A. D. McLean

M. Yoshimine

Computation of Molecular Properties and Structure

Abstract: A discussion of general-purpose computer programs in theoretical chemistry is given, followed by a description of the procedures adopted in one such program written by the authors. Specific details on the use of the program for computing molecular wave functions and properties for closed-shell linear molecules are presented. The details of a method for computing the axial components of the static electric polarizability and shielding factor tensors are given. A "Table of Linear Molecule Wave Functions" is available, on request to the authors, as a supplement to the paper. This tabulation was made with the program described in the paper and is the most extensive compilation of molecular wave functions currently available.

Introduction

The quantitative characteristics of any chemical process, with the exception of nuclear phenomena, can in principle be derived from the masses and charges of the participating nuclei and electrons, using the laws of quantum mechanics and statistics. This is true at least to the accuracy currently achievable in experimental observation. Until the advent of the electronic computer, the effective utilization of this powerful theory to provide information on molecular systems, *ab initio*, was impossible. The most significant computation of this precomputer era was that of James and Coolidge on the hydrogen molecule,[1] which gave a molecular dissociation energy 0.02 eV (161 cm^{-1}) smaller than observed. This result was hailed by chemists as a convincing demonstration of the validity of quantum mechanics to solve problems of molecular structure. Apart from this triumph, and a limited number of less impressive results (but more difficult computations) on other very small systems, the literature through the mid 1950's is essentially devoid of reliable quantitative predictions of molecular properties through computation. The application of electronic computers to chemical problems initiated a revolution, the preliminary effects of which are currently appearing. These include the semiquantitative prediction of properties of molecules and radicals not accessible to measurement, detailed analysis of the nature of chemical bonding, and the determination of energy surfaces for the study of chemical reactions, all from *ab initio* computation.[2–4]

While work in all these areas dates back 30 or more years, *ab initio* computations accurate enough to provide reliable information belong to the age of the computer. Even so, nearly all work so far completed on systems more complex than H_2, has not reached the level of accuracy of the early calculation of James and Coolidge. However, on the basis of work currently in progress, and the projected speeds of future computers, we can confidently predict that work of comparable accuracy on many hundreds of molecular systems with up to 20 electrons will be available by 1975.

To date, the H_2 molecule is still the most complex system on which a convincing demonstration of the validity of quantum-mechanical principles has been attempted in a molecular computation. This careful work of Kolos, Roothaan and Wolniewicz[5,6] has extended the calculations of James and Coolidge to an accuracy such that, as the work currently stands, a discrepancy between theory and experiment has been revealed. (The computed dissociation energy is approximately 3 cm^{-1} larger than the observed, and exhaustive analysis indicates this to be outside the range of any errors or corrections considered so far.[6]) The situation is potentially as important as the difference between theory and experiment in the hyperfine structure of atomic spectra, due to the anomalous magnetic moment of the electron, which led to important advances in quantum electrodynamics.[7] This is an excellent example of the importance of calculations of extremely high accuracy, even though they are (and will be) restricted to very simple systems. The explanation of this discrepancy for H_2 will probably turn out to be not particularly profound, al-

The authors are located at the IBM Research Laboratory, San Jose, California. Mr. Yoshimine is at present on leave at the General Education Department, Kyushu University, Fukuoka, Japan.

though it is certainly very puzzling. Both theoreticians and experimentalists are carefully analyzing their results in an attempt to resolve it.[8] This level of accuracy is not necessary for carrying out quantitative calculations of most chemical data, where the original accuracy of the James and Coolidge calculation with two less significant figures in the computed energy is all that is reasonably desired. This is essentially the limit with which we are concerned in the following discussion.

- *Characteristics of automatic computer programs*

As can be seen by perusing the proceedings of conferences on molecular quantum mechanics,[9,10] real progress in quantitative theoretical chemistry depends on general-purpose computer programs, whose broad characteristics are as follows.

The *input* is at an elemental level, being essentially restricted to the number of nuclei, their charges, and the number of electrons in the system under study. Currently existing programs have not reached this minimum but this is primarily due to restrictions in computer speed. For example, we currently specify the geometry of the nuclei and the electronic configuration and state, because the computation for a single molecular geometry and state of even a small molecule (having less than 20 electrons) runs into many hours of computer time.

The *methods* are those of quantum mechanics at various levels of approximation and, again, the accuracy of the approximations that can be used is a function of computer speed. Most effort to date has been in the single-configuration, self-consistent field molecular orbital approximation, based on the assumption that this would form the best starting point for more accurate procedures which, except for two-electron systems, have been largely restricted to various forms of configuration mixing. Great improvements in the levels of approximation implemented can be expected in the next few years.

The *output* is the wave function and properties of the system under study, examples of which will be given later.

The program must have the characteristics of ease of use (simple input formats) and intelligibility of output, both at the printed and display (visual) level. They must be versatile, in that they can automatically produce results which relate to a wide variety of applications. It must be emphasized that the scope of the programs, meaning the size of the systems that can be studied and the accuracy attainable for a given size, is restricted by the speed of the central processor and the high-speed memory. (Secondarily it is restricted by the total configuration of various elements in the storage hierarchy and the data channels through which data are transferred into high-speed memory, although we assume in these discussions that the total configuration is always adequate to support the central processor.)

The IBM 7094 (2 μs access to high-speed core, 10–20 μs floating point arithmetic operations, 32K high-speed memory, data channels allowing overlap of data transmission with computation) is representative of the class of computers that were first able to adequately support programs of the type described, in the sense that reasonably accurate data can be generated in a reasonable amount of time. The specific program to be described in this paper was in fact written for this computer. The point to be made is that the minimal requirement of general programs in this field is close to the most capable computers even currently available. To significantly extend the scope of these programs requires order-of-magnitude increases in computer speeds, which should be available in the early 1970's. The reason for this requirement is that current techniques for small systems, which are the best designed so far, involve an amount of computation which increases as the fourth or fifth power of the size of the basis set from which the wave function is constructed. Alternative methods which do not have this catastrophic dependence of the amount of computation on the size of the system under study are not efficient enough for small systems to have been implemented in general programs on the computers available to date, even though exploratory programs along these lines have been written. Order-of-magnitude increases in computer speed can require modified algorithms in the implementing of existing techniques and can open up the possibility of new methods, and work in these directions is progressing in a number of laboratories. Factors that would affect computing algorithms are significant changes in the ratio of the cycle time in the central processor to the transmission rate of information from peripheral (tape, disk or drum) storage, or the degree of parallelism (amount of computation that can proceed simultaneously) in the central processing unit.

Possibly the most useful development will be the use of archive storage (characterized by low cost and long access time) to support a library of molecular wave functions and properties. A supervisory program would control the generating and editing of the library and would operate continuously, automatically producing molecular wave functions and properties to increasing levels of accuracy and in a systematic way, from input data taken from the library itself. User intervention could provide input data which override the automatic sequence, or could request display of information from the library. Such programs could run continuously as a back-up job in a time-shared computing system.

To evaluate the potential impact of more powerful computer programs it is instructive to consider the comparison of computed results to measured ones. They are competitive at the levels of ease of obtaining the result, the accuracy of the result, and the cost. They are complementary to the extent that one procedure can obtain results inaccessible to the other. For example, the extremely high accuracy of some experimental techniques, notably spectroscopic and molecular beam, will never be accessible to computation, whereas the detailed accurate information on charge density routinely available from computations will probably continue to be inaccessible to measurement. Computation and measurement can also be complementary in taking some results from both to derive new information, as for example in taking a measured susceptibility with a computed quadrupole moment to derive a rotational magnetic moment. Another element of this discussion is the fact that a transient or unstable species is just as amenable to computation as a stable one.

For some properties, computer programs have already given significantly more information than can be obtained experimentally. On more powerful computers the balance between computation and measurement will be pushed significantly towards computation. Thus, we can anticipate the existence of more powerful computer programs that will create a demand for the computers to support them. It seems quite likely that these programs will prove to be the most important tool in the investigation of molecular structures of the next decade.

- *Methods used in this study*

Returning to the current situation, the purpose of this paper is to describe in some detail the procedures which we have incorporated into a computer program that has been used to significantly advance the level of computation on a number of molecular systems.[11] The description is complete enough for the reader to intelligently use the program, which is being distributed by the Quantum Chemistry Program Exchange at Indiana University. It also includes some procedures which we hope will be routinely incorporated into more advanced programs in the future. While it falls far short of the ultimate program outlined above, which has the computer behave as a molecular system with output immediately intelligible to a non-specialist user, it does incorporate some features that will lie at the heart of such a program. On the basis of a small amount of input data it can evaluate single-configuration, closed-shell ground state wave functions, and a limited number of expectation values computed with these wave functions, for linear molecules. Intermediate results of a previous calculation can be called on in a subsequent one to minimize the amount of computation in the latter. Automatic optimization of various parameters in the calculation can be performed without user intervention. Different phases of the computation are executed by program modules independent of each other, apart from an interface of data. The path of the computation through the program modules is controlled by a short supervisory program easily modified to expand the number of different types of computation which the program can perform. (The path through the program modules for the current computation is set by a single input flag.)

The approximation used for wave functions which can be computed with this program is the self-consistent field matrix Hartree-Fock procedure,[12] which has been elegantly formulated by Roothaan.[13] Thus, the wave function of an electronic state of a molecular system is constructed, from a single configuration of molecular orbitals, as a linear combination of determinants which is an eigenfunction of the total electron spin and one of its components, and also of the covering operations of the symmetry group of the nuclei. (For the closed-shell case, this is a single determinant.) These molecular orbitals are, in turn, expanded from basis functions, which in this program are Slater-type functions having origins on the different nuclei. (These functions are simple products of a polynomial and $\exp(-\zeta r)$, where r is measured from some nucleus and ζ is the so-called orbital exponent of the function.) The accuracy of such wave functions falls far short of the James-Coolidge accuracy discussed earlier and is a severe shortcoming. (For the total energy the error is of the order 1 to 2 eV per electron pair in the molecule.) However, at this point in development, Hartree-Fock wave functions have been useful in many applications and we considered it worth while to significantly expand the range of such calculations.

One extension of the method employed in our computer programs is in the direction of configuration interaction. In particular, configuration interaction expansions in terms of natural orbitals[14] look very promising and may be a practical way of achieving James-Coolidge type accuracy for small systems (less than 20 electrons), although the calculations will need a much faster computer than the IBM 7090. If these more sophisticated wave functions are expanded in the space of a set of Slater basis functions, the major sections of our current program will still be fundamental to future programs. However, for larger systems the use of more complicated basis functions and more direct numerical integration procedures may prove more efficient.[15] In fact, our current work at this laboratory is directed along such lines.

The most significant feature of the computer program discussed in this paper is its capacity, measured in terms of the size of the Slater function basis sets it can handle efficiently. Organizing the computations required the handling of long data lists (10^5 to 10^7 entries) generated by the program and stored on magnetic tape. We were able to do this in a way which essentially completely overlapped computation in the central processing unit with data flow between core and magnetic tape. In other words, we were able to effectively use the computer as an infinite core machine. This can be done only if the average amount of computation per word on the data coming to or from tape through storage buffers in the core (typically assigned capacities of $\sim$500 words) is in excess of the transfer time be-

Table 1 Index to Table of Linear Molecule Wave Functions[a].

Molecule	*Basis*[b] *Set*	*Accuracy*[c] (a.u.)	*Internuclear Separations*[d] (a.u.)	*Page*
FH	BA + P	0.0005	1.7328	1
LiF	DZ + P	0.012	2.85	2
LiF	BA + P	0.0005	2.45; 2.65; 2.7877; 2.8877; 2.9877*; 3.2; 3.55	3
BeO	BA + P	0.0005	1.8; 2.1; 2.35; 2.4377*; 2.5; 2.75; 3.1; 3.8; 5.5	10
BF	DZ + P	0.010	2.391	19
BF	BA + P	0.001	2.0; 2.1; 2.1925; 2.391*; 2.5775; 2.77; 2.9625	20
CO	DZ + P	0.011	2.132	27
CO	BA + P	0.001	1.8; 1.898; 2.015; 2.132*; 2.249; 2.366; 2.483	28
ClH	BA + P	0.001	2.4087	35
LiCl	DZ + P	0.010	3.7228	36
LiCl	BA + P	0.003	3.35; 3.6; 3.66; 3.7228; 3.825*; 3.91; 4.0; 4.1; 4.55	37
NaF	DZ + P	0.020	3.779	46
NaF	BA + P	0.005	3.1; 3.56; 3.62883*; 3.779; 4.35	47
MgO	DZ + P	0.013	2.5; 3.0; 3.2; 3.3052*; 3.4; 3.6; 4.1; 5.1	52
MgO	BA + P	0.002	3.3052	60
AlF	DZ + P	0.014	3.45	61
AlF	BA + P	0.003	2.6; 2.85; 3.05; 3.126*; 3.25; 3.45; 3.7	62
SiO	DZ + P	0.014	2.854	69
SiO	BA + P	0.003	2.304; 2.5; 2.604; 2.75*; 2.854; 3.104; 3.404	70
PN	DZ + P	0.013	2.818	77
PN	BA + P	0.003	2.268; 2.45; 2.568, 2.67*; 2.818; 3.068; 3.368	78
NaCl	DZ + P	0.025	4.4609	85
NaCl	BA + P	0.009	3.7; 4.3; 4.4609*; 4.485; 4.6; 4.75; 5.0	86
KF	DZ + P		4.1035	93
KF	BA + P		3.5; 3.95; 4.04; 4.10348*; 4.15; 4.4; 4.8	94
CaO	DZ + P	0.020	2.9912; 3.1412; 3.2912; 3.4412*; 3.6412; 3.8912; 4.1412	101
LiBr	DZ + P		3.05; 3.55; 3.8; 3.93; 4.0655*; 4.175; 4.26; 4.4; 4.85; 5.6	108
KCl	DZ + P		4.3; 4.7; 5.039*; 5.29; 5.65	128
NaBr	DZ + P		4.728	138
RbF	DZ + P		4.3653	140
SrO	DZ + P		3.1; 3.25; 3.4; 3.525; 3.6283*; 3.78; 4.1	142
HCN	DZ + P	0.007	2.0143, 2.1791	156
HCN	BA + P	0.001	1.81287, 1.96119; 2.00899, 1.76507; 1.93430, 2.1091*; 1.9343, 2.2491; 2.0843, 2.1091; 1.81287, 2.38053; 2.0143, 2.1791; 2.23221, 1.96119; 2.0843, 2.2491; 2.00899, 2.60325; 2.23221 2.38053; 2.47026, 2.14248	157
FHF^-	DZ + P	0.020	1.9, 1.9; 2.0, 2.0; 1.8, 2.3; 1.9, 2.2; 2.0, 2.1; 2.05, 2.05; 1.85, 2.35; 1.95, 2.25; 2.05, 2.15; 2.1, 2.1*; 1.9, 2.4; 2.0, 2.3; 2.1, 2.2; 2.15, 2.15; 2.15, 2.25; 2.2, 2.2; 1.75, 2.75; 1.875, 2.625; 2.0, 2.5; 2.1, 2.4; 2.2, 2.3	169
CO_2	DZ + P	0.021	2.1944, 2.1944	190
CO_2	BA + P	0.002	2.0444, 2.0444; 2.1444, 2.1444*; 2.1944, 2.1944; 2.2944, 2.2944	191
NNO	DZ + P	0.019	2.1273, 2.2418	195
NNO	BA + P	0.007	1.72311, 2.20908; 1.91457, 2.01762; 2.11633, 1.81586; 1.91457, 2.45453; 2.1273, 2.2418*; 2.35148, 2.01762; 2.10603, 2.69998; 2.34003, 2.46598; 2.58663, 2.21938	196
OCN^-	DZ + P	0.021	2.213, 2.281	205
OCN^-	BA + P	0.008	2.213, 2.281	206
FCN	DZ + P	0.021	2.38109, 2.20156	207
FCN	BA + P	0.008	2.38109, 2.20156	208
SCO	DZ + P	0.025	2.9442, 2.2016	209
SCN^-	DZ + P	0.025	2.95, 2.3	210
ClCN	DZ + P	0.025	3.0784, 2.1978	211
C_2H_2	DZ + P	0.006	2.002, 2.281, 2.002	212
C_2H_2	BA + P	0.001	2.002, 2.281, 2.002	213
LiCCH	DZ + P	0.014	3.55, 2.2696, 2.0088	214
LiCCH	BA + P	0.006	3.55, 2.2696, 2.0088	215
FCCH	DZ + P	0.020	2.417, 2.2639, 1.9899	216
FCCH	BA + P	0.008	2.417, 2.2639, 1.9899	217
C_2N_2	DZ + P	0.020	2.186, 2.608, 2.186	218
C_2N_2	BA + P	0.010	2.186, 2.608, 2.186	220
ClCCH	DZ + P	0.025	3.084, 2.2885, 1.988	222
NCCCH	DZ + P	0.020	2.1864, 2.6116, 2.2734, 1.9975	223

[a] The "Tables of Linear Molecule Wave Functions" is available on request to the authors.
[b] Depending on the size of the basis set, it is labeled DZ + P or BA + P. [See M. Yoshimine and A. D. McLean, *Intern. J. Quantum Chem.* (to be published), Slater Symposium issue]. In the DZ + P sets there are two basis functions for each occupied atomic orbital in the separated neutral atoms. The BA + P sets contain a more liberal number of atomic basis functions. Both contain additional functions to help represent polarization of the atoms in a molecule.
[c] This column contains an estimate of the amount the total energy of the tabulated functions is above the Hartree-Fock limit.
[d] Adjacent internuclear separations (for molecules with more than 3 atoms) are given; different nuclear configurations are separated by a semicolon; the nuclear configuration at which the computed total energy is lowest is identified with an asterisk.

tween core and tape. The key to a well-organized computer program is to arrange the sequence of computation so that this is the case. Detailed information on the computer program used to accomplish this data handling is given in a User Manual to the program.[16] The large capacity of this computer program has been important in allowing easy evaluation of single-configuration molecular wave functions to some preset level of accuracy, relative to the Hartree-Fock limit, with a minimum amount of computation.[17]

The following sections of this paper will outline the computational procedures, give precise description of program input blocks, and present examples of input decks to illustrate the type of computations that can be made and to provide examples around which subsequent discussion can be focused. Also included will be a description of some molecular properties which can be routinely computed by the program. The program has been rigorously tested and used extensively by us. Final wave functions for a variety of systems obtained after varying degrees of optimization, have been tabulated in a set of "Tables of Linear Molecule Wave Functions," which form a supplement to the current paper.[18] Table 1 contains the index to these Tables and is presented here both to indicate specifically the contents and to demonstrate the range of systems on which close to Hartree-Fock computations can be made. Compilations of molecular properties evaluated from the wave functions contained in these Tables are being made. Published results contain dissociation energies and dipole moments.[17] Still in preparation are complete tables of molecular quadrupole moments, magnetic susceptibilities and rotational magnetic moments,[19] static electric polarizabilities,[20–21] nuclear electric quadrupole coupling constants, and nuclear electric dipole and quadrupole shielding factors. We would like to draw attention to a particularly valuable source of computed data on molecular energies and properties which includes many results not yet published. This is the comprehensive tabulation of Krauss.[21]

Program organization and procedures

The program McL-YOSH LINEAR MOLECULE PROGRAM 1[11] can determine single-configuration, self-consistent field molecular orbital wave functions for closed-shell electronic states of linear molecular systems. It can also determine a variety of expectation values of one-electron operators with these wave functions. In the present section we will proceed to define the wave function, outline the computational sequence, and give a discussion of the procedures used in the computations.

The total $2N$-electron, closed-shell molecular wave function, Ψ, is an antisymmetrized product of N doubly occupied molecular orbitals, ϕ_n, with electrons in any one orbital having opposed spin. Thus

$$\Psi(1\cdots 2N) = (2N)!^{-1/2} \begin{vmatrix} \phi_1\alpha(1)\phi_1\beta(1) & \cdots & \phi_N\alpha(1)\phi_N\beta(1) \\ \phi_1\alpha(2)\phi_1\beta(2) & \cdots & \phi_N\alpha(2)\phi_N\beta(2) \\ \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots \\ \phi_1\alpha(2N)\phi_1\beta(2N) & \cdots & \phi_N\alpha(2N)\phi_N\beta(2N) \end{vmatrix} \quad (1)$$

The molecular orbitals are orthonormal, $\int dV\phi_i\phi_j = \delta_{ij}$, and are expanded out of normalized Slater-type functions, χ_p, defined by

$$\chi_p(n, l, m, k, \zeta) = [(2\zeta_p)^{2n+1}/(2n)!]^{1/2} r_k^{n-1} \times \exp(-\zeta_p r_k)\, Y_{lm}(\theta_k, \phi_k)\,. \quad (2)$$

The coordinates r_k, θ_k, ϕ_k are spherical polar coordinates defined with respect to nucleus k as origin. For the linear systems under discussion here the z-axis (direction $\theta = 0$) will be along the internuclear line in the direction of increasing k, which numbers the nuclei sequentially along the axis. The integer quantum numbers n, l, m are subject to the condition $n > l \geq |m| \geq 0$ and the parameter ζ_p is the orbital exponent in this p-th basis function. The functions $Y_{lm}(\theta, \phi)$ are normalized complex spherical harmonics,

$$Y_{lm}(\theta, \phi) = \left[\frac{(2l+1)}{4\pi}\,\frac{(l-|m|)!}{(l+|m|)!}\right]^{1/2} \times P_l^{|m|}(\cos\theta)e^{im\phi} \cdots, \quad (3)$$

where the $P_l{}^{|m|}$ are the associated Legendre functions,

$$P_l^{|m|}(x) = \frac{1}{2^l l!}(1-x^2)^{|m|/2} \times \frac{d^{l+|m|}}{dx^{l+|m|}}(x^2-1)^l \cdots. \quad (4)$$

In the program we have imposed the limits l, $|m| \leq 4$; $n + l \leq 9$; $1 \leq k \leq 12$.

Each molecular orbital has the symmetry of one of the irreducible representations of the symmetry group for the problem, and is expanded from functions of the same symmetry type. Thus, for molecules without an inversion center the symmetry labels are σ, π, δ, $\cdots$ and the orbitals are expanded from Slater-type functions with $|m| = 0, 1, 2, \cdots$ respectively. The symmetry of the closed shell electronic configuration is $^1\Sigma$. For molecules with an inversion center the symmetry labels are σ_g, σ_u, π_u, π_g, δ_g, δ_u, $\cdots$; the orbitals are expanded from linear combinations of Slater-type functions, $\chi_p \pm \chi_{p'}$, in which χ_p and $\chi_{p'}$ are centered on symmetrically equivalent nuclei and have the same quantum numbers and orbital exponents. (The center atom in a system with an odd number of nuclei is its own equivalent nucleus, so that single Slater functions on this atom already have the correct symmetry.) In this case, the closed shell electronic configuration has $^1\Sigma^+{}_g$ symmetry. The molecular orbitals are determined, in terms of the basis functions, by the self-consistent field procedure of Roothaan,[13] using a Hamiltonian, $\mathfrak{H}$, in which the potential energy contains all Coulombic interactions. Thus, in atomic units,

$$\mathfrak{H} = -\frac{1}{2}\sum_i \nabla_i^2 + \sum_{k<l}\frac{Z_k Z_l}{r_{kl}} - \sum_{k,i}\frac{Z_k}{r_{ki}} + \sum_{i<j}\frac{1}{r_{ij}}, \quad (5)$$

where i, j sum over electrons; k, l sum over nuclei; Z is a nuclear charge, r an interparticle distance, and the operator ∇^2 is the Laplacian.

The computation breaks down into two main parts: (i) the evaluation of all matrix elements involving members of the basis set with the components of the Hamiltonian, and (ii) the use of these matrix elements in applying the variational principle to determine the set of molecular orbitals which minimizes the total energy of the system. The relative amounts of computation involved in each of these sections depends on the size of the basis set, but even for the largest sets (~50) which can practically be used with this program about 80% of the time is consumed producing matrix elements of the electron interaction operator. In fact, the lack of computers adequate for evaluating these matrix elements has always been a bottleneck in molecular computations. The method in this computer program relies heavily on numerical integration, and the primary consideration in organizing the sequence of computation is to set up all tables of data needed to complete the numerical integration in an order which makes it as efficient as possible.

The elements of the electron interaction matrix have the form

$$(pq|rs) = \int d\tau_1 d\tau_2 \chi_p^*(1)\chi_q(1)\chi_r^*(2)\chi_s(2)/r_{12}$$
$$= \int d\tau_2 U_{pq}(2)\chi_r^*(2)\chi_s(2) \tag{6}$$

where

$$U_{pq}(2) = \int d\tau_1 \chi_p^*(1)\chi_q(1)/r_{12}\,. \tag{7}$$

Asterisks in Eqs. (6) and (7) denote complex conjugates. Equation (6) shows that, after integration over the coordinates of electron 1 has been completed, the integration over electron 2 involves an integrand which is the product of a potential U_{pq} and basis functions $\chi_r{}^*$ and χ_s. The potential U_{pq} at any point is that due to the charge distribution $\chi_p{}^*\chi_q$ as indicated in Eq. (7). Elements of the one electron matrices have the same form as Eq. (6) except that the potential U_{pq} is replaced by a one-electron operator. For linear systems, integration over the axial coordinate, ϕ (the angle between the plane containing an arbitrary point and the internuclear axis and a reference plane containing the internuclear axis), can be disposed of trivially in performing both electron 1 and electron 2 integrations. In general, integration over the remaining two electron 2 coordinates in Eq. (6) is performed numerically. (Exceptions are the one-center integrals in which basis functions indexed by p, q, r, s are all defined relative to a common nucleus.) The basic decisions taken in determining the structure of these programs were to (i) order the matrix elements $(pq|rs)$ so that all nonzero elements for a common pq occur in a block and (ii) that basis functions be tabulated at quadrature points taken NPNT at a time, where NPNT is computed (under program control on the basis of current input data), so that the tabulations for all basis functions at this number of points can fit in the core memory of the computer at the time the numerical integration is to be carried out. Tables of potentials will be constructed in an order corresponding to the order of the matrix elements, the potentials being evaluated NPNT points at a time, using the same ordering of points as for the basis function tabulations. Structuring the tables in this way allows efficient numerical integration, as will be demonstrated below. Implicit in this structure is the use of the same quadrature points for all integrals. For a particular matrix element, contributions from many of the quadrature points may be negligible and the computation time employed in performing the integrations can be decreased by taking advantage of this.

For efficient computation different sections of the program should have essentially the entire core memory available. This has been implemented by constructing program modules which operate on data, available from card input or set up by previous modules, and in turn generate data to be made available as printed output or to be used by succeeding program modules. A short control program (≈200 words) resides permanently in core and directs the path of the computation through the program modules, calling them from a program tape for execution. This structure facilitates extension to new types of molecular computation because the control program needs only simple changes to execute different paths of computation on the basis of an input flag. Since the program modules do not directly communicate but only operate on well-defined lists of data, new modules can be written by other users and be incorporated into the program by making the appropriate change to the control program.

We will now outline the functions of the eleven program modules which are currently incorporated into the program. The reader should keep in mind the basic structure required for efficient numerical integration, outlined earlier in this section. Supplementary information containing more details on implementation is available in Ref. 16.

Module 1 contains the control program and a number of utility programs needed for the interrupt and recovery procedures incorporated into the program. The Module 1 programs are brought into core for execution either by using a bootstrap program contained on a utility card which is read on-line to initiate computation (if the program is used in a stand-alone manner), or by a user program call to a short subroutine, provided by us, which saves the user computation at its current status on a magnetic tape and then reads in the Module 1 subroutines. In this latter case, where the molecular program is called as a subroutine, the core is restored after execution of the molecular program is completed.

Module 2 processes the input data which defines the basis set, the quadrature parameters and the matrix elements to be computed. It then determines the quadrature formula and tabulates the basis functions at the points required by this formula. NPNT, the number of integration points that can be processed at a time in the available core storage, is determined at this point since it defines the structure of all tables used in the numerical integration to be carried out in Module 7). If the current run (as indicated by the value of an input flag, RERUN) is a new one, then NPNT is made equal to the integer part of $C/(B+3)$ where C is the number of core locations available in Module 7 for the basis function tabulations, and B is the number of basis functions. The additional 3 in the denominator is used to allow augmenting of the basis set by up to three functions in a subsequent run, without having to reset NPNT. This means that tables of potentials, structured according to NPNT, can be updated in Module 5 rather than recomputed in the subsequent run. If the current run is to process data updated from that used in a previous run from which the input, wave function output, matrix elements over the basis set and potential tables were saved, the value of NPNT is reset to that used in the previous run provided this is possible. If not, it is set to a new value, according to the formula given above for a new run using a value of B equal to the current size of the basis set. Under these circumstances the potential tables will have to be recomputed in Module 5 rather than updated.

More explicitly, the ordering of the basis function tabulation is: First basis function evaluated at the first NPNT integration points, second basis function evaluated at the first NPNT points, . . ., last basis function evaluated at the first NPNT points. This is followed with similar tables for the second NPNT integration points, and so on until the integration points are exhausted. These tables are written onto magnetic tape. It should be apparent from the way NPNT is computed that with this structure all basis function tabulations taken NPNT points at a time can be read into core in Module 7, where the increment to the accumulated values of the integrals over the basis functions due to contributions from the current NPNT points is made.

Our approximation to the integrals given in Eq. (6) involves integrating two of the coordinates of electron 2 numerically. The coordinates chosen are z, along the molecular axis, and ρ, perpendicular to it. The two-dimensional quadrature is the direct product of two one-dimensional formulas. The one-dimensional formulas are generated by breaking up the range of the variable into segments and obtaining a quadrature formula for each segment as follows.[22] Suppose that the range of a segment is (a, b), the variable denoted by z, and we wish to obtain a quadrature formula that has one-half of the points in the range (a, m), the remainder in the range (m, b). Then

$$\int_a^b dz f(z) = \int_{-1}^1 dt g(t) \approx \sum_i U_i g(t_i) = \sum_i W_i f(z_i)\,, \tag{8}$$

where

$$z_i = \frac{1}{2\beta}\left[-(b-a)+\beta(b+a) + \frac{(b-a)(1-\beta^2)}{(1-\beta t_i)}\right] \tag{9}$$

$$W_i = U_i(b-a)(1-\beta^2)/2(1-\beta t_i)^2 \tag{10}$$

$$\beta = [(b+a)-2m]/(b-a) \tag{11}$$

and t_i, U_i are points and weights for integration in the range $(-1, 1)$. We choose the Gauss-Legendre points and weights. Equation (9) accomplishes the transformation from z to t in the desired way, and we have written the quadrature formula (8) in such a way that the factor dz/dt which comes from transforming the integral has been incorporated into the weight factor W_i. Inspection of Eq. (11) shows that $-1 \le \beta \le 1$ and that when β takes on values 0, $\pm$ 1, limiting forms must be taken.

For $\beta = 0$, which corresponds to $b + a = 2m$,

$$z_i = [(b+a)+(b-a)\,t_i]/2$$
$$W_i = U_i(b-a)/2. \tag{12}$$

For $\beta = 1$, corresponding to infinite b,

$$z_i = [m(1+t_i) - 2at_i]/(1-t_i)$$
$$W_i = 2U_i(m-a)/(1-t_i)^2, \tag{13}$$

while for $\beta = -1$, corresponding to infinite a,

$$z_i = [m(1-t_i)+2bt_i]/(1+t_i)$$
$$W_i = 2U_i(b-m)/(1+t_i)^2. \tag{14}$$

The selection of the segments and the way the segments are divided for the purpose of mapping onto $(-1, 1)$ are performed internally by the program unless specified by the user as input data. In practice, segments should be chosen in the z-coordinate so that integration is performed up to and away from nuclei, to avoid loss of accuracy due to discontinuous derivatives of pertinent basis functions at the nuclei. The way in which segments are divided offers a simple way of controlling the distribution of quadrature points in a way that sensibly reflects the electron distribution, and the mapping onto $(-1, 1)$ enables use of the powerful Gauss-Legendre quadrature formulas.

In the two-dimensional direct product formula the weight corresponding to the point $(z_i\rho_j)$ will be the product W_iW_j where W_i and W_j are the weight factors for the points indexed by i and j on segments in ranges of z and ρ. From Eq. (6) we note that at each point in the numerical integration the values of two basis functions are multiplied together. If the square root of the weight associated with a given point is multiplied into the tabulated values of all basis functions at that point, this will save explicitly multiplying in the weight factor when carrying out the integration. This is an important saving, since putting together the already computed component parts of the integrands of Eq. (6) and accumulating the results is the most time consuming stage in the execution of the program.

Also included in Module 2 is the programming for automatic exponent optimization and for saving useful data in binary form on output tape A6 at the end of the current run, and reading it back in, if required, on the next run.

If the input data have called for the computation of the electronic contribution to the electric field and electric field gradient at the nuclei, then tables of derivatives of the basis functions evaluated at the various nuclei are also computed here, since they are required by our method of computation of the required matrix elements.[23]

Module 3 produces tables of required two-center potentials (Eq. (7) for the case that p, q index basis functions defined with respect to different origins) in a spheroidal coordinate system defined relative to the two origins. These tables cannot be used directly in Module 7 in performing the numerical integration since the points in a spheroidal grid at which the potentials are computed are not the same as the integration points of the previous module. The values of the two-center potentials at the required integration points will be obtained by interpolation into the tables produced in Module 3. Hence, the density of the spheroidal tabulation must be adequate to give the required five-decimal-place accuracy on performing quadratic interpolation as discussed in the section on Module 5.

The table for a single two-center potential can contain up to several thousand entries, and these are generated one at a time and written out onto magnetic tape. They must be ordered on the tape in the way that they will be used subsequently. The ordering is that of the first charge distributions, $\chi_p^*(1)\chi_q(1)$ of Eqs. (6) and (7), which will be discussed in detail in the section on Module 4 and we defer the discussion of ordering until then. It is sufficient to note here that only those potentials needed on the current run are computed. Thus, if the current run uses data from a previous run automatically saved on magnetic tape unit A6 at the end of that run, and NPNT has not been reset because of augmentation of the basis set, then it will not be necessary to compute two-center potentials which are unchanged from the previous run.

The analysis used in deriving formulas to be performed for the two-center potential involves straightforward use of the Neumann expansion of $1/r_{12}$ and a relationship between the associated Legendre functions of the first and second kind.[24] We will outline the analysis at a level sufficient to make the formulas intelligible, since the explicit use of two-center potentials is a departure from the usual evalution of two-center exchange integrals and has not previously been well documented.

The coordinate system is illustrated in Fig. 1, and the relationships between Cartesian, spherical polar and

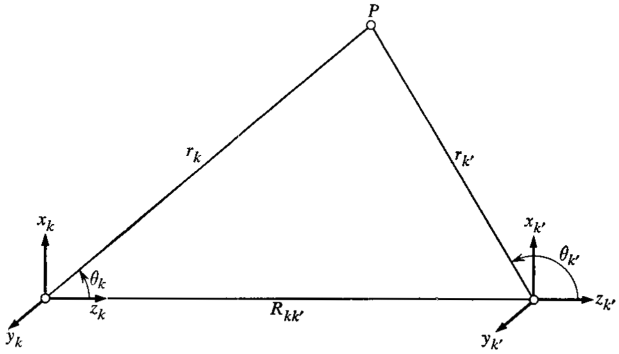

Figure 1 Coordinate systems used in discussion of two-center potential.

spheroidal coordinate systems are as follows. A point P having Cartesian coordinates (x, y, z_k) relative to center k and $(x, y, z_{k'})$ relative to center k', which is a distance $R_{kk'}$ along the positive z-axis, has spheroidal coordinates (ξ, η, ϕ) where

$$\begin{aligned} \xi &= (r_k + r_{k'})/R_{kk'} \\ \eta &= (r_k - r_{k'})/R_{kk'} \\ \phi &= \tan^{-1}(y/x). \end{aligned} \tag{16}$$

This same point P has spherical polar coordinates (r_k, θ_k, ϕ) relative to k, and $(r_{k'}, \theta_{k'}, \phi)$ relative to k' where ϕ is as above and

$$\begin{aligned} r_k &= R_{kk'}(\xi + \eta)/2 \\ \cos\theta_k &= (1 + \xi n)/(\xi + \eta) \\ r_k' &= R_{kk'}(\xi - \eta)/2 \\ \cos\theta_{k'} &= (-1 + \xi\eta)/(\xi - \eta) . \end{aligned} \tag{17}$$

The Cartesian coordinates expressed in terms of the spheroidal are

$$\begin{aligned} x &= R_{kk'}[(\xi^2 - 1)(1 - \eta^2)]^{1/2} \cos\phi/2 \\ y &= R_{kk'}\,[(\xi^2 - 1)(1 - \eta^2)]^{1/2} \sin\phi/2 \\ z_k &= R_{kk'}(1 + \xi\eta)/2 \\ z_{k'} &= R_{kk'}(-1 + \xi\eta)/2. \end{aligned} \tag{18}$$

The two-center potential to be evaluated is

$$\begin{aligned} &N_{nl}N_{n'l'}\int d\tau_1 r_{k1}^{n-1} r_{k'1}^{n'-1} \exp\left(-\zeta r_{k1} - \zeta' r_{k'1}\right) \\ &\quad\times Y_{lm}^*(\theta_{k1}\phi_1)\, Y_{l'm'}(\theta_{k'1}\phi_1)/r_{12} \\ &= N_{nl}N_{n'l'}U_{kk'}^{(nl|m|;n'l'|m'|;\zeta\zeta';|M|)}(\xi_2\eta_2)e^{iM\phi_2} . \end{aligned} \tag{19}$$

In Eq. (19) $N_{nl} = [(2\zeta)^{2n+1}/(2n)!]^{1/2}$, $N_{n'l'} = [(2\zeta')^{2n+1}/(2n')!]^{1/2}$, and a subscript 1 refers to electron 1 coordinates, 2 to electron 2, with the interelectronic distance denoted by r_{12}. The function $U_{kk'}{}^{(nl|m|;n'l'|m'|;\zeta\zeta';|M|)}(\xi_2\eta_2)$ $e^{iM\phi_2}$ is a function of the electron 2 coordinates, and when multiplied by the factor $N_{nl}N_{n'l'}$ gives the value of the potential at points in the space of electron 2 due to the charge distribution of electron 1. In writing Eq. (19) we have anticipated the result of carrying out the integration over the coordinate ϕ_1 which yields the $\exp(iM\phi_2)$ dependence in the potential where $M = -m + m'$. It should be noted that the functions $U_{kk'}{}^{(nl|m|;n'l'|m'|;\zeta\zeta';|M|)}(\xi_2\eta_2)$, in which the subscripts and superscripts denote parameters specifying the potential, are functions only of the absolute magnitudes $|m|$, $|m'|$, $|M|$.

The first step in performing the integration of Eq. (19) is to substitute the expressions for the spherical polar coordinates in terms of spheroidal. Thus

$$e^{-\zeta r_{k1}-\zeta' r_{k'1}} = e^{-\alpha\xi_1-\beta\eta_1} \tag{20}$$

where $\alpha = (\zeta + \zeta')R_{kk'}/2$ and $\beta = (\zeta - \zeta')R_{kk'}/2$, and

$$\begin{aligned}4\pi r_{k1}^{n-1} r_{k'1}^{n'-1} Y_{lm}(\theta_{k1}\phi_1)\, Y_{l'm'}(\theta_{k'1}\phi_1) &= (R_{kk'}/2)^{n+n'-2} \sum_{(i+j)=p(D^{|M|})}^{D^{|M|}}{}' \sum_{j=0}^{i} \omega_{ij}^{(nl|m|;n'l'|m'|;|M|)} \\ &\times \xi_1^i\eta_1^j[(\xi_1^2-1)(1-\eta_1^2)]^{|M|/2} e^{i(-m+m')\phi_1}\end{aligned} \tag{21}$$

where in Eq. (21) the indicated numerical coefficients $\omega_{ij}{}^{(nl|m|;n'l'|m'|;|M|)}$ depend only on the quantum numbers of the basis functions and are straightforwardly derivable by expanding the left-hand expression. Our computer programs perform the laborious algebraic manipulations involved in making the above expansion. Detailed analysis shows that the summation on $(i + j)$ is in steps of 2 (indicated by the prime) from a starting value of either 0 or 1 depending on whether $D^{|M|} = n + n' - 2 + l + l' - 2|M|$ is even or odd. Now substitute into Eq. (19) $d\tau_1 = (R_{kk'}^3/8)(\xi_1^2 - \eta_1^2)d\xi_1 d\eta_1 d\phi_1$ and the Neumann expansion of $1/r_{12}$.

$$\begin{aligned}\frac{1}{r_{12}} &= \frac{4}{R_{kk'}} \sum_{L=0}^{\infty} \sum_{M=-L}^{L} (-1)^M \frac{(2L+1)}{2} \left[\frac{(L-|M|)!}{(L+|M|)!}\right]^2 \\ &\times P_L^{|M|}(\xi_<) Q_L^{|M|}(\xi_>) P_L^{|M|}(\eta_1) \\ &\times P_L^{|M|}(\eta_2) e^{iM(\phi_1-\phi_2)}\,.\end{aligned} \tag{22}$$

In Eq. (22) $\xi_<$ and $\xi_>$ are respectively the lesser and greater of ξ_1, ξ_2 and the functions $P_L^{|M|}$, $Q_L^{|M|}$ are associated Legendre functions of the first and second kinds.[25] The $P_L^{|M|}$ for argument x in the range $-1 \le x \le 1$ are the usual functions defined in Eq. (4). For argument x in the range $1 \le x < \infty$ we simply replace $(1 - x^2)^{|M|/2}$ in Eq. (4) by $(x^2 - 1)^{|M|/2}$. The $Q_L^{|M|}$ for argument x in the range $1 \le x < \infty$ are given by

$$Q_L^{|M|}(x) = (x^2-1)^{|M|/2}(d^M/dx^M)Q_L(x)\,, \quad \text{where}$$

$$Q_L(x) = \frac{1}{2}\int_{-1}^{1} du P_L(u)/(L-u)\,.$$

Making the above substitutions and carrying out the integration over the ϕ_1 coordinate, we have

$$\begin{aligned}&U_{kk'}^{(nl|m|;n'l'|m'|;\zeta\zeta';|M|)}(\xi_2\eta_2) \\ &= \left(\frac{R_{kk'}}{2}\right)^{n+n'} \int_1^{\infty} d\xi_1 \int_{-1}^{1} d\eta_1 \sum_{L=|M|}^{\infty} \frac{(2L+1)}{2} \\ &\times \left[\frac{(L-|M|)!}{(L+|M|)!}\right]^2 e^{-\alpha\xi_1-\beta\eta_1} \\ &\times P_L^{|M|}(\xi_<) Q_L^{|M|}(\xi_>) P_L^{|M|}(\eta_1) P_L^{|M|}(\eta_2) \\ &\times \sum_{(i+j)=p(D^{|M|})}^{D^{|M|}+2}{}' \sum_{j=0}^{i} \omega_{ij}^{(n+1\,l|m|;n'+1\,l|m|;|M|)} \\ &\times \xi_1^i\eta_1^j[(\xi_1^2-1)(1-\eta_1^2)]^{|M|/2}.\end{aligned} \tag{23}$$

In writing Eq. (23) it will be noted that the $(\xi^2 - \eta^2) = (R_{kk'}/2)^2 r_k r_{k'}$ has been absorbed into the polynomial arising from the basis function product, with appropriate adjustment of superscripts and summation limits.

Defining the auxiliary functions

$$\begin{aligned}b_j^{|M|L}(\beta) &= \frac{1}{2}\frac{(L-|M|)!}{(L+|M|)!}\int_{-1}^{1} d\eta\eta^j(1-\eta^2)^{|M|/2} \\ &\times P_L^{|M|}(\eta)e^{\beta\eta}\end{aligned} \tag{24}$$

and

$$\begin{aligned}&\omega^{(nl|m|;n'l'|m'|;\zeta\zeta';|M|L)}(\xi) \\ &= \sum_{(i+j)=p(D^{|M|})}^{D^{|M|}+2}{}' \sum_{j=0}^{i} \omega_{ij}^{(n+1\,l|m|;n'+1\,l'|m'|;|M|)} \\ &\times \xi^i b_j^{|M|L}(-\beta)\end{aligned} \tag{25}$$

and substituting into Eq. (23) yields

$$\begin{aligned}&U_{kk'}^{(nl|m|;n'l'|m'|;\zeta\zeta';|M|)}(\xi_2\eta_2) \\ &= \left(\frac{R_{kk}}{2}\right)^{n+n'} \sum_{L=|M|}^{\infty} (2L+1)\frac{(L-|M|)!}{(L+|M|)!} P_L^{|M|}(\eta_2) \\ &\times \Big[Q_L^{|M|}(\xi_2) \int_1^{\xi_2} d\xi(\xi^2-1)^{|M|/2} e^{-\alpha\xi} \\ &\times P_L^{|M|}(\xi)\omega^{(nl|m|;n'l'|m'|;\zeta\zeta';|M|L)}(\xi) \\ &+ P_L^{|M|}(\xi_2)\int_{\xi_2}^{\infty} d\xi(\xi^2-1)^{|M|/2}e^{-\alpha\xi} \\ &\times Q_L^{|M|}(\xi)\omega^{(nl|m|;n'l'|m'|;\zeta\zeta';|M|L)}(\xi)\Big]\,.\end{aligned} \tag{26}$$

Equation (26) is already in form suitable for use numerically and, in fact, was so used in a previous computer program.[26] However, in the current program we have carried the analysis further as will be described below. The auxiliary functions $b_0^{0l}(\beta)$ are the spherical Bessel functions, and the two-zero indices can be raised by use of recurrence relations which do not lose accuracy. The spherical Bessel functions themselves can be computed extremely rapidly and accurately using a continued fraction expansion for the ratios between successive l values for a given argument.[27]

Continuing with the analysis,[28] we note that for an arbitrary function $f(x)$

$$\int_{\xi_2}^{\infty} dx f(x) Q_L^{|M|}(x)$$
$$= \int_{\xi_2}^{\infty} dx (Q_L^{|M|}(x)/P_L^{|M|}(x)) f(x) P_L^{|M|}(x)$$
$$= \int_{\xi_2}^{\infty} dx (Q_L^{M}(x)/P_L^{|M|}(x) \frac{d}{dx} \int_1^x dy f(y) P_L^M(y) ,$$

which when integrated by parts together with use of

$$P_L^{|M|}(x) \frac{d}{dx} Q_L^{|M|}(x) - Q_L^{|M|}(x) \frac{d}{dx} P_L^{|M|}(x)$$
$$= \frac{(-1)^{|M|+1}}{(x^2-1)} \frac{(L+|M|)!}{(L-|M|)!}$$

yields

$$Q_L^{|M|}(\xi_2) \int_1^{\xi_2} dx f(x) P_L^{|M|}(x)$$
$$+ P_L^{|M|}(\xi_2) \int_{\xi_2}^{\infty} dx f(x) Q_L^{|M|}(x)$$
$$= (-1)^M \frac{(L+|M|)!}{(L-|M|)!} P_L^{|M|}(\xi_2)$$
$$\times \int_{\xi_2}^{\infty} dx [1/(x^2-1)(P_L^{|M|}(x))^2$$
$$\times \int_1^x dy f(y) P_L^{|M|}(y) .$$

Use of the result in Eq. (26) leads to

$$U_{kk'}^{(n\,l|m|;n'\,l'|m|;\zeta\zeta';|M|)}(\xi_2\eta_2)$$
$$= \left(\frac{R_{kk'}}{2}\right)^{n+n'} \sum_{L=|M|}^{\infty} (2L+1) P_L^{|M|}(\eta_2) P_L^{|M|}(\xi_2)$$
$$\times \int_{\xi_2}^{\infty} dx [1/(x^2-1)(P_L^{|M|}(x))^2]$$
$$\times \int_1^x dy (y^2-1)^{|M|/2} e^{-\alpha y} \omega^{(n\,l|m|;n'\,l'|m'|;\zeta\zeta';|M|L)}(y)$$
$$\times P_L^{|M|}(y) . \tag{27}$$

Equation (27) is also in a form suitable for numerical use, and one version of the current program employs it. The double integral is evaluated numerically for values of ξ_2 in the range 1 to ∞ by operating on tabulated values of the integrands, in the same range, using a quadrature formula of the Simpson rule type. However, it is possible to carry the analysis further[29] and derive an expression for the inner $\int_1^x dy$ integration in terms of the auxiliary functions

$$E_n(\alpha, x) = \int_0^{x-1} du u^n e^{-\alpha(u+1)}$$

which can be simply evaluated. The reduction of the inner integral to these auxiliary functions is accomplished by use of[25]

$$P_L^{|M|}(y)/(y^2-1)^{|M|2} = \left[\frac{(L+|M|)!}{2^{|M|}|M|!(L-|M|!)}\right]$$
$$\times \left\{1 + \frac{(L-|M|)(L+|M|+1)}{1(|M|+1)}\left(\frac{y-1}{2}\right)\right.$$
$$+ [(L-|M|)(L-|M|-1)(L+|M|+1)]$$
$$\left. \times \frac{[(L+|M|+2)]}{[1.2(|M|+1)(|M|+2)]}\left(\frac{y-1}{2}\right)^2 + \cdots\right\}$$

and transforming the polynomial in y, $\omega^{(nl|m|;n'l'|m'|;\zeta\zeta';|M|L)}(y)$, into a polynomial in $(y-1)$. The advantages of making use of the analytic expression for the inner integral are improved accuracy, and the possibility of doing the outer $\int_{\xi_2}^{\infty} dx$ integration with a more powerful quadrature formula. A second version of the current program incorporates this analysis.

The actual quantity tabulated by the computer program is

$$N_{n\,l} N_{n'\,l'} U_{kk'}^{(n\,l|m|;n'\,l'|m'|;\zeta\zeta';|M|)}(\xi_2, \eta_2)$$
$$\times \xi^{|M|}/[(\xi^2-1)(1-\eta^2)]^{|M|/2} \tag{28}$$

and the tabulation is made at equal spacing in η_2 in the range $-1 \leq \eta_2 \leq 1$ and t in the range $0 \leq t < 1$ where

$$\xi_2 = [(2-b)+(b-1)]/t . \tag{29}$$

In Eq. (29), b is an adjustable parameter (specified as input data) which is a scale factor controlling the mapping of ξ_2 onto t. Selection of the best value of b gives the highest possible accuracy in the numerical integration. The factor $[(\xi^2-1)(1-\eta^2)]^{|M|/2}$ in the denominator of the tabulated quantity (28) is introduced to provide a suitable function for interpolation. That it is necessary can be seen by considering the behaviour of $U_{kk}^{(nl|m|;n'l'|m'|;\zeta\zeta';|M|}(\xi_2)$ near $\xi_2 = 1$ for odd values of $|M|$, where the square root of (ξ_2^2-1) arising from $P_L^{|M|}(\xi_2)$ controls the behaviour of the function. Since the polynomial expansion in powers of ξ_2 of this square root in the neighborhood of $\xi_2 = 1$ is slowly convergent it is not amenable to interpolation.

Module 4 generates a list which indexes the matrix elements, and contains information on how they are to be computed or whether they are available from a previous run. The list contains one positive entry for each one-electron operator in the one-electron matrix elements, and for each electron 1 charge distribution (see Eq. 6) in the two-electron matrix elements. These positive entries contain either the code for a one-electron operator or the indices of the basis functions contributing to an electron 1 charge

Table 2 Ordering of charge distributions by symmetry.

First Charge Distribution[a]	$\lvert M \rvert$	*Second Charge Distribution*[a,b]
$\sigma\sigma$	0	$\sigma\sigma$
$\pi\sigma$	1	$\pi\sigma$
$\pi\pi$	2	$\pi\pi$
	0	$\sigma\sigma,\pi\pi$
$\delta\sigma$	2	$\pi\pi,\delta\sigma$
$\delta\pi$	3	$\delta\pi$
	1	$\pi\sigma^*,\delta\pi$
$\delta\delta$	4	$\delta\delta$
	0	$\sigma\sigma,\pi\pi,\delta\delta$
$\phi\sigma$	3	$\delta\pi^*,\phi\sigma$
$\phi\pi$	4	$\delta\delta^*,\phi\pi$
	2	$\pi\pi^*,\delta\sigma^*,\phi\pi$
$\phi\delta$	5	$\phi\delta$
	1	$\pi\sigma^*,\delta\pi^*,\phi\delta$
$\phi\phi$	6	$\phi\phi$
	0	$\sigma\sigma,\pi\pi,\delta\delta,\phi\phi$
$\gamma\sigma$	4	$\delta\delta^*,\phi\pi^*,\gamma\sigma$
$\gamma\pi$	5	$\phi\delta^*,\gamma\pi$
	3	$\delta\pi^*,\phi\sigma^*,\gamma\pi$
$\gamma\delta$	6	$\phi\phi^*,\gamma\delta$
	2	$\pi\pi^*,\delta\sigma^*,\phi\pi^*,\gamma\delta$
$\gamma\phi$	7	$\gamma\phi$
	1	$\pi\sigma^*,\delta\pi^*,\phi\delta^*,\gamma\phi$
$\gamma\gamma$	8	$\gamma\gamma$
	0	$\sigma\sigma,\pi\pi,\delta\delta,\phi\phi,\gamma\gamma$

[a] $\lvert m \rvert$ values of 0, 1, 2, 3 . . . are denoted σ, π, δ, ϕ, . . .
[b] An asterisk denotes a second charge distribution which need be combined with the designated first charge distribution for configuration mixing calculations only.

distribution, whose potential will be needed in order to evaluate the electron repulsion matrix elements. Each positive entry is followed by a number (at least one) of negative entries that contain indices specifying the basis functions contributing to a charge distribution. The total number of matrix elements is the number of negative entries, and the full specifications of a matrix element are given by a negative entry and the positive entry preceding it in the list. Since the code for an operator or potential (positive entry) appears only once, it is clear that all matrix elements that are to be computed with that operator or potential are specified by the negative entries immediately following it. This structure is chosen since it simply allows tabulations of operators and potentials, made in the same order as the positive entries, to be used sequentially and with no redundancy.

The matrix elements are in the order of overlap, kinetic energy, nuclear attraction, electron repulsion, followed by blocks of one-electron matrix elements corresponding to one-electron operators which are either selected by input data or, if this is not the case, are automatically generated by the program.

Within blocks of a given type, entries are ordered with respect to the symmetry of the charge distributions involved in the matrix elements, and within each symmetry are ordered according to the indexing of the basis functions. In the case of $C_{\infty v}$ (no inversion center), we have taken full advantage of symmetry simply by breaking the list down according to the symmetry classifications of the charge distributions derived from basis function products. The redundancy which would still remain in this list for $D_{\infty h}$ symmetry has been removed by making additional tests within each symmetry classification. The removal of any redundancy due to the molecular symmetry is important for efficiency, and essential for the algorithms we have used in constructing supermatrices (Modules 8, 10) to be valid.

The symmetry of a charge distribution is specified by the values of $\lvert m \rvert$ and $\lvert m' \rvert$ from the two basis functions involved, and the value of $\lvert M \rvert = \lvert -m + m' \rvert$. In the case of the one-electron matrix elements the $\lvert M \rvert$ dependence of the operator must be the same as the $\lvert M \rvert$ dependence of the charge distribution, while for the two-electron matrix elements the $\lvert M \rvert$ values of the two-charge distributions must be the same. Otherwise, the integral is identically zero. Table 2 lists all possible second charge distributions that give nonzero matrix elements when combined with a given first charge distribution. Entries marked with an asterisk are those which do not need to be computed in obtaining a single-configuration molecular orbital wave function. They are needed for calculations involving configuration mixing and the program was designed to compute them even though in its current state it does not allow configuration mixing calculations. Whether they are computed or not is determined by an input flag, IRFLG, defined in the input data specifications. The order of the symmetry classifications of the two-electron matrix elements generated by the program is simply derived from Table 2 by taking each first charge distribution with the second distributions in the order in which they appear in the Table.

For a given symmetry classification there are many ways to order the nonredundant matrix elements. Our choice, which is strongly dependent on the structure of the IBM 7090 computer word, is not of general interest and will not be given in detail here.[30]

Each entry in the list is tagged with control information necessary for the successful execution of later modules, and to provide efficient computation.

Positive entries are tagged according to which of the following categories describes the operator or potential they specify:

(i) Not to be tabulated numerically, since no matrix element involving it will be computed numerically.

(ii) Same tabulation used as on previous run, since no change in the basis functions involved.

(iii) To replace a tabulation generated in previous run, because of a change in the basis functions involved.

(iv) An added tabulation to those generated in a previous run because of insertion of a basis function into the basis set.

Negative entries are tagged according to which, with the preceding positive entry, of the following categories describes the specified matrix element:

(i) Not computed numerically, and involves a basis function changed from the previous run.
(ii) Not computed numerically, and involves a basis function inserted into the basis set of the previous run.
(iii) Computed numerically, and involves a basis function changed from the previous run.
(iv) Computed numerically, and involves a basis function inserted into the basis set of the previous run.
(v) Same as in previous run, in which it was computed numerically.
(vi) Same as in previous run, in which it was not computed numerically.

It will be recalled that the information retained from the previous run included the tabulations of potentials at the points required for numerical integration, and the matrix elements. Consideration of the categories listed above will show that all information for the updating of the potential tabulations and matrix element list is listed. We note that logically, an operator in a one-electron matrix element is treated in the same way as the potential in a two-electron matrix element. Also, that a run which does not call on previously generated potential tabulations and matrix elements is logically treated as one in which all basis functions are inserted into a previous basis set of zero occupancy, i.e. all operators, potentials and matrix elements are tagged as inserted in terms of the above categories. Runs which cannot call on previously generated potential tabulations are those in which NPNT (see Module 2) was reset because of augmentation of the basis set. Runs which cannot call on previously generated potential tables or matrix elements are new runs, or runs which involve different quadrature points for numerical integration. This latter case is realized, for example, on changing the internuclear configuration without any change in the basis set, other than the change of origin.

In Module 5 tabulations of operators and potentials will be made in the order of the positive entries, for NPNT points at a time, the tables being obtained according to the control information encoded in the positive entries as described above. With reference to Module 3, the ordering of the two-center potential tabulations in spheroidal coordinates is the ordering of the two-center potentials in the positive entries of the Module 4 list which are tagged as containing a changed or inserted basis function.

The entries in the list which indexes the matrix elements over the basis functions are written out onto magnetic tape, 250 at a time. They will later be used sequentially and read in through a double buffer (500 words of core storage) so that entries occupying the first buffer can be processed while the second is being filled and vice versa. Thus, the list can be processed efficiently using only 500 core locations even though the total length of the list can typically be of order 10^5.

Module 5 constructs tables of operators and potentials at the integration points selected in Module 2. As for the basis function tabulations of Module 2, the points are taken NPNT at a time. For each NPNT points the operators and potentials that need be tabulated, as determined by the information in the positive entries of Module 4, are written out onto magnetic tape. The precise ordering of the records on magnetic tape is: First operator or potential (whose tabulation is needed) at first NPNT points, second operator or potential at first NPNT points, third, etc., until we have exhausted the positive entries in the Module 4 list. This is followed by similar tabulations for the next NPNT points, and so on until all points are exhausted. The tabulations can be, (i) a copy of a tabulation from the previous run, (ii) a new two-center potential obtained at the current NPNT points by interpolation into the table produced in Module 3, (iii) a new one-center potential computed from scratch in the current module, (iv) an operator whose tabulation will be made from scratch in this module.

It is of interest at this point to consider the amount of data being processed in the setting up of these tabulations. The Module 4 index list is being read in through a 500-word double buffer, once for each NPNT points. The potential tabulation from the previous run is being read in through a 2(NPNT)-word double buffer. The two-center potential tables produced in Module 3 are being read in through a ~10,000-word double buffer, once for each NPNT points. The potential tabulations being produced are written out through a 2(NPNT)-word double buffer. For typical runs involving ~40 basis functions the Module 4 list contains $\sim 2 \times 10^5$ entries, the other lists $\sim 2 \times 10^6$ entries. Synchronization of the data transmission is implemented through two powerful I–O subroutines which facilitate optimum use of the data channels on the computer.[31] These subroutines are the key to the success of the current program. The use of double buffered data transmission (as described at the end of the section on Module 4) is an extremely simple way of overlapping data transmission with computation. In fact, in the current program, essentially the only stage where it can be I–O bound, is in this Module when a long string of potential tabulations is being copied from the old list to the new, uninterrupted by entries that must be recomputed because of a change or insertion in the previous basis set.

The analysis of the one-center potential involving Slater functions is well documented in the literature, and need not be repeated here. In particular, the analysis given by Wahl, Cade and Roothaan[32] is essentially identical to the one incorporated into the present program.

Module 6 generates the list of values of matrix elements in its final form, except that entries corresponding to all matrix elements to be computed numerically in the current run (as encoded in the Module 4 index list) will be set equal to zero. All other matrix elements will either be copied from the corresponding list produced in a previous run, or will be computed by special subroutines in this Module. The types of matrix elements computed specially are (i) all one-center integrals, (ii) one-electron, two-center overlap, kinetic energy and nuclear attraction integrals. These include the largest matrix elements and are required to more significant figures than the remainder. This is made possible by the use of analytical procedures which to a large extent follow our analysis of the one- and two-center potentials. The matrix elements to be computed numerically will be accumulated, into the appropriate entries set to zero in this Module, by the programs in Module 7 which will complete the evaluation of matrix elements.

In the execution of this Module, the data lists being transmitted into core are the index list of Module 4 and the matrix element list from the previous run. The outgoing data list is the new matrix element list. All lists are transmitted through double buffers.

The key formulas for the matrix elements computed in this module are listed below. $\chi(nlmk\zeta)$ is used to denote a Slater function and center (Eq. (2)), and different Slater functions are denoted by using combinations of bars and primes.

(i) One-center overlap integral.

$$\int d\tau \chi(nlmk\zeta)^* \chi'(n'l'm'k\zeta') = \delta_{ll'}\delta_{-m+m',0} N_{nl} N_{n'l'} \frac{(n+n')!}{(\zeta+\zeta')^{n+n'+1}} \tag{30}$$

with $N_{nl}N_{n'l'}$ as defined following Eq. (19), and δ denoting a Kronecker delta.

(ii) One-center, two-electron integral.

$$\begin{aligned}
\int d\tau_1 d\tau_2 \chi(nlmk\zeta)_1^* \chi'(n'l'm'k\zeta')_1 & \\
\times \bar{\chi}(\bar{n}\bar{l}\bar{m}k\bar{\zeta})_2^* \bar{\chi}'(\bar{n}'\bar{l}'\bar{m}'k\bar{\zeta}')_2 / r_{12} & \\
= C \sum_{L=p(l+l')+|M|}^{(l+l',\bar{l}+\bar{l}')<}{}' A_{L|M|} \bar{A}_{L|M|} \frac{(L+|M|)!}{(L-|M|)!} & \\
\times \left\{ \frac{(N-L)!(\bar{N}+L+1)!}{(\zeta+\zeta')} \right. & \\
\times B_{\bar{N}+L+1,N-L}\left(\frac{\zeta+\zeta'+\bar{\zeta}+\zeta'}{\zeta+\zeta'}\right) & \\
+ \frac{(\bar{N}-L)!(N+L+1)!}{(\bar{\zeta}+\bar{\zeta}')} & \\
\times B_{N+L+1,\bar{N}-L}\left. \left(\frac{\zeta+\zeta'+\bar{\zeta}+\bar{\zeta}'}{\bar{\zeta}+\bar{\zeta}'}\right)\right\}, &
\end{aligned} \tag{31}$$

where $p(l+l') = 0$ if the sum is even, $=1$ otherwise. The upper limit of the sum is the smaller of the two quantities in parentheses and the index L advances in steps of 2. $|M| = |-m+m'|$, $N = n+n'-1$, $\bar{N} = n+\bar{n}'-1$, and the remaining quantities used in Eq. (31) are defined below.

$$C = \delta_{-m+m'-\bar{m}+\bar{m}',0};\ \delta_{p(l+l'+\bar{l}+\bar{l}'),0} \times FF'\bar{F}\bar{F}'/[(\zeta+\zeta')^{N+2}(\bar{\zeta}+\bar{\zeta}')^{\bar{N}+2}]$$

$$F = \frac{(2\zeta)^{n+1/2}}{[(2n)!]^{1/2}}\left[(2l+1)\frac{(l-|m|)!}{(l+|m|)!}\right]^{1/2}$$

and similar expressions for F', $\bar{F}$, $\bar{F}'$.

$A_{L|M|}$ are the expansion coefficients in

$$P_l^{|m|}(\cos\theta) P_{l'}^{|m'|}(\cos\theta) = \sum_{L=p(l+l')+|M|}^{(l+l')}{}' (2L+1) A_{L|M|} P_L^{|M|}(\cos\theta) \tag{32}$$

with a similar expansion defining $\bar{A}_{L|M|}$.

The functions B in Eq. (31) are defined by

$$B_{pq}(x) = \sum_{k=0}^{q} \frac{(p+k)!}{p!k!} x^{q-k}$$

The result contained in Eq. (31) follows from routine use of the Laplace expansion of $1/r_{12}$ and the orthogonality of the spherical harmonics.

(iii) Two-center overlap integral.

$$\begin{aligned}
\int d\tau_1 \chi(nlmk\zeta)^* \chi'(n'l'm'k'\zeta') & \\
= \delta_{-m+m',0} \frac{N_{nl}N_{n'l'}}{2}\left(\frac{|R_{kk'}|}{2}\right)^{n+n'+1} \sum_{(i+j)=p(D^0)}^{D^0+2} & \\
\times \sum_{j=0}^{i} \omega_{ij}^{(n+1\,l|m|;n'+1\,l'|m'|;0)} & \\
\times \int_1^\infty d\xi \xi^i e^{-\alpha\xi} \int_{-1}^{1} d\eta \eta^j e^{-\beta\eta}. &
\end{aligned} \tag{33}$$

The notation of Eq. (33) is that developed in the discussion of the two-center potential in Module 3. We note that the η integration of Eq. (33) is, apart from a factor 1/2, the function $b_j^{00}(-\beta)$ as defined in the two-center potential. The ξ integration gives $A_i(\alpha)$, where $A_0(\alpha) = e^{-\alpha}/\alpha$ and $\alpha A_i(\alpha) = iA_{i-1}(\alpha) + e^{-\alpha}$, which can be used to recur up on index i without loss of accuracy.

The remaining integrals are all simply related to the above three, with the exception of the two-center nuclear attraction integrals involving a one-center charge distribution on center k and an operator $1/r_{k'}$.[32] These latter integrals drop out the analysis of the one-center potential, since they are in fact the value of the potential at k' due to the charge distribution.

We have included these formulas, since they closely reflect the way in which they are programmed. In particular, our result for the two-center overlap integral using the functions $b_j^{00}(-\beta)$ is of note, since the same subroutines developed for the two-center potential are used.

Module 7 completes the evaluation of the matrix elements by carrying out the numerical integration, for all matrix elements so tagged in the Module 4 index list, and accumulating the result into the corresponding entries in the matrix element list set up in Module 6. The integration is carried out NPNT points at a time. It will be recalled that the value of NPNT was chosen so that all tabulated basis functions, which because of the matrix element ordering are needed nonsequentially, could reside in core storage in Module 7. Operator and potential tabulations, produced in Module 5, are used sequentially, and can, therefore, be read through a double buffer of dimension 2(NPNT) with the potential table in one buffer being used while the other is read in.

The data flow for the first NPNT points is as follows. The basis function tabulations are read into core in their entirety. Input data lists read in through double buffers are: the index list of Module 4 (necessary to determine which integrals are computed numerically and which basis functions are involved), the potential tabulations of Module 5 which were set up for the same ordering of points as the basis functions and the same ordering of potentials as in the index list, and the matrix element list produced in Module 6 into which the numerically evaluated integrals will be accumulated. The output data list is a matrix element list identical in structure to the input list, but with the accumulation due to the current NPNT points included in the entries corresponding to numerically evaluated matrix elements. For the next NPNT points, the basis function tabulations are read into core, the Module 4 index list is read again, the reading of the potential tabulations (which now correspond to the current NPNT points) is continued, and the input matrix element list is the output list from the previous NPNT points and vice versa. This procedure is continued until all integration points are exhausted.

In the case that matrix elements of the operators z_k/r_k^3 and $(3z_k^2 - r_k^2)/r_k^5$ are being computed, the entries in the basis function tables are modified to remove the singularities in the integrands. The details of this have been reported elsewhere.[33]

Module 7 completes the evaluation of the matrix elements, and the remaining Modules determine a single configuration self-consistent molecular orbital wave function for the system. This is done in two stages. First, the matrix elements over the basis set are combined into a supermatrix, defined in the following discussion. For $C_{\infty v}$ symmetry this is done in Module 8, and for $D_{\infty h}$ in Module 10. Secondly, the supermatrix is contracted in constructing the Fock matrix used in the iterative solution of a pseudoeigenvalue problem which yields the self-consistent molecular orbitals. For $C_{\infty v}$ symmetry this is done in Module 9, and for $D_{\infty h}$ in Module 11. The theory of these procedures has been expounded at length[13,34–37] and we will not repeat it here; rather, we will restrict the discussion to computational procedures which have not previously been reported.

Modules 8 and 10 construct the one-electron kinetic energy and nuclear attraction matrices and the two-electron supermatrix, required for the determination of molecular orbitals, for $C_{\infty v}$ and $D_{\infty h}$ symmetry, respectively.

First, it is of interest to outline what is involved. The elements in the list of matrix elements over the basis functions, which were ordered in a way required for their efficient computation, are to be combined with certain coefficients and placed into positions in a supermatrix according to an indexing scheme unrelated to the indexing of the matrix elements over the basis functions. In other words, supermatrix elements are linear combinations of more or less random entries in the list of matrix elements over the basis functions. Since, for typical cases, both supermatrix and matrix elements over the basis functions can have $>10^5$ entries and only $\sim$20,000 words of storage are available to accomplish the transformation, it is clear that some thought has to go into making the process efficient.

Either of two approaches can be made. The first approach would be to read in blocks of the input list (matrix elements over the basis set) as large as possible at a time, and have the output list (supermatrix) written out from a double buffer of modest ($\sim$500) size. Elements of the output list would be determined sequentially by picking up the components currently available in the input list resident in core. After going through the entire output list, the next section of the input list would be read in and the process repeated. For this approach a directory of the input list which would make it easy to pick up a required matrix element is necessary. The ordering of the matrix elements chosen in Module 4 is not particularly well chosen with respect to computing the position of a specific entry, which is why we chose to implement the second approach described below. In regard to execution time the two approaches would be comparable.

The second approach is to have the input list read through a modest sized ($\sim$500) double buffer and to have the largest possible block of core assigned to the output list (supermatrix). Elements of the input list are processed sequentially, each one being placed with the appropriate factor into as many places as it is required in the section of supermatrix currently resident in core. After the entire input list has been read through, the current supermatrix block will have been completed and the process can be repeated for the next supermatrix block. In this procedure the computation of the location of a specific supermatrix element is extremely simple, but care is required because it only works for a nonredundant list of matrix elements over the basis functions. The program is structured according to symmetry classifications, so that in running through the input list we only run through that section which, by symmetry, can make a contribution to the block of the supermatrix currently resident in core. The actual details of the algorithm used for placing a given matrix element over the basis functions are nontrivial and we will present the results for the case of $C_{\infty v}$ symmetry. The procedures for constructing the $D_{\infty h}$ supermatrix are significantly more complicated but not essentially different and we will present no further discussion of this case here. It will be noted that in the $C_{\infty v}$ case, the maximum number of supermatrix elements to which a given matrix element over the basis functions can contribute is 3, while in the $D_{\infty h}$ case it is 48. For this reason, the computer program in Module 8 is I–O bound, i.e., the total execution time is the time required to read the matrix element list in its core as many times as there are blocks of supermatrix, while the Module 10 program is very much compute bound, i.e., the amount of arithmetic to be performed on each entry in the matrix element list is considerably in excess of the average time required to read it in from tape.

The elements in the supermatrix $\mathfrak{P}_{\lambda pq,\mu rs}$ for closed shell molecules are defined by[36]

$$\mathfrak{P}_{\lambda pq,\mu rs} = \mathfrak{J}_{\lambda pq,\mu rs} - \tfrac{1}{2}\mathfrak{K}_{\lambda pq,\mu rs} \tag{34}$$

where

$$\mathfrak{J}_{\lambda pq,\mu rs} = (d_\lambda d_\mu)^{-1} \sum_{\alpha,\beta} \int \chi^*_{p\lambda\alpha}(1)\chi^*_{r\mu\beta}(2) \times (r_{12})^{-1}\chi_{q\lambda\alpha}(1)\chi_{s\mu\beta}(2)d\tau_1 d\tau_2, \text{ and} \tag{35}$$

$$\mathfrak{K}_{\lambda pq,\mu rs} = (d_\lambda d_\mu)^{-1} \sum_{\alpha,\beta} \frac{1}{2}\left[\int \chi^*_{p\lambda\alpha}(1)\chi^*_{s\mu\beta}(2)(r_{12})^{-1} \times \chi_{r\mu\beta}(1)\chi_{q\lambda\alpha}(2)d\tau_1 d\tau_2 + \int \chi^*_{p\lambda\alpha}(1)\chi^*_{r\mu\beta}(2)(r_{12})^{-1}\chi_{s\mu\beta}(1)\chi_{q\lambda\alpha}(2)d\tau_1 d\tau_2\right]. \tag{36}$$

In Eqs. (34) to (36) p, q, r, s index symmetrized basis functions λ, μ, index irreducible representations of the point group, and α, β index the subspecies in an irreducible representation. d_λ, d_μ are the degeneracies of the λ, μ irreducible representations respectively. Asterisks denote complex conjugates and 1, 2 label the two electrons. In the case of $C_{\infty v}$ symmetry our basis functions are already symmetrized (have $C_{\infty v}$ symmetry) and the symmetry index is the axial quantum number $|m|$.

The program of Module 8 stores the supermatrix, $\mathfrak{P}_{\lambda pq,\mu rs}$, in triangular form with row index λpq and column index μrs. Ordering of these indices is to increasing values of $\lambda(=|m|)$, and for a given value of λ, $p \geq q$; similarly for μrs. Thus, the rows and columns are in the order 011 021 022 031 032 033 $\cdots$ 111 121 122 131 132 133 $\cdots$. The elements stored for $\lambda = \mu$ have $pq \geq rs$, i.e., the lower triangle of the full supermatrix is stored. This is the same ordering described in detail in Ref. 34. Thus, the elements in the supermatrix in reading order would start out with indices 011011 021011 021021 022011 022021 022022 $\cdots$.

The entries in the list of matrix elements over the basis set, to be placed into appropriate positions in the supermatrix are

$$(pq|rs)_I = \int \chi^*_p(1)\chi^*_r(2)(r_{12})^{-1}\chi_q(1)\chi_s(2)d\tau_1 d\tau_2. \tag{37}$$

These matrix elements will be labelled by an additional superscript $M_<$ or $M_>$ depending on whether the axial dependence in the charge distributions pq, rs of Eq. (37) is $\big||m| - |m'|\big|$ or $|m| + |m'|$ respectively. If all four m values represented in Eq. (37) are nonzero, $M_< \neq M_>$; otherwise they are equal. The subscript I in Eq. (37) is used to designate that the quantity is an entry in the matrix elements over the basis set. If we denote the element $\mathfrak{P}_{\lambda pq,\mu rs}$ in the supermatrix by $(pq|rs)_S$, then for the case of $C_{\infty v}$ symmetry, Eqs. (34) to (36) can be rewritten as

$$(pq|rs)_S = (pq|rs)_I^{M_<} - \tfrac{1}{8}(pr|qs)_I^{M_<} - \tfrac{1}{8}(ps|qr)_I^{M_<} - \tfrac{1}{8}(pr|qs)_I^{M_>} - \tfrac{1}{8}(ps|qr)_I^{M_>}. \tag{38}$$

Equation (38) yields the correct result when $M_< = M_>$ as well as for the more general case. The algorithm for placing a member of a nonredundant list of matrix elements over basis functions into the supermatrix satisfying Eq. (38) is given below.

1. The triangular sections of the supermatrix corresponding to $\lambda = \mu$ are constructed from matrix elements over basis functions which all have the same $|m|$ value, i.e., symmetry classification $(\lambda\lambda|\lambda\lambda)$. These are placed into the supermatrix in the following ways:

a) $p = q = r$ and/or $q = r = s$

$$-\tfrac{1}{4}(pq|rs)_I^{M_>} \to (pq|rs)_S$$
$$\tfrac{3}{4}(pq|rs)_I^{M_<} \to (pq|rs)_S$$

b) $p = q$ and/or $r = s$ but not including 1(a) above.

$$-\tfrac{1}{8}(pq|rs)_I^{M_>} \to (pr|qs)_S$$
$$(pq|rs)_I^{M_<} \to (pq|rs)_S$$
$$-\tfrac{1}{8}(pq|rs)_I^{M_<} \to (pr|qs)_S$$

Note that in computing the supermatrix address of $(pr|qs)_S$ it may be necessary to transpose q and s to conform with the triangular indexing convention.

c) $p = r$ and/or $q = s$ but not including 1(a) above

$$-\tfrac{1}{8}(pq|rs)_I^{M_>} \to (pq|rs)_S$$
$$-\tfrac{1}{4}(pq|rs)_I^{M_>} \to (pr|qs)_S$$
$$\tfrac{7}{8}(pq|rs)_I^{M_<} \to (pq|rs)_S$$
$$-\tfrac{1}{4}(pq|rs)_I^{M_<} \to (pr|qs)_S$$

d) $q = r$ but not including case 1(a) above

$$-\tfrac{1}{8}(pq|rs)_I^{M_>} \to (pq|rs)_S$$
$$-\tfrac{1}{4}(pq|rs)_I^{M_>} \to (ps|qr)_S$$
$$\tfrac{7}{8}(pq|rs)_I^{M_<} \to (pq|rs)_S$$
$$-\tfrac{1}{4}(pq|rs)_I^{M_<} \to (ps|qr)_S$$

e) $p \neq q \neq r \neq s$

$$
\begin{aligned}
-\tfrac{1}{8}\,(pq|rs)_I^{M>} &\longrightarrow (pr|qs)_S \\
-\tfrac{1}{8}\,(pq|rs)_I^{M>} &\longrightarrow (ps|qr)_S \\
(pq|rs)_I^{M<} &\longrightarrow (pq|rs)_S \\
-\tfrac{1}{8}\,(pq|rs)_I^{M<} &\longrightarrow (pr|qs)_S \\
-\tfrac{1}{8}\,(pq|rs)_I^{M<} &\longrightarrow (ps|qr)_S
\end{aligned}
$$

2. The rectangular sections of the supermatrix corresponding to $\lambda > \mu$ are contributed to by matrix elements over the basis functions of symmetry classifications $(\lambda\lambda|\mu\mu)$ and $(\lambda\mu|\lambda\mu)$. These are placed as follows:

a) Matrix element symmetry classification is $(\lambda\lambda|\mu\mu)$.

$$(pq|rs)_I \longrightarrow (pq|rs)_S$$

b) Matrix element symmetry classification is $(\lambda\mu|\lambda\mu)$ and $p = r$ and/or $q = s$.

$$
\begin{aligned}
-\tfrac{1}{4}\,(pq|rs)_I^{M>} &\longrightarrow (pr|qs)_S \\
-\tfrac{1}{4}\,(pq|rs)_I^{M<} &\longrightarrow (pr|qs)_S
\end{aligned}
$$

c) Matrix element symmetry classification is $(\lambda\mu|\lambda\mu)$ for all cases not covered by 2(b) above.

$$
\begin{aligned}
-\tfrac{1}{8}\,(pq|rs)_I^{M>} &\longrightarrow (pr|qs)_S \\
-\tfrac{1}{8}\,(pq|rs)_I^{M<} &\longrightarrow (pr|qs)_S
\end{aligned}
$$

We should emphasize that in the case that $M_< = M_>$ all steps listed under 1 and 2 above must be executed. Thus, for example, in case 1(a) if $\lambda = 0$ the procedure would simplify to $\tfrac{1}{2}\,(pq|rs)_I \longrightarrow (pq|rs)_S$. In placing an integral, care must be taken to cover the possible reordering of the second index pair to compute the proper address in the supermatrix.

Modules 9 and 11 construct self-consistent field-molecular orbitals for a closed-shell single configuration wave function. Excellent discussions of the theory and computational procedures are available elsewhere[13,34] and need not be recapitulated here. In fact, our programs for these modules were adapted from the atomic self-consistent field program of Roothaan and Bagus[34], whose cooperation in making their program available is gratefully acknowledged.

These modules also produce the expectation values of the list of one-electron operators supplied as input data or, in the event that none was supplied, the expectation values of a standard set of operators. The standard set comprises those necessary for determining the dipole and quadrupole moments of the system.

In writing this program we chose to make it completely independent of the system monitor. This was done partly to ensure that the program would work on any stand-alone IBM 7094 without any trouble, thereby avoiding the distribution problems which plague most large-scale programs. In addition, on the IBM 7094 it is possible to use the basic input, output and IO subroutines from the system programs within the framework of one's own control program with ease, and we considered this to be simpler and more efficient than working within the framework of the system monitor. This approach will no longer be satisfactory for the current generation of computers if one wishes the program to work in a computer partitioned to work simultaneously on independent programs.

Input specifications and examples

Details of the input data to the program McL-YOSH LINEAR MOLECULE PROGRAM 1 will be given in this section, along with some examples. This will serve the function of demonstrating the simplicity and flexibility of the input data, and the level of control that the user has in specifying program parameters. It will also be complete enough to guide potential users of the program in the preparation of input data. Additional details are included beyond what has appeared in the User's Manual.[16]

The execution time for any run is a function of the size of the basis set and is essentially proportional to the number of matrix elements and the number of quadrature points for numerical integration, at least for typical runs with this program. A typical time per matrix element averaged over the entire computation is in the range 50 to 100 milliseconds. Thus, computation of a wave function for a molecule with no inversion center using a basis of 21 σ and 9 π functions with 1240 quadrature points takes approximately 3500 seconds, while if 26 σ and 14 π basis functions and a 1024-point quadrature formula are used the execution time increases to approximately 11,000 seconds. A repeat calculation in which one basis function is added or changed, takes approximately 25 percent of the time to carry out the calculation from scratch.

The format used in presenting the input blocks is that punched into cards constituting the input data decks.[11,38] Each card contains three types of information, punched into the location field (columns 2 to 7), the operation field (columns 8 to 10) and the data field (columns 12 to 72). The location field, if present, gives the name of the current input block and symbolically identifies the address into which the data commences loading. If not present, the data load consecutive to the data on the previous card. The operation field specifies the conversion mode for the storing of the data. The three entries used in the operation field are:

BCI — The data field begins (anywhere in columns 12 to 16) with either a comma or with a count digit N ($N = 0, 1, \cdots 9$) followed by a comma. The data words begin immediately following the comma. If N is present, N words (each of six BCD characters) will be set into storage; if N is absent, 10 such words will be set. ($N = 0$ is interpreted as $N = 1$.)

OCT — The data words consist of signed or unsigned octal integers separated by commas. The first of these begins anywhere within columns 12 to 16, and the last is signalled by a following blank column.

DEC Data words are separated by commas; the first of these begins anywhere within columns 12 to 16, and the last one is signalled by a following blank column. Each word can be any one of (1) a signed or unsigned decimal integer (no decimal point); such an integer is stored as a binary integer, but may have an absolute value no greater than $2^{35} - 1$; any exceeding this value will be left unstored, as will all subsequent data on the same card; (2) signed or unsigned decimal numbers written with a decimal point, and optionally followed by a power of 10 scaling factor. Such a number is converted to a binary floating point word, which may not exceed the approximate range 10^{-38} to 10^{38}. The entries 1.234E−1, 0.1234, 0.01234E1 would all be converted to the same floating point number, for example.

The key input flag determining the nature of the current computation is RERUN. The meaning of this flag will be described first and, for each value that it can take on, a listing of the remaining mandatory and optional input data blocks will be given. These remaining data blocks will then be described in an order which corresponds to the type of input. The program reads the data in two sections, one in Module 2 before the integral computation and the other in Module 8 or 10 before the construction of the supermatrix. The first data essentially contain the nuclear geometry, the basis set and a list of one-electron operators, while the second essentially contain the atomic numbers of the nuclei, the electron configuration and SCF trial vectors. The two sections of input are separated by a blank card as can be seen in the examples of Tables 3–6 which can be used to illustrate and clarify the following descriptions. The dimension, given in the description of each input block, is the maximum number of memory locations assigned to the designated block and is, therefore, a limitation on the program. The dimensions chosen essentially place no limitation on calculations practicable (from the point of view of execution time) on the IBM 7094. Actual information to be punched into input cards is shown boxed.

RERUN Dimension 1. Required input on all runs.

RERUN DEC X

X is an integer $0 \leq X \leq 7$ and its possible values have the following meanings.

$X = 0$ denotes a new run, using no information from previous runs. The wave function and one-electron expectation values called for by the input data will be computed. Intermediate results (matrix elements over the Slater basis, and potentials of charge distributions arising from products of Slater functions) together with the computed wave function and input data that generated it are stored in a file on tape unit A6 at the completion of the run.

Mandatory input blocks are TITLE, SERNO, RERUN, RKK, NSF, TSF, ATMNO, NCOB, NZCOE. Optional input blocks are SYM, IRFLG, ZINTP, RINTP, OEOP, CODEW, SFOPT, CWOPT, SCFCW. Input blocks not explicitly listed above are not allowed. Examples are given in Tables 3 and 4.

$X = 1$ computes a wave function and expectation values using a basis set which is changed from that used for a previous computation in a way specified by the input deck. The output tape, A6, from the previous computation is remounted and becomes an input tape for this run. Use of this type of run can lead to considerable saving of time since, for most cases, no significant time consuming computation which was also needed for the previous run is repeated. At the completion of the run, a file on the A6 output tape is written (over the previous one unless a new tape was mounted) which, for future runs, is the same as would have been generated in a RERUN $= 0$ computation with the current basis set. This tape can, therefore, be used as an input tape on future computations with RERUN > 0.

Mandatory input blocks are TITLE, SERNO, RERUN. Optional input blocks are CSF, ASF, OEOP, CODEW, SFOPT, CWOPT, ATMNO, NCOB, NZCOE, SCFCW. Input blocks not explicitly listed are not allowed.

Examples are given in Tables 3 and 4.

$X = 2$ is the same as $X = 1$ except that additional integrals between Slater functions, not needed for a single configuration SCF calculation, will be computed if this was not the case previously. Details of this are given in the discussion of Module 4 in a previous section of the paper.

$X = 3$ is used to recompute a wave function with the same basis set as a previous run whose output tape is remounted on either tape drive A5 or A6. Such runs may be necessary in case that the convergence level reached in a previous computation was unsatisfactory. At completion of the run an A6 output tape is generated which, for future runs, is the same as would have been generated in a RERUN $= 0$ computation with the current basis set. The result of a RERUN $= 3$ computation could also be achieved, with longer execution time, by a RERUN $= 1$ with no ASF or CSF input blocks. This latter procedure is more satisfactory if one-electron expectation values are required, since the OEOP input block is not allowed if RERUN $= 3$.

Mandatory input blocks are TITLE, SERNO, RERUN. Optional input blocks are NZCOE, SCFCW, CODEW. All input blocks not explicitly listed are not allowed.

$X = 4$ is used to compute expectation values of one-electron operators with a previously computed wave function available as an input tape on either tape drive A5 or A6 (see CODEW input block). Such runs will be necessary in case that matrix elements of the operators, which are computed by numerical integration, are needed to higher accuracy than is obtainable with the integration grids used for

Table 3 Input data decks for generating C_2H_2 (DZ + P) wave function.[a]

TITLE	BCI ,	C2H2 DZ BASIS SET
SERNO	DEC	1661
RERUN	DEC	0
SYM	DEC	1
RKK	DEC	4,2.002,2.281,2.002
NSF	DEC	9,3
TSF	DEC	1,0,1,0.97493,1,0,1,1.2029,2,1,1,1.72338,1,0,2,5.2309
	DEC	1,0,2,7.96897,2,0,2,1.16782,2,0,2,1.82031,2,1,2,1.25572
	DEC	2,1,2,2.72625,2,1,1,0.7901,2,1,2,1.25572,2,1,2,2.72625
RINTP	DEC	−1,16
ZINTP	DEC	−1,16,8,8,8,8,8,8,16
OEOP	OCT	21,401000,401101,401104,401102,401103,401202,401203
	OCT	403202,403203,404102,404103,404202,404203
	OCT	412001,412004,412002,412003
		(Blank card)
ATMNO	DEC	1.0,6.0,6.0,1.0
NCOB	DEC	3,2,1
NZCOE	DEC	0,1,4,1.,0,2,7,1.,0,3,1,−.3,0,3,2,.7,0,3,8,−.3,1,1,4,1.
	DEC	1,2,2,.5,1,2,7,.4,2,1,2,1.
		(Blank card)
TITLE	BCI ,	C2H2 DZ + 3D BASIS SET
SERNO	DEC	1661
RERUN	DEC	1
ASF	DEC	0,3,3,2,1,1.65,0,9,3,2,1,1.99175,1,1,3,2,1,2.31071
	DEC	1,3,3,2,2,2.13462
		(2 blank cards)
TITLE	BCI ,	C2H2 DZ + P BASIS SET
SERNO	DEC	1661
RERUN	DEC	1
ASF	DEC	0,11,4,3,2,1.86851,1,5,4,3,2,2.14558
		(2 blank cards)

[a] The C_2H_2 (DZ + P) wave function generated by the third of these input decks is on page 212 of the "Tables of Linear Molecule Wave Functions," which supplements this paper. (Ref. 18.)

Table 4 Input data decks for generating HCN (DZ + P) wave function.[a]

TITLE	BCI ,	HCN DZ BASIS SET
SERNO	DEC	167
RERUN	DEC	0
RKK	DEC	3,2.0143,2.1791
NSF	DEC	15,5
TSF	DEC	1,0,1,0.97155,1,0,1,1.23206,2,1,1,1.37568,1,0,2,5.2309
	DEC	1,0,2,7.96897,2,0,2,1.16782,2,0,2,1.82031,2,1,2,1.25572
	DEC	2,1,2,2.72625,1,0,3,6.11863,1,0,3,3.93843,2,0,3,1.39327
	DEC	2,0,3,2.22157,2,1,3,1.50585,2,1,3,3.26741,2,1,1,0.79006
	DEC	2,1,2,1.25572,2,1,2,2.72625,2,1,3,1.50585,2,1,3,3.26741
RINTP	DEC	−1,16
ZINTP	DEC	−1,16,8,8,8,8,16
		(Blank card)
ATMNO	DEC	1.0,6.0,7.0
NCOB	DEC	5,1
NZCOE	DEC	0,1,10,.8,0,1,11,.2,0,2,4,.84,0,2,5,.16,0,3,10,−.28
	DEC	0,3,12,.3,0,3,13,.76,0,4,1,.4,0,4,7,.8,0,4,8,−.8
	DEC	0,5,13,.7,0,5,14,.7,1,1,2,.3,1,1,4,.7,1,1,5,.2
		(Blank card)
TITLE	BCI ,	HCN DZ + 3D BASIS SET
SERNO	DEC	167
RERUN	DEC	1
ASF	DEC	0,3,3,2,1,1.7,0,9,3,2,2,2.27993,0,15,3,2,3,2.00387
	DEC	1,1,3,2,1,2.2,1,3,3,2,2,2.27222,1,5,3,2,3,2.18731
		(2 blank cards)
TITLE	BCI ,	HCN DZ + P BASIS SET
SERNO	DEC	167
RERUN	DEC	1
ASF	DEC	0,11,4,3,2,1.69214,0,18,4,3,3,2,2.7,1,5,4,3,2,2.24432
	DEC	1,8,4,3,3,2.53488
		(2 blank cards)
TITLE	BCI ,	HCN DZ + P BASIS SET, CHANGE GEOMETRY
SERNO	DEC	167
RERUN	DEC	5
RKK	DEC	3,2.2143,2.3791
		(2 blank cards)

[a] The HCN (DZ + P) wave function generated by the third of these input decks is on page 156 of the "Tables of Linear Molecule Wave Functions" which supplements this paper. (Ref. 18.)

Table 5 Input data decks for computing expectation values of one-electron operators with FCN (BA + P) wave function.

TITLE	BCI ,	FCN BA + P WAVE FUNCTION RERUN 4−1
SERNO	DEC	9670
RERUN	DEC	4
RKK	DEC	3,2.38109,2.20156
CODEW	DEC	0,0,0,0,0,0,88
OEOP	OCT	14,401000,401101,401102,401103,401202,403102,403202
	OCT	404102,404202,412001,412002,412003
		(2 blank cards)
TITLE	BCI ,	FCN BA + P WAVE FUNCTION RERUN 4−2
SERNO	DEC	9670
RERUN	DEC	4
RKK	DEC	3,2.38109,2.20156
RINTP	DEC	−1,28
ZINTP	DEC	−1,28,14,14,14,14,28
OEOP	OCT	7,401000,413001,413002,413003,414001,414002,414003
		(2 blank cards)

the wave function computation. Computation of the electronic component of the electric field and electric field gradient at the nuclei must be done with RERUN = 4 or RERUN = 7 since the required matrix elements cannot be evaluated on any other type of run.

Mandatory input blocks are TITLE, SERNO, RERUN, RKK, OEOP. Optional input blocks are ZINTP, RINTP, CODEW. All blocks not explicitly listed are not allowed.

Examples are given in Table 5.

$X = 5$ is used to compute a wave function and expectation values with the same basis set as used in a previous run, but with a different quadrature formula. The quadrature formula can be for the same or for different nuclear geometry. This is the standard type of run for generating points on a potential surface with a basis set generated for a previous point. The output tape generated on A6 for the previous point is input to the RERUN = 5 run and must be mounted on A6.

Mandatory input blocks are TITLE, SERNO, RERUN, RKK. Optional input blocks are ZINTP, RINTP, OEOP, SFOPT, CWOPT, ATMNO, NCOB, NZCOE, SCFCW, CODEW. Input blocks not explicitly listed are not allowed.

$X = 6$ is used as a preliminary run to a RERUN = 7, and is only allowed for a run of $C_{\infty v}$ symmetry. It operates in exactly the same way as a RERUN = 3 except that it terminates after the construction of the $C_{\infty v}$ supermatrix in Module 8.

Mandatory input blocks are TITLE, SERNO, RERUN. No other input is allowed. An example is given in Table 6.

$X = 7$ is designed for the calculation of the axial components of the molecular polarizability tensor. It is restricted to $C_{\infty v}$ symmetry and must follow immediately after a run which computes the two-electron supermatrix. This could be a RERUN = 0, 1, 2, 3, 5, 6. (To do calculations on molecules of $D_{\infty h}$ symmetry, the run must be set up as if the molecule had the lower $C_{\infty v}$ symmetry.) The run is set up as if there were two additional nuclei symmetrically placed with respect to the center of mass on the molecular

Table 6 Input data decks for calculating molecular polarizabilities and nuclear electric shielding factors with HCN (BA+P) wave function.

TITLE	BCI , HCN BA+P WAVE FUNCTION POLARIZABILITY CALCULATION	
SERNO	DEC	1670
RERUN	DEC	6
		(2 blank cards)
TITLE	BCI , HCN BA+P WAVE FUNCTION POLARIZABILITY CALCULATION	
SERNO	DEC	1670
RERUN	DEC	7
RKK	DEC	5,16.93116,2.0143,2.1791,18.87544
RINTP	DEC	−1,32
ZINTP	OCT	6,10,777777777777,0
	DEC	−6.0,32,0,16.93116,15.93116
	DEC	32,16.93116,18.94546,17.93831
	DEC	32,18.94546,21.12456,20.03501
	DEC	32,21.12456,50.0,22.12456,8,40.0
	OCT	377777777777
	DEC	46.0
OEOP	OCT	22,1000,0,2101,2102,2103,2104,2105,403102,403202,404202
	OCT	413001,413002,413003,413004,414001,414002,414003,414004
		(blank card)
ATMNO	DEC	0,1.0,6.,7.,0
POLCH	DEC	7,3.06884,0,0
	DEC	0,2.,0,−2.,2.,2.,−2.,−2.,−2.,2.,2.,−2.
SCFCW	DEC	30,0,0,0,0,0,0.5E−7,0,0,1.0E−5
		(blank card)

nuclei. A series of molecular wave functions will be calculated for different charges placed at the positions of these false nuclei, and the polarizabilities are computed from them as described later in this paper.

Mandatory input blocks are TITLE, SERNO, RERUN, RKK, ZINTP, OEOP, ATMNO, POLCH. Optional input blocks RINTP, CODEW, SCFCW, NZCOE. All input blocks not explicitly listed above are not allowed.

TITLE Dimension = 10. Required input on all runs.

```
TITLE  BCI  , ANY TITLE TO APPEAR ON
EACH OUTPUT PAGE → column 72.
```

This card enables simple identification of each run by heading each output page with a title provided by the user of the program.

SERNO Dimension = 1. Required input on all runs.

```
SERNO  DEC  X
```

X is the serial number of the current sequence of runs. Any run in which RERUN $\neq$ 0, which takes information from a tape, A5 or A6 (see CODEW), written at the end of a previous run, must have the same SERNO entry as the previous run. The SERNO entry in the input deck is always checked against that saved from the previous run, and if they do not match the program halts on an alarm.

SYM Dimension = 1. Required input for $D_{\infty h}$ symmetry.

```
SYM  DEC  X
```

$X = 0$ denotes $C_{\infty v}$ symmetry

$X = 1$ denotes $D_{\infty h}$ symmetry

If this input card is not present SYM is automatically set to zero. The format of input blocks NSF, TSF, ASF, CSF, NZCOE depends on the value of SYM.

SYM = 1 is not allowed for RERUN = 6, 7.

IRFLG Dimension = 1.

```
IRFLG  DEC  X
```

$X = 0$ implies that the index list generated in Module 4 will contain only matrix elements needed for single-configuration calculations.

$X = 1$ implies that the index list will contain all possible matrix elements over the basis set.

IRFLG is set internally to zero; hence this card is needed only if the internally set value is to be overwritten.

The following input blocks NSF, TSF, CSF, ASF are used to specify the basis set.

NSF Dimension = 5. Required input if RERUN = 0.

NSF DEC $n_\sigma, n_\pi, n_\delta, n_\phi, n_\gamma$

The entries of the NSF input block are automatically set to zero by the program, and will be overwritten by the entries punched into the input card.

For $C_{\infty v}$, the entries are the number of basis functions of the indicated symmetry type.

For $D_{\infty h}$, the entries are the number of nonredundant basis functions of the indicated symmetry type, i.e., basis functions on nuclei which are equivalent to one previously counted are not included.

TSF Dimension 300. Required input if RERUN = 0.

TSF DEC $(n, l, k, \zeta)_\sigma$

$(n, l, k, \zeta)_\pi$

.

.

.

Each basis function is defined by four entries n, l, k, ζ, as in Eq. (2). Basis functions are ordered to increasing $|m|$ values, i.e. all σ before all π before all δ, etc. The position in which they appear in this table is the index of the basis function. For $C_{\infty v}$, an entry into this input block is made for each basis function. For $D_{\infty h}$, we include only entries corresponding to the nonequivalent nuclei, i.e., the ones counted in NSF. (This truncated table is expanded internally to include all basis functions.)

The maximum number of basis functions is 75. In the case of $D_{\infty h}$ this applies to the expanded table.

CSF Dimension = 92. Optional input if RERUN = 1, 2.

CSF DEC ($|m|, p, n, l, k, \zeta$)

This input block allows the changing of basis functions. Each changed basis function is specified by six entries. The first two, $|m|$ and p, specify the index of the basis function to be changed, namely the p-th function in the block of symmetry type specified by $|m|$. The new parameters of this basis function are given by the remaining four entries, n, l, k, ζ.

For both $C_{\infty v}$ and $D_{\infty h}$ the indexing refers to that in the input block TSF.

The ordering of entries in CSF must be to increasing $|m|$, and for a given $|m|$ value to increasing p.

The maximum number of changed functions in a run is 15.

ASF Dimension 92. Optional input if RERUN = 1, 2.

ASF DEC ($\pm|m|, p, n, l, k, \zeta$)

This input block allows the adding of basis functions. Each added basis function is specified by six entries. The first two, $|m|$ and p, specify the basis function in the previous basis set relative to which an insertion is to be made. $-|m|$ means insert before the p-th basis function in the block of symmetry $|m|$; $+|m|$ means insert after. A special case is when a symmetry was previously unoccupied. This is indicated by setting $p = -0$ for the basis function which populates a previously unoccupied symmetry. If more than one added function has the same $\pm|m|, p$ value then the functions are added in the order in which they appear in this input block. The parameters of the added basis function are given by the remaining four entries n, l, k, ζ.

For both $C_{\infty v}$ and $D_{\infty h}$ the indexing refers to that in the input block TSF, which will be augmented by the program. At the completion of this run TSF will look as if the run had been made from scratch, and a subsequent addition must be made relative to the indexing of TSF at the end of this run. Tables 3 and 4 give examples.

The ordering of entries in ASF must be to increasing $|m|$; for a given $|m|$ negative entries appear before positive; for a given $\pm|m|$ entries are in the order of increasing p.

The maximum number of added functions in a run is 15.

SFOPT Dimension = 60. Optional input on any run except RERUN = 3, 4, 6, 7.

SFOPT DEC ($|m|, p$, inc)

This input block specifies the basis functions which are to be optimized with respect to their orbital exponent. Each basis function to be optimized is designated by three entries in which the first two, $|m|$ and p (integers), identify the basis function and the third, inc (floating point), is the increment by which the orbital exponent is changed in performing the optimization.

If this input deck is present in a RERUN = 1, 2 deck in which there are added Slater functions specified in the input block ASF, then the value of p in this input deck must be the serial number *after the added functions have been inserted into* TSF, *as described in* ASF.

If more than one Slater function is specified in this input, then one-dimensional optimizations are done with respect to the orbital exponents in the order in which they appear in the input block SFOPT. The procedure for a single optimization is as follows. The initial orbital exponent is changed in steps of inc until three energies span a minimum. Assuming a functional relationship between energy and orbital exponent, specified by a code word in the input block CWOPT, the value of the exponent corresponding to minimum energy is computed. A computation is done for this value of the exponent, and then the procedure is repeated for the next basis function specified in SFOPT.

The optimization procedure for a given exponent is terminated under any of the following conditions:

1. Five computations are done and a minimum has not yet been spanned. The lowest energy computation of the five will be selected, ζ reset to the value appropriate to that computation; and optimization with respect to the next Slater function listed in SFOPT is commenced.
2. The first run attempted, in optimizing an exponent, fails. The computation, including all optimization not yet attempted, is terminated.
3. A run, subsequent to the first, in optimizing an exponent fails. The exponent is reset to the value which gave the lowest energy in the previous successful runs performed in optimizing this exponent and that computation repeated. Optimization with respect to the next basis function listed in SFOPT is then commenced.
4. After three successful runs for a given exponent, energy differences between runs are computed. If these differences are less than a threshhold, ΔE_2, described in the input block CWOPT, then the same procedure is followed as in (3) above.

The maximum number of functions that can be listed in this input block is 20.

CWOPT Dimension = 5. Optional input on any run containing the SFOPT input block.

CWOPT	DEC	Curve type, E_{est}, ΔE_1, Con_1, ΔE_2

This input block contains the code word for the functional relation between E and ζ assumed in the optimization circuit, and the criteria to be used to establish whether a particular run failed or not, for purposes of optimization.

The entries in CWOPT are set as follows. If explicitly given in the input block above then the value given is used. If not, then either the last value used (in case that RERUN $\neq$ 0) or a standard value (in case that RERUN = 0) is set for the parameter.

Curve type = 0 Assumed relation between ζ and E is $E = A + B\zeta + C\zeta^2$ where A, B, C are constants determined by the program.

Curve type = 1 Assumed relation between ζ and E is $E = A + B\zeta + C/\zeta$ where A, B, C are constants determined by the program.

E_{est} Estimated total energy for the state being computed. It is used in combination with the next entry ΔE_1 to determine whether the current computation has produced an acceptable result (i.e., has not converged on different state). If $E_{est} = 0$ this test is bypassed.

ΔE_1 If $|E_{\text{current computation}} - E_{est}| < \Delta E_1$ then the run is successful. Otherwise unsuccessful. The test is not applied if $E_{est} = 0$.

Con_1 If the convergence threshhold achieved in the SCF computation (and available as output from the SCF program) is less than Con_1, then the run is successful. If Con_1 is zero, then this test is made against a standard of 0.0001.

ΔE_2 After three successful runs with a given exponent, the differences between the one with lowest energy and the others are computed. If they are $< \Delta E_2$, then we recompute the energy and wave function for the case that yielded the lowest energy, and go on to the next Slater function to be optimized. If $\Delta E_2 = 0$, then this test is made against a standard value of 0.00005.

The next three input blocks RKK, ZINTP and RINTP specify the nuclear geometry and the quadrature formula to be used.

RKK. Dimension 12. Required input for RERUN = 0, 4, 5, 7.

RKK	DEC	$N, R_{12}, R_{23}, \cdots R_{N-1N}$

N = Number of nuclei (integer). In the case of RERUN = 7 it is the number of nuclei in the molecule +2 corresponding to the positions where point charges will be introduced for polarizability and shielding calculations.

$R_{12}, R_{23} \cdots$ are the internuclear distances (floating point) between adjacent nuclei, counting left to right. If RERUN = 7, positions 1 and N are located at distances $\pm R$ relative to the center of mass of the molecular nuclei which is given in input block POLCH.

For $D_{\infty h}$ molecules all nuclei are counted in N, and all adjacent internuclear distances are entered. The maximum value of N is 12.

ZINTP Dimension = 60. Required input for RERUN = 0, 4, 5, 7 if automatic setting by program is not satisfactory (see format 2 below). Not allowed on other runs.

There are three acceptable input formats for ZINTP, distinguished by the first word in the block.

Format 1. First word of block > 0.

ZINTP	DEC	N	(n, a, b, m)

Following the discussion of Module 2 in a previous section of the paper, N (integer) is the number of segments into which the z integration is broken. Each segment is specified by four entries. The first, n (integer), is the number of quadrature points for the segment. The next three, a, b and m (floating point numbers), have the meaning indicated in the Module 2 discussion. (If one of the limits is infinite it is punched into an input card with an OCT rather than a DEC format: OCT 777777777777 codes $-\infty$, while OCT 377777777777 codes $+\infty$.) Values of n are restricted to 2, 4, 6, 8, 10, 12, 14, 16, 20, 24, 28, 32, 40, 48, 64.

The first nucleus (or left hand position of a perturbing charge for RERUN = 7) corresponds to $z = 0$. Format 1 should always be used for RERUN = 7 (see Table 6) since the other two formats are unsatisfactory. Format 1 is the most flexible of the three, giving the user great flexibility in setting up the quadrature points.

Format 2 is elected in RERUN = 0, 4, 5 if there is no input card present, or if the first word on the input card is zero. In this case input will be set up internally as if format 1 had been used. The segments will be automatically chosen by the program. The first segment has $n = 20$, $a = -\infty$, $b = 0$, $m = -1.0$ and the last has $n = 20$, $a = R_N$, $b = \infty$,

$m = R_N + 1.0$ where R_N is the value of z at the rightmost nucleus. For distances less than 3.0, each adjacent internuclear separation is broken up into two equal segments each having $n = 10$, the first segment has m positioned 0.4 of the segment length from a and the second segment has m positioned 0.4 of the segment length from b. For distances greater than or equal to 3.0, each adjacent internuclear separation is broken into three segments, each with $n = 10$, the first and third covering distances of 1.0 from the two nuclei with m values positioned 0.4 from each nucleus. The remaining segment covers the rest of the internuclear separation with m positioned in the center. The program generates the segments in order of increasing z, i.e., from left to right across the range $-\infty$ to ∞.

This method of selecting segments partially reflects our experience and is designed to provide a conservative quadrature formula for molecules containing first row atoms.

Format 3. First word of block is -1.

```
ZINTP   DEC   −1, n1, n2, n3, ···
```

Format 3 produces the same result as format 2 except that the number of points for each segment is explicitly given by the entries after the first in the above input block. Thus, the segments are those selected in format 2, in order of increasing z, with the number of points for each segment given by $n_1, n_2, n_3 \cdots$.

It must be noted that for a large number of nuclei, the dimension of ZINTP can be overflowed if format 2 or 3 is selected. If this is the case, format 1 must be used in a way that does not use more than 60 words.

RINTP Dimension 60. Required input for RERUN = 0, 4, 5, 7 if automatic setting by program is not satisfactory (see format 2 below). As with ZINTP there are three acceptable formats for the input block RINTP.

Format 1. First word of block > 0.

```
RINTP   DEC   N   (n, a, b, m)
```

The entries in this input block have the same meaning as for ZINTP, except that the integration variable ρ is being covered rather than z.

Format 2. Same meaning as in ZINTP. A single segment will be generated with $n = 20, a = 0, b = \infty, m = 1.0$.

Format 3. First word of block is -1.

```
RINTP   DEC   −1, n1
```

Same meaning as in ZINTP. The single segment automatically selected is spanned with n_1 points.

OEOP Dimension = 21. Optional input on all runs except RERUN = 3, 6 where it is not allowed.

```
OEOP   OCT   N, Op1, Op2, Op3, ··· OpN
```

This input block defines the one-electron operators whose matrix elements will be computed. N (octal integer) is the number of operators, and $\mathrm{Op}_1, \mathrm{Op}_2 \cdots$ (octal integers) are code words defining the operators, one per operator. These code words are right-adjusted integers with the format MFOONKK. M specifies the axial dependence, $\exp(iM\phi)$, of the operator and F specifies the manner of computation. $F = 0$ specifies that the special circuits of Module 6 are used (restricted to kinetic energy, overlap and nuclear attraction) and $F = 4$ specifies that the numerical integration circuits will be used (except for one-center matrix elements of operators z/r^3 and $(3z^2 - r^2)/r^5$ which are computed analytically.) The values of the remaining integers in MFOONKK can be read from the following list.

Operator	OO	N	KK
Kinetic Energy	00	—	—
r_k^n	01	n	k
$(1/r_k)^n$	02	n	k
z_k^n	03	n	k
$\rho_k^n \exp(im\phi)$	04	n	k
$\xi_{kk'}^n$	05	n	kk'
$(1/\xi_{kk'})^n$	06	n	kk'
$\eta_{kk'}^n$	07	n	kk'
$(1/\eta_{kk'})^n$	10	n	kk'
$\sin^2\theta_k/r_k$	11	—	k
$\cos^2\theta_k/r_k$	12	—	k
z_k/r_k^3	13	—	k
$(3z_k^2 - r_k^2)/r_k^5$	14	—	k

The operators z/r^3 and $(3z^2 - r^2)/r^5$ can be requested only if RERUN = 4, 7.

For a given run the OEOP block is set as follows.

RERUN = 3, 6. OEOP input is not allowed.

RERUN = 7. OEOP input deck must appear and must contain the codes for overlap, kinetic energy, nuclear attraction for all nuclei, including positions of perturbing charges, z, z^2 and ρ defined with respect to the left-hand molecular nucleus ($k = 2$) in exactly this order (see Table 6.) Additional operators can be in any order.

RERUN = 0, 1, 2, 4, 5. If no OEOP input card is in the deck a list is generated internally. This list contains operators sufficient to compute the dipole and quadrupole moments of the molecule and is OEOP OCT 4, 401000, 403101, 403201, 401201. If the OEOP input cards are in the input deck, then matrix elements of the specified operators are computed.

CODEW. Dimension 20. Optional input on all runs.

```
CODEW  DEC  LMAX, TH2CP, NXI, NETA
       DEC  b, ITRSH, FILE
```

Only the first seven entries in this input block are defined. The first six specify miscellaneous program parameters which would only be set from input data under exceptional circumstances. Normally, they would be set internally to standard values; this will be the case if the input block is not present or the entry in the input block is zero. For all practical purposes they can be disregarded; details are contained in Ref. 16. The seventh entry in this input block, FILE, specifies the position of binary data saved from a previous run which is to be used as input on the current run.

FILE = 0 must be used for RERUN = 1, 2 and indicates that the binary input is in the first file of the tape on A6. FILE = N can be used if RERUN = 3, 5, 6 and indicates that the required binary input data is contained in file N on tape unit A5.

The remaining input blocks, if present, follow a blank card in the input deck, and are read in Module 8 or 10 depending on whether SYM = 0 or 1.

ATMNO Dimension 12. Required input for RERUN = 0, 7

ATMNO DEC $Z_1, Z_2, \cdots$

$Z_1, Z_2, \cdots$ (floating point) are the nuclear charges on the centers defined in RKK. In the case of RERUN = 7, where the first and last centers are the positions of perturbing charges, the corresponding entries in ATMNO must be entered as zero.

NCOB. Dimension 10. Required input on RERUN = 0.

NCOB DEC $NS_1, NS_2, NS_3, \cdots$

$NS_1, NS_2, \cdots$ (integers) are the number of filled molecular orbitals of symmetry number 1, 2, $\cdots$. For $C_{\infty v}$ symmetry the symmetries are in the order $\sigma, \pi, \delta, \phi, \gamma$. For $D_{\infty h}$ they are in the order $\sigma_g, \sigma_u, \pi_u, \pi_g, \delta_g, \delta_u, \phi_u, \phi_g, \gamma_g, \gamma_u$.

NZCOE Dimension 2500. Required input for RERUN = 0.

NZCOE DEC (s, i, p, c)

This input block specifies the trial vectors used to initiate the SCF computation. In an actual run trial vectors are taken from the input deck if this input block is present, otherwise are taken from the output vectors of a previous run which were automatically saved at the end of that run. (In the case of an unsuccessful previous run, the trial vectors for that run were saved and these will be used on the current run if no NZCOE block is present in the input deck.)

In the input block NZCOE each nonzero coefficient in the trial vectors is specified by four entries. s (integer) identifies the symmetry of the molecular orbital. For $C_{\infty v}$ symmetry, $s = 0, 1, 2, 3, 4$ denotes $\sigma, \pi, \delta, \phi, \gamma$, while for $D_{\infty h}$ symmetry $s = 0, 1, 2, 3, 4 \cdots$ denotes $\sigma_g, \sigma_u, \pi_u, \pi_g \cdots$)the same ordering discussed in NCOB). i (integer) identifies the molecular orbital ($i = 1$ for the first of any symmetry). p (integer) identifies the symmetrized basis function which is contributing to the molecular orbital with coefficient c (floating point). For $C_{\infty v}$ symmetry the symmetrized basis functions are the basis functions themselves, and the value of s, p in this input block identifies the basis function in TSF of the same p value with $s = |m|$. For $D_{\infty h}$ symmetry it is necessary to construct a table of symmetrized basis functions by making the appropriate linear combinations from TSF, in the order of appearance in TSF, for each of the $D_{\infty h}$ symmetry classifications. p then indexes entries in this symmetrized list. The example of Table 3 should serve to make this clear.

POLCH Dimension 81. Required input for RERUN = 7 not permitted on other runs.

POLCH DEC $NP, CM, 0., 0., (0., Q, 0., -Q, Q, Q, -Q, -Q, -Q, Q, Q, -Q)$

This input block serves to define the perturbing charges which are introduced in order to perform the polarizability calculations described in the last section of the paper. NP (integer) is the number of pairs of entries following CM (floating point) which is the distance of the center of mass of the nuclei from the left-hand molecular nucleus. (The previous input RKK, ZINTP, has set up a system of two false nuclei symmetrically placed with respect to CM. Alternatively CM need not be the center of mass, but can be chosen as any desired expansion point relative to which the polarizabilities will be defined. Either way, the false nuclei set up by RKK and ZINTP must be symmetrically placed relative to CM.) The pairs of entries following CM are the values of the charges that are placed on the false nuclei. For each pair an SCF calculation will be performed. The first pair of 0., 0. calls for the normal unperturbed SCF result. The remaining pairs are in groups of six as indicated, corresponding to the six different perturbing fields for which SCF calculations will be done.

SCFCW Dimension 10. Optional input for any run.

SCFCW DEC SITMX, BIAS1, BIAS2, JITMX, NISVD, SCTH1, SCTH2, JCTH, DGATH, DGSTH

This input block is a set of parameters which defines convergence criteria and extrapolation information for the SCF computation, and convergence criteria for the diagonalization subroutines.

The value of a parameter listed in this input block is set in the same way as entries in CODEW and CWOPT. If the entry is not +0, then the entry is used as the current parameter value. If it is =0 then take the last value of the parameter used (if RERUN $\neq$ 0) or a standard value (if RERUN = 0).

SITMX Maximum number of SCF iterations which will be performed before jumping out of SCF program (integer). Standard = 50.

BIAS1 The number of iterations before an extrapolation (on the last three to produce the next set of trial vectors) is equal to BIAS1 + 2. BIAS1 is only used to set the number of iterations before extrapolation if the level of SCF convergence is worse than SCTH1. Standard value of BIAS1 is 1.

BIAS2 The number of iterations before extrapolation, after convergence of the density matrices, is better than SCTH1. Standard value is 3.

JITMX This is a signed integer. If negative, it takes unit vectors as trial vectors for each entry into the Jacobi diagonalization. If positive, it uses the output vectors of the previous diagonalization. The absolute value is the maximum number of iterations allowed in the diagonalization program.

For near-degenerate eigenvectors a positive entry should be used, otherwise the near-degenerate vectors may flip from one iteration to the next, making convergence impossible. Standard value is 1000.

NISVD This is a signed integer. If negative (*but not* -0), single-vector diagonalization will be bypassed. If positive the maximum number of iterations in the single-vector diagonalization before loosening the convergence threshhold (see DGATH, DGSTH writeups) will be set equal to NISVD. Standard value is 20.

SCTH1 See BIAS1, BIAS2 above. The purpose of this threshhold is to allow extrapolation to be used more heavily at the start of a SCF computation than in its later stages. Our observation is that extrapolation is helpful in preventing divergence at an early stage, but can hinder convergence at a later stage. Standard is 10^{-4}.

SCTH2 This is the convergence threshhold for the SCF procedure. These convergence threshholds SCTH1 and SCTH2 on the SCF procedure are used in tests on the density matrix. Comparison of the vectors obtained in the last two iterations (standard printed output) will show how this convergence level is reflected in vector convergence. Standard for SCTH2 is 10^{-8}.

JCTH Convergence threshhold for Jacobi diagonalization. If the single-vector diagonalization is also used, we can think of the Jacobi as being used to provide input vectors for the single-vector routine. Standard is 10^{-13}.

DGATH Final threshhold for single-vector diagonalization. If convergence to better than this threshhold cannot be obtained, the program jumps out on an alarm exit. Standard value is 10^{-4}.

DGSTH Initial convergence threshhold for single vector diagonalization. If after NISVD iterations convergence to better than a threshhold, initially set equal to DGSTH, is not obtained, the convergence threshhold for a single-vector diagonalization is loosened by a factor of 2. This loosening is repeated until either convergence is achieved, or the current threshhold is >DGATH which will cause an alarm exit. Standard value is 10^{-6}.

Molecular properties

A good example of the range of molecular properties which can routinely be computed with this linear molecule computer program is given in a recent publication on HF and HCl,[39] and discussion of some of these properties is covered in our previous work.[17–21] In this section of the paper we will restrict our discussion to molecular polarizabilities and nuclear electric shielding factors since the methods we use for computing these properties have not previously been described in detail.

Static electric polarizabilities and shielding factors are properties of an electronic charge distribution, in an atom or molecule, which are determined by its distortion under the action of an external electric field. We proceed to define these quantities for the case of an atom or linear molecule in an axially symmetric external field whose axis coincides with the molecular axis.

An external electric field, in the region of interaction with an atomic or molecular system, can be defined by the numerical values of the field potential and its Cartesian derivatives at some point in the system. (These are the coefficients in a Taylor expansion of the field potential around this point.) These numerical values can be treated as parameters, in terms of which effects due to the interaction of the atom or molecule, with the external field, can be expanded. For axially symmetric systems in which the point at which these field parameters are evaluated is chosen on the symmetry axis (z-axis), the number of independent field parameters reduces to the values of ϕ, $\partial\phi/\partial z$, $\partial^2\phi/\partial z^2$, $\partial^3\phi/\partial z^3$, $\cdots$, determined at that point. The potential, ϕ, can be eliminated from this list in discussing induction effects due to electronic distortion in a system of fixed nuclei, since there can be no distortion in an external field of constant potential. The above derivatives are the values of the electric field and its derivatives at the expansion point; $F_z = -\partial\phi/\partial z$, $F'_{zz} = -\partial^2\phi/\partial z^2$, $F_{zzz} = -\partial^3\phi/\partial z^3$, $\cdots$. For the axial systems under study, we drop the subscripts which denote the differentiation variables and list the field parameters as $\phi, F, F', F'', \cdots$.

- *Molecular polarizabilities*

We now discuss the energy of interaction of the electronic charge distribution with the external field. If $E^{(0)}$ is the total electronic energy in the field of the fixed nuclei of the molecule, and $E^{(f)}$ is the total electronic energy in the field of the fixed nuclei plus the external field, then the interaction energy $E^{(f)} - E^{(0)}$ can be expanded as a power series in the field parameters,[21]

$$\begin{aligned} E^{(f)} - E^{(0)} = q\phi - \mu F - \frac{1}{2}\Theta F' - \frac{1}{6}\Omega F'' - \frac{1}{24}\Phi F''' \\ - \frac{1}{2}\alpha F^2 - \frac{1}{2}AFF' - \frac{1}{4}CF'^2 \\ - \frac{1}{6}EFF'' - \frac{1}{6}\beta F^3 - \frac{1}{4}BF^2F' \\ - \frac{1}{24}\gamma F^4 - \cdots . \end{aligned} \tag{39}$$

The linear terms in the field parameters in Eq. (39) are the energies of interaction of the permanent electric multipole moments of the electronic charge distribution with the external field, and the higher order terms are contributions to the energy of induction caused by the external field inducing a change in the charge distribution. Terms involving ϕ do not appear in the induction energy, which exists only because of the distortion of the electronic cloud. The coefficients of the field parameters in Eq. (39) are components of the electric multipole moment and polarizability tensors for the system and will be identified below. The numerical factors occur naturally in the detailed treatment of the interaction of an arbitrary charge distribution with an arbitrary external field, which has been given elsewhere.[21] In Eq. (39) we have written only the lead terms in an infinite expansion. Use of the expansion in a truncated form depends on convergence, which in practice remains rapid until the external field comes within an order of magnitude of fields present internally in the molecule. For atomic or molecular systems it is usual to make the expansion (Eq. (39)) in terms of the field parameters $\phi, F, F', F'', \cdots$ evaluated at the center of mass of the system. It is apparent that, in general, choice of a different expansion point (and therefore different field parameters, even though the external field is the same) will lead to different coefficients in the expansion. The transformation of multipole moments and polarizabilities under change of the expansion point has been considered in detail elsewhere,[21] where the quantities q, α, β, γ are shown to be invariant under the transformation. Individually identifying the coefficients of Eq. (39), we have that q is the total electronic charge, $\mu = \mu_z$ is the z component of the permanent electronic dipole moment, $\Theta = \Theta_{zz}$ is the zz component of the permanent electronic quadrupole moment, $\Omega = \Omega_{zzz}$ is the zzz component of the permanent electronic octopole moment, and $\Phi = \Phi_{zzzz}$ the designated component of the permanent electronic hexadecapole moment. $\alpha, A, C, E, \beta, B, \gamma \cdots$, are components of the static electric polarizability tensors of the electron distribution. $\alpha = \alpha_{zz}$ is a component of the dipole polarizability; $A = A_{z:zz}$ is a component of the polarizability tensor which measures the contribution to the interaction energy due to a nonzero value of FF' (or equivalently, the contribution to the induced dipole moment for nonzero F, and to the induced quadrupole moment for nonzero F'); $C = 3/2\ C_{zz:zz}$ is equal to 3/2 times the $zz{:}zz$ component of the fourth-rank field gradient quadrupole polarizability tensor; $E = E_{z:zzz}$ is a component of another fourth-rank tensor which gives the contribution to the induced octopole moment in the presence of a nonzero F. $\beta = \beta_{zzz}$ and $\gamma = \gamma_{zzzz}$ are components of tensors which describe the nonlinear behavior of the induced dipole moment in the presence of a nonzero F, and $B = B_{zz:zz}$ has been called (in the case of atoms)[40] a uniform-field quadrupole polarizability.

For linear molecules with an inversion center ($D_{\infty h}$ symmetry) the quantities μ, Ω, A and β are identically zero if the expansion point for evaluating the external field parameters is the inversion center. For atomic systems, μ, Θ, Ω, Φ, A, and β are identically zero if the expansion point is the atomic nucleus.

The induced electric multipole moments in the charge distribution can also be expanded as a power series in the field components. The results from the general theory[21] are that the multipole moments of the electronic charge distribution, in the field of the nuclei plus the external field, which we represent by $\mu^{(f)} = \mu_z^{(f)}, \Theta^{(f)} = \Theta_{zz}^{(f)}, \Omega^{(f)} = \Omega_{zzz}^{(f)}, \Phi^{(f)} = \Phi_{zzzz}^{(f)}, \cdots$, are given by

$$\begin{aligned} \mu^{(f)} &= \mu + \alpha F + \tfrac{1}{2} AF' + \tfrac{1}{6} EF'' + \tfrac{1}{2}\beta F^2 \\ &\quad + \tfrac{1}{2} BFF' + \tfrac{1}{6}\gamma F^3 + \cdots \end{aligned} \tag{40}$$

$$\Theta^{(f)} = \Theta + AF + CF' + \tfrac{1}{2} BF^2 + \cdots \tag{41}$$

$$\Omega^{(f)} = \Omega + EF + \cdots \tag{42}$$

$$\Phi^{(f)} = \Phi + \cdots . \tag{43}$$

In Eq. (40) to (43) the expansions are given explicitly through terms corresponding to those given in Eq. (39). The coefficients of the field components are the same polarizabilities introduced in Eq. (39).

Our method of computation, implemented as RERUN = 7 described in the previous section, is to directly compute self-consistent field-molecular orbital wave functions in the presence of different electric fields and to calculate the corresponding dipole and quadrupole moments. These are then substituted into truncated forms of Eqs. (40) to (43) to obtain sets of linear equations which are solved for the polarizabilities. Specifically, wave functions and dipole moments are computed for six different fields which arise from point charges placed symmetrically with respect to the expansion point relative to which the polarizabilities are defined. The charges are placed on the molecular axis at distances $-R$, $+R$ relative to the expansion point. The pairs of charges giving rise to the six fields are (1) 0, Q (2) 0, $-Q$ (3) Q, Q (4) $-Q$, $-Q$ (5) $-Q$, Q (6) Q, $-Q$. If the corresponding dipole moments are denoted $m_1 - m_6$ and quadrupole moments $T_1 - T_6$, then substitution into Eqs. (40), (41) truncated to include all terms explicitly written in Eqs. (40), (41) except the term containing the polarizability E yields

$$\begin{aligned}
m_1 &= \mu - \alpha \frac{Q}{R^2} - A \frac{Q}{R^3} + \frac{1}{2}\beta \frac{Q^2}{R^4} + B \frac{Q^2}{R^5} \\
&\quad - \frac{1}{6}\gamma \frac{Q^3}{R^6} \\
m_2 &= \mu + \alpha \frac{Q}{R^2} + A \frac{Q}{R^3} + \frac{1}{2}\beta \frac{Q^2}{R^4} + B \frac{Q^2}{R^5} \\
&\quad + \frac{1}{6}\gamma \frac{Q^3}{R^6} \\
m_3 &= \mu - 2A \frac{Q}{R^3} \\
m_4 &= \mu + 2A \frac{Q}{R^3} \\
m_5 &= \mu - 2\alpha \frac{Q}{R^2} + 2\beta \frac{Q^2}{R^4} - \frac{4}{3}\gamma \frac{Q^3}{R^6} \\
m_6 &= \mu + 2\alpha \frac{Q}{R^2} + 2\beta \frac{Q^2}{R^4} - \frac{4}{3}\gamma \frac{Q^3}{R^6}
\end{aligned} \tag{44}$$

and

$$\begin{aligned}
T_1 &= \Theta - A \frac{Q}{R^2} - 2C \frac{Q}{R^3} + \frac{1}{2} B \frac{Q^2}{R^4} \\
T_2 &= \Theta + A \frac{Q}{R^2} + 2C \frac{Q}{R^3} + \frac{1}{2} B \frac{Q^2}{R^4} \\
T_3 &= \Theta - 4C \frac{Q}{R^3} \\
T_4 &= \Theta + 4C \frac{Q}{R^3} \\
T_5 &= \Theta - 2A \frac{Q}{R^2} + 2B \frac{Q^2}{R^4} \\
T_6 &= \Theta + 2A \frac{Q}{R^2} + 2B \frac{Q^2}{R^4}.
\end{aligned} \tag{45}$$

The unique solutions of Eqs. (44) yielding the axial polarizability components are

$$\begin{aligned}
\mu &= \tfrac{1}{2}(m_4 + m_3) \\
\alpha \frac{Q}{R^2} &= \frac{2}{3}(m_2 - m_1) - \frac{1}{3}(m_4 - m_3) \\
&\quad - \frac{1}{12}(m_6 - m_5) \\
A \frac{Q}{R^3} &= \frac{1}{4}(m_4 - m_3) \\
\beta \frac{Q^2}{R^4} &= -\frac{1}{4}(m_4 + m_3) + \frac{1}{4}(m_6 + m_5) \\
B \frac{Q^2}{R^5} &= \frac{1}{2}(m_2 + m_1) - \frac{3}{8}(m_4 + m_3) \\
&\quad - \frac{1}{8}(m_6 + m_5) \\
\gamma \frac{Q^3}{R^6} &= -(m_2 - m_1) + \frac{1}{2}(m_4 - m_3) \\
&\quad + \frac{1}{2}(m_6 - m_5).
\end{aligned} \tag{46}$$

Equations (45) contain only four polarizabilities to be determined and are therefore overdetermined. The program computes Θ, A, C, B from the first, second, fifth, and sixth of Eqs. (45) whose solution is

$$\begin{aligned}
\Theta &= \tfrac{2}{3}(T_2 + T_1) - \tfrac{1}{6}(T_6 + T_5) \\
A \frac{Q}{R^2} &= \tfrac{1}{4}(T_6 - T_5) \\
C \frac{Q}{R^3} &= \tfrac{1}{4}(T_2 - T_1) - \tfrac{1}{8}(T_6 - T_5) \\
B \frac{Q^2}{R^4} &= -\tfrac{1}{3}(T_2 + T_1) + \tfrac{1}{3}(T_6 + T_5)
\end{aligned} \tag{47}$$

and computes a second value for Θ, C from the third and fourth of Eqs. (45) according to

$$\begin{aligned}
\Theta &= \tfrac{1}{2}(T_4 + T_3) \\
C \frac{Q}{R^3} &= \tfrac{1}{8}(T_4 - T_3).
\end{aligned} \tag{48}$$

The agreement between the two values of Θ, C obtained from Eqs. (47), (48) and between the values of A, B obtained from Eqs. (46), (47) can be used to check the accuracy of the computation and the validity of the truncation.

Nuclear electric shielding factors

It appears that no general theory of nuclear electric shielding in molecules is available. However, the results for axially symmetric systems can be written down from simple considerations which we now outline. The change in the electric field at a nucleus due to the electrons, caused by the external field, is expanded in a power series in the external field parameters F, F', F'', $\cdots$. We note that the potential of the external field, ϕ, is missing from the parameter list because there is no effect on the field at the nuclei due to the electrons in an external field of constant potential. Also, for any nucleus, the expansion will be in the external field parameters evaluated at that nucleus. Suppose that the z component of the electric field and its derivatives with respect to the z coordinate, at a nucleus in a linear molecule, due to the electronic motions in the presence of an external field are $F_{el}^{(f)}, F_{el}'^{(f)}, F_{el}''^{(f)}, \cdots$. In the absence of the external field the same quantities will be $F_{el}^{(0)}$, $F_{el}'^{(0)}$, $F_{el}''^{(0)}$, $\cdots$. The external field is parametrized by its axial field components F, F', F'', $\cdots$ evaluated at the nucleus under consideration. Then, the differences $(F_{el}^{(f)} - F_{el}^{(0)})$, $(F_{el}'^{(f)} - F_{el}'^{(0)})$, $(F_{el}''^{(f)} - F_{el}''^{(0)})$, $\cdots$, which are the changes in the electronic contribution to the electric field at the nucleus because of the external field, can be expanded as a power series in the field parameters, as shown in the following equations:

$$\begin{aligned} F_{el}^{(f)} - F_{el}^{(0)} = & -\gamma_1^{(d)}F - \gamma_2^{(d)}F^2 - \gamma_3^{(d)}F^3 - \gamma_{10}^{(d)}F' \\ & - \gamma_{11}^{(d)}FF' - \gamma_{100}^{(d)}F'' - \gamma_{20}^{(d)}F'^2 \\ & - \gamma_{12}^{(d)}F^2F' - \cdots \end{aligned} \tag{49}$$

$$\begin{aligned} F_{el}'^{(f)} - F_{el}'^{(0)} = & -\gamma_1^{(q)}F - \gamma_2^{(q)}F^2 - \gamma_3^{(q)}F^3 - \gamma_{10}^{(q)}F' \\ & - \gamma_{11}^{(q)}FF' - \gamma_{100}^{(q)}F'' - \gamma_{20}^{(q)}F'^2 \\ & - \gamma_{12}^{(q)}F^2F' - \cdots \end{aligned} \tag{50}$$

$$\begin{aligned} F_{el}''^{(f)} - F_{el}''^{(0)} = & -\gamma_1^{(o)}F - \gamma_2^{(o)}F^2 - \gamma_3^{(o)}F^3 - \gamma_{10}^{(o)}F' \\ & -\gamma_{11}^{(o)}FF' - \gamma_{100}^{(o)}F'' - \gamma_{20}^{(o)}F'^2 \\ & -\gamma_{12}^{(o)}F^2F' \cdots, \end{aligned} \tag{51}$$

and similar equations for higher derivatives of the field due to the electrons. In Eqs. (49 to 51) we have introduced a notation adequate to cover the axially symmetric case. The γ's are all known as shielding factors. The superscript (d), (q), (o), $\cdots$, denotes dipole, quadrupole, octopole, $\cdots$, and the subscript denotes the powers of the external field components in the current term. If one integer appears in the subscript this is the power of F; if two integers appear then they give, in reading order, the power of F' and F; if three, then they are the powers of F'', F', F; and so on. Of course, in a general theory, for arbitrary symmetry, these would be components in shielding factor tensors and a considerably more complex notation would be required. The quantity $\gamma_1^{(d)}$ is normally called the dipole shielding factor, denoted in Dalgarno's review[(41)] by β_∞. The quantity $\gamma_2^{(q)}$ is, in atomic calculations, known as the uniform field quadrupole shielding factor,[40] while $\gamma_{10}^{(q)}$ is known as the field gradient quadrupole shielding factor. This latter, $\gamma_{10}^{(q)}$, is the γ_∞ of Sternheimer.[42]

A comment on two possible independent ways of computing the Sternheimer shielding factor, γ_∞, may be enlightening. Consider the interaction between an externally produced field gradient with a nucleus having a nonzero electric quadrupole moment. One method of computing the interaction would be to calculate the distortion of the electronic cloud of the molecule due to the nuclear electric quadrupole, thereby evaluating the induced quadrupole moment in the electronic cloud due to the nuclear quadrupole. The interaction of the external field with both the nuclear and nuclear induced quadrupole moments is then evaluated. This is the Sternheimer procedure. An alternative method would calculate the distortion of the electronic cloud due to the external field. The change in the field gradient at the nucleus due to this distortion will then be evaluated, and the required interaction is that between the nuclear quadrupole moment and the modified field gradient. Thus, in the first procedure the interaction is computed as an unperturbed field gradient interacting with a perturbed quadrupole moment, while in the second it is computed as a perturbed field gradient interacting with an unperturbed quadrupole moment. The equivalence of the two procedures has been proved by Das and Bersohn.[43]

As described above, the interaction of the external field with the distortion in the electronic cloud due to a nuclear quadrupole yields the quadrupole shielding factor $\gamma_{10}^{(q)}$. The interaction of the external field with the distortion in the electronic cloud due to the external field itself is, of course, an entirely different effect already discussed in terms of polarizabilities.

Calculations of shielding factors with the computer program described in this paper are done using the second of the above two procedures. For the same configurations of external charges described in the polarizability calculations, the field gradients at each of the nuclei are computed. When compared with the values of the field gradients at the nuclei with no external field, we have enough information to solve sets of linear equations truncated from Eqs. (49) to (51) to obtain values of shielding factors. The program as it currently stands simply computes the field gradients, and the processing of these to obtain the shielding factors is still in progress. We note that Eqs. (49) to (51) are infinite expansions and, as with the polarizabilities, their utility depends on rapid convergence which will be the case for external fields at least an order of magnitude less than those present internally in molecules.

From the point of view of assessing the accuracy of computed wave functions, polarizabilities and shielding factors offer a severe test because of their sensitivity to the wave function. Polarizabilities are determined largely by the valence shells of the molecule and are therefore sensitive to the wave function far from the nuclei, where they are often seriously deficient. This is especially so if inadequate basis sets have been used and arises primarily because the limited basis functions are used to best represent the inner part of the wave function, which is most important in minimizing the total energy. Shielding factors are especially sensitive to changes in the wave function in the neighborhood of the nuclei, and therefore complement the polarizabilities in checking the adequacy of a basis set. A basis set large enough to stabilize the computed values of these properties is larger than that needed to stabilize the total energy. In reporting calculations of these properties it is important to establish the sensitivity of the results to changes in the basis set.

References

1. H. M. James and A. S. Coolidge, *J. Chem. Phys.* **1,** 825 (1933).
2. B. M. Gimarc and R. G. Parr, *Ann. Rev. Phys. Chem.* **16,** 451 (1965); A. Golebiewski and H. S. Taylor, *ibid,* **18,** 353 (1967).
3. E. Clementi and D. R. Davis, *J. Chem. Phys.* **15,** 2595 (1966); E. Clementi, *ibid,* **46,** 3851 (1967).
4. R. S. Mulliken, Lecture delivered when receiving the 1966 Nobel Prize in Chemistry, reprinted in *Science* **157,** 13 (1967).
5. W. Kolos and C. C. J. Roothaan, *Revs. Mod. Phys.* **32,** 205 (1960); W. Kolos and L. Wolniewicz, *J. Chem. Phys.* **43,** 2429 (1965).
6. L. Wolniewicz, *J. Chem. Phys.* **45,** 515 (1965).
7. N. F. Ramsey, *Molecular Beams,* Oxford (1965) Chapter 12; P. Kusch, *Physics Today* **19,** 23 (1966).
8. W. Kolos and L. Wolniewicz, *Phys. Rev. Letters* **20,** 243 (1968).
9. Conference on Molecular Quantum Mechanics, Boulder, Colorado, June 1959, published in *Revs. Mod. Phys.* **23,** 169–476 (1960).
10. International Symposium on Atomic and Molecular Quantum Theory, January 1963 published in *Revs. Mod. Phys.* **35,** 415–736 (1963); January 1965 published in *J. Chem. Phys.* **43,** S1-S272 (1965).
11. "McL-YOSH LINEAR MOLECULE PROGRAM 1" written by A. D. McLean and M. Yoshimine and available as QCPE 104 from the Quantum Chemistry Program Exchange at Indiana University. It is coded in FAP Assembly Language for an IBM 7094.
12. R. K. Nesbet in *Advances in Chemical Physics* Volume 9, Interscience Publishers, New York, 1966.
13. C. C. J. Roothaan, *Revs. Mod. Phys.* **23,** 69 (1951); **32,** 179 (1960).
14. G. Das and A. C. Wahl, *J. Chem. Phys.* **44,** 87 (1966); E. Clementi and A. Veillard, *ibid.,* **44,** 3050 (1966); C. F. Bender and E. R. Davidson, *ibid.,* **47,** 360 (1967); J. Hinze and C. C. J. Roothaan, Supplement No. 40 to *Progress of Theoretical Physics,* p. 37 (1967).
15. S. F. Boys and P. Rajagopal, *Advances in Quantum Chemistry* Ed. P. O. Lowdin, Academic Press, New York, 1965, Volume 2, p. 1.
16. A. D. McLean and M. Yoshimine, McL-YOSH LINEAR MOLECULE PROGRAM 1—USER MANUAL, available from QCPE (Ref. 11) or from the authors.
17. M. Yoshimine and A. D. McLean, *Intern. J. Quant. Chem.* (in press), Slater Symposium Issue (1967).
18. A. D. McLean and M. Yoshimine, "Tables of Linear Molecule Wave Functions" published as a Supplement to the current paper. These Tables are available on request to the authors.
19. Some basic theory is reviewed, and preliminary results on H_2, CO, N_2O given by A. D. McLean and M. Yoshimine, *J. Chem. Phys.* **45,** 3676 (1966).
20. A. D. McLean and M. Yoshimine, *J. Chem. Phys.* **46,** 3682 (1967); **47,** 1927 (1967).
21. M. Krauss, "Compendium of ab initio Calculations of Molecular Energies and Properties," Tech. Note #438 of the National Bureau of Standards, Washington, D. C. (1968).
22. A. D. McLean and M. Yoshimine, *IBM J. Res. Develop.* **9,** 203 (1965).
23. A. D. McLean and M. Yoshimine, *J. Chem. Phys.* **46,** 1812 (1967).
24. This is the essential feature of an analysis of the two center exchange integrals. K. Ruedenberg, *J. Chem. Phys.* **19,** 1459 (1951).
25. E. W. Hobson, *The Theory of Spherical and Ellipsoidal Harmonics,* Cambridge University Press, Cambridge, 1931.
26. A. D. McLean, *J. Chem. Phys.* **32,** 1595 (1960).
27. F. J. Corbato, *J. Chem. Phys.* **24,** 452 (1956).
28. This step follows J. R. Hoyland, *J. Chem. Phys.* **40,** 3540 (1964).
29. We believe that this procedure was first used by P. Merryman in unpublished work incorporated into a computer program in Whirlwind at MIT in 1957.
30. The details of this ordering are available in Ref. 16.
31. The two subroutines are TRIGR written by W. Worley, H. Kuki and C. C. J. Roothaan of the University of Chicago Computation Center, and DBUFF written by the current authors. TRIGR is the basic routine which enables the programmer to use the data channels in a simple powerful way. DBUFF, written around TRIGR, enables the programmer to simply use double buffered data transmission. Both subroutines are described fully in Ref. 16.
32. A. C. Wahl, P. E. Cade and C. C. J. Roothaan, *J. Chem. Phys.* **41,** 2578 (1964).
33. A. D. McLean and M. Yoshimine, *J. Chem. Phys.* **46,** 1812 (1967).
34. C. C. J. Roothaan and P. S. Bagus, *Methods in Computational Physics,* Academic Press, New York, 1963, Volume 2, p. 47.
35. A. C. Wahl, *J. Chem. Phys.* **41,** 2600 (1964).
36. W. M. Huo, *J. Chem. Phys.* **43,** 624 (1965).
37. R. K. Nesbet, in *Advances in Chemical Physics,* Interscience Publishers, New York, 1966, Volume 9.
38. The "source controlled input data program" SCINP which processes this data was written by L. Monheit and C. C. J. Roothaan of the University of Chicago Computation Center.
39. A. D. McLean and M. Yoshimine, *J. Chem. Phys.* **47,** 3256 (1967).
40. J. D. Lyons, P. W. Langhoff and R. P. Hurst, *Phys. Rev.* **151,** 60, (1966).
41. A. Dalgarno, *Adv. in Phys.* **11,** 281 (1962).
42. R. M. Sternheimer, *Phys. Rev.* **80,** 102 (1950); **84,** 244 (1951); **86,** 316 (1953); **95,** 736 (1954); H. M. Foley, R. M. Sternheimer and D. Tycko, *ibid.,* **93,** 734 (1954).
43. T. P. Das and R. Bersohn, *Phys. Rev.* **102,** 733 (1956).

Received December 26, 1967.