffuse functions improves the situation slightly (errors about 15%), but even calculations using higher angular functions are not properly converged. We

should also note here that electron correlation is very significant for this quantity, the experimental value being $\pm 73.5 \pm 2.1$ Hz (that is, the sign is not known from experiment).

The calculation of spin–spin coupling constants clearly requires extension of energy-optimized basis sets beyond the addition of diffuse functions (see e.g. Ref. 115). Since the perturbation operators (the Fermi contact, spin–dipole and orbital hyperfine operators) mainly affect electrons close to the nuclei, it is reasonable to extend the basis sets such that the description of the nuclear regions is improved. In segmented basis sets, this may be achieved by uncontracting an additional primitive orbital from the inner-shell CGTO. We may recall that the reason for contracting the core and inner valence in the first place was that they were unaffected by molecule formation: this approach to contraction is harder to justify when we deal with perturbations that affect the electronic structure close to the nuclei.

We have therefore recalculated the magnetic properties with one more set of uncontracted *s* and *p* orbitals on oxygen, and an additional uncontracted *s* on hydrogen (keeping the number of primitive functions constant). This results in significant improvements for the spin–spin coupling constant. We do not list the results, but the mean errors in $J(\text{OH})$ are reduced from about 15% to 10% at the extended *spd/sp* level, and from 11% to only 3% at the *spdf/sp* or *spdf/spd* levels. For shieldings, the improvements are more modest, but the convergence to the basis set limit is smoother. As expected, no improvements are observed for magnetic susceptibilities, which do not depend critically on the electronic structure in the nuclear regions.

We conclude that molecular magnetic susceptibilities and nuclear magnetic shieldings are reasonably well behaved, and that *spd/sp* basis sets may be sufficient for first-row atoms, provided that they are extended by diffuse functions. The spin–spin couplings clearly require extensive basis sets and demonstrate that even large energy-optimized basis sets do not necessarily yield satisfactory properties.

6.5. Basis Set Recommendations

In view of our discussions here, the basis set recommendations for property calculations can be formulated rather simply. For properties in which the basis set response to the perturbation is explicitly incorporated into the computational approach (nuclear displacements, magnetic field properties

with London orbitals), we can expect the basis set convergence to be similar to the energy itself. Thus geometries and force constants will converge fairly rapidly using energy-optimized basis sets alone, and this behavior is clear in Tables 3 and 4. The convergence of diamagnetic susceptibilities and nuclear magnetic shieldings is very similar. If the perturbation response is not included explicitly, the basis set must be extended to account for it. For electric field properties, this will involve the addition of diffuse functions. Ideally, a diffuse function for each symmetry occupied in the atom, and one degree higher for a homogeneous field, two degrees higher for a field gradient, etc., should be added, but a diffuse function of the highest angular type according to this rule is indispensable. The convergence of electric field properties in energy-optimized basis sets without diffuse extensions is poor, with the polarizability converging even slower than the dipole moment. Internal field properties, such as electric field gradients at nuclei, are often fairly well described by energy-optimized basis sets provided the angular quantum number is high enough.

We have not explicitly considered properties that are third- or higher-order derivatives of the energy here. These can be analyzed according to the same principles. Thus we may expect that cubic or quartic force constants will display similar basis set convergence to the energy, whereas “electric anharmonicity” or Raman intensities, which involve one or two levels of differentiation with respect to an applied electric field, will converge similarly to polarizabilities. Hyperpolarizabilities, which involve third or higher derivatives with respect to an applied field, can be expected to display poor convergence with respect to the basis set, and indeed, multiple diffuse high angular momentum functions are required to obtain reliable values for these quantities in atoms and small molecules (see Ref. 98 for an extensive basis set study).

7. Basis Set Superposition Error

One of the consequences of using finite, incomplete basis sets is that the description of an atom or molecular fragment will be improved by basis functions on other atoms or fragments in the molecule. Since the energy of each fragment described in the full basis of the molecule will be lower than in the fragment basis alone, there will be a spurious attractive interaction between fragments that arises only from improving the individual fragment descriptions. This is termed *basis set superposition error* (BSSE). It is

important to realize that BSSE appears in every calculation performed with finite basis sets. The question is then to what extent it needs to be analyzed and corrected for.

7.1. Counterpoise Correction

One way to estimate the BSSE in a calculation is via a prescription suggested by Boys and Bernardi.¹¹⁶ Suppose we are interested in the interaction energy between fragments *A* and *B* in the molecule *AB*. Naively, this would be computed as

$$\Delta E = E_{AB} - E_A - E_B, \quad (33)$$

where *A* is described in its own basis set and similarly for *B*. Of course, since the basis set on *A* is incomplete, the energy of *A* will be improved in the molecule by the basis set on *B*. Thus a better estimate of the interaction energy would be

$$\Delta E = E_{AB} - E_A(AB) - E_B(AB), \quad (34)$$

where the notation implies that *all three* calculations: molecule, fragment *A*, and fragment *B*, are performed in the full basis set that is used for *AB*. We note that since in a finite basis $E_A(AB) < E_A$ and $E_B(AB) < E_B$, ΔE from Eq. (34) will be less than the value from Eq. (33). The use of Eq. (34) is termed the *counterpoise correction*. One sometimes speaks of including the basis functions of *B* in the calculation on *A* as employing a “ghost basis.”

Under what circumstances is it desirable to compute counterpoise-corrected interaction energies? It should be obvious that the smaller the interaction energy, the more we must be concerned with BSSE. The classic example is the helium dimer, for which even very extended basis sets have BSSE’s larger than the interaction energy itself.¹¹⁷ Of course, this is for an interaction energy of only 10 K, and unfortunately BSSE is frequently perceived as a problem that need be considered only for such weak interactions. As a counterexample from the other end of the scale, consider the binding energy of the molecule N₂, which experimentally is of the order of 228 kcal/mol. In large ANO basis sets, the computed binding energy is within 3 kcal/mol of experiment, but the BSSE contribution to this binding energy can be almost 1 kcal/mol (Ref. 84). Hence BSSE can be significant on the scale of 1 kcal/mol for chemical binding energies.

It therefore appears that BSSE should at least be estimated (using a counterpoise correction) if there is any pretence to accurately estimating binding energies. Whether the BSSE-corrected binding energy is more reliable an estimate than the directly computed value can be debated as a philosophical issue; the more important point is to provide the BSSE so that a reader can judge the quality of the calculation and the degree of basis set incompleteness. The BSSE has been used empirically to provide an estimate for basis set incompleteness in very large basis sets: instead of subtracting the counterpoise correction, it is scaled by a factor of 1.5 and then *added* to the computed interaction energy to obtain an estimated “basis set limit” value.¹¹⁸

Another issue concerning counterpoise corrections is that of whether the ghost basis should comprise all the basis functions from the other fragment(s). After the question of whether to account for BSSE at all, this is the second most contentious issue related to BSSE. It has been argued repeatedly that in computing the counterpoise correction to E_A , the occupied orbitals on B should be excluded from the ghost basis since they are not available in AB because of the Pauli exclusion principle. Empirically, the convergence of the BSSE estimated in this way with improvements in the basis is more erratic than when the full ghost basis is used. More formally, this question has been addressed in great detail by Liu and McLean,¹¹⁷ building on earlier work by Gutowski *et al.*¹¹⁹ Their conclusion is that the full ghost basis is the appropriate formal choice for estimating the counterpoise correction.

7.2. Basis Set Improvements and BSSE

As we have defined it, BSSE is a consequence of basis set incompleteness. Presumably, as the basis set approaches completeness, BSSE will diminish and eventually vanish, and we may ask whether by suitable improvements and modifications to the basis set we can eliminate BSSE, or make it negligible, in practical calculations with finite basis sets. It is convenient to consider this separately for the SCF energy and for the correlation energy.

We have seen in Sec. 4 that convergence of the SCF energy with respect to the basis set is fairly rapid. Basis sets with $2d1f$ polarization sets are already close to the Hartree–Fock limit. Similarly, many properties are well converged at this level, and, with some extension of the basis, satisfactory results for multipole moments and polarizabilities can be obtained (Sec. 6).

We can thus expect that extended basis sets that provide a good description of the fragments and their properties should result in rather small BSSE's. However, calculations in smaller basis sets may need attention, especially if sets that provide a rather poor description of the fragment energies are augmented with diffuse functions to improve the polarizabilities. Diffuse functions in the ghost basis will obviously be more effective at improving fragment energies than more localized basis functions.

As we have reiterated throughout this work, the convergence of correlation energies with respect to basis set is slow. This suggests that BSSE will be slower to diminish with basis set improvements for the correlation energy than for the SCF energy, and this is what is observed. For rather weak interactions, this makes some attempt to estimate the BSSE mandatory. We should note that this is another situation in which a potential cancellation of errors (BSSE partially cancelling basis set incompleteness) makes comparison of directly computed interaction energies with experiment useless as calibration. The slow convergence of the correlation energy also argues against recommending a (necessarily limited) extension of the basis set as a better use of resources than performing counterpoise correction. It is very seldom that the former strategy will improve the reliability of the calculations more than the latter, unless there is a cancellation of errors.

A prerequisite for minimizing BSSE in correlated calculations is to obtain good fragment energies. For atomic fragments, there is then a considerable advantage in the use of ANO's, since they provide very accurate atomic energies. ANO basis sets do indeed yield small (although rarely negligible!) BSSE's (Refs. 39, 45 and 84). However, for small ANO sets this inevitably leads to a bias towards the fragments in the computed interaction energy, and these can be significantly underestimated. The small ANO sets provide a very good atomic description when a large primitive basis is used, so that BSSE is minimized. But the smaller the contracted set, the less flexible it will be for molecular calculations, so the molecular description will be poorer than the atomic description. This is a strong argument against the use of a [3s 2p 1d] ANO basis set for first-row atoms, and is why such a set was not advocated in the original ANO work.⁴⁵ The corresponding cc-pVDZ set⁴⁶ is less biased towards the atom and is more satisfactory for molecular use.

8. Basis Sets: Miscellaneous Topics

8.1. *Excited and Ionized States*

We have only briefly mentioned basis set requirements for other than ground-state systems so far. Valence-like excited states can be expected to have the same basis set requirements as the ground state. Rydberg excited states, on the other hand, pose a completely different problem. Here the need is to describe an extended, diffuse excited orbital, and the basis set must therefore be extended with diffuse functions of the appropriate angular momentum. We need hardly point out that in situations in which the nature of the desired excited state is not known in advance, the only safe strategy is to include diffuse functions, since a basis that describes the ground state well and a potential Rydberg state well will certainly be adequate for a valence excited state, while the converse is not true. Typical exponents for Rydberg orbitals are similar to those for molecular polarizabilities.

In positively charged ionic states, the electron density can be expected to contract compared to the neutral species. Most medium-sized or large basis sets developed for neutral atoms are quite capable of describing such ionized states of atoms or molecules, that is, they are flexible enough to describe the contraction of the charge cloud. On the other hand, negative ions, like Rydberg states, usually have charge distributions that are much more diffuse than for neutral states. Just as for Rydberg states, most basis sets must be supplemented with diffuse functions to properly describe negative ions (a recent detailed discussion is given by Kendall *et al.*¹⁰⁰). Exponents for the diffuse functions are fairly well given by an even-tempered extrapolation of the existing basis. Calculations on negative ions in which there is no extension of a neutral atom basis set should be viewed with suspicion.

8.2. *Core Correlation*

The H₂O examples presented in Sec. 5 all involved correlation of the eight valence electrons only, and all of the discussion and recommendations were also confined to valence correlation. For very accurate results for first-row atoms, however, we may have to explicitly account for correlation effects involving the 1s electrons. Further, for heavier elements the energy separation between the valence shell and the next inner shell is smaller than for

the first row, and we can expect that correlation involving these inner shell electrons will become more important. Core correlation is a rather difficult problem, and even the question of what is an appropriate treatment of the many-electron problem is not straightforward to answer.¹²⁰ Here we shall be concerned only with the basis set requirements.

In Sec. 2.4, we established that basis functions for describing correlation should have the same extent as the orbitals being correlated, but should have more radial and/or angular nodes. Most basis sets lack such functions when it comes to correlating core orbitals, since they are not required for SCF or valence-correlated calculations. It is true that some of the primitive p GTO's in larger basis sets for first-row atoms would be suitable for describing core correlation, but they are usually combined into CGTO's with weights that do not allow them to fulfill any role in describing core correlation. Describing L shell correlation in second-row atoms will require d functions with similar exponents to the $2p$ orbitals — again, such functions are not normally part of the d sets added to second-row sets to describe polarization or valence correlation. It is crucial to realize that most basis sets must be augmented to describe core correlation adequately. Including core electrons in the correlation treatment without extending the basis set is a waste of time, since the core correlation effects will not be properly described and may lead to incorrect predictions. Correlating inner-shell electrons without extension of the basis can also exacerbate BSSE problems.

The selection of exponents for functions to describe core correlation is not difficult¹²⁰ — the GTO exponents that describe the appropriate occupied inner-shell orbitals provide an immediate guide to the exponents of the correlating functions. For a large first-row primitive basis, such as van Duijneveldt's ($13s\ 8p$) set, the innermost p GTO's, which would usually be contracted, form an excellent set of $1s$ correlating functions. If higher angular functions are to be included, the usual factor of 1.2 for each step in l can be used to estimate the exponents. The problem is not so much in developing the basis, it is that the size of the resulting basis set rapidly becomes unmanageably large. Attempts to develop ANO sets capable of describing core and valence correlation have not been very successful, at least for light atoms, and the resulting sets are still very large.^{45,120} It is difficult to see how core correlation can be treated properly without the use of large basis sets.

Related to the issue of core correlation and inner shells is the proper description of relativistic effects in heavier systems. Relativistic quantum chemistry is beyond the scope of this work, but it is pertinent to note that the basis set requirements for full four-component electronic structure calculations dwarf those of nonrelativistic calculations. Several discussions are to be found in a volume edited by Wilson.¹²¹ There is now a considerable literature on the subject, to which the interested reader is referred. More approximate treatments, such as effective one-component formulations, have similar basis set requirements to their nonrelativistic kin.

8.3. Bond Functions

In this work, we have been exclusively concerned with developing basis sets for molecular calculations in which all the functions are centered on the various nuclei in the molecule. If the main aim is to obtain the most rapid convergence towards basis set completeness, it can be argued that there is no reason to restrict the functions to the nuclei. For example, an *s* function in the center of the bond in H₂ can serve as a polarization function, rather than using *p* functions on the nuclei. This should reduce the size of the basis set, and hence reduce the cost of the calculation. Such “bond function” polarization sets have their advantages and disadvantages.

The major advantage of bond functions is their apparent cost-effectiveness: lower angular functions can be used for polarization as compared to atom-centered functions. From the point of view of integral evaluation alone this is not necessarily an enormous gain, but any reduction in the number of CGTO's will reduce the work in the other steps of a calculation. The optimum exponents and positions for the bond functions are not immediately obvious, but a suitable range of test calculations could be used to define rules for both exponents and positions. A significant disadvantage is the potential for substantial BSSE. In the simplest case of a function added in the center of a bond and remaining at that point, it is obvious that the function will not contribute to the atoms at infinite separation, but could provide substantial energy lowering by superposition when the atoms are near the equilibrium bond length. Several good binding energy predictions based on bond function calculations have resulted from BSSE, rather than a better description of the molecule proper, as shown by Bauschlicher, for instance.¹²² This is a particular problem with correlated wave functions, since the convergence of the atomic energies themselves is

slow. Another disadvantage is related to the correlation cusp itself — as we have said, the convergence of the atomic correlation energy is $(l + \frac{1}{2})^{-4}$, and accurate correlation energy estimates thus require high angular momentum basis functions. It is difficult to believe that the need for such functions is not also present for accurate molecular predictions, and several analyses attest to the importance of high angular momentum atom-centered basis functions in molecular applications.^{69,84} This need cannot be fulfilled by bond functions alone. In this context, a rather complete analysis of using bond functions in conjunction with fairly large atom-centered basis sets to improve binding energy predictions has been presented by Martin and coworkers.^{123–125}

9. Molecular Integrals — Introduction

We shall now consider the evaluation of the one- and two-electron molecular integrals like those that appear in Eqs. (17) and (18), namely

$$O_{ab} = \langle \chi_a(r) | O(r) | \chi_b(r) \rangle \quad (35)$$

and

$$g_{abcd} = \left\langle \chi_a(r_1) \chi_b(r_1) \left| \frac{1}{r_{12}} \right| \chi_c(r_2) \chi_d(r_2) \right\rangle, \quad (36)$$

where $\chi_a(r)$ are atomic orbitals, $O(r)$ a one-electron operator and r_{12} the separation between the electrons. The following integrals are discussed in detail: one-electron overlap and multipole moment integrals, momentum and kinetic energy integrals, one-electron Coulomb potential, field, and field gradient integrals as well as two-electron Coulomb interaction integrals. We also consider the evaluation of integrals differentiated with respect to the coordinates of the orbital centers, since these are important for the evaluation of molecular geometrical derivatives, and some integrals involving London orbitals or spin-orbit operators that are related to differentiated integrals. For simplicity, in much of our exposition the atomic orbitals are assumed to be contracted Cartesian Gaussian orbitals, but we shall devote some attention to the use of spherical harmonic basis functions.

We have devoted little space to reviewing the history of methods for Gaussian integrals. The availability of two splendid reviews of the subject from Saunders^{4,5} makes such an effort largely redundant. In particular, Saunders' 1975 review⁴ covers the subject up to that time very well. But it

is valuable to trace briefly here some of the seminal contributions to integral evaluation methods, most of which directly bear on our presentation.

In the beginning was Boys,¹⁷ who derived formulae for the required integrals for electronic structure calculations over GTO's. In his work, all integrals were evaluated either in closed form, or in terms of one-dimensional integrals related to the gamma function, and this approach is still followed in many modern formulations. Detailed formulae for GTO integrals were also derived by Taketa and coworkers.²⁶ In the early 1970's a number of groups pointed out that enormous benefits arose from computing integrals over *shells* of functions^{126,127} (see also Ref. 128), that is, GTO's or CGTO's with the same exponents, contraction coefficients and center, but different angular quantum numbers. For example, if we have four *p* shells each on different centers, they will give rise to 81 (3^4) integrals. However, as we shall see, these 81 integrals have very many terms in common, and it is much more efficient, computationally, to compute these common terms once and use them to build all 81 integrals.

A radically different approach to GTO integral evaluation was introduced in 1976, when Dupuis, Rys, and King¹²⁹ showed that the integrals that could not be done in closed form could be computed by an exact numerical quadrature procedure. This approach was especially important for use with higher angular momentum functions, since it complemented a very different formulation for the cases of *s* and *p* functions only, developed by Pople and Hehre.¹²⁸ Another novel insight into GTO integral evaluation was provided by McMurchie and Davidson,¹³⁰ who employed expansions in Hermite functions to obtain efficient integral expressions. Hermite functions had already been suggested and used by Maksić and coworkers (see, for example, Ref. 131). Like Boys, McMurchie and Davidson related the non-closed form integrals to forms related to the gamma function.

More recently, a number of efficient recursive procedures have been introduced, driven primarily by work by Obara and Saika.¹³² Their recursive approach has many points of similarity with Boys' original formulation, which may give the impression that the wheel has come full circle. Perhaps the one certain fact about the methodology of GTO integral evaluation is that improvements to existing schemes, whether incremental or gross, have continually been made and will continue to be made (see, for instance, the proceedings of a 1990 workshop on Gaussian integral evaluation, reported in the *International Journal of Quantum Chemistry*, Vol. 40, issue 6).

In this review, we shall concentrate first on a general exposition that contains most of the mathematics as well as some details relevant to computer implementation. Our presentation follows closely the treatment of McMurchie and Davidson.¹³⁰ We shall then briefly consider other formulations and discuss some recent developments that are of particular computational importance. Finally, we shall review various aspects peripheral to integral evaluation itself, including the treatment of molecular symmetry, and the calculation of energy derivatives. We identify no particular "recommended method" for the evaluation of Gaussian molecular integrals in this work. No single method will be optimum in every situation, so that any method will be a compromise (unless one is willing to program a variety of different methods for different cases). Rapid evaluation of integrals is a high priority, but it should not come at the expense of the simplicity and flexibility that allows an existing program to be readily extended to compute new properties or differentiated integrals.

10. Charge Distributions

10.1. Properties of GTO's

For our purpose here, it is desirable to extend our previous notation somewhat. We write a Cartesian Gaussian as

$$G_{ijk}(\mathbf{r}, a, \mathbf{A}) = x_A^i y_A^j z_A^k \exp(-ar_A^2), \quad (37)$$

where a is the orbital exponent, \mathbf{r} the electronic coordinates, \mathbf{A} the origin of the Gaussian (which usually coincides with some nuclear position) and

$$\mathbf{r}_A = \mathbf{r} - \mathbf{A}. \quad (38)$$

The non-negative integers i , j , and k are loosely referred to as the angular quantum numbers. Cartesian Gaussians may be factorized:

$$G_{ijk}(\mathbf{r}, a, \mathbf{A}) = G_i(x, a, A_x)G_j(y, a, A_y)G_k(z, a, A_z), \quad (39)$$

where for example

$$G_i(x, a, A_x) = x_A^i \exp(-ax_A^2). \quad (40)$$

As we shall see, this factorization simplifies the calculation of integrals over Cartesian Gaussians significantly.

For a number of purposes, we shall need the differentiation property of the Gaussians. The first derivative (Eq. 32) is given by (omitting arguments for clarity)

$$\frac{\partial G_i}{\partial A_x} = -\frac{\partial G_i}{\partial x} = 2aG_{i+1} - iG_{i-1}. \quad (41)$$

The derivative depends on the exponent and is a linear combination of two undifferentiated Gaussians with incremented and decremented quantum numbers. We may obtain higher derivatives by further differentiation, but in practice recursion is more useful. Using Eq. (41), we find that

$$\frac{\partial^{q+1} G_i}{\partial A_x^{q+1}} = \left(\frac{\partial}{\partial A_x} \right)^q [2aG_{i+1} - iG_{i-1}] = 2a \frac{\partial^q G_{i+1}}{\partial A_x^q} - i \frac{\partial^q G_{i-1}}{\partial A_x^q}. \quad (42)$$

Using the notation

$$G_i^q = \frac{\partial^q G_i}{\partial A_x^q}, \quad (43)$$

we therefore have the following recurrence relation

$$G_i^{q+1} = 2aG_{i+1}^q - iG_{i-1}^q. \quad (44)$$

In this way, we may construct higher derivatives from those of lower order. This is useful since we normally need all derivatives up to a given order. We should also note the trivial recurrence

$$x_A G_i = G_{i+1}, \quad (45)$$

which is used on several occasions later.

10.2. Hermite Gaussian Functions

The Hermite Gaussians of exponent p and centered on \mathbf{P} are defined by

$$\Lambda_{tuv}(\mathbf{r}, p, \mathbf{P}) = (\partial/\partial P_x)^t (\partial/\partial P_y)^u (\partial/\partial P_z)^v \exp(-pr_P^2), \quad (46)$$

where

$$\mathbf{r}_p = \mathbf{r} - \mathbf{P}. \quad (47)$$

Like the GTO's, these functions are separable in factors such as

$$\Lambda_t(x, p, P_x) = (\partial/\partial P_x)^t \exp(-px_P^2); \quad (48)$$

they differ from the Cartesian Gaussians, Eq. (37), only in the polynomial factors, which are generated by differentiation. Like Cartesians, Hermite Gaussians are also symmetric or antisymmetric with respect to inversion through the origin, but their nodal structure is quite different. The number of nodes in the Hermite Gaussian Λ_t is t . This follows from their definition by differentiation: Λ_{t+1} has nodes where Λ_t has extrema because the gradient vanishes at extrema. Since Λ_0 has one extremum only and no nodes we conclude that Λ_1 has one node and therefore two extrema, that Λ_2 has two nodes and therefore three extrema, and so on.

It is possible to consider the use of Hermite Gaussians themselves as basis functions. Much work has been done in this direction by Maksić and coworkers, and the interested reader is referred to the original literature^{131,133} (see also Refs. 134–136). Here, however, we shall consider the use of Hermite Gaussians only as intermediates in the calculation of integrals over Cartesian Gaussians. Their usefulness stems from the fact that they are defined by differentiation, which leads to many simplifications when calculating integrals. The use of Hermite Gaussians as intermediates in molecular integral calculations was pioneered by McMurchie and Davidson,¹³⁰ although again some of expressions were derived by Maksić and coworkers.¹³¹ In this section, we derive some important relationships which are used repeatedly in the remainder of this chapter.

From the definition of Hermite Gaussians, Eq. (46), we obtain the simple differentiation formula

$$\frac{\partial \Lambda_t}{\partial P_x} = -\frac{\partial \Lambda_t}{\partial x} = \Lambda_{t+1}, \quad (49)$$

which may be compared with the slightly more complicated formula for Cartesian Gaussians, Eq. (41). We also need to know what happens when Hermite Gaussians are multiplied by x_P . We note that

$$\Lambda_{t+1} = (\partial/\partial P_x)^t \frac{\partial \Lambda_0}{\partial P_x} = 2p(\partial/\partial P_x)^t x_P \Lambda_0. \quad (50)$$

In order to move x_P to the left of the differential operators, we use the commutator

$$[(\partial/\partial P_x)^t, x_P] = -t(\partial/\partial P_x)^{t-1}. \quad (51)$$

Inserting this into Eq. (50), we obtain

$$\Lambda_{t+1} = 2p(x_P \Lambda_t - t \Lambda_{t-1}). \quad (52)$$

This gives us the following important recurrence relation for Hermite Gaussians

$$x_P \Lambda_t = \frac{1}{2p} \Lambda_{t+1} + t \Lambda_{t-1}. \quad (53)$$

Note that the corresponding formula for Cartesian Gaussians, Eq. (45), is slightly simpler. To establish the relationship between Hermite Gaussians and Hermite polynomials we introduce the polynomials H_t by the equation

$$\Lambda_t = p^{t/2} H_t(\sqrt{p}x_P) \exp(-px_P^2), \quad (54)$$

and insert this in the recursion formula for the Hermite Gaussians, Eq. (53). Rearranging, we find that the $H_t(\sqrt{p}x_P)$ fulfil the well-known recursion formula for Hermite polynomials:

$$H_{t+1}(\sqrt{p}x_P) = 2\sqrt{p}x_P H_t(\sqrt{p}x_P) - 2t H_{t-1}(\sqrt{p}x_P), \quad (55)$$

and since

$$H_0(\sqrt{p}x_P) = 1, \quad (56)$$

we have indeed established that $H_t(\sqrt{p}x_P)$ are the Hermite polynomials in $\sqrt{p}x_P$. Hence apart from a normalization constant, the Hermite Gaussians consist of Hermite polynomials multiplied by a Gaussian exponential term. It is therefore of some interest to compare with the so-called Hermite functions

$$\psi_n(x) = (\sqrt{\pi}2^n n!)^{-1/2} H_n(x) \exp(-x^2/2) \quad (57)$$

from harmonic oscillator theory.⁶ The Hermite Gaussians and Hermite functions differ by a factor of two in the exponents. Since the Hermite functions are orthonormal,

$$\int_{-\infty}^{+\infty} \psi_m(x) \psi_n(x) dx = \delta_{mn}, \quad (58)$$

we conclude that the Hermite Gaussians are not orthogonal. The integral over a *single* Gaussian may, however, be expressed in a simple form:

$$\begin{aligned} \int_{-\infty}^{+\infty} \Lambda_t(x) dx &= \int_{-\infty}^{+\infty} (\partial/\partial P_x)^t \exp(-px_P^2) dx \\ &= (\partial/\partial P_x)^t \int_{-\infty}^{+\infty} \exp(-px_P^2) dx. \end{aligned} \quad (59)$$

Since the integral on the right-hand side is independent of P_x , it vanishes when differentiated:

$$\int_{-\infty}^{+\infty} \Lambda_t(x) dx = \delta_{t0} \int_{-\infty}^{+\infty} \exp(-px_P^2) dx = \delta_{t0} \sqrt{\frac{\pi}{p}}. \quad (60)$$

As we shall see, Eq. (60) simplifies the calculation of many one-electron integrals. Incidentally, this equation may also be proved by considering Λ_t as the product of two Hermite functions of degree t and 0 and invoking the orthogonality of Hermite functions. Equation (59) illustrates an important technique which we shall use on several occasions: since the Hermite Gaussians are defined as derivatives, we may take their differential operators outside the integration sign. In this way, we turn integrals over Hermite Gaussians into differentiated integrals over “spherical” Gaussians, that is, functions with no angular terms.

10.3. Overlap Distributions

In molecular integrals, orbitals always occur in pairs. It is therefore important to investigate the properties of products of Cartesian Gaussians. We show here how such products may be expanded linearly in Hermite Gaussians positioned on the line segment connecting the two Cartesian Gaussians. Consider first the product of two s -type GTO's, centered on **A** and **B**, which may be written in terms of a single Gaussian located on the line segment connecting the two centers. This is the *Gaussian product rule*

$$\exp(-ax_A^2) \exp(-bx_B^2) = \exp(-qQ_x^2) \exp(-px_P^2), \quad (61)$$

where

$$pP_x = aA_x + bB_x, \quad (62)$$

$$Q_x = A_x - B_x, \quad (63)$$

$$p = a + b, \quad (64)$$

and

$$q = \frac{ab}{a+b}. \quad (65)$$

Note that only the last factor on the right-hand side of Eq. (61) depends on the electronic coordinate x . The first factor

$$K_{AB} = \exp(-qQ_x^2) \quad (66)$$

is constant, and is sometimes referred to as the “pre-exponential factor”: it will be small when the separation between the centers is large. The Gaussian product rule greatly simplifies the calculation of integrals. It is illustrated in Fig. 11 below for normalized Gaussians of exponents 1.0 and separation 2.0.

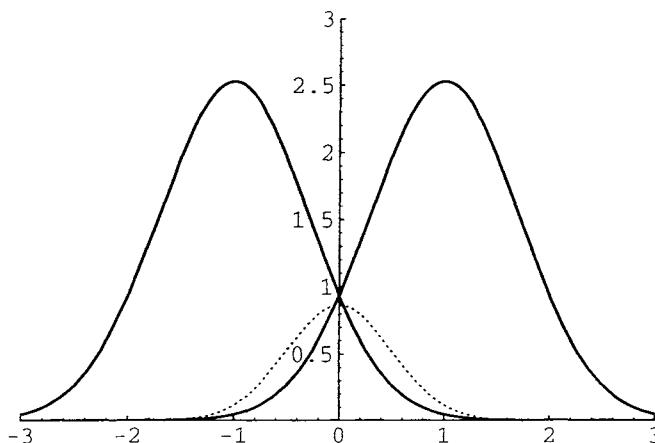


Fig. 11. The Gaussian product rule.

A product of two Gaussians on the same or different centers is referred to as an *overlap distribution*:

$$\Omega_{ij}(x, a, b, A_x, B_x) = G_i(x, a, A_x)G_j(x, b, B_x). \quad (67)$$

According to the Gaussian product rule Eq. (61), a Cartesian overlap distribution may be written in terms of a single Gaussian centered on \mathbf{P} ,

$$\Omega_{ij} = K_{AB}x_A^i x_B^j \exp(-px_P^2). \quad (68)$$

This expression is still not useful since it involves Cartesian monomials $x_A^i x_B^j$. One approach would be to expand this term in Cartesian polynomials about the center \mathbf{P} , and this approach was commonly used in early

schemes for GTO molecular integrals.^{4,26} However, we shall see in the next section that Hermite Gaussians are better suited for integration, and we shall therefore expand the Cartesian overlap distribution directly in Hermite Gaussians centered on \mathbf{P} . Such a strategy is possible because *any* polynomial of degree $i + j$ (as in Eq. (68)) may be expanded in Hermite polynomials of degree $t \leq i + j$ (Ref. 6). We therefore write

$$\Omega_{ij} = \sum_{t=0}^{i+j} E_t^{ij} \Lambda_t, \quad (69)$$

where the expansion coefficients E_t^{ij} are constant (that is, independent of the electronic coordinates). It is quite difficult to derive explicit expressions for these coefficients, but by using the properties of the Cartesian and Hermite Gaussians it is easy to arrive at a set of recurrence relations for them.^{5,130} We consider the incremented distribution

$$\Omega_{i+1,j} = \sum_{t=0}^{i+j+1} E_t^{i+1,j} \Lambda_t. \quad (70)$$

To relate this expansion to that of Ω_{ij} , we use the recurrence relation Eq. (45) for Cartesian Gaussians and obtain

$$\Omega_{i+1,j} = x_A \Omega_{ij} = x_P \Omega_{ij} - \frac{qQ_x}{a} \Omega_{ij}. \quad (71)$$

We expand the first term on the right-hand side in Hermite Gaussians, and eliminate x_P using the Hermite recurrence relation Eq. (53):

$$\begin{aligned} x_P \Omega_{ij} &= \sum_{t=0}^{i+j} E_t^{ij} \left[t \Lambda_{t-1} + \frac{1}{2p} \Lambda_{t+1} \right] \\ &= \sum_{t=0}^{i+j+1} \left[(t+1) E_{t+1}^{ij} + \frac{1}{2p} E_{t-1}^{ij} \right] \Lambda_t. \end{aligned} \quad (72)$$

The second term in Eq. (71) is expanded straightforwardly. Inserting these expansions into Eq. (71) and identifying terms of same order in t , we arrive at the desired recurrence relations. Repeating this for j , we obtain a starting coefficient of

$$E_0^{00} = K_{AB}, \quad (73)$$

and the relations

$$E_t^{i+1,j} = \frac{1}{2p} E_{t-1}^{ij} - \frac{qQ_x}{a} E_t^{ij} + (t+1) E_{t+1}^{ij}, \quad (74)$$

and

$$E_t^{i,j+1} = \frac{1}{2p} E_{t-1}^{ij} + \frac{qQ_x}{b} E_t^{ij} + (t+1) E_{t+1}^{ij} \quad (75)$$

From these expressions, the Hermite expansion coefficients are easily generated and the overlap distribution may then be expanded in Hermite Gaussians. The Hermite expansion coefficients are functions of \mathbf{Q} but are independent of \mathbf{P} . The following equations generate expansion coefficients differentiated with respect to Q_x :

$$E_0^{00;n+1} = -2q \left[Q_x E_0^{00;n} + n E_0^{00;n-1} \right], \quad (76)$$

$$E_t^{i+1,j;n} = \frac{1}{2p} E_{t-1}^{ij;n} - \frac{q}{a} \left[Q_x E_t^{ij;n} + n E_t^{ij;n-1} \right] + (t+1) E_{t+1}^{ij;n}, \quad (77)$$

and

$$E_t^{i,j+1;n} = \frac{1}{2p} E_{t-1}^{ij;n} + \frac{q}{b} \left[Q_x E_t^{ij;n} + n E_t^{ij;n-1} \right] + (t+1) E_{t+1}^{ij;n}, \quad (78)$$

where

$$E_t^{ij;n} = \frac{\partial^n E_t^{ij}}{\partial Q_x^n}. \quad (79)$$

Equation (76) is used to increment n , and is derived from Eq. (66) in the same way as the recursions for Hermite Gaussians, Eq. (53). The other recursions, Eqs. (77) and (78), are derived by differentiating the recursions for the undifferentiated coefficients Eqs. (74) and (75), and are identical to these except for an extra term in the middle. Differentiated and undifferentiated coefficients are therefore obtained in the same way.

10.4. Overlap Distributions by Recursion

We have seen how Cartesian overlap distributions may be calculated from Hermite Gaussians by expansion. The distributions may also be generated recursively. To see this, we introduce the auxiliary distributions

$$\Omega_{ij}^t = K_{AB} x_A^i x_B^j \Lambda_t(x_P), \quad (80)$$

which contain as special cases the Cartesian overlap distributions ($t = 0$) and the Hermite Gaussians ($i = j = 0$). To relate these by recursion, we first multiply the recurrence relation for Hermite Gaussians Eq. (53) by $K_{AB}x_A^i x_B^j$ to give

$$x_P \Omega_{ij}^t = \frac{1}{2p} \Omega_{ij}^{t+1} + t \Omega_{ij}^{t-1}. \quad (81)$$

Next we substitute

$$x_P = x_A + \frac{qQ_x}{a} = x_B - \frac{qQ_x}{b}, \quad (82)$$

and noting (cf. the recurrence relation Eq. (45)) that

$$x_A \Omega_{ij}^t = \Omega_{i+1,j}^t, \quad (83)$$

and

$$x_B \Omega_{ij}^t = \Omega_{i,j+1}^t, \quad (84)$$

we arrive at

$$\Omega_{i+1,j}^t = t \Omega_{ij}^{t-1} - \frac{qQ_x}{a} \Omega_{ij}^t + \frac{1}{2p} \Omega_{ij}^{t+1}, \quad (85)$$

and

$$\Omega_{i,j+1}^t = t \Omega_{ij}^{t-1} + \frac{qQ_x}{b} \Omega_{ij}^t + \frac{1}{2p} \Omega_{ij}^{t+1}. \quad (86)$$

These provide relations from which we may generate Cartesian overlap distributions from Hermite Gaussians recursively. We therefore have at our disposal two methods for obtaining Cartesian overlap distributions from Hermite Gaussians. One involves expansion, the other recursion. Recursion has the advantage of generating all distributions of a given total quantum number $i + j \leq n$ simultaneously. Thus, for example, when we generate Ω_{pp} , an overlap distribution over two p -type GTO's, we obtain the distributions Ω_{ss} , Ω_{sp} , and Ω_{ps} at no extra cost. Expansion is preferable when the distributions of lower quantum numbers $i + j < n$ are not needed.

11. Simple One-Electron Integrals

From a computational point of view integrals come in two classes: those that involve Coulomb interactions and those that do not. Non-Coulomb

integrals may be expressed in closed analytical form, whereas Coulomb integrals involve at some stage a numerical procedure. We treat non-Coulomb integrals first.

11.1. Multipole Moments

We consider a general Cartesian multipole moment integral

$$S_{ab}^{efg} = \langle G_a | x_C^e y_C^f z_C^g | G_b \rangle, \quad (87)$$

where

$$G_a = G_{ikm}(\mathbf{r}, a, \mathbf{A}), \quad (88)$$

$$G_b = G_{jln}(\mathbf{r}, b, \mathbf{B}), \quad (89)$$

and integration is over the coordinates of one electron. Special cases of the multipole integral are the overlap, dipole and quadrupole integrals.

Since both the operator and the orbitals are separable (when Cartesian GTO's are used) the integral may be calculated as

$$S_{ab}^{efg} = S_{ij}^e S_{kl}^f S_{mn}^g, \quad (90)$$

where, for example,

$$S_{ij}^e = \langle G_i | x_C^e | G_j \rangle. \quad (91)$$

We consider this component only since the other two are obtained in the same manner. Inserting the Hermite expansion of the overlap distribution Eq. (69) into Eq. (91), we obtain

$$S_{ij}^e = \sum_{t=0}^{i+j} E_t^{ij} \int_{-\infty}^{+\infty} x_C^e \Lambda_t dx. \quad (92)$$

The coefficients are calculated according to Eqs. (73)–(74), and we now discuss the Hermite moment integrals

$$M_t^e = \int_{-\infty}^{+\infty} x_C^e \Lambda_t dx. \quad (93)$$

Consider first the case when $e = 0$. From Sec. 10.2, we know that the integral then reduces to (see Eq. (60))

$$M_t^0 = \delta_{t0} \sqrt{\frac{\pi}{p}}. \quad (94)$$

Note that this integral vanishes for $t > 0$. The remaining integrals may be obtained by recursion. Incrementing e , we obtain

$$M_t^{e+1} = \int_{-\infty}^{+\infty} x_C^e x_C \Lambda_t dx. \quad (95)$$

The integrand is rewritten in the usual way (see Eq. (53))

$$x_C \Lambda_t = x_P \Lambda_t + X_{PC} \Lambda_t = \frac{1}{2p} \Lambda_{t+1} + t \Lambda_{t-1} + X_{PC} \Lambda_t. \quad (96)$$

Substituting this expression in the Hermite integral Eq. (95), we obtain

$$M_t^{e+1} = t M_{t-1}^e + X_{PC} M_t^e + \frac{1}{2p} M_{t+1}^e, \quad (97)$$

which allows us to generate higher-order moments from those of lower orders. Note that all integrals for which $t > e$ vanish. This is seen by partial integration of Eq. (93):

$$\begin{aligned} M_t^e &= \int_{-\infty}^{+\infty} x_C^e \left[\frac{\partial^t \exp(-px_P^2)}{\partial P_x^t} \right] dx \\ &= (-1)^t \int_{-\infty}^{+\infty} x_C^e \left[\frac{\partial^t \exp(-px_P^2)}{\partial x^t} \right] dx \\ &= \int_{-\infty}^{+\infty} \left(\frac{\partial^t x_C^e}{\partial x^t} \right) \exp(-px_P^2) dx = 0 \quad (t > e), \end{aligned} \quad (98)$$

since the derivative inside the last integral vanishes for $t > e$. The final expression for the Cartesian multipole moment therefore becomes

$$S_{ij}^e = \sum_{t=0}^{\min(i+j, e)} E_t^{ij} M_t^e, \quad (99)$$

where the Hermite expansion coefficients are calculated recursively from Eqs. (73)–(77) and the Hermite multipole moment integrals are calculated recursively from Eqs. (94) and (97). The multipole moments are thus easily generated from expansion coefficients and Hermite integrals.

Let us consider two examples. According to the above discussion, the x component of the overlap integral is simply

$$S_{ij}^0 = E_0^{ij} \sqrt{\frac{\pi}{p}}, \quad (100)$$

and the total integral is therefore

$$S_{ab}^{000} = E_0^{ij} E_0^{kl} E_0^{mn} \left(\frac{\pi}{p} \right)^{3/2} \quad (101)$$

If we assume *s* orbitals, this becomes (see Eq. (73))

$$S_{ab}^{000} = \exp(-qR_{AB}^2) \left(\frac{\pi}{p} \right)^{3/2}, \quad (102)$$

and we see that the overlap behaves like a Gaussian with respect to variations in the distance between the orbital centers. The *x* component of the dipole integral consists of two terms (the first vanishes when both orbitals are of *s* type):

$$S_{ij}^1 = \left[E_1^{ij} + X_{PC} E_0^{ij} \right] \sqrt{\frac{\pi}{p}}, \quad (103)$$

which is easily established from the recurrence relation of the Hermite multipole moment integrals Eq. (97).

Closely related to the multipole moment integrals are integrals of the form

$$\langle G_i | \exp(ik_x x_C) | G_j \rangle = \sum_{t=0}^{i+j} E_t^{ij} K_t^x, \quad (104)$$

where the Hermite integrals are given by

$$K_t^x = \int_{-\infty}^{+\infty} \exp(ik_x x_C) \Lambda_t dx. \quad (105)$$

Taking the differential operator from the Hermite Gaussian outside the integration, we obtain

$$\begin{aligned} K_t^x &= \left(\frac{\partial}{\partial P_x} \right)^t \int_{-\infty}^{+\infty} \exp(ik_x x_C) \exp(-px_P^2) dx \\ &= \left(\frac{\partial}{\partial P_x} \right)^t \exp(ik_x X_{PC}) \int_{-\infty}^{+\infty} \exp(ik_x x_P - px_P^2) dx. \end{aligned} \quad (106)$$

The integral in Eq. (106) is independent of P_x ,

$$\int_{-\infty}^{+\infty} \exp(ik_x x_P - px_P^2) dx = \sqrt{\frac{\pi}{p}} \exp\left(-\frac{k_x^2}{4p}\right), \quad (107)$$

and we therefore obtain for the Hermite integrals

$$K_t^x = \sqrt{\frac{\pi}{p}} \exp\left(ik_x X_{PC} - \frac{k_x^2}{4p}\right) (ik_x)^t. \quad (108)$$

The Cartesian integrals may therefore be calculated from the expression

$$\langle G_i | \exp(ik_x x_C) | G_j \rangle = \sqrt{\frac{\pi}{p}} \exp\left(ik_x X_{PC} - \frac{k_x^2}{4p}\right) \sum_{t=0}^{i+j} E_t^{ij} (ik_x)^t. \quad (109)$$

For vanishing k_x , this expression reduces to the overlap integral Eq. (100); to first order we recover the dipole moment integral Eq. (103) times ik_x .

11.2. Momentum and Kinetic Energy Integrals

We now consider one-electron integrals involving differential operators:

$$D_{ab}^{qrs} = \langle G_a | (\partial/\partial x)^q (\partial/\partial y)^r (\partial/\partial z)^s | G_b \rangle. \quad (110)$$

These may be factorized in the same way as the multipole moment integrals

$$D_{ab}^{qrs} = D_{ij}^q D_{kl}^r D_{mn}^s, \quad (111)$$

where for example

$$D_{ij}^q = \langle G_i | (\partial/\partial x)^q | G_j \rangle. \quad (112)$$

The differential operator generates a linear combination of undifferentiated Gaussians as discussed above, and the integral reduces to a combination of overlap integrals. From the differentiation formula Eq. (44), we obtain

$$D_{ij}^0 = S_{ij}^0, \quad (113)$$

and

$$D_{ij}^{q+1} = j D_{i,j-1}^q - 2b D_{i,j+1}^q, \quad (114)$$

from which differentiated integrals of arbitrary order may be calculated. The explicit formulae for the lowest orders are

$$D_{ij}^1 = j S_{i,j-1}^0 - 2b S_{i,j+1}^0, \quad (115)$$

$$D_{ij}^2 = j(j-1) S_{i,j-2}^0 - 2b(2j+1) S_{ij}^0 + 4b^2 S_{i,j+2}^0. \quad (116)$$

Alternatively, we note that

$$D_{ij}^q = (-\partial/\partial B_x)^q S_{ij}^0 = (\partial/\partial A_x)^q S_{ij}^0, \quad (117)$$

since (except for a change of sign) differentiation with respect to the electronic coordinates is equivalent to differentiation with respect to coordinates of the orbital centers. Therefore, the above integrals may be calculated as differentiated overlap integrals, as discussed in Sec. 14.1.

Once these basic integrals are available, we may calculate linear momentum and kinetic energy integrals

$$\mathbf{P}_{ab} = -i\langle G_a | \nabla | G_b \rangle, \quad (118)$$

$$T_{ab} = \left\langle G_a \left| -\frac{1}{2} \nabla^2 \right| G_b \right\rangle. \quad (119)$$

The x component of the linear momentum becomes

$$P_{ab}^x = -i D_{ij}^1 D_{kl}^0 D_{mn}^0, \quad (120)$$

and the kinetic energy integral is

$$T_{ab} = -\frac{1}{2} [D_{ij}^2 D_{kl}^0 D_{mn}^0 + D_{ij}^0 D_{kl}^2 D_{mn}^0 + D_{ij}^0 D_{kl}^0 D_{mn}^2]. \quad (121)$$

Combining the techniques for multipoles and linear momentum, we may also calculate the integrals for angular momentum

$$\mathbf{L}_{ab} = -i\langle G_a | \mathbf{r} \times \nabla | G_b \rangle. \quad (122)$$

For the z component, we find

$$L_{ab}^z = -i [S_{ij}^1 D_{kl}^1 S_{mn}^0 - D_{ij}^1 S_{kl}^1 S_{mn}^0], \quad (123)$$

and similarly for the other components. This completes our discussion of non-Coulomb integrals.

12. Coulomb Integrals

Coulomb interaction integrals cannot be expressed in closed analytical form.¹⁷ They may, however, be reduced to one-dimensional integrals whose

evaluation is relatively straightforward. We begin by considering the electrostatics of spherical Gaussian charge distributions. We show that the potential arising from a spherical Gaussian distribution and the interaction between such distributions are described by the so-called incomplete gamma function. We then show how Coulomb integrals over Hermite Gaussians may be obtained by differentiating the incomplete gamma function, and we derive a set of recursion formulae for obtaining these derivatives. Finally, we discuss how Coulomb integrals over Cartesian Gaussians may be obtained from integrals over Hermite Gaussians by expansion or recursion.

12.1. Electrostatics for Gaussian Charge Distributions

The Gaussian distribution of unit charge around \mathbf{P} is given by

$$\rho_p = \left(\frac{p}{\pi}\right)^{3/2} \exp(-pr_p^2). \quad (124)$$

We consider the electrostatic potential at \mathbf{C} due to such a charge distribution:

$$V_p(\mathbf{C}) = \int \frac{\rho_p}{r_C} d\tau, \quad (125)$$

(integration over all space) and the energy of interaction between two such distributions

$$V_{pp'} = \iint \frac{\rho_p(r_1)\rho_{p'}(r_2)}{r_{12}} d\tau_1 d\tau_2. \quad (126)$$

We expect the electric potential and the interaction energy to depend on the exponents and the separations R_{PC} and $R_{PP'}$. For large separations, they should reduce to the expressions for point charges.

We treat the Coulomb potential first. Unlike the one-electron integrals discussed in Sec. 11, the potential integrals cannot be factorized in the Cartesian directions. The complications arise from the presence of the inverse operator. Since the charge density is described by a Gaussian we also write the operator $1/r_C$ in terms of a Gaussian, using the identity

$$\int_{-\infty}^{+\infty} \exp(-r_C^2 t^2) dt = \frac{\sqrt{\pi}}{r_C}. \quad (127)$$

This is the key point in treating integrals involving Coulomb interactions and Gaussian orbitals. We obtain the following expression for the potential $V_P(\mathbf{C})$:

$$V_p(\mathbf{C}) = \frac{p^{3/2}}{\pi^2} \int \exp(-pr_P^2) \left[\int_{-\infty}^{+\infty} \exp(-t^2 r_C^2) dt \right] d\tau. \quad (128)$$

We now invoke the Gaussian product rule Eq. (61),

$$\begin{aligned} V_p(\mathbf{C}) &= \frac{p^{3/2}}{\pi^2} \int_{-\infty}^{+\infty} \left\{ \int \exp[-(p + t^2)r_S^2] d\tau \right\} \\ &\quad \times \exp\left(-\frac{pt^2}{p+t^2} R_{CP}^2\right) dt, \end{aligned} \quad (129)$$

where \mathbf{S} is some point on the line connecting \mathbf{P} and \mathbf{C} , and carry out the integration over all space, giving

$$V_p(\mathbf{C}) = \frac{p^{3/2}}{\sqrt{\pi}} \int_{-\infty}^{+\infty} (p + t^2)^{-3/2} \exp\left(-pR_{CP}^2 \frac{t^2}{p+t^2}\right) dt. \quad (130)$$

To simplify further, we make the substitution

$$u^2 = \frac{t^2}{p+t^2}, \quad (131)$$

and obtain the final expression

$$\int \frac{\rho_p}{r_C} d\tau = \sqrt{\frac{4p}{\pi}} \int_0^1 \exp(-pR_{CP}^2 u^2) du. \quad (132)$$

We have replaced integration over all space by a one-dimensional integration over a finite interval. The integral appearing on the right-hand side of Eq. (132) is of central importance in the calculation of Coulomb interaction integrals over Gaussian distributions. It belongs to a class of functions referred to as the *incomplete gamma functions*:

$$F_n(x) = \int_0^1 \exp(-xt^2) t^{2n} dt, \quad (133)$$

which are discussed in more detail in Sec. 12.2 below. The potential from the Gaussian charge distribution may therefore be written

$$\int \frac{\rho_p}{r_C} d\tau = \sqrt{\frac{4p}{\pi}} F_0(pR_{CP}^2), \quad (134)$$

in terms of the incomplete gamma function. We shall see that this function also arises when we treat the interaction between two Gaussian charge distributions.

The interaction between two charge distributions Eq. (126) may be calculated as the electrostatic energy of the second distribution in the potential due to the first distribution:

$$V_{pp'} = \int V_p(r) \rho_{p'}(r) d\tau. \quad (135)$$

Since we have expressed the potential due to the first distribution in terms of the incomplete gamma function Eq. (134), this interaction becomes

$$V_{pp'} = \sqrt{\frac{4p}{\pi}} \left(\frac{p'}{\pi}\right)^{3/2} \int F_0(pr_P^2) \exp(-p'r_{P'}^2) d\tau. \quad (136)$$

We now insert the definition of the incomplete gamma function Eq. (133), use the Gaussian product rule Eq. (61) once again, and integrate over all space. We obtain

$$V_{pp'} = \sqrt{\frac{4pp'}{\pi}} \int_0^1 \frac{p'}{(pt^2 + p')^{3/2}} \exp\left(-\frac{pp't^2 R_{PP'}^2}{pt^2 + p'}\right) dt. \quad (137)$$

If we now substitute

$$u^2 = \frac{p + p'}{pt^2 + p'} t^2 \quad (138)$$

we obtain the final result

$$V_{pp'} = \sqrt{\frac{4\alpha}{\pi}} F_0(\alpha R_{PP'}^2), \quad (139)$$

where α is the reduced exponent

$$\alpha = \frac{pp'}{p + p'}. \quad (140)$$

The interaction between two spherical Gaussian distributions may therefore also be expressed in terms of the incomplete gamma function.

To summarize, we have found that the following relationships hold for spherical Gaussian distributions of unit charge:

$$\int \frac{\rho_p}{r_C} d\tau = \sqrt{\frac{4p}{\pi}} F_0(p R_{PC}^2), \quad (141)$$

and

$$\iint \frac{\rho_p(r_1)\rho_{p'}(r_2)}{r_{12}} d\tau_1 d\tau_2 = \sqrt{\frac{4\alpha}{\pi}} F_0(\alpha R_{PP'}^2). \quad (142)$$

These equations replace the usual expressions $1/R_{PC}$ and $1/R_{PP}$ for point charges. Figure 12 below illustrates the potentials due to Gaussian distributions of exponent 2 and 0.4 compared to the potential from a point charge. At large distances both Gaussian potentials approach closely the point charge potential, at short distances they are drastically reduced. At the center of charge the potential is finite and the field vanishes.

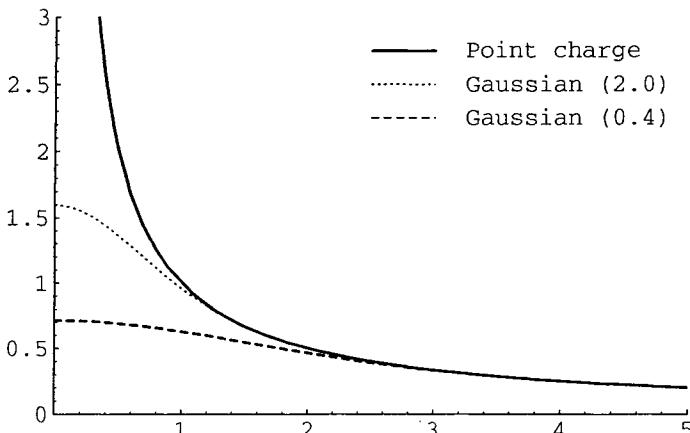


Fig. 12. Potentials due to a point charge and to two Gaussian charge distributions.

12.2. The Incomplete Gamma Function

We have seen that Coulomb interactions between Gaussian distributions are described by the incomplete gamma function. In this section, we discuss in more detail the properties and evaluation of this function. The incomplete gamma function is defined by

$$F_n(x) = \int_0^1 \exp(-xt^2) t^{2n} dt, \quad (143)$$

for $x \geq 0$. We note that it is a strictly positive function, since the integrand is positive:

$$F_n(x) > 0. \quad (144)$$

Since its derivatives are

$$\frac{dF_n}{dx} = -F_{n+1} < 0, \quad (145)$$

we conclude that the incomplete gamma function is a strictly decreasing function. We also note that

$$F_n(x) - F_{n+1}(x) = \int_0^1 \exp(-xt^2)t^{2n}(1-t^2)dt > 0, \quad (146)$$

which implies that

$$F_n(x) > F_{n+1}(x) \quad (147)$$

for all n . The values at $x = 0$ may be expressed in closed form

$$F_n(0) = \int_0^1 t^{2n} dt = \frac{1}{2n+1}, \quad (148)$$

which implies that

$$F_n(x) \leq \frac{1}{2n+1}. \quad (149)$$

For large values of x we may determine the incomplete gamma function approximately, since

$$F_n(x) = \int_0^1 \exp(-xt^2)t^{2n} dt \approx \int_0^\infty \exp(-xt^2)t^{2n} dt \text{ (}x \text{ large).} \quad (150)$$

This integral has the exact solution

$$F_n(x) \approx \frac{(2n-1)!!}{2^{n+1}} \sqrt{\frac{\pi}{x^{2n+1}}} \text{ (}x \text{ large),} \quad (151)$$

from which we conclude that the incomplete gamma function goes to zero as x goes to infinity. The properties discussed above are illustrated for $n \leq 3$ in Fig. 13.

In Sec. 12.1, we saw that the incomplete gamma function gives the potential from a Gaussian charge distribution at the origin according to the formula

$$V(r) = \sqrt{\frac{4p}{\pi}} F_0(pr^2). \quad (152)$$

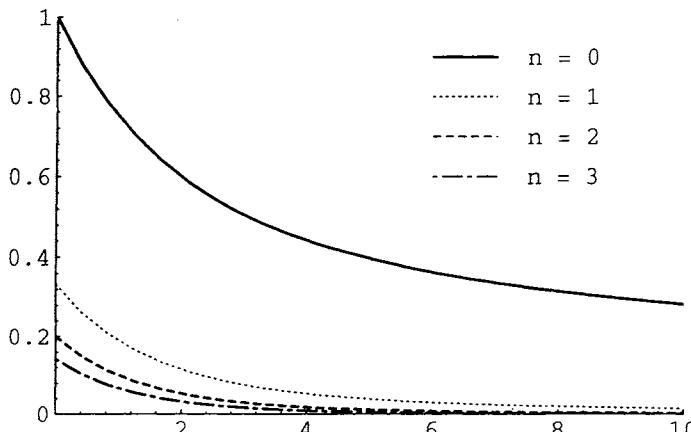


Fig. 13. $F_n(x)$ for $n = 0, 1, 2, 3$.

Inserting Eq. (151) into Eq. (152), we see that at large separations

$$V(r) \approx \frac{1}{r} \quad (r \text{ large}), \quad (153)$$

which is the potential from a point charge. At the center of the charge distribution the potential is

$$V(0) = 2\sqrt{\frac{p}{\pi}}, \quad (154)$$

which approaches infinity when p goes to infinity, that is, when the distribution approaches that of a point charge. We may also easily verify that the field and field gradient due to a Gaussian charge distribution are respectively

$$E(0) = 0 \quad (155)$$

and

$$E'(0) = \frac{4}{3}\sqrt{\frac{p^3}{\pi}} \quad (156)$$

at the center of the charge.

Before discussing the evaluation of the incomplete gamma function, we establish its relationship to the gamma function. Substituting $u = xt^2$, the incomplete gamma function Eq. (143) may be written

$$F_{n-1/2}(x) = \frac{1}{2x^n} \int_0^x \exp(-u) u^{n-1} du \quad (157)$$

if we allow non-integer values of n . The gamma function is defined by

$$\Gamma(n) = \int_0^\infty \exp(-t)t^{n-1}dt, \quad (158)$$

from which we conclude that

$$\Gamma(n) = 2 \lim_{x \rightarrow \infty} x^n F_{n-1/2}(x). \quad (159)$$

The incomplete gamma function Eq. (157), as a function of n rather than x , differs from the gamma function by a normalization constant and a finite rather than an infinite integration limit. We note that in the mathematical literature the functions

$$\gamma(n, x) = \int_0^x \exp(-t)t^{n-1}dt \quad (160)$$

and

$$\Gamma(n, x) = \int_x^\infty \exp(-t)t^{n-1}dt \quad (161)$$

are referred to as the incomplete gamma functions.^{137,138} Comparing Eqs. (157) and (160), we find that

$$\gamma(n, x) = 2x^n F_{n-1/2}(x) \quad (162)$$

Therefore, the term “incomplete gamma function” for the $F_n(x)$ of Eq. (143) in quantum chemistry is something of a misnomer.

The zeroth-order incomplete gamma function of quantum chemistry (Eq. (143)) is also closely related to the error function⁶

$$\text{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x \exp(-t^2)dt, \quad (163)$$

which is of importance in statistics. The relationship between Eq. (143) and the error function is seen to be

$$F_0(x) = \sqrt{\frac{\pi}{4x}} \text{erf}(\sqrt{x}), \quad (164)$$

by substitution of variables (see also Ref. (139)). Returning to Eqs. (141) and (142), we find that these equations may be cast in the simple form

$$\int \frac{\rho_p}{r_C} d\tau = \frac{\operatorname{erf}(\sqrt{p}R_{PC})}{R_{PC}}, \quad (165)$$

$$\iint \frac{\rho_p(r_1)\rho_{p'}(r_2)}{r_{12}} d\tau_1 d\tau_2 = \frac{\operatorname{erf}(\sqrt{\alpha}R_{PP'})}{R_{PP'}}, \quad (166)$$

which makes their relationship to Coulomb's law particularly clear. As the exponents p and α tend to infinity, the error functions here tend to one, since

$$\lim_{x \rightarrow \infty} \operatorname{erf}(x) = 1, \quad (167)$$

and we thus recover the usual expressions for point charges.

We now consider the calculation of the incomplete gamma function. Since this function is at the center of integral evaluation, it is important that it be calculated efficiently. Different methods of evaluating the incomplete gamma function have been suggested in the literature.^{4,130} We do not attempt to survey all methods here. Instead we discuss one possible approach.

We first note that, for large values of x , Eq. (151) provides a good estimate. For small values of x this approximation breaks down, but since we know both the function and its derivatives at $x = 0$ in closed form (Eq. (148)) we may construct a Taylor expansion for small values:

$$F_n(x) = \sum_{k=0}^{\infty} \frac{(-x)^k}{k!(2n+2k+1)}. \quad (168)$$

These two approximations, Eq. (151) for large values of x , and Eq. (168) truncated at $k = 6$ for small values, are illustrated for $F_0(x)$ in Fig. 14. While we clearly have a reasonable approximation to the incomplete gamma function by these two methods, they are inadequate for use in integral calculations, in which we would like errors of the order 10^{-10} or even smaller. This accuracy is attained only in the regions $x > 25$ and $x < 0.06$ using the above two approximations. More accurate methods are therefore needed in practice.

An alternative is to pretabulate the function at regular intervals for small arguments. During the calculation of the integrals, the incomplete

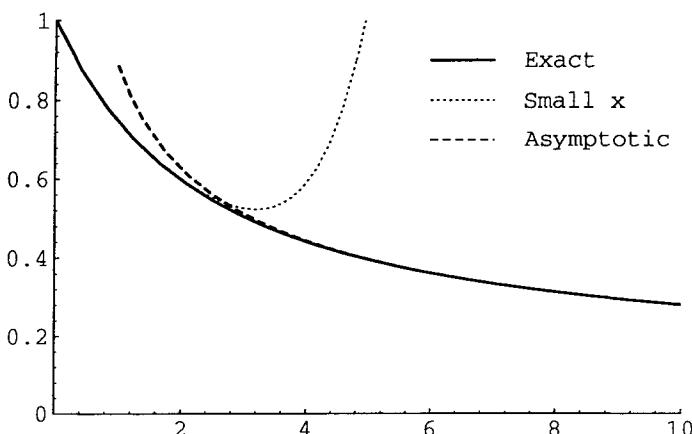


Fig. 14. Approximations for $F_0(x)$.

gamma function is expanded around the nearest tabulated point:

$$F_n(x_0 + \Delta x) = \sum_{k=0}^{\infty} \frac{F_{n+k}(x_0)(-\Delta x)^k}{k!}. \quad (169)$$

Using intervals of 0.1 satisfactory convergence is obtained after six terms. For large arguments, the asymptotic formula (Eq. (151)) is used.

The incomplete gamma functions of different orders n are related by recursion. Integrating the incomplete gamma function by parts, we obtain for upward recursion

$$F_{n+1}(x) = \frac{(2n+1)F_n(x) - \exp(-x)}{2x}, \quad (170)$$

and for downward recursion

$$F_n(x) = \frac{2xF_{n+1}(x) + \exp(-x)}{2n+1}. \quad (171)$$

We therefore only need to calculate the incomplete gamma function for the highest or the lowest order needed, obtaining the others by downward or upward recursion. For small values of x , upward recursion is unstable, since it involves the difference between two almost equal numbers, making downward recursion the preferred strategy.

From the expression for upward recursion Eq. (170), we may infer that the incomplete gamma functions may be written as an expression involving $\exp(-x)$ and the zeroth-order incomplete gamma function. The closed form for this expression is

$$F_n(x) = \frac{(2n-1)!!}{(2x)^n} \left[\sqrt{\frac{\pi}{4x}} \operatorname{erf}(\sqrt{x}) - \exp(-x) \sum_{k=1}^n \frac{(2x)^{k-1}}{(2k-1)!!} \right], \quad (172)$$

which — although not useful for calculations — neatly describes the relationship between the incomplete gamma functions of Eq. (143) and the error function of Eq. (163).

12.3. Hermite Integrals

We now go one step further and consider nonspherical electron distributions as described by Hermite Gaussians. The one-electron Coulomb integral is then

$$V_{tuv}^{qrs} = \int \Lambda_{tuv} (\partial/\partial C_x)^q (\partial/\partial C_y)^r (\partial/\partial C_z)^s r_C^{-1} d\tau, \quad (173)$$

where for example V_{tuv}^{000} is the potential and V_{tuv}^{100} , V_{tuv}^{010} , and V_{tuv}^{001} are the three components of the electric field. The two-electron Coulomb interaction integral is

$$V_{tuv; t'u'v'} = \iint \frac{\Lambda_{tuv}(r_1) \Lambda_{t'u'v'}(r_2)}{r_{12}} d\tau_1 d\tau_2, \quad (174)$$

where Λ_{tuv} is a Hermite Gaussian of exponent p centered on \mathbf{P} and $\Lambda_{t'u'v'}$ a Gaussian of exponent p' centered on \mathbf{P}' . Inserting the definition of Hermite Gaussians (Eq. (46)) and taking the differential operators outside the integration sign, we obtain

$$\begin{aligned} V_{tuv}^{qrs} &= \left(\frac{\partial}{\partial P_x} \right)^t \left(\frac{\partial}{\partial P_y} \right)^u \left(\frac{\partial}{\partial P_z} \right)^v \left(\frac{\partial}{\partial C_x} \right)^q \left(\frac{\partial}{\partial C_y} \right)^r \left(\frac{\partial}{\partial C_z} \right)^s \\ &\times \int \frac{\exp(-pr_P^2)}{r_C} d\tau, \end{aligned} \quad (175)$$

and

$$\begin{aligned} V_{tuv; t'u'v'} &= \left(\frac{\partial}{\partial P_x} \right)^t \left(\frac{\partial}{\partial P_y} \right)^u \left(\frac{\partial}{\partial P_z} \right)^v \left(\frac{\partial}{\partial P'_x} \right)^{t'} \left(\frac{\partial}{\partial P'_y} \right)^{u'} \left(\frac{\partial}{\partial P'_z} \right)^{v'} \\ &\times \iint \frac{\exp(-pr_{1P}^2) \exp(-p'r_{2P'}^2)}{r_{12}} d\tau_1 d\tau_2. \end{aligned} \quad (176)$$

The integrals appearing in these expressions are the Coulomb potential and interaction integrals for spherical charge distributions treated in Sec. 12.1. Equations (141) and (142) may now be used to replace these integrals by the incomplete gamma function. Taking the normalization of the charge distributions Eq. (124) into account, we obtain

$$V_{tuv}^{qrs} = \frac{2\pi}{p} \left(\frac{\partial}{\partial P_x} \right)^t \left(\frac{\partial}{\partial P_y} \right)^u \left(\frac{\partial}{\partial P_z} \right)^v \left(\frac{\partial}{\partial C_x} \right)^q \left(\frac{\partial}{\partial C_y} \right)^r \left(\frac{\partial}{\partial C_z} \right)^s \\ \times F_0(pR_{PC}^2), \quad (177)$$

and

$$V_{tuv; t'u'v'} = \frac{2\pi^{5/2}}{pp'\sqrt{p+p'}} \left(\frac{\partial}{\partial P_x} \right)^t \left(\frac{\partial}{\partial P_y} \right)^u \left(\frac{\partial}{\partial P_z} \right)^v \left(\frac{\partial}{\partial P'_x} \right)^{t'} \\ \times \left(\frac{\partial}{\partial P'_y} \right)^{u'} \left(\frac{\partial}{\partial P'_z} \right)^{v'} F_0(\alpha R_{PP'}^2). \quad (178)$$

The definition of Hermite Gaussians in terms of differentiation has thus allowed us to express integrals over non-spherical distributions as derivatives of integrals over spherical distributions. These derivatives may be further simplified, since the incomplete gamma function depends on the relative separation of the two centers. We now obtain the following major simplification¹³⁰:

$$V_{tuv}^{qrs} = (-1)^{q+r+s} \frac{2\pi}{p} \left(\frac{\partial}{\partial P_x} \right)^{t+q} \left(\frac{\partial}{\partial P_y} \right)^{u+r} \left(\frac{\partial}{\partial P_z} \right)^{v+s} F_0(pR_{PC}^2), \quad (179)$$

and

$$V_{tuv; t'u'v'} = (-1)^{t'+u'+v'} \frac{2\pi^{5/2}}{pp'\sqrt{p+p'}} \left(\frac{\partial}{\partial P_x} \right)^{t+t'} \left(\frac{\partial}{\partial P_y} \right)^{u+u'} \left(\frac{\partial}{\partial P_z} \right)^{v+v'} \\ \times F_0(\alpha R_{PP'}^2). \quad (180)$$

This occurs because we use Hermite Gaussians to describe non-spherical distributions. Since the derivatives of the incomplete gamma function play such an important role, we introduce the integrals¹³⁰

$$R_{tuv} = (\partial/\partial P_x)^t (\partial/\partial P_y)^u (\partial/\partial P_z)^v F_0. \quad (181)$$

The Hermite Coulomb integrals may now be written

$$\begin{aligned} & \int \Lambda_{tuv} \left(\frac{\partial}{\partial C_x} \right)^q \left(\frac{\partial}{\partial C_y} \right)^r \left(\frac{\partial}{\partial C_z} \right)^s r_C^{-1} d\tau \\ &= \frac{(-1)^{q+r+s} 2\pi}{p} R_{t+q, u+r, v+s}(p, \mathbf{R}_{PC}), \end{aligned} \quad (182)$$

and

$$\begin{aligned} & \iint \frac{\Lambda_{tuv}(r_1) \Lambda_{t'v'u'}(r_2)}{r_{12}} d\tau_1 d\tau_2 \\ &= \frac{(-1)^{t'+u'+v'} 2\pi^{5/2}}{pp' \sqrt{p+p'}} R_{t+t', u+u', v+v'}(\alpha, \mathbf{R}_{PP'}). \end{aligned} \quad (183)$$

Hence, to calculate Coulomb integrals over Hermite Gaussians, we simply take the derivatives of the incomplete gamma function. We also see that field and field gradient integrals may be calculated in the same way as potential integrals.¹³⁰ It remains, however, to develop a method for calculating the derivatives of the incomplete gamma function. We do this by recursion.

We note that the first derivative involves the first-order incomplete gamma function

$$\begin{aligned} R_{100} &= -2pX_{PC} \int_0^1 \exp(-pR_{PC}^2 t^2) t^2 dt \\ &= -2pX_{PC} F_1(pR_{PC}^2). \end{aligned} \quad (184)$$

Higher derivatives must therefore be linear combinations of incomplete gamma functions of different orders, and our task is to develop a recursive scheme by which the Hermite integrals R_{tuv} for $t+u+v \leq N$ may be calculated from the incomplete gamma functions F_n of orders $n \leq N$. We introduce the auxiliary integrals

$$R_{tuv}^n = (\partial/\partial P_x)^t (\partial/\partial P_y)^u (\partial/\partial P_z)^v R_{000}^n, \quad (185)$$

where

$$R_{000}^n = (-2p)^n F_n(pR_{CP}^2). \quad (186)$$

Note that the definition Eq. (185) includes the integrals R_{tuv} and the scaled incomplete gamma function F_n as special cases. We must now relate the integrals R_{tuv}^n by recursion. Incrementing t , we obtain

$$R_{t+1, u, v}^n = (\partial/\partial P_x)^t (\partial/\partial P_y)^u (\partial/\partial P_z)^v \frac{\partial R_{000}^n}{\partial P_x}. \quad (187)$$

Using Eq. (186) and differentiating the n th order incomplete gamma function, we find

$$R_{t+1, u, v}^n = (\partial/\partial P_x)^t X_{PC} R_{0uv}^{n+1}. \quad (188)$$

The operator on the right-hand side may be written as

$$\begin{aligned} (\partial/\partial P_x)^t X_{PC} &= [(\partial/\partial P_x)^t, X_{PC}] + X_{PC} (\partial/\partial P_x)^t \\ &= t(\partial/\partial P_x)^{t-1} + X_{PC} (\partial/\partial P_x)^t. \end{aligned} \quad (189)$$

Inserting this in Eq. (188), we obtain the recursion relation for the Hermite integrals. The formulae for increments in the three indices are

$$R_{t+1, u, v}^n = t R_{t-1, u, v}^{n+1} + X_{PC} R_{t, u, v}^{n+1}, \quad (190)$$

$$R_{t, u+1, v}^n = u R_{t, u-1, v}^{n+1} + Y_{PC} R_{t, u, v}^{n+1}, \quad (191)$$

and

$$R_{t, u, v+1}^n = v R_{t, u, v-1}^{n+1} + Z_{PC} R_{t, u, v}^{n+1}. \quad (192)$$

In this way, all Hermite integrals of order $t + u + v \leq N$ may be calculated from the incomplete gamma functions of order $n \leq N$ by recursion. The same recursions are found in the two-electron case, except that $X_{PP'}$ replaces X_{PC} , and so on.

To summarize, we have shown that the Coulomb integrals over Hermite Gaussians Eqs. (182) and (183) may be obtained from the incomplete gamma function by three simple recursion formulae, Eqs. (190)–(192). The one- and two-electron integrals follow essentially the same scheme, and field and field gradient integrals may be calculated in the same way as potential integrals. This includes two-electron operators obtained by differentiating the two-electron Coulomb interaction, as arise in certain fine-structure integrals, as well as derivatives of the nuclear attraction operator. The

remaining task is to obtain integrals over Cartesian Gaussians from the Hermite integrals.

12.4 Cartesian Integrals

We are now in a position to calculate the Cartesian Coulomb integrals

$$V_{ab}^{qrs} = \langle G_a | (\partial/\partial C_x)^q (\partial/\partial C_y)^r (\partial/\partial C_z)^s r_C^{-1} | G_b \rangle, \quad (193)$$

and

$$g_{abcd} = \left\langle G_a(r_1) G_b(r_1) \left| \frac{1}{r_{12}} \right| G_c(r_2) G_d(r_2) \right\rangle. \quad (194)$$

Special cases of the one-electron integral are the electrostatic potential, field, and field gradient at some point C in space. The potential integrals enter the one-electron Hamiltonian to describe the nuclear attraction interaction. Field and field gradient integrals are used in molecular gradient and Hessian calculations, and field gradient integrals are also used to describe the interaction with nuclear quadrupole moments. The two-electron integrals account for the electrostatic interactions between electrons.

The Coulomb integrals may be calculated from the Hermite integrals discussed in Sec. 12.3 by expansion or recursion. We consider expansion first. In terms of the overlap distribution, these integrals are given by

$$V_{ab}^{qrs} = \int \Omega_{ab} (\partial/\partial C_x)^q (\partial/\partial C_y)^r (\partial/\partial C_z)^s r_C^{-1} d\tau \quad (195)$$

$$g_{abcd} = \iint \frac{\Omega_{ab}(r_1) \Omega_{cd}(r_2)}{r_{12}} d\tau_1 d\tau_2. \quad (196)$$

Inserting the Hermite expansion of the overlap distribution Eq. (69), we obtain

$$V_{ab}^{qrs} = \sum_{tuv} E_{tuv}^{ab} \int \Lambda_{tuv} (\partial/\partial C_x)^q (\partial/\partial C_y)^r (\partial/\partial C_z)^s r_C^{-1} d\tau \quad (197)$$

$$g_{abcd} = \sum_{tuv} E_{tuv}^{ab} \sum_{t'u'v'} E_{t'u'v'}^{cd} \iint \frac{\Lambda_{tuv}(r_1) \Lambda_{t'u'v'}(r_2)}{r_{12}} d\tau_1 d\tau_2, \quad (198)$$

using the notation

$$E_{tuv}^{ab} = E_t^{ij} E_u^{kl} E_v^{mn}. \quad (199)$$

We recall that Λ_{tuv} is a Hermite Gaussian of exponent p centered on \mathbf{P} :

$$p = a + b \quad (200)$$

$$p\mathbf{P} = a\mathbf{A} + b\mathbf{B}, \quad (201)$$

and $\Lambda_{t'u'v'}$ is a Gaussian of exponent p' centered on \mathbf{P}'

$$p' = c + d, \quad (202)$$

$$p'\mathbf{P}' = c\mathbf{C} + d\mathbf{D}. \quad (203)$$

The Hermite Coulomb integrals were treated in Sec. 12.3 above. Inserting Eqs. (182) and (183) in Eqs. (197) and (198), respectively, we obtain the final expressions for Cartesian Coulomb integrals:

$$V_{ab}^{qrs} = (-1)^{q+r+s} \frac{2\pi}{p} \sum_{tuv} E_{tuv}^{ab} R_{t+q, u+r, v+s}(p, \mathbf{R}_{PC}), \quad (204)$$

$$\begin{aligned} g_{abcd} = & \frac{2\pi^{5/2}}{pp' \sqrt{p+p'}} \sum_{tuv} E_{tuv}^{ab} \sum_{t'u'v'} (-1)^{t'+u'+v'} \\ & \times E_{t'u'v'}^{cd} R_{t+t', u+u', v+v'}(\alpha, \mathbf{R}_{PP'}). \end{aligned} \quad (205)$$

We see that the Cartesian integrals may be calculated very simply from the Hermite integrals, the basic manipulation is transformation of an array from one basis (Hermite Gaussian overlap distributions) to another (Cartesian Gaussian overlap distributions). However, in practice this transformation, Eq. (205), is the time-consuming step in the evaluation of two-electron integrals. Note that once the Hermite Coulomb integrals have been generated, the calculation of field and field gradient integrals is no more expensive than the calculation of the potential integrals since the summations are the same. The nuclear attraction part of the one-electron Hamiltonian (h^{NA}) contains one contribution from each nucleus, obtained by multiplying the charge of the nucleus with the potential at the position of the nucleus. We find

$$h_{ab}^{NA} = \sum_K Z_K V_{ab}^{000}(\mathbf{C}_K) = \frac{2\pi}{p} \sum_{tuv} E_{tuv}^{ab} \sum_K Z_K R_{tuv}(\mathbf{C}_K), \quad (206)$$

where the summation is over all nuclei. Note that the contributions from all the nuclei may be added together before we transform from Hermite to Cartesian integrals.

Having derived formulae for calculating the Coulomb integrals by expansion, we now discuss their calculation by recursion. To avoid complication in the notation, we consider only the one-electron potential integrals. One-electron field and field gradient integrals, as well as two-electron integrals, are obtained by a straightforward modification of the scheme. We introduce the integrals

$$\Phi_{ijklmn}^{tuv} = \int \frac{\Omega_{ij}^t \Omega_{kl}^u \Omega_{mn}^v}{r_C} d\tau, \quad (207)$$

where we use the mixed Cartesian and Hermite overlap distributions of Eq. (80). These integrals contain as special cases the Cartesian integrals

$$\Phi_{ijklmn}^{000} = V_{ab}^{000}, \quad (208)$$

and the Hermite integrals

$$\Phi_{000000}^{tuv} = K_{AB}^{xyz} \int \frac{\Lambda_{tuv}}{r_C} d\tau. \quad (209)$$

Here K_{AB}^{xyz} is the product of three pre-exponential factors, one for each Cartesian direction (Eq. (66)). Using the recurrence relation for the mixed overlap distributions Eqs. (85) and (86), we obtain the following recursions

$$\Phi_{i+1,j,k,l,m,n}^{tuv} = \frac{1}{2p} \Phi_{ijklmn}^{t+1,u,v} - \frac{qQ_x}{a} \Phi_{ijklmn}^{tuv} + t \Phi_{ijklmn}^{t-1,u,v}, \quad (210)$$

$$\Phi_{i,j+1,k,l,m,n}^{tuv} = \frac{1}{2p} \Phi_{ijklmn}^{t+1,u,v} + \frac{qQ_x}{b} \Phi_{ijklmn}^{tuv} + t \Phi_{ijklmn}^{t-1,u,v}, \quad (211)$$

and similarly for increments in the other indices. Starting from Hermite integrals where $t \leq N$, $u \leq N$, and $v \leq N$, one may generate all Cartesian integrals for which $i + j \leq N$, $k + l \leq N$, and $m + n \leq N$. Recurrence relations have been developed and used extensively for the calculation of Gaussian integrals, as we shall now discuss.

12.5. The Obara-Saika Scheme

In the McMurchie-Davidson scheme, we start by calculating the incomplete gamma functions, next we generate from these the Hermite integrals by

recursion, and finally we arrive at the Cartesian integrals by recursion or expansion. One advantage of this scheme — the generation of intermediate Hermite functions — is that to a large extent it allows us to focus on one electron at a time. Nevertheless, it is natural to ask whether it is possible to eliminate the intermediates and generate the Cartesian integrals directly from the incomplete gamma functions. Such a scheme was proposed by Obara and Saika,¹³² who presented a set of recursion relations for the generation of Cartesian integrals from incomplete gamma functions.

We shall derive here the Obara–Saika equations for one-electron Coulomb integrals. These are simpler than the two-electron integrals, but still display the essential features of the Obara–Saika scheme. As is usual for recursions, we first construct a set of integrals, θ_{ijklmn}^N , that includes as special cases the target integrals (the one-electron Coulomb integrals)

$$\theta_{ijklmn}^0 = \int \frac{\Omega_{ij}\Omega_{kl}\Omega_{mn}}{r_C} d\tau, \quad (212)$$

and the source integrals (the incomplete gamma functions)

$$\theta_{000000}^N = \frac{2\pi}{p} K_{AB}^{xyz} F_N(pR_{CP}^2). \quad (213)$$

Note that the incomplete gamma functions in Eq. (213) are scaled such that, for $N = 0$, they correspond to the Coulomb integrals for s functions (see Eqs. (186) and (204)). In the following, we shall assume that k, l, m , and n are all zero and omit these from the expressions. This simplifies our notation but does not affect the recursion relations in any way.

To see how to set up the auxiliary integrals θ_{ij}^N , we investigate the structure of θ_{00}^N more closely. From the relationship (see Eq. (145))

$$F_{N+1}(pR_{CP}^2) = -\frac{1}{2pX_{CP}} \frac{\partial F_N(pR_{CP}^2)}{\partial C_x}, \quad (214)$$

we obtain from Eq. (213)

$$\theta_{00}^1 = -\frac{1}{2pX_{CP}} \frac{\partial \theta_{00}^0}{\partial C_x}. \quad (215)$$

We write the Coulomb integral in the form (see Sec. 12.1)

$$\begin{aligned}\theta_{00}^0 &= K_{AB}^{xyz} \int \frac{\exp(-pr_P^2)}{r_C} d\tau \\ &= \frac{K_{AB}^{xyz}}{\sqrt{\pi}} \int \exp(-pr_P^2) \exp(-t^2 r_C^2) dt d\tau \\ &= \frac{K_{AB}^{xyz}}{\sqrt{\pi}} \int \exp\left(-\frac{pt^2}{p+t^2} R_{CP}^2\right) \exp[-(p+t^2)r_S^2] dt d\tau, \quad (216)\end{aligned}$$

where

$$\mathbf{S} = \frac{p\mathbf{P} + t^2\mathbf{C}}{p+t^2}. \quad (217)$$

The second equality in Eq. (216) follows from Eq. (127), and the third by application of the Gaussian product rule (Eq. (61)). All the integration variables run over all real numbers. Equation (215) may now be written

$$\begin{aligned}\theta_{00}^1 &= -\frac{K_{AB}^{xyz}}{2pX_{CP}\sqrt{\pi}} \frac{\partial}{\partial C_x} \int \exp\left(-\frac{pt^2 R_{CP}^2}{p+t^2}\right) \exp[-(p+t^2)r_S^2] dt d\tau \\ &= \frac{K_{AB}^{xyz}}{\sqrt{\pi}} \int \left(\frac{t^2}{p+t^2}\right) \exp\left(-\frac{pt^2 R_{CP}^2}{p+t^2}\right) \exp[-(p+t^2)r_S^2] dt d\tau. \quad (218)\end{aligned}$$

Comparing the last expressions in Eqs. (216) and (218), we conclude that the *ss* integrals may be written in the following way

$$\theta_{00}^N = \frac{K_{AB}^{xyz}}{\sqrt{\pi}} \int \left(\frac{t^2}{p+t^2}\right)^N \exp\left(-\frac{pt^2 R_{CP}^2}{p+t^2}\right) \exp[-(p+t^2)r_S^2] dt d\tau. \quad (219)$$

To simplify this expression, we introduce

$$u^2 = \frac{t^2}{p+t^2}, \quad (220)$$

and obtain

$$\theta_{00}^N = \frac{K_{AB}^{xyz}}{\sqrt{\pi}} \int u^{2N} \exp(-pu^2 R_{CP}^2) \exp\left(-\frac{p}{1-u^2} r_S^2\right) dv, \quad (221)$$

where

$$\mathbf{S} = (1 - u^2)\mathbf{P} + u^2\mathbf{C}, \quad (222)$$

$$dv = dt d\tau = (1 - u^2)^{-3/2} p^{1/2} du d\tau. \quad (223)$$

In Eq. (221), integration over u is from minus one to one. From this expression, we conclude that the general form of the auxiliary integrals should be chosen as

$$\theta_{ij}^N = \frac{K_{AB}^{xyz}}{\sqrt{\pi}} \int x_A^i x_B^j u^{2N} \exp(-pu^2 R_{CP}^2) \exp\left(-\frac{pr_S^2}{1-u^2}\right) dv. \quad (224)$$

This expression neatly displays the dependency of the integrals on the Cartesian quantum numbers ij and on the incomplete gamma function index N . These integrals are in a form suitable for deriving the recursion relations of the Coulomb integrals. Since we will not carry out the integrations in Eq. (224) explicitly, the detailed form of dv (Eq. (223)) is not needed in the following.

We now increment the Cartesian index i by one and try to express the result in terms of integrals of lower order in the Cartesian indices and possibly higher orders in the gamma function index. The incremented integral may be written as

$$\theta_{i+1,j}^N = \frac{K_{AB}^{xyz}}{\sqrt{\pi}} \int x_A^{i+1} x_B^j u^{2N} \exp(-pu^2 R_{CP}^2) \exp\left(-\frac{pr_S^2}{1-u^2}\right) dv. \quad (225)$$

From Eq. (222), we find that

$$x_A = x_S + X_{PA} - u^2 X_{PC}. \quad (226)$$

Inserting this expression in the incremented integral, we obtain

$$\begin{aligned} \theta_{i+1,j}^N &= X_{PA} \theta_{i,j}^N - X_{PC} \theta_{i,j}^{N+1} \\ &+ \frac{K_{AB}^{xyz}}{\sqrt{\pi}} \int x_A^i x_B^j u^{2N} \exp(-pu^2 R_{CP}^2) x_s \exp\left(-\frac{pr_S^2}{1-u^2}\right) dv. \end{aligned} \quad (227)$$

We now use the relationship

$$x_S \exp\left(-\frac{pr_S^2}{1-u^2}\right) = \frac{u^2 - 1}{2p} \frac{d}{dx} \exp\left(-\frac{pr_S^2}{1-u^2}\right) \quad (228)$$

and partial integration to obtain

$$\begin{aligned} & \int x_A^i x_B^j u^{2N} \exp(-pu^2 R_{CP}^2) x_S \exp\left(-\frac{pr_S^2}{1-u^2}\right) dv \\ &= \int \frac{d(x_A^i x_B^j)}{dx} u^{2N} \exp(-pu^2 R_{CP}^2) \frac{1-u^2}{2p} \exp\left(-\frac{pr_S^2}{1-u^2}\right) dv. \end{aligned} \quad (229)$$

Differentiation of the product $x_A^i x_B^j$ now yields

$$\begin{aligned} \theta_{i+1,j}^N &= X_{PA} \theta_{i,j}^N - X_{PC} \theta_{i,j}^{N+1} \\ &+ \frac{1}{2p} (i \theta_{i-1,j}^N - i \theta_{i-1,j}^{N+1} + j \theta_{i,j-1}^N - j \theta_{i,j-1}^{N+1}), \end{aligned} \quad (230)$$

which is the final recursion relation for the Coulomb integrals. Using this equation, and its analogs for recursion in the other indices, we may generate all Cartesian Coulomb integrals starting from the incomplete gamma function.

The two-electron case is similar, although more complicated. Using i and j to denote the Cartesian angular quantum numbers in the x direction associated with electron one, and k and l correspondingly for electron two, the recursion relations may be written in the form

$$\begin{aligned} \theta_{i+1,j,k,l}^N &= X_{PA} \theta_{i,j,k,l}^N - \frac{\alpha}{p} X_{PP'} \theta_{i,j,k,l}^{N+1} \\ &+ \frac{i}{2p} \left(\theta_{i-1,j,k,l}^N - \frac{\alpha}{p} \theta_{i-1,j,k,l}^{N+1} \right) \\ &+ \frac{j}{2p} \left(\theta_{i,j-1,k,l}^N - \frac{\alpha}{p} \theta_{i,j-1,k,l}^{N+1} \right) \\ &+ \frac{k}{2(p+p')} \theta_{i,j,k-1,l}^{N+1} + \frac{l}{2(p+p')} \theta_{i,j,k,l-1}^{N+1}, \end{aligned} \quad (231)$$

where α was defined in Eq. (140). The similarity with the one-electron case is apparent. In particular, we note that the two-electron recursions reduce to those for one-electron Eq. (230) as p' (the exponent for the second electron) tends to infinity. Thus it is possible to set up a recursion relation which yields the Cartesian integrals directly from the incomplete gamma functions. However, it is apparent that the Obara-Saika recursions

Eqs. (230) and (231) are more complicated than the McMurchie–Davidson recursions Eqs. (190)–(192), (210) and (211). Also, the original Obara–Saika scheme does not allow us to treat the two electrons separately, as is done in the second and time-consuming part of the McMurchie–Davidson scheme. Such a separate treatment is useful for highly contracted basis sets, since — as we shall see later — it allows us to contract the orbitals of the first electron before attacking the second electron, reducing the number of operations significantly.

We do not derive the Obara–Saika two-electron integral recursion equations¹³² here, but an important special case is easily obtained. If we assume that the second electron is described by s functions, so that k and l are zero, the Obara–Saika recursion equations become

$$\begin{aligned}\theta_{i+1,j,0,0}^N &= X_{PA} \theta_{i,j,0,0}^N - \frac{\alpha}{p} X_{PP'} \theta_{i,j,0,0}^{N+1} \\ &\quad + \frac{i}{2p} \left(\theta_{i-1,j,0,0}^N - \frac{\alpha}{p} \theta_{i-1,j,0,0}^{N+1} \right) \\ &\quad + \frac{j}{2p} \left(\theta_{i,j-1,0,0}^N - \frac{\alpha}{p} \theta_{i,j-1,0,0}^{N+1} \right),\end{aligned}\quad (232)$$

i.e., the last two terms of Eq. (231) vanish. The recursions (Eq. (232)) are similar to the one-electron recursions, and indeed may be derived the same way. The only differences are that in Eq. (216) we replace the point charge Coulomb potential by the potential from a spherical distribution, as given by Eq. (132), and obtain

$$\theta_{0000}^0 = K \int \exp(-pr_P^2) \exp(-p't^2 r_{P'}^2) dt d\tau,\quad (233)$$

where K is some constant and integration over t is now from zero to one. Note the similarity with Eq. (216). Following the same procedure as for the one-electron case, we are led to the conclusion that the desired two-electron integrals are of the form

$$\theta_{ij00}^N = K \left(\frac{p}{\alpha} \right)^N \int x_A^i x_B^j u^{2N} \exp(-pu^2 R_{PP'}^2) \exp\left(-\frac{pr_S^2}{1-u^2}\right) dv,\quad (234)$$

where

$$u^2 = \frac{p't^2}{p+p't^2},\quad (235)$$

$$\mathbf{S} = (1 - u^2)\mathbf{P} + u^2\mathbf{P}', \quad (236)$$

and we have again collapsed the volume element for integration into dv , as we did in Eq. (223). Finally, writing

$$x_A = x_S + X_{PA} - u^2 X_{PP'}, \quad (237)$$

we obtain the special two-electron recursions Eq. (232).

After the publication of the Obara–Saika scheme, much activity has been directed towards deriving alternative recursion schemes. We illustrate this by presenting an approach that generates the final Cartesian integrals from the incomplete gamma functions in three steps. First, a set of integrals with only $i > 0$ is generated by using the following simplified version of the Obara–Saika recursions:

$$\begin{aligned} \theta_{i+1, 0, 0, 0}^N &= X_{PA}\theta_{i, 0, 0, 0}^N - \frac{\alpha X_{PP'}}{p}\theta_{i, 0, 0, 0}^{N+1} \\ &\quad + \frac{i}{2p}\theta_{i-1, 0, 0, 0}^N - \frac{i\alpha}{2p^2}\theta_{i-1, 0, 0, 0}^{N+1}. \end{aligned} \quad (238)$$

Next, we transfer Cartesian powers from the first to the second electron by the recursion relations:

$$\begin{aligned} \theta_{i, 0, k+1, 0} &= \frac{bX_{BA} + dX_{DC}}{p'}\theta_{i, 0, k, 0} + \frac{i}{2p'}\theta_{i-1, 0, k, 0} \\ &\quad + \frac{k}{2p'}\theta_{i, 0, k-1, 0} - \frac{p}{p'}\theta_{i+1, 0, k, 0}. \end{aligned} \quad (239)$$

We have omitted the superscript N here, since $N = 0$ for the Cartesian integrals. Finally, we may transfer Cartesian powers between the orbitals of the same electron using the simple recursions

$$\theta_{i, j+1, k, l} = \theta_{i+1, j, k, l} + X_{AB}\theta_{i, j, k, l}, \quad (240)$$

$$\theta_{i, j, k, l+1} = \theta_{i, j, k+1, l} + X_{CD}\theta_{i, j, k, l}. \quad (241)$$

In this way, we may build up the full set of Cartesian integrals by using three sets of recursion relations, each of which is considerably simpler than the full Obara–Saika scheme. In particular, the recursions of Eqs. (240) and (241) do not involve orbital exponents, and may therefore be applied

after the transformation of the integrals to the contracted basis, as we shall discuss later. The recursions of Eqs. (240) and (241) follow directly from the relationships

$$x_A = x_B - X_{AB}, \quad (242)$$

$$x_C = x_D - X_{CD}. \quad (243)$$

The electron transfer relation Eq. (239) follows from the translational invariance of the two-electron integrals

$$\frac{\partial \theta_{i,0,k,0}}{\partial A_x} + \frac{\partial \theta_{i,0,k,0}}{\partial B_x} + \frac{\partial \theta_{i,0,k,0}}{\partial C_x} + \frac{\partial \theta_{i,0,k,0}}{\partial D_x} = 0, \quad (244)$$

to be discussed in Sec. 14.1. Using Eq. (41), this equation may be rewritten as

$$2a\theta_{i+1,0,k,0} + 2b\theta_{i,1,k,0} + 2c\theta_{i,0,k+1,0} + 2d\theta_{i,0,k,1} \\ - i\theta_{i-1,0,k,0} - k\theta_{i,0,k-1,0} = 0, \quad (245)$$

and, applying Eqs. (242) and (243), we obtain

$$2p\theta_{i+1,0,k,0} + 2bX_{AB}\theta_{i,0,k,0} + 2p'\theta_{i,0,k+1,0} + 2dX_{CD}\theta_{i,0,k,0} \\ - i\theta_{i-1,0,k,0} - k\theta_{i,0,k-1,0} = 0, \quad (246)$$

which may be rearranged to yield the electron transfer recursion.

There has been considerable recent interest in developing recursive methods for computing Coulomb integrals. In addition to the work by Obara and Saika,¹³² we note (among others) work by Head-Gordon and Pople,¹⁴⁰ Gill and coworkers,^{141–143} Hamilton and Schaefer,¹⁴⁴ and Lindh and coworkers.¹⁴⁵

12.6. Integral Prescreening

In all but the very smallest molecules, a considerable fraction of the two-electron integrals will be small because the orbitals involved are far apart. If a reliable approximation or, better, a true upper bound to these integrals is available, their calculation may be avoided when it is determined that their contribution to the energy or SCF Fock matrix will be negligible.

Such *integral prescreening* is especially important in direct calculations on large molecular systems, where the majority of integrals may be neglected. Ramifications of this issue are discussed in Chapter 3 by Almlöf; we discuss here an efficient upper bound to integrals based on Schwarz's inequality.

Roothaan¹² proved that the two-electron integrals are elements of a positive definite matrix with diagonal elements

$$g_{abab} \geq 0. \quad (247)$$

Therefore, the two-electron integrals satisfy the conditions for inner products, in the metric defined by r_{12}^{-1} , and, as noted by Häser and Ahlrichs,¹⁴⁶ we may apply Schwarz's inequality and write

$$\|g_{abcd}\| \leq \sqrt{g_{abab}}\sqrt{g_{cdcd}}. \quad (248)$$

If we calculate the square root of all the diagonal integrals

$$G_{ab} = \sqrt{g_{abab}} \quad (249)$$

prior to the calculation of the two-electron integrals, we can easily compute upper bounds to individual integrals as these are generated. Integrals below a desired threshold may be identified in advance and their evaluation avoided. Note that Schwarz's inequality is optimal in the sense that it is exact for all diagonal elements. For example, an accuracy of the order 10^{-16} in g_{abab} gives an error of the order 10^{-8} in G_{ab} . From Secs. 12.1, 12.3, and 12.4 above, we note that the diagonal two-electron integrals may be evaluated in closed form, since $R_{PP'} = 0$ and Eq. (148) applies. Their evaluation is further simplified since most of the Hermite integrals vanish according to Eqs. (190)–(192), and the nonvanishing terms are trivial. For more details on the use of Schwarz's inequality, we refer to Häser and Ahlrichs¹⁴⁶ and to Häser *et al.*¹⁴⁷

13. Integrals over CGTO's

13.1. *Cartesian Gaussian Integrals*

In our discussions so far, we have considered primitive GTO integrals only. In most cases, we are interested in integrals over CGTO's, and we will assume here that we are interested in general, as well as segmented, contractions. The CGTO integrals may be obtained by a straightforward transformation after the primitive GTO integrals have been calculated, but this

procedure is seldom the most efficient. In the following, we discuss how the transformation may be carried out more efficiently. We restrict ourselves to two-electron integrals, since these are the most expensive computationally. We first outline the “obvious” scheme and then describe how it may be improved.

A two-electron integral may be written

$$g_{abcd} = \sum_{tuv} E_{tuv}^{ab} \sum_{t'u'v'} F_{t'u'v'}^{cd} \mathfrak{R}_{t+t', u+u', v+v'}(\alpha, \mathbf{R}_{PP'}) , \quad (250)$$

where

$$F_{t'u'v'}^{cd} = (-1)^{t'+u'+v'} E_{t'u'v'}^{cd} , \quad (251)$$

$$\mathfrak{R}_{tuv}(\alpha, \mathbf{R}_{PP'}) = \frac{2\pi^{5/2}}{pp'\sqrt{p+p'}} R_{tuv}(\alpha, \mathbf{R}_{PP'}) . \quad (252)$$

The transformation of Eq. (250) from Hermite to Cartesian GTO integrals may clearly be carried out in two steps:

$$g_{tuv}^{cd} = \sum_{t'u'v'} F_{t'u'v'}^{cd} \mathfrak{R}_{t+t', u+u', v+v'} , \quad (253)$$

and

$$g_{abcd} = \sum_{tuv} E_{tuv}^{ab} g_{tuv}^{cd} . \quad (254)$$

The CGTO integrals may then be obtained from the GTO integrals by a four-index transformation³⁹

$$g_{\alpha\beta\gamma\delta} = \sum_{abcd} b_{\alpha a} b_{\beta b} b_{\gamma c} b_{\delta d} g_{abcd} , \quad (255)$$

where \mathbf{b} denotes the transformation matrix from primitive GTO’s to CGTO’s (see also Sec. 3.4).

Since the number of CGTO integrals is smaller than the number of GTO integrals, however, the optimum strategy is to contract as early as

possible. The following four-step scheme can be used:

$$g_{tuv}^{cd} = \sum_{t' u' v'} F_{t' u' v'}^{cd} \mathfrak{R}_{t+t', u+u', v+v'} , \quad (256)$$

$$g_{tuv}^{\gamma\delta} = \sum_{cd} b_{\gamma c} b_{\delta d} g_{tuv}^{cd} , \quad (257)$$

$$g_{ab\gamma\delta} = \sum_{tuv} E_{tuv}^{ab} g_{tuv}^{\gamma\delta} , \quad (258)$$

$$g_{\alpha\beta\gamma\delta} = \sum_{ab} b_{\alpha a} b_{\beta b} g_{ab\gamma\delta} . \quad (259)$$

The second step, Eq. (257), reduces the number of integrals, making the second half-transformation from Hermite to Cartesian integrals (Eq. (258)) less expensive.

The following modified recursion technique, which is applicable to all integrals, also takes advantage of the reduction in the number of integrals after contraction. Consider, for example, the overlap integral

$$S_{ijklmn} = \int \Omega_{ij} \Omega_{kl} \Omega_{mn} d\tau . \quad (260)$$

As we have described above, one way to obtain CGTO integrals is to first calculate all integrals over a given quadruplet of shells in the primitive basis, and then to contract. As an alternative, consider the recursion relation

$$\Omega_{i+1,j} = \Omega_{i,j+1} - X_{AB} \Omega_{ij} \quad (261)$$

obtained by subtracting Eqs. (83) and (84). The recursion coefficients are independent of the exponents, and these recursions may therefore be applied *after* transformation from the primitive GTO to the CGTO basis. Hence, we may first calculate all primitive integrals of the type

$$S_{i+j,0,k+l,0,m+n,0} = \int \Omega_{i+j,0} \Omega_{k+l,0} \Omega_{m+n,0} d\tau . \quad (262)$$

We next transform these to the CGTO basis and obtain the final Cartesian integrals using the recursion

$$S_{i,j+1,k,l,m,n} = S_{i+1,j,k,l,m,n} + X_{AB} S_{i,j,k,l,m,n} . \quad (263)$$

This procedure is faster, since the number of Cartesian integrals to be computed (the time-consuming step) is reduced: there are fewer integrals $S_{i+j,0,k+l,0,m+n,0}$ than S_{ijklmn} . For example, there are only six Cartesian $\langle d|s \rangle$ integrals but nine $\langle p|p \rangle$ integrals. The final recursion Eq. (263) is fast, since the integrals have been reduced in number by contraction.

13.2. Spherical Harmonic Gaussian Integrals

Exactly the same strategy as we have described in the previous section can be used to obtain integrals over spherical harmonic, rather than Cartesian, CGTO's. As we have discussed in Sec. 5.2, it is usually preferable to work in a basis of spherical harmonic functions rather than Cartesians, to reduce linear dependence problems and to reduce the final number of integrals. We write the transformation to spherical harmonics as

$$\phi_a = \sum_{\alpha} s_{a\alpha} \psi_{\alpha}, \quad (264)$$

where ϕ_a is a spherical harmonic basis function and ψ_{α} is a Cartesian CGTO. The transformation matrix s mixes functions only within a given Cartesian angular shell, of course. Now, transforming a block of integrals over Cartesian shells can be viewed as an explicit four-index transformation, but, just as for contraction, this is not the most efficient procedure. Instead, we perform the transformation of the first two Cartesian indices as soon as the integrals are available. That is, we follow the formation of $g_{tuv}^{\gamma\delta}$ from Eq. (257) with an additional transformation to give

$$g_{tuv}^{cd} = \sum_{\gamma\delta} s_{c\gamma} s_{d\delta} g_{tuv}^{\gamma\delta}. \quad (265)$$

Since the transformation to spherical harmonics reduces the number of CGTO overlap distributions within a shell pair, there are fewer terms to be processed in the remaining steps. In this way, it becomes faster to generate spherical harmonic integrals than Cartesian integrals. See also the discussion by Lindh and coworkers.¹⁴⁵ Another possibility would be to rederive the various integral formulae based on recursion for spherical harmonic functions, and full details of this approach can be found in Saunders' review.⁵

From the above discussion, it should be clear that integrals can only be calculated efficiently if they are processed in batches. For example,

the transformation from primitive to contracted integrals clearly requires that all primitive integrals contributing to the same contracted integrals be present at the same time. Also, to be able to generate spherical harmonic integrals efficiently, we must have all the Cartesian integrals available simultaneously. Furthermore, all recursion schemes by their very nature imply the simultaneous evaluation of batches of integrals. Hence all modern programs generate large batches of integrals simultaneously. Such batches comprise, for example, all shell components and all contracted integrals of a given angular momentum, for each of the four orbital indices. This batch processing has the additional benefit that overhead is substantially reduced, because the information that is common to all integrals in a batch is not recomputed.

14. Property Integrals

In this section, we treat various integrals related to the calculation of molecular properties, in particular the calculation of integrals differentiated with respect to nuclear positions (geometrical derivatives), derivatives of London orbital integrals with respect to an external magnetic field, and spin-orbit integrals. Of course, several important property integrals have been discussed in Secs. 11 and 12, including multipole moment integrals and field and field gradient integrals.

14.1. Geometrical Derivative Integrals

Integrals differentiated with respect to the nuclear positions are needed for the calculation of geometrical derivatives of molecular energies and properties, in particular molecular gradients and Hessians.¹⁴⁸ They also enter the calculation of other property integrals, such as spin-orbit integrals.¹⁴⁹

In molecular calculations, the Gaussian orbitals are nearly always centered on the nuclei, and therefore each integral depends on the molecular geometry. For example, the derivative of a two-electron integral with respect to the x coordinate of nucleus N is

$$\frac{\partial g_{abcd}}{\partial N_x} = \frac{\partial g_{abcd}}{\partial A_x} \delta_{AN} + \frac{\partial g_{abcd}}{\partial B_x} \delta_{BN} + \frac{\partial g_{abcd}}{\partial C_x} \delta_{CN} + \frac{\partial g_{abcd}}{\partial D_x} \delta_{DN}, \quad (266)$$

and there is one contribution from each orbital centered on this nucleus. If none of the orbitals are positioned on the nucleus, the derivative vanishes.

Also, if all four orbitals are positioned on the same nucleus the derivative vanishes, since in this case we have a one-center integral and these are independent of the molecular geometry. In general, from *translational invariance*^{150–152} we infer that

$$\frac{\partial g_{abcd}}{\partial A_x} + \frac{\partial g_{abcd}}{\partial B_x} + \frac{\partial g_{abcd}}{\partial C_x} + \frac{\partial g_{abcd}}{\partial D_x} = 0, \quad (267)$$

since the integral must be invariant to an overall translation of all four orbitals. Therefore, we never need to calculate more than three derivatives in each Cartesian direction explicitly; the remaining derivative is given by translational symmetry according to Eq. (267). We may also use rotational symmetry to reduce the number of derivatives to be calculated explicitly.¹⁵³ However, the equations relating derivatives by rotation are more involved and also depend on the GTO exponents. They must therefore be applied to primitive GTO integrals rather than CGTO integrals.

In the case of one-electron Coulomb potential integrals, we must include contributions from derivatives with respect to the orbital centers and the derivatives of the operator. The derivative of the nuclear attraction part of the one-electron Hamiltonian (Eq. (206)) therefore becomes

$$\begin{aligned} \frac{\partial h_{ab}^{NA}}{\partial N_x} &= Z_N V_{ab}^{100}(\mathbf{C}_N) + \delta_{AN} \sum_K Z_K \frac{\partial V_{ab}^{000}(\mathbf{C}_K)}{\partial A_x} \\ &\quad + \delta_{BN} \sum_K Z_K \frac{\partial V_{ab}^{00}(\mathbf{C}_K)}{\partial B_x}. \end{aligned} \quad (268)$$

The first term on the right-hand side is the contribution from the electrostatic field at the nucleus (the “Hellmann–Feynman force”), the remaining two contributions are from the orbital centers. Second derivatives are obtained in the same way, with the field gradient replacing the electric field.

Derivative integrals are calculated in much the same way as undifferentiated integrals. For example, the first derivative of the two-electron integral with respect to the position of the first orbital is

$$\frac{\partial g_{abcd}}{\partial A_x} = \iint \frac{(\partial \Omega_{ab}(r_1)/\partial A_x)\Omega_{cd}(r_2)}{r_{12}} d\tau_1 d\tau_2, \quad (269)$$

and we may simply replace undifferentiated overlap distributions by differentiated ones in the calculation. The most straightforward way to obtain

the differentiated distributions is to note that differentiation turns each Cartesian Gaussian into a linear combination of undifferentiated Gaussians (Eq. (41)). We may therefore obtain differentiated overlap distributions by combining undifferentiated distributions. From Eq. (41), we have

$$\Omega_{ij}^{q+1, r} = 2a\Omega_{i+1, j}^{qr} - i\Omega_{i-1, j}^{qr}, \quad (270)$$

$$\Omega_{ij}^{q, r+1} = 2b\Omega_{i, j+1}^{qr} - j\Omega_{i, j-1}^{qr}, \quad (271)$$

using the notation

$$\Omega_{ij}^{qr} = \frac{\partial^{q+r} \Omega_{ij}}{\partial A_x^q \partial B_x^r}. \quad (272)$$

In this way, we may generate n th derivative overlap distributions of angular quantum number $i + j$ from derivatives of order $n - 1$ and quantum number $i + j + 1$. The calculation of derivative integrals is evidently more expensive than the calculation of undifferentiated integrals, because higher quantum numbers are needed. Another way to generate differentiated overlap distributions is to expand them directly in Hermite Gaussians in the same way as we expand undifferentiated distributions:

$$\Omega_{ij}^{qr} = \sum_{t=0}^{i+j+q+r} E_t^{ij; qr} \Lambda_t. \quad (273)$$

The summation is now over more terms than for undifferentiated distributions since the distribution contains x_A and x_B to higher powers. Note that the coefficients in this expansion are not differentiated coefficients, but rather the coefficients of the differentiated overlap distribution expanded in undifferentiated Hermite Gaussians.

To obtain the expansion coefficients, we expand each of the three overlap distributions in the recursions Eqs. (270) and (271) in Hermite Gaussians and identify terms of same order in t . We arrive at the following expressions:

$$E_t^{ij; q+1, r} = 2aE_t^{i+1, j; qr} - iE_t^{i-1, j; qr}, \quad (274)$$

$$E_t^{ij; q, r+1} = 2bE_t^{i, j+1; qr} - jE_t^{i, j-1; qr}. \quad (275)$$

Once these are available, the calculation of derivative distributions (and hence derivative integrals) follows the same scheme as the calculation of

undifferentiated integrals, except that there are now more terms in the summation.

A disadvantage of the methods discussed above is that they lead to longer expansions than for undifferentiated distributions: one term extra for each order of differentiation. This is unfortunate, because the longer expansions for differentiated distributions lead not only to more Hermite integrals (which are rather cheap to compute), but more work in the transformation from Hermite to Cartesian integrals (an expensive step). In fact, it is possible to avoid this extra work.¹⁵⁴ We recall first (see Sec. 10.3) that in the Hermite expansions, the expansion coefficients are functions of Q_x only and the Gaussians are functions of P_x only. The derivatives with respect to these coordinates therefore take a particularly simple form. Using the notation

$$\Omega_{ij}^{mn} = \frac{\partial^{m+n} \Omega_{ij}}{\partial P_x^m \partial Q_x^n}, \quad (276)$$

we obtain

$$\Omega_{ij}^{mn} = \sum_{t=0}^{i+j} E_t^{ij; n} \Lambda_{t+m}, \quad (277)$$

where the differentiated coefficients are those given by Eqs. (76)–(78). Note that the summation range is the *same* as for undifferentiated distributions since differentiation simply increments the quantum number of a Hermite Gaussian (see Eq. (49) and compare with the more complicated expression for Cartesian Gaussians Eq. (41)). Thus the transformation from Hermite integrals to Cartesian integrals is now no more expensive than for undifferentiated integrals. This is especially important for higher derivatives.

It still remains to determine how these derivatives are related to the derivatives with respect to orbital centers. The first derivatives are related as

$$\frac{\partial}{\partial A_x} = \frac{a}{p} \frac{\partial}{\partial P_x} + \frac{\partial}{\partial Q_x}, \quad (278)$$

$$\frac{\partial}{\partial B_x} = \frac{b}{p} \frac{\partial}{\partial P_x} - \frac{\partial}{\partial Q_x}, \quad (279)$$

and

$$\frac{\partial}{\partial P_x} = \frac{\partial}{\partial A_x} + \frac{\partial}{\partial B_x}, \quad (280)$$

$$\frac{\partial}{\partial Q_x} = \frac{q}{a} \frac{\partial}{\partial A_x} - \frac{q}{b} \frac{\partial}{\partial B_x}. \quad (281)$$

It is best to calculate derivatives with respect to A_x according to Eq. (278) first and then obtain derivatives with respect to B_x using Eq. (280), since the latter equation is independent of the exponents.¹⁵⁴ Higher derivatives are transformed according to the binomial expansion

$$(\partial/\partial A_x)^q = \sum_{k=0}^q \binom{q}{k} \left(\frac{a}{p}\right)^k (\partial/\partial Q_x)^{q-k} (\partial/\partial P_x)^k, \quad (282)$$

$$(\partial/\partial B_x)^r = (\partial/\partial P_x)^r - \sum_{k=0}^{r-1} \binom{r}{k} (\partial/\partial A_x)^{r-k} (\partial/\partial B_x)^k. \quad (283)$$

Note that these transformations are independent of Cartesian quantum numbers.

We should note that the coordinates of the nuclear centers are not the only differentiation variables of interest. We can differentiate a basis function with respect to the exponent, which allows us to calculate the gradient (or higher derivatives) of the energy with respect to the exponents.^{30–32} This can be useful in basis set optimization studies. If the basis functions are CGTO's, we can differentiate with respect to contraction coefficients, which is also useful in basis set optimization.³² Finally, we should also note that special formulae have been developed for the calculation of differentiated integrals over CGTO's with low angular quantum numbers (s and p functions).¹⁵⁵

14.2. Derivatives of London Orbitals

In calculations of properties involving an external magnetic field, it is common to attach to the Gaussian orbitals a set of field-dependent phase factors given as

$$\omega_A(\mathbf{r}; \mathbf{A}_A) = \exp(-i\mathbf{A}_A \cdot \mathbf{r}) G(\mathbf{r}, a, \mathbf{A}) \quad (284)$$

in atomic units. Here \mathbf{A}_A is the vector potential at the center of the Gaussian distribution $G(\mathbf{r}, a, \mathbf{A})$:

$$\mathbf{A}_A = \frac{1}{2} \mathbf{B} \times \mathbf{R}_{AO}, \quad (285)$$

$$\mathbf{R}_{AO} = \mathbf{R}_A - \mathbf{R}_O . \quad (286)$$

In Eq. (285), \mathbf{B} is the magnetic field and in Eq. (286) \mathbf{R}_O is the origin of the vector potential. These *London orbitals* are used since they incorporate some of the physics of the electrons in the field and also ensure gauge invariance of the magnetic properties (see Sec. 6.4). The London orbitals are complex at nonvanishing magnetic fields \mathbf{B} . The integrals over such orbitals may be calculated following the same scheme as for our usual GTO's or CGTO's at zero field, but the calculation becomes more involved and standard codes cannot be used. However, if analytical derivative techniques are used for calculating magnetic properties, all we need are the derivatives of the London orbitals, with respect to the magnetic field, *at zero field*. Such derivatives are either real or pure imaginary, and thus may be calculated within the usual framework for real orbitals. All one-electron integrals involving London orbitals may be written in the form

$$O_{ab} = \left\langle G_a \left| \exp \left(\frac{1}{2} i \mathbf{B} \cdot \mathbf{R}_{AB} \times \mathbf{r} \right) O(\mathbf{r}) \right| G_b \right\rangle , \quad (287)$$

where $O(\mathbf{r})$ is some one-electron operator. (We should perhaps note that for nonlocal one-electron operators, the effect of moving the phase factor from the ket to before the operator changes the original one-electron operator, see Ref. 88.) We now introduce the matrix

$$\mathbf{Q}_{AB} = \frac{1}{2} \begin{bmatrix} 0 & -Z_{AB} & Y_{AB} \\ Z_{AB} & 0 & -X_{AB} \\ -Y_{AB} & X_{AB} & 0 \end{bmatrix} \quad (288)$$

so that we may write cross products involving \mathbf{R}_{AB} as matrix–vector multiplications

$$\mathbf{R}_{AB} \times \mathbf{r} = 2\mathbf{Q}_{AB}\mathbf{r} . \quad (289)$$

The first and second derivatives of Eq. (288) may now be written as

$$\left. \frac{\partial O_{ab}}{\partial \mathbf{B}} \right|_{\mathbf{B}=0} = i\mathbf{Q}_{AB} \langle G_a | \mathbf{r} O(\mathbf{r}) | G_b \rangle , \quad (290)$$

$$\left. \frac{\partial^2 O_{ab}}{\partial \mathbf{B}^2} \right|_{\mathbf{B}=0} = \mathbf{Q}_{AB} \langle G_a | \mathbf{r} \mathbf{r}^T O(\mathbf{r}) | G_b \rangle \mathbf{Q}_{AB} , \quad (291)$$

where the superscript T denotes transposition. Similar expressions may be derived for the two-electron integrals.

It is clear that the first derivative one-electron integrals may be calculated by taking linear combinations of contracted integrals of the form

$$\mathbf{O}_{ab}^r = \langle G_a | \mathbf{r} O(\mathbf{r}) | G_b \rangle. \quad (292)$$

The x component of this integral may be rewritten as

$$\begin{aligned} O_{ab}^x &= \langle G_a | x O(\mathbf{r}) | G_b \rangle \\ &= \langle G_a | x_A O(\mathbf{r}) | G_b \rangle + A_x \langle G_a | O(\mathbf{r}) | G_b \rangle. \end{aligned} \quad (293)$$

Hence if the undifferentiated integrals are calculated according to the expression

$$O_{ab} = \sum_t E_t^{ij} O_t, \quad (294)$$

we may calculate Eq. (293) as

$$O_{ab}^x = \sum_t \left(E_t^{i+1,j} + A_x E_t^{ij} \right) O_t, \quad (295)$$

where the expansion coefficients may be precalculated. This scheme is easily extended to higher derivatives and to two-electron integrals.

14.3. Spin-Orbit Integrals

The leading relativistic corrections to the electrostatic interactions in a many-electron system yield in the Pauli approximation an electronic spin-orbit interaction operator of the form

$$H_{SO} = \frac{\alpha^2}{2} \left[\sum_{iA} Z_A \frac{(\mathbf{r}_{iA} \times \mathbf{p}_i) \cdot \mathbf{s}_i}{r_{iA}^3} - \sum_{i \neq j} \frac{(\mathbf{r}_{ij} \times \mathbf{p}_i) \cdot (\mathbf{s}_i + 2\mathbf{s}_j)}{r_{ij}^3} \right], \quad (296)$$

in atomic units, where α is the fine structure constant, Z_A the charge of nucleus A and \mathbf{s}_i the spin of electron i . The evaluation of the spatial part of the spin-orbit operator involves the calculation of spatial one-electron integrals of the type

$$\mathbf{h}_{ab}^{so(C)} = \left\langle G_a \left| \frac{\mathbf{r}_C \times \nabla}{r_C^3} \right| G_b \right\rangle, \quad (297)$$

and spatial two-electron integrals of the type

$$\mathbf{g}_{abcd}^{\text{so}} = \left\langle G_a(r_1)G_c(r_2) \left| \frac{\mathbf{r}_{12} \times \nabla_1}{r_{12}^3} \right| G_b(r_1)G_d(r_2) \right\rangle. \quad (298)$$

The one-electron integrals may be rewritten as

$$\begin{aligned} \left\langle G_a \left| \frac{\mathbf{r}_C \times \nabla}{r_C^3} \right| G_b \right\rangle &= - \left\langle G_a \left| \frac{\mathbf{r}_C}{r_C^3} \times \nabla_B \right| G_b \right\rangle \\ &= \nabla_B \times \langle G_a | \nabla_C r_C^{-1} | G_b \rangle \\ &= -\nabla_B \times (\nabla_A + \nabla_B) \langle G_a | r_C^{-1} | G_b \rangle, \end{aligned} \quad (299)$$

which implies that the one-electron spin-orbit integrals may be calculated as linear combinations of second derivative integrals

$$\mathbf{h}_{ab}^{\text{so}(C)} = \nabla_A \times \nabla_B \langle G_a | r_C^{-1} | G_b \rangle. \quad (300)$$

For the two-electron integrals, we obtain a similar expression

$$\mathbf{g}_{abcd}^{\text{so}} = \nabla_A \times \nabla_B \langle G_a(r_1)G_b(r_1) | r_{12}^{-1} | G_c(r_2)G_d(r_2) \rangle. \quad (301)$$

Thus the evaluation of spin-orbit integrals does not present any difficulties, provided the geometrical derivative integrals are available.^{149,156}

15. Rys Quadrature

Let us consider again the expression of Eq. (205) for a two-electron integral over Cartesian Gaussians:

$$g_{abcd} = \frac{2\pi^{5/2}}{pp' \sqrt{p+p'}} \sum_{tuv} E_{tuv}^{ab} \sum_{t'u'v'} (-1)^{t'+u'+v'} E_{t'u'v'}^{cd} R_{t+t', u+u', v+v'}, \quad (302)$$

where the E coefficients give the expansion of the Cartesian overlap distributions in Hermite Gaussians, and the R integrals and numerical factors correspond to integrals over the Hermite Gaussians. We have already shown that the Hermite Gaussian integrals $R_{t+t', u+u', v+v'}$ are linear combinations

of incomplete gamma functions of order 0 to $j_{\max} = t + t' + u + u' + v + v'$ (see Sec. 12.3), so we can write the integral as

$$g_{abcd} = \sum_{n=0}^{j_{\max}} c_n F_n(X), \quad (303)$$

where for convenience we have collapsed the expansion coefficient products, etc., into c_n , and introduced $X = \alpha R_{PP'}^2$. Now,

$$F_n(X) = \int_0^1 t^{2n} \exp(-Xt^2) dt, \quad (304)$$

so Eq. (303) can be rewritten as

$$g_{abcd} = \int_0^1 \left[\sum_{n=0}^{j_{\max}} c_n t^{2n} \exp(-Xt^2) \right] dt. \quad (305)$$

In this form, we can see that the integrand is a polynomial in t^2 times an exponential in t^2 . The integral can therefore be evaluated *exactly* by numerical quadrature over m points as¹²⁹

$$g_{abcd} = \sum_{k=1}^m \left(\sum_{n=0}^{j_{\max}} c_n t_k^{2n} \right) W_k. \quad (306)$$

Here W_k is a quadrature weight, and t_k^2 is a root of the *Rys polynomial of order m*, denoted $P_m(t^2, X)$. These polynomials¹⁵⁷ are orthogonal in the sense

$$\int_0^1 P_m(t^2, X) P_n(t^2, X) \exp(-Xt^2) dt = k_m \delta_{mn}, \quad (307)$$

where the normalization constant k_m is given by King and Dupuis.¹⁵⁷ The number of quadrature points in Eq. (306) should exceed $[m/2]$ for an exact integration: the minimum value used is thus $[m/2] + 1$.

The original development of this *Rys quadrature* procedure was based on Cartesian Gaussian overlap distributions, rather than Hermite Gaussians. Expressions for the coefficients c_n that appear in Eq. (306) are the basis of some early GTO integral methods,²⁶ but the expressions are cumbersome and lengthy. As Rys, Dupuis and King pointed out,¹⁵⁸ for the quadrature

scheme all that is required is the value of the polynomial $\sum_{n=0}^{j_{\max}} c_n t_k^{2n}$ at the zeroes t_k of the Rys polynomial $P_m(t^2, X)$. They utilized an identity similar to that of Eq. (127) to arrive at an expression for the final integral that is factorized into Cartesian directions:

$$g_{abcd} = \sqrt{\frac{4\alpha}{\pi}} \sum_k^m I_x(t_k) I_y(t_k) I_z(t_k) W_k, \quad (308)$$

where

$$\alpha = \frac{(a+b)(c+d)}{a+b+c+d} \quad (309)$$

involves only the exponents of the original four Gaussians. Their “two-dimensional” integrals I_x , etc., can be evaluated by a recursive scheme, as described in Ref. 158. As we indicated above, they can also be evaluated as a expansion in Hermite Gaussians.⁵

In order to use this as the basis for a practical scheme for Gaussian integrals, the quadrature roots and weights must be obtained. This replaces the calculation of incomplete gamma function values in the McMurchie-Davidson and other earlier schemes. King and Dupuis¹⁵⁷ gave a very extensive discussion of determining the roots and weights in 1976, and additional work has been performed by others.^{145,159} As a result, the roots and weights can be obtained very efficiently and accurately. Rys quadrature can obviously also be applied to the calculation of differentiated integrals (see, for example, Refs. 160 and 161), and to fine-structure integrals.¹⁶²

16. Molecular Symmetry and Integrals

The exploitation of any symmetry present in a molecule can provide very significant reductions in the computational effort. This affects the calculation of CGTO integrals in two ways. First, it is desirable to ensure that only those integrals that are nonredundant by symmetry are actually computed, and that contributions to the energy, Fock matrix, etc., from integrals related to these distinct integrals are evaluated using only the latter. Second, many manipulations involving CGTO or MO integral arrays can be decomposed into the manipulation of subarrays if the integrals are expressed in terms of a symmetry-adapted basis. This leads to the use of symmetry-adapted linear combinations of CGTO’s as the actual one-electron basis functions. We will sketch qualitatively here some aspects of

these two symmetry issues. For full details, the interested reader is referred to a recent review.¹⁶³

16.1. Symmetry-Distinct Integrals over CGTO's

Let the molecular symmetry group be denoted \mathcal{G} , with elements G and order g . We assume that the atomic basis χ is closed under \mathcal{G} . This is not usually a restriction in practice, since, as we have repeatedly observed, shells of functions of a given angular type are invariably used in molecular calculations. We thus have

$$G\chi = \chi G, \quad (310)$$

where the matrix \mathbf{G} provides a reducible, nonunitary representation of \mathcal{G} . \mathbf{G} has a relatively simple block structure, since each shell of basis functions on a given center A will map into an identical shell on the image center $G(A)$ under the operation G .

Consider now as an example a *totally symmetric* operator O , that is, an operator for which

$$G^{-1}OG = O \quad \forall G \in \mathcal{G}. \quad (311)$$

In terms of the matrix of elements of O over a given AO basis we have

$$\mathbf{G}^\dagger \mathbf{O} \mathbf{G} = \mathbf{O}, \quad (312)$$

where \mathbf{G} is given by Eq. (310). Let M and N be arbitrary shells of AO's: the block of \mathbf{O} over these shells is denoted \mathbf{O}_{MN} . Dupuis and King¹⁶⁴ then define a *skeleton matrix* $\bar{\mathbf{O}}_{MN}$ with the property that

$$\bar{\mathbf{O}}_{MN} = \lambda_{MN} \mathbf{O}_{MN}, \quad (313)$$

where λ_{MN} is a proportionality constant. They then proved that \mathbf{O} can be recovered by symmetrization of this skeleton matrix, defined through

$$\mathbf{O} = \bar{\mathbf{O}}_{sym} = (2g)^{-1} \sum_G \mathbf{G}^\dagger (\bar{\mathbf{O}} + \bar{\mathbf{O}}^\dagger) \mathbf{G}. \quad (314)$$

The numerical factor is $(2g)^{-1}$, since in addition to the factor of g from the spatial symmetry the operator is Hermitian and there is a two-fold degeneracy from index permutations.

Clearly, one possible choice for λ_{MN} in Eq. (313) is $\lambda_{MN} = 1$ for all MN . The symmetrization in Eq. (314) is then trivial. But we can do much better than this. Consider the elements of \mathcal{G} : some operations will leave the shell pair MN fixed in space, while some will generate a new shell pair on transformed centers. We enumerate the subset of the elements of \mathcal{G} that generate new shell pairs $G(MN)$, that is, shell pairs that are distinct from MN and from one another, and denote the dimension of this subset by $q_2(MN)$. There are thus $q_2(MN)$ shell pairs that are distinct by symmetry. Let $M'N'$ label shell pairs that are *equivalent* to MN by symmetry. Requiring that

$$\sum_{(M'N')} \lambda_{M'N'} = q_2(MN), \quad (315)$$

we can see that another choice for λ_{MN} can be based on

$$\lambda_{M'N'} = q_2(MN) \delta_{MM'} \delta_{NN'}. \quad (316)$$

In this case, the skeleton matrix comprises only those blocks $\bar{\mathbf{O}}_{MN}$ that are distinct by symmetry, multiplied by the factor $q_2(MN)$. Such an array is termed a *petite list* by Dupuis and King,¹⁶⁴ as opposed to the *grande list* that comprises all elements, symmetry-distinct or not. The petite and grande one-electron integral lists are denoted $P2$ and $G2$ respectively. Exactly the same general strategy can be employed for the two-electron integral list: reducing the grande list $G4$ to a $P4$ petite list, whose elements will be weighted by a factor $q_4(MNPQ)$ defined analogously to $q_2(MN)$ (see also Ref. 165).

How is the petite list constructed? While powerful group-theoretical methods can be employed,¹⁶⁶ a very simple strategy termed the “principle of procrastination”¹⁶⁴ is the easiest approach to implement. It requires only the information as to how each shell behaves under the operations of the point group. A loop over (all) products of shells generates overlap distributions, and for each distribution a list of image overlap distributions for all elements G of \mathcal{G} is examined. If any of the image distributions is equivalent to an overlap distribution that is later in the $G2$ list of products, the current distribution is skipped. Thus the work of evaluating integrals for a particular shell product is put off as late as possible. A similar strategy can be used for generating the $P4$ list.

Using petite lists obviously yields a very significant reduction in the work required to generate and manipulate integrals when symmetry is present. The symmetrization formula Eq. (314) we presented was defined for totally symmetric operators, which covers not only the overlap and one-electron Hamiltonian matrices of Eqs. (16) and (17), but also the Fock matrix of Eq. (15). The formula can also be extended without difficulty to the case of nontotally symmetric operators that do not satisfy Eq. (311) (Ref. 167). Typically, the use of symmetry in this way reduces the number of AO integrals that must be calculated by approximately the order of the point group: this not only reduces the computational effort required to generate or manipulate the integral list, but also reduces the storage requirements by the same factor.

16.2. Symmetry-Adapted Basis Functions

A step beyond the use of petite lists of AO integrals is to explicitly employ symmetry-adapted linear combinations of atomic orbitals as a basis — *symmetry orbitals* (SO's). The advantages are again obvious: all succeeding steps of a quantum chemical calculation can be blocked by symmetry, with a reduction in effort in these steps. Of course, there would be little point to using SO's if computing integrals over the SO basis required much greater effort than computing AO integrals, since this would offset the advantages seen later in the calculation. In fact, it is possible to compute SO integrals very efficiently, as we shall now discuss. In this section, we confine ourselves to the point group D_{2h} and its subgroups: Abelian groups with real one-dimensional representations. The strategy we shall discuss can be derived for arbitrary finite point groups, but the formulae are extremely complicated,¹⁶⁶ and computer implementations have mainly been restricted to the case of D_{2h} and its subgroups.

This is another situation where it is advantageous to begin with a discussion of what is *not* done. We define a symmetry orbital by projecting on the AO ψ_a to obtain a basis function for irreducible representation (irrep) α :

$$\psi_a^\alpha = g^{-1} \sum_G \chi^\alpha(G) G\psi_a. \quad (317)$$

Here $\chi^\alpha(G)$ is the character of symmetry operation G in irrep α . Clearly,

a two-electron integral over SO's can be written as

$$g_{abcd}^{\alpha\beta\gamma\delta} = g^{-4} \sum_G \sum_H \sum_K \sum_L \chi^\alpha(G) \chi^\beta(H) \chi^\gamma(K) \chi^\delta(L) \\ \times \left\langle G\psi_a(r_1) H\psi_b(r_1) \left| \frac{1}{r_{12}} \right| K\psi_c(r_2) L\psi_d(r_2) \right\rangle. \quad (318)$$

However, it should be clear that this will not form the basis for a very effective procedure. First, it involves a four-index transformation of blocks of integrals, a step that will behave as g^5 here. Second, the summations over symmetry operations generate all quadruplets of AO indices, whether the integrals are distinct or symmetry-related. Eq. (318), therefore, is not suitable as it stands.

One of the defects of Eq. (318) can be rectified fairly straightforwardly. We note that the integral appearing inside the summations is

$$\left\langle G\psi_a(r_1) H\psi_b(r_1) \left| \frac{1}{r_{12}} \right| K\psi_c(r_2) L\psi_d(r_2) \right\rangle. \quad (319)$$

This is a scalar object, and is thus unaffected by the application of any $G \in \mathcal{G}$. In particular, if we apply G^{-1} (which is identical to G for D_{2h} and its subgroups, of course) we obtain

$$\left\langle \psi_a(r_1) G^{-1} H\psi_b(r_1) \left| \frac{1}{r_{12}} \right| G^{-1} K\psi_c(r_2) G^{-1} L\psi_d(r_2) \right\rangle. \quad (320)$$

However, if H runs over all elements of \mathcal{G} , then so does $G^{-1}H$ for any $G \in \mathcal{G}$. Thus by multiplying through by G^{-1} we arrive at the formula

$$g_{abcd}^{\alpha\beta\gamma\delta} = g^{-4} \mathcal{M}^{\alpha\beta\gamma\delta} \sum_H \sum_K \sum_L \chi^\beta(H) \chi^\gamma(K) \chi^\delta(L) \\ \times \left\langle \psi_a(r_1) H\psi_b(r_1) \left| \frac{1}{r_{12}} \right| K\psi_c(r_2) L\psi_d(r_2) \right\rangle. \quad (321)$$

Here the quantity $\mathcal{M}^{\alpha\beta\gamma\delta}$ incorporates numerical factors and selection rules arising from the summation over G in Eq. (318). The symmetrization procedure of Eq. (321) no longer involves a four-index transformation, reducing the work required by a factor of g . This is more satisfactory, but we still have not addressed the second problem with Eq. (318): the occurrence of

AO integrals that are redundant by symmetry. The reduction from four summations to three helps to some extent, but not completely.

What is needed, in effect, is a method for reducing the symmetrization formula to involve only distinct integrals — the P4 list of Dupuis and King¹⁶⁴ — and for obtaining the numerical factors ($\mathcal{M}^{\alpha\beta\gamma\delta}$, $\chi^\beta(H)$, etc.) with which a given AO integral contributes to the SO integrals. Pitzer obtained a formula for this,¹⁶⁸ in work that was later extended by Davidson using *double coset decompositions*.¹⁶⁶ In general, we can perceive that an overlap distribution from two AO's $\psi_a\psi_b$ will be invariant to some symmetry operations in \mathcal{G} . Such operations form a subgroup of \mathcal{G} , of order h , say. There must then be g/h symmetry operations that generate distinct distributions from $\psi_a\psi_b$. This quantity is the same as the q_2 values in the treatment of Dupuis and King. A list of these nontrivial operations forms a set of *double coset representatives* for the overlap distribution. Of course, a two-electron integral comprises two overlap distributions. To obtain the shortest list of AO integrals we must determine double coset representatives not only for each overlap distribution, but also for the product of the two overlap distributions.¹⁶⁶ Thus we have three sets of double coset representatives, and hence three sums of symmetry operations, consistent with our discussion of Eq. (321). The final expression for a two-electron integral over SO's can be written as

$$g_{abcd}^{\alpha\beta\gamma\delta} = \sum_R \sum_S \sum_T \mathcal{J}_{\alpha\beta\gamma\delta}^{RST} \left\langle \psi_a(r_1) R \psi_b(r_1) \left| \frac{1}{r_{12}} \right| T \psi_c(r_2) T S \psi_d(r_2) \right\rangle. \quad (322)$$

Here R , S , T denote operators from three different sets of double coset representatives: R generates distinct overlap distributions from $\psi_a\psi_b$; S generates distinct distributions from $\psi_c\psi_d$; and T generates distinct products of $\psi_a R \psi_b$ and $\psi_c S \psi_d$. This corresponds exactly to elements of the P4 list. Finally, $\mathcal{J}_{\alpha\beta\gamma\delta}^{RST}$ contains various numerical factors and the selection rule (that the product of all four irreps must contain the totally symmetric irrep). Complete details are given by Davidson¹⁶⁶ and by Taylor.¹⁶³

16.3. Computational Implementations

The implementation of the petite list-based method is evidently rather straightforward. The only information required is the behavior of different basis functions (or shells) under the operations of the group, and this is easily pretabulated for the one-center case, to be combined with information

as to how individual centers are transformed as required. The symmetry behavior of the shells is then used to apply the principle of procrastination, and the shell blocks of distinct integrals are computed according to this principle. This requires essentially trivial modification of the loop structure of the integral program. Finally, when arrays such as the skeleton Fock matrix have been computed, they may be symmetrized according to Eq. (314). Again, once the transformation properties of the shells are available this becomes trivial. The petite list methods have also been extended to cover the general case of transforming integrals from an AO to a (symmetry-adapted) MO basis.^{169,170}

The method of double coset decompositions is somewhat more demanding of computational implementation, but not much. Once again, information about the symmetry properties of basis functions (or shells) and centers is required. In addition, some method of implementing the group algebra, multiplying symmetry operations and calculating products of irreps is required. Almlöf¹²⁷ pointed out some time ago that for the case of D_{2h} and its subgroups, the symmetry operations and their multiplication properties could be represented very efficiently by bit patterns and bitwise (logical) operations like exclusive OR. The computer time and storage requirements for these group-theoretical activities then become completely negligible. The loop structure of the integral program is modified to involve loops over shells, and then loops over the three sets of double coset representatives appropriate for each shell quadruplet. The numerical factors required for the contributions of each distinct shell quadruplet to the SO integrals are easily found. It may also be noted that since the generation of SO integrals is normally done for shell blocks, there are several possibilities for implementing this step in terms of vector operations. Again, the overheads associated with generating SO integrals is not high. On the other hand, it should be obvious that, at least for multicenter SO's, there will be fewer integrals that are small numerically (that is, for reasons related to distance rather than to symmetry). For SCF calculations on large systems (especially direct SCF calculations) it may be preferable to work in terms of the petite list of AO integrals rather than the list of SO integrals, since the former can exploit numerically small integrals better than the latter.

Whichever method is used to exploit symmetry in an integral program, it adds little or no overhead to the calculation in cases where there is no symmetry, and leads to very significant gains in the integral calculation and

later steps when symmetry is present. It is thus hard to see a good reason against the inclusion of a symmetry treatment.

16.4. Differentiated Integrals and Symmetry

As we noted in Secs. 6.2 and 14.1, for the purposes of calculating molecular properties efficiently, it is necessary to compute integrals that are differentiated with respect to some parameters (such as the coordinates of the nuclear centers). Both of the above strategies for handling symmetry and integral calculations can be extended to cover these cases. For the gradient of the energy, only totally symmetric elements of the gradient vector can be nonzero, so the symmetrization procedure for the petite list/skeleton matrix approach can be applied directly.¹⁶⁰ For higher derivatives, other less symmetric quantities appear, since information concerning differentiation with respect to nontotally symmetric coordinates is required. Different $P4$ lists will be associated with different differentiation variables.¹⁶¹ Where SO's are used, it will usually be advantageous to use symmetry-adapted differentiation variables (e.g., symmetry-adapted nuclear coordinates). Treatments based on double coset decompositions can then be rather easily extended to generate lists of symmetry-distinct differentiated integrals and to combine them to yield differentiated SO integrals.¹⁷¹

As a final note, we comment on the argument occasionally heard that “even symmetric molecules have no symmetry when they vibrate.” This comment is becoming less and less relevant as analytical derivative methods are used to compute frequencies. Analytical derivatives are computed at the equilibrium geometry, and therefore the calculation displays the symmetry of that geometry. Just as for the case of symmetry and integrals themselves, it is hard to see a good reason against including symmetry in the calculation of differentiated integrals.

17. Conclusions

We have reviewed at some length the reasons behind particular choices of Gaussian basis functions and the evaluation of molecular integrals over these functions. Our intention has been to provide a detailed background so that the interested reader can explore these topics in more depth in the literature. The choice of basis set determines at the outset of a calculation how reliable the results will be, and is therefore deserving of considerable attention.

We have emphasized on several occasions that the two-electron correlation cusp is treated poorly by expansions in products of one-electron functions. In fact, for small molecules, at least, we may already be approaching the limits of what can be accomplished using one-electron basis functions, given the slow convergence of the correlation energy with angular quantum numbers. The alternative to traditional one-electron basis sets is to consider two-particle basis functions that explicitly include factors in r_{ij} in the wave function. Methods that use two-electron basis functions are beyond the scope of this work, but we mention here recent work by Kutzelnigg and coworkers¹⁷²⁻¹⁷⁴ that shows considerable promise as a way of better describing the two-electron cusp. The integrals required for their method can all be derived using the techniques we have described here.¹⁷⁵

Acknowledgments

Many helpful discussions with colleagues too numerous to list here are gratefully acknowledged.

This work was supported in part by National Science Foundation Cooperative Agreement DASC-8902825, and by a grant of computer time from the San Diego Supercomputer Center.

References

1. E. R. Davidson and D. Feller, *Chem. Rev.* **86**, 681 (1986).
2. D. Feller and E. R. Davidson, in *Reviews in Computational Chemistry*, eds. K. B. Lipkowitz and D. B. Boyd (VCH, New York, 1990).
3. R. Poirier, R. Kari and I. G. Csizmadia, *Handbook of Gaussian Basis Sets* (Elsevier, Amsterdam, 1985).
4. V. R. Saunders, in *Computational Methods in Quantum Chemistry and Molecular Physics*, eds. G. H. F. Diercksen, B. T. Sutcliffe and A. Veillard (D. Reidel, Dordrecht, 1975).
5. V. R. Saunders, in *Methods in Computational Molecular Physics*, eds. G. H. F. Diercksen and S. Wilson (D. Reidel, Dordrecht, 1983).
6. I. M. Sneddon, *Special Functions of Mathematical Physics and Chemistry* (Oliver and Boyd, Edinburgh, 1961).
7. R. G. Parr, *The Quantum Theory of Molecular Electronic Structure* (Benjamin, New York, 1964).
8. E. Hylleraas, *Zeit. Phys.* **48**, 469 (1928).
9. H. M. James and A. S. Coolidge, *J. Chem. Phys.* **1**, 825 (1933).
10. R. Ahlrichs, M. Hoffmann-Ostenhof, T. Hoffman-Ostenhof and J. D. Morgan, *Phys. Rev.* **A23**, 2106 (1981).
11. P.-O. Löwdin, *Adv. Chem. Phys.* **2**, 207 (1959).

12. C. C. J. Roothaan, *Rev. Mod. Phys.* **23**, 69 (1951).
13. G. G. Hall, *Proc. Roy. Soc. A* **230**, 541 (1951).
14. L. C. Allen and A. M. Karo, *Rev. Mod. Phys.* **32**, 275 (1960).
15. P. E. Cade and W. H. Huo, *J. Chem. Phys.* **44**, 1973 (1967).
16. A. D. McLean and M. Yoshimine, *Tables of Linear Molecule Wave Functions* (IBM Research Laboratory, San Jose, 1967).
17. S. F. Boys, *Proc. Roy. Soc. A* **200**, 542 (1950).
18. S. F. Boys, G. B. Cook, C. M. Reeves and I. Shavitt, *Nature* **178**, 1207 (1956).
19. H. Preuss, *Z. Naturforsch.* **A11**, 323 (1956).
20. E. R. Davidson, *Reduced Density Matrices in Quantum Chemistry* (Academic, London, 1976).
21. P.-O. Löwdin and H. Shull, *Phys. Rev.* **101**, 1730 (1956).
22. W. Kutzelnigg, *Theoret. Chim. Acta* **68**, 445 (1985).
23. F. Sasaki and M. Yoshimine, *Phys. Rev. A* **9**, 17 (1974).
24. S. Huzinaga, D. McWilliams and B. Domsky, *J. Chem. Phys.* **54**, 2283 (1971).
25. D. Davis and E. Clementi, *J. Comput. Phys.* **1**, 223 (1966).
26. H. Taketa, S. Huzinaga and K. O-Ohata, *J. Phys. Soc. Japan* **21**, 2313 (1966).
27. S. Huzinaga, *J. Chem. Phys.* **42**, 1293 (1965).
28. R. F. Stewart, *J. Chem. Phys.* **52**, 431 (1970).
29. W. J. Hehre, R. F. Stewart and J. A. Pople, *J. Chem. Phys.* **51**, 2657 (1969).
30. K. Fægri and J. Almlöf, *J. Comput. Chem.* **7**, 396 (1986).
31. G. Tonachini and H. B. Schlegel, *J. Chem. Phys.* **87**, 514 (1987).
32. A. Schäfer, H. Horn and R. Ahlrichs, *J. Chem. Phys.* **97**, 2571 (1992).
33. F. B. Van Duijneveldt, *Gaussian Basis Sets for the Atoms H–Ne for Use in Molecular Calculations*, RJ945 (IBM Research Laboratory, San Jose, 1971).
34. S. Huzinaga, *Approximate Atomic Wave Functions. I* (University of Alberta, Edmonton, 1971).
35. S. Huzinaga, *Approximate Atomic Wave Functions. II* (University of Alberta, Edmonton, 1971).
36. R. C. Raffenetti, *J. Chem. Phys.* **58**, 4452 (1973).
37. T. H. Dunning, *J. Chem. Phys.* **66**, 1382 (1977).
38. M. W. Schmidt and K. Ruedenberg, *J. Chem. Phys.* **71**, 3951 (1979).
39. J. Almlöf and P. R. Taylor, *Adv. Quantum Chem.* **22**, 301 (1992).
40. R. D. Bardo and K. Ruedenberg, *J. Chem. Phys.* **59**, 5956 (1973).
41. D. Feller and R. Ruedenberg, *Theoret. Chim. Acta* **52**, 231 (1979).
42. S. Huzinaga, M. Klobukowski and H. Tatewaki, *Can. J. Chem.* **63**, 1812 (1985).
43. R. Ahlrichs and P. R. Taylor, *J. Chim. Physique* **78**, 315 (1981).
44. R. Krishnan, J. S. Binkley, R. Seeger and J. A. Pople, *J. Chem. Phys.* **72**, 650 (1980).
45. J. Almlöf and P. R. Taylor, *J. Chem. Phys.* **86**, 4070 (1987).

46. T. H. Dunning, *J. Chem. Phys.* **90**, 1007 (1989).
47. SIRIUS is a second-order MCSCF program written by H. J. Jensen, H. Ågren and J. Olsen; ABACUS is an MCSCF energy derivatives program written by T. Helgaker, K. L. Bak, H. J. Jensen, P. Jørgensen, R. Kobayashi, H. Koch, K. Mikkelsen, J. Olsen, K. Ruud, P. R. Taylor and O. Vahtras.
48. W. S. Benedict, N. Gailar and E. K. Plyler, *J. Chem. Phys.* **24**, 1139 (1956).
49. A. A. Frost, in *Modern Theoretical Chemistry*, ed. H. F. Schaefer (Plenum, New York, 1977).
50. J. S. Binkley, J. A. Pople and W. J. Hehre, *J. Am. Chem. Soc.* **102**, 939 (1980).
51. R. Ditchfield, W. J. Hehre and J. A. Pople, *J. Chem. Phys.* **54**, 724 (1971).
52. W. J. Hehre, R. Ditchfield and J. A. Pople, *J. Chem. Phys.* **56**, 2257 (1972).
53. T. H. Dunning, *J. Chem. Phys.* **53**, 2823 (1970).
54. H. Tatewaki and S. Huzinaga, *J. Chem. Phys.* **71**, 4339 (1979).
55. H. Tatewaki, Y. Sakai and S. Huzinaga, *J. Comput. Chem.* **2**, 278 (1981).
56. B. Roos and P. Siegbahn, *Theoret. Chim. Acta* **17**, 199 (1970).
57. T. H. Dunning, *J. Chem. Phys.* **55**, 3958 (1971).
58. P. C. Hariharan and J. A. Pople, *Theoret. Chim. Acta* **28**, 213 (1973).
59. A. D. McLean and G. S. Chandler, *J. Chem. Phys.* **72**, 5639 (1980).
60. P. J. Hay, *J. Chem. Phys.* **66**, 4377 (1977).
61. A. J. H. Wachters, *J. Chem. Phys.* **52**, 1033 (1970).
62. B. Roos, A. Veillard and G. Vinot, *Theoret. Chim. Acta* **20**, 1 (1971).
63. S. P. Walch and C. W. Bauschlicher, *Chem. Phys. Lett.* **105**, 171 (1984).
64. H. Partridge, *J. Chem. Phys.* **87**, 6643 (1987).
65. H. Partridge, *J. Chem. Phys.* **90**, 1043 (1989).
66. H. Partridge and K. Fægri, *Theoret. Chim. Acta* **82**, 207 (1992).
67. B. O. Roos, *Int. J. Quantum Chem., Quantum Chem. Symp.* **14**, 175 (1980).
68. L. T. Redmon, G. D. Purvis and R. J. Bartlett, *J. Am. Chem. Soc.* **101**, 2856 (1979).
69. K. Jankowski, R. Becherer, P. Scharf, H. Schiffer and R. Ahlrichs, *J. Chem. Phys.* **82**, 1413 (1985).
70. K. Raghavachari, G. W. Trucks, J. A. Pople and M. Head-Gordon, *Chem. Phys. Lett.* **157**, 479 (1989).
71. MOLECULE-SWEDEN is an electronic structure program system written by J. Almlöf, C. W. Bauschlicher, M. R. A. Blomberg, D. P. Chong, A. Heiberg, S. R. Langhoff, P.-Å. Malmqvist, A. P. Rendell, B. O. Roos, P. E. M. Siegbahn and P. R. Taylor.
72. TITAN is a set of electronic structure programs written by T. J. Lee, A. P. Rendell and J. E. Rice.
73. C. W. Bauschlicher and P. R. Taylor, *Theoret. Chim. Acta* **74**, 63 (1988).
74. J. Almlöf and P. R. Taylor, *J. Chem. Phys.* **92**, 551 (1990).
75. C. W. Bauschlicher, *Chem. Phys. Lett.* **142**, 71 (1987).
76. S. R. Langhoff, C. W. Bauschlicher and P. R. Taylor, *J. Chem. Phys.* **88**, 5715 (1988).

77. P.-O. Widmark, P.-Å. Malmqvist and B. O. Roos, *Theoret. Chim. Acta* **77**, 291 (1990).
78. P.-O. Widmark, B. J. Persson and B. O. Roos, *Theoret. Chim. Acta* **79**, 419 (1991).
79. C. W. Bauschlicher and P. R. Taylor, *Theoret. Chim. Acta* **86**, 13 (1993).
80. J. M. L. Martin, *J. Chem. Phys.* **97**, 5012 (1992).
81. P. R. Taylor, in *Lecture Notes in Quantum Chemistry*, ed. B. O. Roos (Springer, Berlin, 1992).
82. C. W. Bauschlicher, S. R. Langhoff, H. Partridge and L. A. Barnes, *J. Chem. Phys.* **91**, 2399 (1989).
83. V. Parasuk and J. Almlöf, *J. Chem. Phys.* **91**, 1137 (1989).
84. J. Almlöf, B. J. DeLeeuw, P. R. Taylor, C. W. Bauschlicher and P. Siegbahn, *Int. J. Quantum Chem., Quantum Chem. Symp.* **23**, 345 (1989).
85. R. D. Amos, *Adv. Chem. Phys.* **67**, 99 (1987).
86. P. Pulay, in *Modern Theoretical Chemistry*, ed. H. F. Schaefer (Plenum, New York, 1977).
87. B. O. Roos and A. J. Sadlej, *Chem. Phys. Lett.* **94**, 43 (1985),
88. T. Helgaker and P. Jørgensen, *J. Chem. Phys.* **95**, 2595 (1991).
89. A. J. Sadlej, *Chem. Phys. Lett.* **47**, 50 (1977).
90. M. Zeitz, S. D. Peyerimhoff and R. J. Buenker, *Chem. Phys. Lett.* **64**, 243 (1979).
91. T. Helgaker and J. Almlöf, *J. Chem. Phys.* **89**, 4889 (1988).
92. G. Fogarasi and P. Pulay, *Ann. Rev. Phys. Chem.* **35**, 191 (1984).
93. H. P. Huber, *Theoret. Chim. Acta* **55**, 117 (1980).
94. H. Nakatsuji, M. Hada and T. Yonezawa, *J. Chem. Phys.* **77**, 3109 (1982).
95. J. Almlöf and T. Helgaker, *Chem. Phys. Lett.* **83**, 125 (1981).
96. H.-J. Werner and W. Meyer, *Mol. Phys.* **31**, 855 (1976).
97. D. Feller, C. M. Boyle and E. R. Davidson, *J. Chem. Phys.* **86**, 3424 (1987).
98. P. R. Taylor, T. J. Lee, J. E. Rice and J. Almlöf, *Chem. Phys. Lett.* **163**, 359 (1989).
99. M. A. Spackman, *J. Phys. Chem.* **93**, 7594 (1989).
100. R. A. Kendall, T. H. Dunning and R. J. Harrison, *J. Chem. Phys.* **96**, 6796 (1992).
101. R. Ditchfield, *J. Chem. Phys.* **56**, 5688 (1972).
102. K. Wolinski, J. F. Hinton and P. Pulay, *J. Am. Chem. Soc.* **112**, 8521 (1990).
103. W. Kutzelnigg, *Israel J. Chem.* **19**, 193 (1980).
104. M. Schindler and W. Kutzelnigg, *J. Chem. Phys.* **76**, 1919 (1982).
105. A. E. Hansen and T. D. Bouman, *J. Chem. Phys.* **82**, 5035 (1985).
106. T. D. Bouman and A. E. Hansen, *Chem. Phys. Lett.* **197**, 59 (1992).
107. P. J. Stephens, *J. Phys. Chem.* **91**, 1712 (1987).
108. K. Ruud, T. Helgaker, K. L. Bak, P. Jørgensen and H. J. Jensen, *J. Chem. Phys.* **99**, 3847 (1993).
109. K. L. Bak, P. Jørgensen, T. Helgaker, K. Ruud and H. J. Jensen, *J. Chem. Phys.* **98**, 8873 (1993).

110. K. L. Bak, P. Jørgensen, T. Helgaker, K. Ruud and H. J. Jensen, *J. Chem. Phys.* **100**, 6620 (1994).
111. R. Weast, *CRC Handbook of Chemistry and Physics, 64th Edition* (CRC Press, Boca Raton, Florida, 1983).
112. W. T. Raynes, *Nucl. Mag. Res.* **7**, 1 (1977).
113. R. E. Wasylishen, S. Mooibroek and J. B. Macdonald, *J. Chem. Phys.* **81**, 1057 (1984).
114. W. B. Jenkins and W. McFarlane, *J. Chem. Soc., Chem. Commun.* **1977**, 922 (1977).
115. O. Vahtras, H. Ågren, P. Jørgensen, H. J. Jensen and T. Helgaker, *Chem. Phys. Lett.* **209**, 201 (1993).
116. S. F. Boys and F. Bernardi, *Mol. Phys.* **19**, 553 (1970).
117. B. Liu and A. D. McLean, *J. Chem. Phys.* **91**, 2348 (1989).
118. C. W. Bauschlicher, S. P. Walch, S. R. Langhoff, P. R. Taylor and R. L. Jaffe, *J. Chem. Phys.* **88**, 1743 (1988).
119. M. Gutowski, J. H. Van Lenthe, J. Verbeek, F. B. Van Duijneveldt and G. Chalasinski, *Chem. Phys. Lett.* **124**, 370 (1986).
120. C. W. Bauschlicher, S. R. Langhoff and P. R. Taylor, *J. Chem. Phys.* **88**, 2540 (1988).
121. S. Wilson, *The Effects of Relativity in Atoms, Molecules and the Solid State* (Plenum, New York, 1991).
122. C. W. Bauschlicher, *Chem. Phys. Lett.* **122**, 572 (1985).
123. J. M. L. Martin, J. P. François and R. Gijbels, *J. Comput. Chem.* **10**, 152 (1989).
124. J. M. L. Martin, J. P. François and R. Gijbels, *J. Comput. Chem.* **10**, 875 (1989).
125. J. M. L. Martin, J. P. François and R. Gijbels, *Theoret. Chim. Acta* **76**, 195 (1989).
126. G. Van der Velde, Ph. D. thesis, Groningen, 1974.
127. The MOLECULE integral program, J. Almlöf, University of Stockholm Institute of Physics Report 74-29 (1974).
128. J. A. Pople and W. J. Hehre, *J. Comput. Phys.* **27**, 161 (1978).
129. M. Dupuis, J. Rys and H. F. King, *J. Chem. Phys.* **65**, 111 (1976).
130. L. E. McMurchie and E. R. Davidson, *J. Comput. Phys.* **26**, 218 (1978).
131. T. Živković and Z. B. Maksić, *J. Chem. Phys.* **49**, 3083 (1968).
132. S. Obara and A. Saika, *J. Chem. Phys.* **84**, 3963 (1986).
133. Z. B. Maksić, K. Kovačević and M. Primorac, *Pure and Appl. Chem.* **61**, 2075 (1989).
134. A. Golebiewski and J. Mrozek, *Int. J. Quantum Chem.* **7**, 623 (1973).
135. A. Golebiewski and J. Mrozek, *Int. J. Quantum Chem.* **7**, 1023 (1973).
136. R. Colle and O. Salvetti, *Theoret. Chim. Acta* **53**, 55 (1979).
137. G. Arfken, *Mathematical Methods of Physics* (Academic, San Diego, 1985).
138. M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions* (Dover, New York, 1970).

139. R. Ahlrichs, *Theoret. Chim. Acta* **33**, 157 (1974).
140. M. Head-Gordon and J. A. Pople, *J. Chem. Phys.* **89**, 5777 (1988).
141. P. M. W. Gill, M. Head-Gordon and J. A. Pople, *J. Phys. Chem.* **94**, 5564 (1990).
142. P. Gill and J. Pople, *Int. J. Quantum Chem.* **40**, 753 (1991).
143. B. G. Johnson, P. M. W. Gill and J. A. Pople, *Chem. Phys. Lett.* **206**, 229 (1993).
144. T. P. Hamilton and H. F. Schaefer, *Chem. Phys.* **150**, 163 (1991).
145. R. Lindh, U. Ryu and B. Liu, *J. Chem. Phys.* **95**, 5889 (1991).
146. M. Häser and R. Ahlrichs, *J. Comput. Chem.* **10**, 104 (1989).
147. M. Häser, R. Ahlrichs, H. P. Baron, P. Weis and H. Horn, *Theoret. Chim. Acta* **83**, 455 (1992).
148. T. Helgaker and P. Jørgensen, *Adv. Quantum Chem.* **19**, 183 (1988).
149. P. W. Abegg, *Mol. Phys.* **30**, 579 (1975).
150. A. Komornicki, K. Ishida, K. Morokuma, R. Ditchfield and M. Conrad, *Chem. Phys. Lett.* **45**, 595 (1977).
151. K. Ishida, K. Morokuma and A. Komornicki, *J. Chem. Phys.* **66**, 2153 (1977).
152. L. R. Kahn, *J. Chem. Phys.* **75**, 3962 (1981).
153. M. A. Vincent, P. Saxe and H. F. Schaefer, *Chem. Phys. Lett.* **94**, 351 (1983).
154. T. Helgaker and P. R. Taylor, *Theoret. Chim. Acta* **83**, 177 (1992).
155. H. B. Schlegel, *J. Chem. Phys.* **77**, 3676 (1982).
156. O. Vahteras, H. Ågren, P. Jørgensen, H. J. Jensen, T. Helgaker and J. Olsen, *J. Chem. Phys.* **96**, 2118 (1992).
157. H. F. King and M. Dupuis, *J. Comput. Phys.* **21**, 144 (1976).
158. J. Rys, M. Dupuis and H. F. King, *J. Comput. Chem.* **4**, 154 (1983).
159. R. P. Sagar and V. H. Smith, *Int. J. Quantum Chem.* **42**, 827 (1992).
160. M. Dupuis and H. F. King, *J. Chem. Phys.* **68**, 3998 (1978).
161. T. Takada, M. Dupuis and H. F. King, *J. Chem. Phys.* **76**, 332 (1981).
162. H. F. King and T. R. Furlani, *J. Comput. Chem.* **9**, 771 (1988).
163. P. R. Taylor, in *Lecture Notes in Quantum Chemistry*, ed. B. O. Roos (Springer, Berlin, 1992).
164. M. Dupuis and H. F. King, *Int. J. Quantum Chem.* **11**, 613 (1977).
165. P. D. Dacre, *Chem. Phys. Lett.* **7**, 47 (1970).
166. E. R. Davidson, *J. Chem. Phys.* **62**, 400 (1975).
167. P. R. Taylor, *Int. J. Quantum Chem.* **27**, 89 (1985).
168. R. M. Pitzer, *J. Chem. Phys.* **58**, 3111 (1973).
169. M. Häser, J. Almlöf and M. W. Feyereisen, *Theoret. Chim. Acta* **79**, 115 (1991).
170. E. Hollauer and M. Dupuis, *J. Chem. Phys.* **96**, 5220 (1992).
171. P. R. Taylor, *Theoret. Chim. Acta* **69**, 447 (1986).
172. W. Kutzelnigg and W. Klopper, *J. Chem. Phys.* **94**, 1985 (1991).

173. V. Termath, W. Klopper and W. Kutzelnigg, *J. Chem. Phys.* **94**, 2002 (1991).
174. W. Klopper and W. Kutzelnigg, *J. Chem. Phys.* **94**, 2020 (1991).
175. W. Klopper and R. Rohse, *Theoret. Chim. Acta* **83**, 441 (1992).