ELECTRON-PAIR THEORIES OF MOLECULAR ELECTRONIC STRUCTURE

by

Peter James Carrington B.Sc.

A thesis submitted for the degree of Doctor
of Philosophy at the University of Glasgow

Chemistry Department
University of Glasgow

March 1971

ACKNOWLEDGEMENTS

I should like to express my deep gratitude to Dr. Graham Doggett for his invaluable and continual guidance throughout the course of this work and for the many helpful suggestions and discussions on quantum chemistry in general and on the problems of this work in particular which have made this thesis possible.

I also wish to thank Professor Roy McWeeny for his forbearance in allowing the completion of the writing of this thesis during my stay in the theoretical chemistry group at the University of Sheffield. My sincerest thanks are due to Sue Rogers for her patience and skill in typing a difficult manuscript.

I wish to acknowledge financial support from the Science Research Council during the tensure of this work.

CONTENTS

INTRODUCTION 1

CHAPTER 1

Section 1 Molecular Wavefunctions and the Non-Orthogonality Problem 4

Section 2 The Theory of the Non-Orthogonal Pair Function Method and its Application to some Few Electron Systems 42

Section 3 Approximate Solutions of the Non-Orthogonality Problem 82

Section 4 Valence Theory and the Non-Orthogonal Pair Function Wavefunction 123

CHAPTER 2 Spin Theory for Electron Pair Bonds 143

CONCLUSION 204

APPENDIX 1 Notation 206

APPENDIX 2 Energy Expression and Variational Equations for Two Pair Functions 209

APPENDIX 3 Expansion of the Two-Particle Density Function 218

APPENDIX 4 Matrix Elements for the Approximate Variational Methods 226

INTRODUCTION

INTRODUCTION

Wave mechanics presents a picture of the electronic structure of molecules far removed from that of the experimental chemist. Thus the quantum theorist describes the molecule in terms of electrons, nuclei and probability distributions, and the molecular energetics in terms of Coulomb forces and kinetic and potential energy, whereas the chemist thinks of a molecule as composed of bonds, lone pairs and inert cores and in terms of ill-defined but nonetheless useful concepts such as electronegativity, bond order and electron pair repulsions. It is the task of the quantum chemist to bridge this gap; on the one hand, to justify, clarify and quantify the qualitative chemical concepts, and, on the other, to extract from the mathematical complexities of the wavefunction, quantities reflecting the electronic characteristics of the molecules.

The notion of the two-electron entities, electron pair bonds and lone pairs, plays a large part in the qualitative description of many molecules. It is therefore natural to seek methods which build up the molecular wavefunction from units describing the individual bonds and lone pairs, and hence provide a link between quantum mechanics and chemistry. The simplest such method is based on the pair function wavefunction which is an antisymmetrised product of localised two-electron, or pair, functions. The currently most important version of this approach - the Separated Pair method of Hurley, Lennard-Jones and Pople - has been shown to be a valuable tool for obtaining molecular wavefunctions and for relating the mathematics of the wavefunctions to chemical concepts. The pair functions are, in the Separated Pair method, constrained by stringent orthogonality requirements, which although producing a method with considerable practical advantages, restricts the generality and accuracy of the wavefunction. By removing the orthogonality constraints, these deficiencies are avoided, and a potentially very accurate and widely applicable wavefunction - the Non-Orthogonal Pair Function (NPF) wavefunction - is obtained, which up till now has received little attention and, that it has, of a very restricted nature. The NPF wavefunction bears a close resemblance to the Valence Bond method, which is the fore-runner of all electron pair theories, with bonds being represented quantum mechanically by Heitler-London two-electron functions. The main reason why little work has been directed towards the NPF method is that it encounters the same practical problem as Valence bond theory - the molecular non-orthogonality problem - due to the use of non-orthogonal basis orbitals, which makes the evaluation of the wavefunction and expectation values an extremely complex task. The first part of this work is devoted to the non-orthogonality problem and, in particular, the NPF method.

Current electron pair theories are limited to singlet states since the electrons of each pair function are constrained to be coupled to a zero spin state. Although many molecules of interest do exist in singlet spin states, there are also many situations where the present pair theories are inapplicable due to this restriction. In the second part of this work, the scope of the Separated Pair theory is extended by the introduction of a new basis of spin functions which allows wavefunctions of arbitrary multiplicity to be constructed for even-electron molecules and molecular positive ions.

CHAPTER 1

Section 1

Molecular Wavefunctions and the Non-Orthogonality Problem

Many different methods have been proposed for calculating approximations to the unobtainable exact solutions of the Schrödinger equation for polyatomic systems. The most important of these are based on the orbital expansion of the exact n-electron wavefunction

$$\Psi = \sum_{i}^{m} c_i \Psi_i \tag{1.1}$$

where the structure Ψ_i is given by

$$\Psi_i = \hat{A}\,\psi_i = \sum_{j}^{n_i} d_{ij}\,\hat{A}\,\psi_{ij}$$

and the configuration ψ_{ij} by

$$\psi_{ij} = \phi_1^{ij}(1)\phi_2^{ij}(2)\ \dots\ \phi_n^{ij}(n)\ \theta_{ij}(1\dots n)$$

The orbitals ϕ_k^{ij} are one-electron spatial functions, and the spin functions $\theta_{ij}(1\dots n)$ simple products of the one-electron spin functions, α and β. The configuration coefficients d_{ij} are fixed by the requirement that the wavefunction is an eigenfunction of the square of the spin angular momentum operator, $\hat{S}^2$, and is symmetry adapted to the point group of the molecule.

As m, the number of structures, and hence the size of the basis of one-electron functions, tends to infinity, so the wavefunction Ψ tends to the exact wavefunction. The ϕ_k^{ij} are generally taken as atomic orbitals or linear combinations of atomic orbitals since the potential energy terms in the molecular Hamiltonian ensure that the molecule retains, to a large extent, especially near the nuclei, the characteristics of the free atoms.

Since neither an infinite basis nor an infinite set of structures can be dealt with in practice, approximations, by the truncation of the basis and number of structures to a finite size, are inevitable. The structure coefficients c_i, and any free parameters that the orbitals ϕ_k^{ij} may contain, are then determined so that the approximate wavefunction is as close to the exact wavefunction as possible: in practice, this means that the energy of the approximate wavefunctions is as close as possible to the true energy. The various methods of orbital wavefunction construction differ in the way that the orbitals are set up and the number and type of structure included in the wavefunction: two approaches to the problem may be distinguished which we shall call the "mathematical" and the "chemical" approaches. In the "mathematical" approach, no cognisance is taken of any conceptions we might have, from experimental sources, of the electronic structure of the molecule. Rather, the molecular wavefunction is generated, from the knowledge of the geometry of the molecule only, by insertion of the chosen set of atomic orbitals into a general prescription for the wavefunction. The advantage of such an approach is that, by concentrating on the mathematical formulation of the method, i.e. the general prescription, the wavefunction can be made to incorporate features which from a practical point of view, are highly desirable. The most important of these is orthogonality of the one-electron orbitals,

$$\int \phi_i(1)\phi_j(1)\, d\tau_1 = \delta_{ij}$$

since the expressions for the energy and the variational equations that determine the optimum values of any free parameters that the wavefunction may contain, are then usually, as we shall see, greatly simplified and well suited to practical applications of the theory. The best known example of this approach is the Hartree-Fock (HF) wavefunction for 2N electron singlet states which is essentially the simplest possible orbital wavefunction, consisting of a single antisymmetrised configuration of doubly occupied orbitals

$$\Psi = \hat{A}\, \phi_1(1)\phi_1(2)\phi_2(3) \ldots \phi_N(2N)\alpha(1)\beta(2)\alpha(3) \ldots \beta(2N)$$

where the orbitals ϕ_i, the molecular orbitals (MO), are expanded as linear combinations of all the atomic orbitals in the molecule. The prescription for the MO's ϕ_i, and hence the molecular wavefunction, is the eigenvalue equation

$$h^F(1)\ \phi_i(1) \quad = \quad \varepsilon_i\ \phi_i(1)$$

which the MO's satisfy, where ε_i is the energy of the orbital ϕ_i and $h^F(1)$ the well known Hartree-Fock Hamiltonian operator[1].

On the other hand, the "chemical" approach depends on having a detailed qualitative picture of the molecular electronic structure, in which the molecule is broken down into small independent units, and set of rules for converting this empirical picture, unit by unit, into a molecular wavefunction. Valence Bond (VB) theory belongs to this category. Valence theory interprets the structure of most molecules in terms of two-electron two-centre bonds. The two-electron bond between atoms A and B is represented quantum mechanically, in the first approximation, by a Heitler-London function

$$(\phi_a(1)\phi_b(2) + \phi_b(1)\phi_a(2))(\alpha(1)\beta(2) - \beta(1)\alpha(2))$$

where the orbitals ϕ_a and ϕ_b are valence atomic orbitals or fixed combinations of the atomic orbitals on atoms A and B respectively, directed along the bond direction. These "covalent" bond functions, one for each of the M bonds in the molecule, are augmented by "ionic" bond functions, in which both electrons are localised on one atom of the bond, and "long bond" functions or alternative spin couplings, in which orbitals in different bonds are paired together. A structure is constructed by taking an antisymmetrised product of M bond functions; the total wavefunction is a linear combination of such structures corresponding to different choices of the M bond functions[2]. The disadvantage of the "chemical" approach is that, because the wavefunction is built up from wavefunctions for the individual parts of the molecule, no account is taken of the practical aspects of the resulting wavefunction. The price that has to be paid is the non-orthogonality problem: this becomes clear when we consider the optimisation of the structure coefficients, c_i, in the VB wavefunction. Application of the variational theorem to equation (1) gives the best coefficients as solutions of the secular equation

$$(\underline{H} - E\,\underline{S})\underline{c} = 0$$

where E is the electronic energy, the matrix elements

$$H_{ij} = \int \psi_i \hat{H} \hat{A} \psi_j \, d\tau \qquad \text{and}$$

$$S_{ij} = \int \psi_i \hat{A} \psi_j \, d\tau \; ,$$

and $\hat{H}$ is the molecular Hamiltonian. The antisymmetriser

$$\hat{A} = \sum_{P}^{n!} (-1)^P P$$

is the sum of all signed permutation operators of order n, and sends ψ_j into a linear combination of $n_j \times n!$ orbital product terms with electron labels permuted between the orbitals. Since $\hat{H}$ contains at most operators dependent on the coordinates of two electrons, the integration over all coordinates gives for each term in the expansion of ψ_j an energy integral times at least (n-2) overlap integrals, $S_{k\ell} = \int \phi_k(1)\phi_\ell(1)\,d\tau$. In VB theory, the orbitals on which the configurations are based are atomic orbitals: hence, orbitals on different atoms are not orthogonal and, in general, none of the overlap integrals vanish. The number of terms contributing to H_{ij} therefore increases with the number of electrons as approximately n! so that the work involved in evaluating the matrix elements, and hence the VB wavefunction itself, rapidly becomes prohibitive. Herein lies the whole non-orthogonality problem of molecular quantum mechanics, in the rapid escalation of the labour required to construct wavefunctions based on non-orthogonal orbitals. And the problem is intimately linked to the "chemical" wavefunction method through its demand that wavefunctions are constructed unit by unit. (In practice, the situation may be alleviated to some extent[3], but the rate of increasing complexity of the matrix elements is not in general greatly reduced). In wavefunctions based on orthogonal orbitals, however, any permutation, P, which results in mismatch of orbitals gives a zero result on integration since the overlap integral $S_{k\ell}$ is zero unless $k=\ell$. Thus, in the MO method, only 2N terms in the expansion of an expectation value of a one-electron operator, and 2N(2N-1) for a two-electron operator, are non-zero; and, it has been estimated, the work required for the complete evaluation of the MO wavefunction increases as approximately the cube of the number of atomic orbitals in the basis[4], a much less rapid increase than in the VB method.

The disparity between the practical aspect of the MO and VB methods is underlined when we consider the current situation regarding their application to molecular systems. The larger the molecule, the more difficult it is to obtain an approximation to the wavefunction. With current computing facilities, it is possible to calculate crude, usually semi-empirical, wavefunctions for large molecules, by quantum mechanical standards, with less than 100 electrons, simple ab initio wavefunctions for medium sized molecules, with less than 40 electrons and accurate wavefunctions only for small molecules, with less than 15 electrons[5]. A molecular wavefunction calculation falls naturally into two stages - the setting up of the atomic orbital basis and evaluation of the integrals, kinetic energy, nuclear attraction, electron repulsion and overlap, over this basis, and the optimisation of the free parameters in the wavefunction. A balance must be struck between the stages, firstly for the sake of efficiency, since the combination of a very accurate basis, which requires a long integral computation time, with a wavefunction, which is very simple and easily optimised but which can never approach an accurate wavefunction, will not give results which merit the work involved; and vice versa, for accurate wavefunctions and simple bases. Also, from a practical point of view, the two stages must be kept in step for, if one stage becomes much more time consuming than the other, this stage, and hence the whole wavefunction construction, will be in danger of becoming completely intractable.

In the MO method, the optimisation step is, in most situations, less time comsuming than the computation of the integrals. It has however been possible to introduce approximations at the integral evaluation stage to bring the two steps into balance, and to make calculations on large molecules feasible. These range from the crudest empirical methods of pi-electron theory, through semi-theoretical approaches such as the neglect-of-differential-overlap methods as used for example in the INDO method[6], up to the Gaussian orbital methods in which Slater-type orbitals are expanded in linear combinations of Gaussian orbitals[7], and which may be considered more as integration aids than approximations.

But, in the VB method, the situation is reversed, and the wavefunction optimisation process is the more complex, it being at this stage that the non-orthogonality problem, in the matrix elements evaluation, makes itself felt. The integral evaluation is a matter of secondary importance. Unfortunately, no satisfactory approximation method has been found to overcome the problem, and make VB theory applicable to the size of molecule tractable by MO theory. Certainly, a semi-empirical method has been used in which all the (n!-1) interchange terms are combined into a single parameter, the exchange integral, which is then partitioned into bond and interbond interaction terms to be disposed of as empirical parameters[8]. Apart from the crudity of such an approximation, the partition is derived on the basis of the neglect of all overlap integrals and a single structure "perfect pairing" wavefunction. These have been shown to be totally unfounded, leading for example to the conclusion that the hydrogen molecule is not bound[9]. In an attempt to circumvent directly the practical difficulties of VB theory, McWeeny has combined the advantages of a "mathematical" approach with the spirit of the VB method, by orthogonalising the atomic orbital basis before constructing the VB wavefunction[10]. The matrix elements then take fairly simple forms, but, unfortunately, the convergence of wavefunction as more structures are added is disappointingly slow[11]. Such is the seriousness of the non-orthogonality problem that complete calculations by the VB method are only possible for very small molecules[12], and in default of any approximate solution, has resulted in almost total neglect of the VB theory as a practical method.

This is the current state of the molecular non-orthogonality problem. Is it worth considering further, or would it not be better to concentrate on the more practical, i.e. "mathematical" methods? For several reasons, the answer is no. Firstly, there are areas of interest where the practical difficulties of traditional VB theory are justified by the accuracy of the results. These we shall try to identify in this Section. Secondly, the relation between quantum mechanics and valence theory needs to be better understood: a "chemical" wavefunction is best suited for this task. And thirdly, the non-orthogonality problem must be overcome, for, when we come to very large molecules, "mathematical" methods will be of little use, the vital orthogonalisation process itself becoming totally impractical. It may be that the problem is surmountable or may be alleviated to a great extent by approximation techniques, but that the VB wavefunction is not the most suitable starting point for finding the solution. These points will be considered later in this Section, and in greater detail in remaining Sections of the Chapter.

Let us now consider the contention that, in certain circumstances, the VB wavefunction may have advantages over "mathematical" types of wavefunction, and hence that the current view that it is an impractical method is unfounded. We must first look in more detail at these "mathematical" methods. Just as VB theory has, because of the non-existence of approximation methods, an upper limit on its applicability, so MO theory has a lower limit, but, in this case, for theoretical reasons. The price that has to be paid for the pleasant practical characteristics of the wavefunction is that it can never approach the exact wavefunction. As the size of the basis increases, so the energy tends to a limit, the Hartree-Fock limit, above the true energy. The origin of this limit is the neglect of correlation between the motions of the electrons: in the MO representation of reality, the electrons are allowed to collide with one another because, instead of moving in the instantaneous coulombic field of the other electrons, each electron moves in an average field[13]. This is most readily described in terms of the one- and two-particle density functions. Thus the one particle density function is

$$\rho(1,1') = 2N \int \Psi(1\ldots 2N)\Psi(1'2\ldots 2N)d\tau_{2\ldots 2N}$$

$$= \rho^{\alpha}(1,1') + \rho^{\beta}(1,1')$$

where, for example $\rho^{\alpha}(1,1')$ gives the probability of finding an electron at point r_1 with spin α. The two-particle density function, whose diagonal elements give the probability of finding simultaneously electrons ar r_1 with spin s_1 and at r_2 with spin s_2, is, in the MO approximation[14],

$$\rho(1,1';2,2') = N(2N-1)\int\Psi(12\ldots 2N)\Psi(1'2'3\ldots 2N)d\tau_{3\ldots 2N}$$

$$= \rho^{\alpha}(1,1')\rho^{\beta}(2,2') + (\rho^{\alpha}(1,1')\rho^{\alpha}(2,2') - \rho^{\alpha}(2,1')\rho^{\alpha}(1,2'))$$

$$+ (\rho^{\beta}(1,1')\rho^{\beta}(2,2') - \rho^{\beta}(2,1')\rho^{\beta}(1,2'))$$

Thus, electrons with different spin are uncorrelated, since the probability of finding two electrons simultaneously at r_1 and r_2 with spins α and β is the same as finding them at these positions independently, whilst those with the same spin spin are correlated, for, as $r_1 \to r_2$, the probability goes to zero, i.e. the electrons cannot collide. Since the same-spin, or Fermi, correlation is present in all antisymmetric wavefunctions, it is the different-spin, or Coulomb, correlation that is the important factor in molecular wave mechanics. The seriousness of this neglect may be judged from the fact that the difference between the HF and the exact energies is of the same order as molecular dissociation energies[15], and hence large compared to the energy differences of interest to chemists. In general, the current computational facilities are such that MO theory is the only practical method of obtaining wavefunctions for medium-sized and large molecules. But, for smaller molecules, more accurate wavefunctions may be found.

The search for ways of transcending the MO approximation has been concentrated on "mathematical" methods. This is natural since the problem is a very difficult one and it is vital to have as practical a method as possible, and also the obvious starting point for a method introducing correlation is the uncorrelated MO wavefunction. The first step in this direction is to add extra structures to the single structure MO wavefunction. But from what orbitals are these to be constructed? The Hartree-Fock operator, which determines the MO's, has as many solutions as atomic orbitals. Since the number of atomic orbitals is usually greater than the number of occupied orbitals, i.e. those used in building up the MO wavefunction, the redundant solutions - the virtual or unoccupied orbitals - provide a ready-made additional basis from which to build further structures, by replacing occupied MO's in the MO wavefunction by unoccupied ones. Since the occupied and virtual MO's are eigenfunctions of the same operator, they are automatically orthogonal, making the matrix element construction an easy task. It has become apparent however that, despite the attractiveness of this technique, known as Configuration Interaction (CI), the wavefunction, as more structures are added, converges very slowly. This is due to the fact that the virtual orbitals do not possess the correct shape to improve the wavefunction: they are small in regions of high electron density[16]. To overcome this deficiency, the Multi-Configuration Self-Consistent-Field (MCSCF) method has been proposed[17], in which, instead of using the fixed, unsuitable virtual orbitals, which are determined by orthogonality conditions rather than by an energy criterion, in the CI expansion, all the MO's, occupied and "unoccupied", are optimised simultaneously with the structure coefficients. This ensures that the "unoccupied" orbitals have the correct form to contribute significantly to the wavefunction.

A closely related approach is the General Separated Pair (GSP) method, in which the wavefunction is written as an antisymmetrised product of two-electron, or pair, functions

$$\Psi = \hat{A}\, \Delta^1(12)\Delta^2(34) \ldots \Delta^N(2N-1,2N)$$

For a singlet state, the pair functions $\Delta^I(12)$ are expanded in terms of one-electron orbitals ϕ_i^I

$$\Delta^I(12) = \sum_i c_i^I\, \phi_i^I(1)\ \phi_i^I(2)(\alpha(1)\beta(2) - \beta(1)\alpha(2))$$

where the orbitals ϕ_i^I are further expanded in terms of all atomic orbitals in the basis, as in the MO method[18]. Formally simple variational equations and energy expression are ensured by constraining the pair functions to be strongly orthogonal[19], that is,

$$\int \Delta^I(12)\Delta^J(13)d\tau_1 = 0, \qquad I \neq J$$

which, in effect, requires that the orbitals ϕ_i^I are orthogonal to the orbitals in all other pair functions ϕ_j^J, $J \neq I$[20]. It is usual to also demand that the orbitals within a pair function are orthogonal. It is found, in practice, that each pair function contains one strongly occupied orbital, ϕ_1^I, with coefficient, c_1^I, approximately unity and a set of weakly occupied orbitals with coefficients c_i^I, $i = 2,3,\ldots$, approximately zero, and that the ϕ_1^I closely resemble the MO's of the Hartree-Fock wavefunction[21]. The wavefunction is thus, on expansion, very similar to a MCSCF wavefunction, but constructed only from doubly, quadruply etc. excited structures and with each set of weakly occupied orbitals - equivalent to the "unoccupied" orbitals of MCSCF theory - correlating only one occupied MO. Also, the coefficients, for example, of a quadruply excited structure is the product of the coefficients of the related doubly excited states, (divided by a constant, $\prod_{I=1}^{N} c_1^I$, which is close to unity): this "unlinked cluster" approximation[22] has been shown, in practice, to yield satisfactory results[23]. The GSP wavefunction has conceptual advantages in that the correlation it introduces is evident from inspection of the density functions. The one-particle density function reduces, under the strong orthogonality condition, to the sum of the probability distributions of the individual pairs

$$\begin{aligned}\rho(1,1') &= \sum_{I} (\rho_I^{\alpha}(1,1') + \rho_I^{\beta}(1,1') \\ &= \sum_{I}\sum_{i} (c_i^I)^2\ \phi_i^I(1)\phi_i^I(1')(\alpha(1)\alpha(1') + \beta(1)\beta(1'))\end{aligned} \tag{1.3}$$

The two-particle density function is[24]

$$\begin{aligned}\rho(1,2,1',2') &= \sum_{I} \rho_I^{\alpha\beta}(1,2,1',2') \\ &+ \sum_{I,J}{}' (\rho_I^{\alpha}(1,1')\rho_J^{\beta}(2,2') + \rho_I^{\beta}(1,1')\rho_J^{\alpha}(2,2')) \\ &+ \sum_{I,J}{}' (\rho_I^{\alpha}(1,1')\rho_J^{\alpha}(2,2') - \rho_I^{\alpha}(2,1')\rho_J^{\alpha}(1,2')) \\ &+ \sum_{I,J}{}' (\rho_I^{\beta}(1,1')\rho_I^{\beta}(2,2') - \rho_I^{\beta}(2,1')\rho_J^{\beta}(1,2'))\ ,\end{aligned}$$

where, as in equation (1.2), we have omitted the terms with spin parts $\alpha(1')\beta(1)\alpha(2)\beta(2')$ and $\alpha(2')\beta(2)\alpha(1)\beta(1')$ which do not contribute to the energy since they vanish on spin integration. The intra-pair term is, on expansion,

$$\rho_I^{\alpha\beta}(1,1',2,2') = \sum_{i,j} c_i^I c_j^I \phi_i^I(1)\phi_j^I(1')\phi_i^I(2)\phi_j^I(2')$$

$$\times\ (\alpha(1)\alpha(1')\beta(2)\beta(2') + \beta(1)\beta(1')\alpha(2)\alpha(2'))$$

The GSP wavefunction thus contains Coulomb correlation between the electrons within each pair function but none between those in different pair functions. This is an improvement on the HF approximation, and is capable of giving very accurate wavefunctions when the electron pairs are well separated spatially, for example, in beryllium[25]. The orbitals that diagonalise the one-particle density function are the natural orbitals (NO) of the wavefunction: they derive their importance from a theorem which states that the approximate wavefunction which is closest to the exact wavefunction is one based on the NO's of the density function[26]. The density function of the GSP wavefunction is already in diagonal form, equation (1.3); the natural orbitals ϕ_i^I are therefore determined directly and a rapid convergence of the wavefunction is expected. One of the most recent and most important of "mathematical" wavefunction methods utilises this theorem directly. The natural orbitals, on which the expansion of the wavefunction should be based, are the eigenfunctions of the exact density function, which is, of course, unknown. The Natural-Orbital Configuration Interaction (NO-CI) method[27] locates the NO's by an iterative process in which a correlated wavefunction, usually a simple CI wavefunction is analysed into its approximate natural orbitals. These are then used as the basis for a further CI calculation. By obtaining the NO's of this improved wavefunction and continuing the cycling process, the natural orbitals converge, within the limits of the basis, to those of the exact wavefunction.

Apart from the lack of correlation, other serious shortcomings are inherent in the MO wavefunction. In the first place, MO theory predicts the wrong dissociation products: on the separation of two bonded atoms to infinity, the MO wavefunction goes over to a mixture of neutral atom, and positive and negative ion, product wavefunctions instead of the correct dissociation products which are the neutral atoms[28]. Although, in practice, this is not apparently a catastrophic deficiency for small deviations from the equilibrium geometry, it does cast doubts on the accuracy of energy surfaces determined by the MO method. Secondly, the wavefunction is characterised by doubly occupied MO's from which only a singlet state may be constructed. The extension of the wavefunction to doublet, triplet and higher states is not easily accomplished: the restricted HF wavefunction[29], in which the orbitals are divided into a doubly occupied closed shell set and a singly occupied open shell set is inadequate since it neglects the different interaction of the α and β spin closed shell orbitals with the open shell. This results in the prediction of positive spin densities only, in doublet state molecules, whereas experimentally both positive and negative spin densities are observed[30]. The unrestricted HF wavefunction[31], in which all the α and β spin orbitals are allowed to be different is not an eigenfunction of $\hat{S}^2$, and a complex spin projection operator must be applied to the singlet determinant to give a proper wavefunction[32]. Also, despite the delightfully simple picture, given by MO theory, of excitation whereby an electron "jumps" from an occupied to an unoccupied orbital, it is exceedingly difficult to obtain variational, that is, MO coefficient optimised, wavefunctions for the excited singlet states due to the requirement of orthogonality to the ground state to prevent collapse of the variational process[33].

A certain amount of polarisation has become apparent in the applications of the "mathematical" type of wavefunction. On the one hand, we have the many successes of MO theory in the study of medium-sized and large molecules at the various, fairly low, levels of accuracy, from the vast numbers of applications of semi-empirical pi-electron theory, to the many recent ab initio Gaussian orbital studies on such diverse systems as the water dimer[34], sulphur dioxide[35] and nucleic acid bases[36]. On the other, very accurate wavefunctions have been obtained for small atoms and diatomic molecules, using large extended bases, with the correlated "mathematical" wavefunctions, for example, Ruedenberg and co-workers' applications of the GSP method[37], and Bender and Davidson's extensive studies of first-row hydrides[38]. There has, however, been less success in finding wavefunctions of intermediate accuracy. But several reasons make the search for methods of obtaining simple correlated wavefunctions for small molecules a necessary one. The first is that there remains a great deal of interesting work to be done on small molecular systems for which accurate wavefunctions will be required. Such fields include atomic and molecular interactions; reaction paths; barriers to conformational changes; excited molecular states, singlet and higher multiplicities; open shell molecules; and energy surfaces for force constant calculations. Continual advances in experimental techniques are making more accurate experimental data available, so that the theoretical results are no longer of academic interest only. One current area of interest is the search for exotic light molecular species for use as rocket fuels[39]: also the spectra of small molecules are of astrophysical importance[40].

The energy quantities that are of greatest interest are often very small, of the order of K calories, for example, the barrier to inversion in ammonia, and much smaller than the correlation error the HF wavefunction. An energy difference, arising from a geometrical change in the system, comes from three sources - a direct energy change, present even if the wavefunction is unchanged, a correlation energy change and a change due to the different contributions from atomic orbitals above the valence shell. Often, the first factor is the dominant one so that even simple MO wavefunctions are capable of reasonably reproducing the correct results; but it is becoming increasingly obvious that the latter two can constitute a considerable proportion of the energy difference, although the relative importance of these two effects is as yet unknown. For example, Pipano has estimated that half the barrier to inversion in ammonia is due to correlation energy changes[41], whilst Clementi has found that the addition of a 3d orbital to the basis, in a HF wavefunction, greatly improves the inversion barrier[42]. Also, two-thirds of the dimerisation energy of BH_3 has been attributed to correlation energy differences[43]. Correlation is thus expected to play an important part in many of the areas listed above.

But it is to be hoped that simple, but well chosen, correlated wavefunctions will be sufficiently accurate for many purposes, firstly from a purely practical point of view, since the very accurate wavefunctions, for example, of Bender and Davidson, involve for a single calculation, a very considerable amount of labour. It cannot be forseen, in the near future, that these methods will be feasible for routine studies of molecular systems, especially considering that many areas will involve energy surfaces, rather than a single equilibrium geometry calculation. On the other hand, it is well known that very accurate fully correlated wavefunctions are an unnecessary luxury since a major proportion of the correlation energy, for example, core correlation energy, is unchanged in different geometrical situations. What is vital is that partially correlated wavefunctions are found which are sufficiently flexible to pick out those parts of the correlation energy that do change. A further incentive is that the explanation of basic electronic phenomena, such as bonding, valency and equilibrium molecular geometries, in quantum mechanical terms is exciting increasing interest[44]: the importance of simple but correlated wavefunctions, in this field, is unquestionable[45].

Such wavefunctions should be within the scope of current computational facilities, but how are they to be obtained? Simplification of the very accurate "mathematical" methods appears to be the obvious answer. But such an approach encounters many difficulties, some of them due to the MO origin of the theories, some from the pure "mathematical" nature of the wavefunction. Firstly, there is the choice of structure to include in the wavefunction: it is usually impossible to include all structures, yet there is no a priori way of picking out the most important. For very small molecules, a brute force approach is possible in which a very large number of structures is used to avoid missing the dominant ones, but, for larger systems, such an approach is out of the question. The Optimised Valence Configuration (OVC) method[46] - in effect, a MCSCF wavefunction - in which only a few configurations are included, has had some success in dealing with diatomics[47], but here the situation is well researched and understood so that it is possible to pick out configurations fairly readily, especially those that lead to the correct dissociation products. For polyatomic molecules, the problem will be much greater. Secondly, the MCSCF and GSP wavefunctions, although, on the face of it, possessing highly convenient practical characteristics, involve some practical problems apart from the choice of configuration. Thus, these wavefunctions require a double iteration process since the structure and orbital expansion coefficients must be optimised simultaneously. In general, the methods used have no guarantee of convergence or a rapid rate of convergence. Orbital orthogonality must be maintained in the variational process, and all variational equations are the same order as the size of the basis, hence increasing rapidly as we go to larger molecules and extended bases. Also, many basis transformations, notoriously time-consuming[49], are usually required.

An interesting attempt to avoid these difficulties whilst retaining the accuracy of this type of wavefunction, is the Separated Bond Pair (SBP) version of the GSP wavefunction in which instead of constructing the pair function orbitals by a variational expansion technique, the pair functions are localised within the bonds of the molecule, and the orbitals are fixed by localisation[50] or hybridisation[51] criteria. A single iteration process and one basis transformation only are needed, and the variational equations are usually of a small, more or less fixed, dimension[52]. However, several disadvantages are attached to this development. As indicated above, the GSP, and hence the SBP, wavefunction neglects all interpair correlation, an approximation that is certainly valid for well separated pairs but on which some doubt has been cast recently for larger systems, when the wavefunction is in localised bond form. For example, the valence intrapair correlation energy in methane has been estimated to be approximately equal to the valence interbond correlation energy[53]. The complete neglect of a possibly large and important proportion of the correlation energy is a serious drawback. And finally, the methods are, in the main, directed towards singlet ground states: the difficulties that are present in the extension of MO theory to excited states and higher multiplicities are still present, especially in the separated pair methods. It should also be noted that the complexity of such wavefunctions due to the delocalised nature of the orbitals or the orthogonality of the basis orbitals makes them unable to contribute greatly to valence theory.

In the light of the foregoing difficulties encountered by "mathematical" methods, a reappraisal of VB theory, and "chemical" wavefunctions in general, is of the utmost importance. Here we shall simply cite some of the more outstanding advantages of

VB theory, but this will be sufficient to show how much might be gained by reconsidering VB theory as a practical method. Thus, correlation is present in the wavefunction, due to the non-orthogonality of the atomic orbitals, at all orders of approximation and of all types, intra- and inter-bond; the choice of structures is, in a minimal basis certainly, a much easier task due to the "chemical" nature of the wavefunction - the dominant structure may be picked out by simple empirical physical rules[48]; the correct dissociation products are ensured; any spin multiplicity may be constructed without difficulty; upper bounds to excited states result automatically from the variational process which is non-iterative; and the optimisation of the spin coupling of the electrons, of great importance in reaction paths where bonds are being broken or formed[54], is much more readily achieved than in the "mathematical" type of wavefunction. These attributes are of vital importance in the fields of interest already listed, yet are, as we have seen, so difficult to fit into the non-"chemical" wavefunction methods. The practical problems of VB theory - the non-orthogonality problem - must not be minimised, but it is certain that larger molecules than have, up to now, been tackled by VB theory are, in fact, tractable. Most of the work on applications of VB theory were carried out in the early 1960's, before the availability of large computers. The attractiveness of MO theory then drew most of the attention away from the VB method, yet if the same amount of work had been put into studying the VB method for small molecules, as has been put into the correlated "mathematical" methods, an accurate widely applicable wavefunction construction method would now be available. In fact, the main problems of the implementation of VB theory, the case of small molecules, have been solved. Rapid methods of evaluating the matrix elements between single determinants of non-orthogonal orbitals, have been devised[55], and the problem of storage of large quantities of intermediate data should be, with the development of disk and tape data storage techniques, no longer troublesome. The limiting factor is that, if the VB wavefunction is expanded into a linear combination of determinants, each structure results in 2^M determinants, where M is the number of Heitler-London bond pairs in the structure, and, as M increases, the number of determinantal matrix elements required increases very rapidly. However, polyatomic molecular systems with up to 20 electrons and 5 bonds should certainly be within the reach of the present computing facilities. Thus, from the point of view of accuracy and generality, the VB method has much to recommend it, compared to the "mathematical" methods, whose deficiencies are less apparent but none the less present; the relegation of VB theory to the state of an historical curio is therefore unfounded, and a revival of interest in the VB method long overdue.

It is an interesting speculation that the VB wavefunction might provide a useful method for starting off a very accurate calculation on a small molecule. It seems certain that the best means of obtaining very accurate wavefunctions is the NO-CI method. The main problem, apart from those of a purely practical nature, and the choice of structure, is the choice of starting correlated wavefunction, from which the initial set of natural orbitals is derived. A straightforward CI wavefunction is usually chosen, but considering the slow convergence of such a wavefunction, advantages would be gained from using a VB wavefunction as starting point. Since a VB wavefunction is likely to be often better than a CI wavefunction, the number of iterations required may be reduced. Since correlation is present in the VB wavefunction, no matter how few structures, an estimate of the NO's is always available. In a CI calculation, the addition of an extra orbital to the basis requires a new MO calculation to bring it into the wavefunction, but, in the VB method, only one extra structure need be added to the wavefunction. In fact, it is an easy matter to obtain a starting wavefunction if a large "exponent" extended basis, i.e. one with more than one of each orbital occupied in the atoms, is to be used. Clementi[56] has performed a VB calculation on neon consisting of only three structures, built from the best 1s, 2s and 2p orbitals and a set of orbitals contracted and expanded with respect to these. Thus, the size of the basis, and hence the number of NO's, is trebled, yet only a three term VB wavefunction is needed to introduce them. The choice of structures to include in the NO-CI wavefunction is a question of some nicety: it is common practice to use a perturbation technique of selecting configurations[57], but it has been suggested that this may be unfavourable to certain types of structure, for example, single excitations[58]. The natural orbital expansion theorem[26], alluded to earlier, is, stated explicitly, that if the wavefunction is to be expanded in a finite set of orbital products

$$\Psi = \sum_{i}^{x} \sum_{j}^{y} a_{ij} \, f_i(1) \, g_j(2 \ldots n) \, ,$$

(ignoring the antisymmetry which may be imposed at a later stage), then the approximate wavefunction is closest to the true wavefunction, for a given x, y, when the f_i are the NO's with the highest occupation numbers, and the wavefunction

$$\Psi = \sum_{i}^{x} c_i \, f_i(1) \, g_i(2\ldots n)$$

where

$$g_i(2\ldots n) = c_i^{-1} \int \Psi(1\ldots n) f_i(1) d\tau_1$$

This is of no direct help in the choice of structure since the wavefunction is expanded in terms of an n-fold product of NO's rather than the natural functions of the one- and (n-1)-particle density functions, f_i and g_i. But the g_i may in turn be expanded in terms of antisymmetrised products of the natural orbitals, $h_{ij}(2\ldots n)$

$$\Psi = \sum_{i}^{x} c_i \, f_i(1) \sum_{j}^{z} d_{ij} \, h_{ij}(2\ldots n)$$

$$= \sum_{ij} c_i \, d_{ij} \, \psi_{ij}(1\ldots n)$$

$$= \sum_{k} e_k \, \psi_k(1\ldots n)$$

where $\psi_k(1\ldots n)$ is a structure based on the NO's. By truncation of the sum over i to x' < x, we obtain the wavefunction which is the closest approximation to the choice of x' to Ψ. Of course, Ψ is never in practice the exact wavefunction, but we assume that it is a good approximation to it, and hence, so are the natural orbitals and their occupation numbers. One might alternatively consider the truncation of the sum over k: the structures with e_k greater than a chosen limit could then be used as the basis for a CI calculation to obtain a better wavefunction. The question is how accurately does a crude wavefunction give the order of importance of the structures: for the purposes of selecting the dominant configurations, it does not matter if the magnitudes of the e_k of the approximate wavefunction are grossly in error, only that the order is correct. In Table 1, the expansion in NO structures of various approximate wavefunctions for the pi-electrons of butadiene are listed. (These wavefunctions are studied in Section 2, where further details may be found). The surprising feature that is evident from these results is that even the crudest "perfect pairing" VB wavefunction gives an order for the structuresremarkably similar to that of the very accurate five parameter wavefunction. Should this be true of non-orthogonal orbital wavefunctions in general, NO-CI calculations based on VB wavefunctions with structures chosen by the expansion of the VB wavefunction in terms of NO's should be capable of giving very accurate wavefunctions indeed.

Table 1 Expansion of Various Wavefunctions for Butadiene in Terms of their Natural Orbitals

Natural Orbital Configuration	Wavefunction				
	MO	NPF	NPF	VB	SBP
$1\bar{1}\ 2\bar{2}$	1.0	.5306	.5461	.5907	.343
$1\bar{1}\ 3\bar{3}$	0.0	-.0445	-.0716	-.1708	-.0301
$1\bar{1}\ 4\bar{4}$	0.0	-.0258	-.0226	-.1003	-.0301
$1\bar{2}\ 3\bar{4}$	0.0	.0271	.0221	.1038	.0301
$2\bar{2}\ 3\bar{3}$	0.0	-.0258	-.0190	-.0980	-.0301
$1\bar{1}\ 2\bar{4}$	0.0	.0259	.0079	.0160	0.0
$2\bar{2}\ 4\bar{4}$	0.0	-.0164	-.0069	-.0631	-.0301
$3\bar{3}\ 4\bar{4}$	0.0	.0057	.0037	.0746	.0106
$4\bar{4}\ 1\bar{3}$	0.0	-.0016	-.0027	-.0188	0.0
$3\bar{3}\ 2\bar{4}$	0.0	.0010	.0020	.0173	0.0
$1\bar{3}\ 2\bar{4}$	0.0	-.0012	-.0015	-.0047	0.0
$2\bar{2}\ 1\bar{3}$	0.0	-.0195	-.0014	-.0122	0.0
Number of Free Parameters	1	2	5	0	2
Energy (eV)	-83.18	-83.30	-84.08	-78.86	-82.53

We have shown so far that the VB wavefunction may have a valuable part to play in the search for simple correlated wavefunctions for small molecules and interacting systems. But there appears to be no hope of extending the VB method beyond these limits, because no valid approximation scheme for overcoming the non-orthogonality problem has been forthcoming. The main difficulty, in looking for an approximation technique which simplifies sufficiently the VB matrix element evaluation to make wavefunctions for large molecules accessible, is the vague formulation of the VB wavefunction - it is not possible to write down a general VB wavefunction for an arbitrary molecule because the structures that are included in the wavefunction depend to a large extent on the atoms that make up the molecule, for example, on their electronegativity, and the molecular geometry. For this reason, it is difficult to carry through the theory of any approximation technique that may appear reasonable, for instance, the approximation of the neglect of overlap greater than a certain order: short of a separate analysis of each VB wavefunction, it is not easy to see how a systematic approximation method may be introduced. It is however vital, in the long run, that a solution to the non-orthogonality problem is found since, in large molecules, very often, only a small part of the molecule is of chemical interest. "Mathematical" methods, which must deal with the molecule as a whole, will eventually become impractical, beyond a very crude level of approximation; the "chemical" approach, which builds up the wavefunction for small units, affords the hope of separating out the important chemical groups for quantum mechanical treatment.

The original attractiveness of the VB wavefunction lay in its link with chemical ideas and concepts, the justification and quantification of which are an important part of quantum chemistry. Although VB theory has, because of its practical drawbacks, had little chance to contribute to this field in a proper way, it is evident that it will in this respect also suffer from the vagueness of the formulation. In spite of being explicitly built on (Heitler-London) bond functions, the possible introduction of alternative spin couplings, ionic and charge transfer states to the wavefunction make the completely general partition of any observable, for example, the electron density, the energy, into quantities referring to the bonds and their interactions quite impossible.

In the search for answers to these problems, we have been lead to consider wavefunctions, which are constructed on the "chemical" principle, but which sacrifice some of the generality of VB theory for a more formal approach which might make the derivation of approximation methods an easier task. One of the simplest and most interesting is the Non-Orthogonal Pair Function (NPF) wavefunction.

$$\Psi = \hat{A}\, \Delta^1(12)\Delta^2(34) \ldots \Delta^N(2N-1,2N)$$

where, in contrast to the GSP and SBP methods, no orthogonality conditions whatsoever are placed on the pair functions Δ^I. Each pair function, as in the SBP method, will be assumed to be completely localised within a small region, usually a bond. The strict formulation of the wavefunction, a product of N pair functions, is gained at the expense of being able to treat 2N-electron singlet ground states only: however, such a wide variety of molecular systems fall into this category that the restrictions are not, for the most part, very important.

If each antisymmetrised singlet pair function is expanded in terms of a pair of non-orthogonal hybrids,

$$\begin{aligned}\Delta^I(12) = {} & [c^I_{11}\, \phi^I_1(1)\, \phi^I_1(2) + c^I_{22}\, \phi^I_2(1)\, \phi^I_2(2) \\ & + c^I_{12}\, (\phi^I_1(1)\, \phi^I_2(2) + \phi^I_2(1)\, \phi^I_{\cancel{2}}(2))]\,(\alpha(1)\beta(2) - \beta(1)\alpha(2)),\end{aligned}$$

1/

the link with VB theory is evident. Included in the wavefunction are all structures that can be constructed from the set of "covalent" and "ionic" bond functions, but instead of each having a completely free coefficient, as in VB theory, the coefficients of structures are related: thus, for a molecule with N bonds, 3^N structures are included but with only 3N-1 free coefficients. The effect is to make the optimisation of the pair function coefficients, C^I_{ij}, an iterative process, rather than the solution of a secular problem, as in VB theory.

The SBP and NPF methods differ only in the orthogonality constraint of SBP theory: this similarity prompts some comparisons. The SBP wavefunction, although it yields results which are consistently better than the comparable MO wavefunction, only yields a very small proportion of the correlation energy[59]. This may be traced to two factors - the neglected interpair correlation energy, which may be quite large, and the strong orthogonality restriction which forces hybrids on the same atom to be orthogonal and hence fixes the s, p, d orbital mixing ratios of the hybrids. From an energetic point of view, this may be highly undesirable. The non-orthogonality of the NPF basis ensures that both of these constraints are avoided: the neglect of interpair correlation is achieved by the orthogonality conditions which force "charge-transfer-between-bonds" structures to be mixed in with fixed erroneous weights. By omitting, in the first approximation, the charge transfer structures, the NPF wavefunction introduces interpair correlation, and with no hybrid orthogonality constraints, the weights of the different orbitals within the hybrids may be optimised freely. A much more accurate wavefunction is therefore to be expected.

The range of molecular systems that may be studied by the SBP method is limited by strong orthogonality constraint which does not allow the same orbital to be in different pair functions. Thus, for example, the hydrogen bonded system, FHF^-, where the hydrogen atom is bonded equally to both atoms, is outside the range of the SBP, but not the NPF, method, where an orbital, the hydrogen 1s orbital in this case, may contribute to more than one pair function. Also, in this way, charge transfer structures may be brought variationally into the wavefunction, unlike in the SBP method. The dropping of the strong orthogonality constraint opens up many interacting systems to the pair function method. And, finally, the NPF wavefunction which is based directly on localised bond functions will have an important role to play in the quantitative study of valence problems: the non-orthogonality of the basis orbitals will make the results conceptually more appealing than those obtained by the SBP method which is based on orthogonal hybrids.

The remainder of this chapter is devoted to the detailed study of the NPF wavefunction: in the next section, the general theory is given with some ab initio applications to small molecular systems. In the third section, the solution of the non-orthogonal problem via approximation methods and the NPF wavefunction is considered, and finally, valence theory and the contribution the NPF method has to make are discussed, in section four.

[illegible]

[illegible]

[illegible]

[illegible]

See [illegible]

(19[illegible]) J.Chem.Phys. [illegible]

McWeeny, R. & Sutcliffe, B.T. (1969) "Methods of Molecular Quantum Mechanics" p. 100, Academic Press.

REFERENCES

1. Roothaan, C.C.J. (1951) Rev.Mod.Phys. 23, 69

2. Eyring, H., Walter, J. & Kimball, G.E. (1944) "Quantum Chemistry" p. 212, Wiley

3. Prosser, F. & Hagstrom, S. (1968) Int.J.Quant.Chem. 2, 89

4. Cook, D.B., Hollis, P.C. & McWeeny, R. (1967) Mol.Phys. 13, 553

5. For a more optimistic estimate, see Clementi, E. (1969) Int.J.Quant.Chem. 3S, 179

6. Pople, J.A., Beveridge, D.L. & Dobosh, P.A. (1967) J.Chem.Phys. 47, 2026

7. (a) Cook, D.B. & Palmieri, P. (1969) Mol.Phys. 17, 271
   (b) Hehre, W.J., Stewart, R.F. & Pople, J.A. (1969) J.Chem.Phys. 51, 2657

8. Coulson, C.A. (1961) "Valence" p. 174, Oxford University Press

9. Slater, J.C. (1951) J.Chem.Phys. 19, 220

10. McWeeny, R. (1954) Proc.Roy.Soc.London A223, 63 and 306

11. Van der Lugt, W.Th.A.M. & Oosterhoff, L.J. (1970) Mol.Phys. 18, 177

12. See, for example, Moore, P.L., Browne, J.C. & Matsen, F.A. (1965) J.Chem.Phys. 43, 903

13. McWeeny, R. & Sutcliffe, B.T. (1969) "Methods of Molecular Quantum Mechanics" p. 100, Academic Press

14. Reference 13, p. 84

15. Sinanoglu, O. (1964) "Advances in Chemical Physics" Volume 6, p. 315, Wiley-Interscience

16. (a) Grimaldi, F., Lecourt, A. & Moser, C. (1967) Int.J.Quant.Chem. 1S, 153

(b) Dunning, T.H., Hunt, W.J. & Goddard III, W.A. (1969) Chem.Phys.Letters 4, 147

17. Clementi, E. & Veillard, A. (1967) Theoret.chim.Acta, 7, 133

18. Ruedenberg, K. & Miller, K.J. (1968) J.Chem.Phys. 48, 3414

19. Hurley, A.C., Lennard-Jones, J.E. & Pople, J.A. (1953) Proc.Roy.Soc.London, A220, 446

20. Arai, T. (1960) J.Chem.Phys. 33, 95

21. Reference 18, p. 3422

22. Reference 15, p. 353

23. Cizek, J., Paldus, J. & Sroubkova, L. (1969) Int.J.Quant.Chem. 3, 149

24. Reference 13, p. 175

25. Kutzelnigg, W. (1965) Theoret.chim.Acta, 3, 241

26. Coleman, A.J. (1963) Rev.Mod.Phys. 35, 668

27. Bender, C.F. & Davidson, E.R. (1966) J.Phys.Chem. 70, 2675

28. Green, S. (1970) J.Chem.Phys. 52, 3100

29. Roothaan, C.C.J. (1960) Rev.Mod.Phys. 32, 179

30. Hall, G.G. & Amos, A.T. (1965) "Advances in Atomic and Molecular Physics" p. 30, Academic Press

31. Berthier, G. (1964) "Molecular Orbitals in Chemistry, Physics and Biology" Ed. Lowdin, P.-O. & Pullman, B. p. 71, Academic Press

32. Sando, K.M. & Harriman, J.E. (1967) J.Chem.Phys. 47, 180

33. Hunt, W.J. & Goddard III, W.A. (1969) Chem.Phys.Letters 3, 414

34. Kraemer, W.P. & Diercksen, G.H.F. (1970) Chem.Phys.Letters, 5, 463

35. Rothenberg, S. & Schaeffer III, H.F. (1970) J.Chem.Phys. 53, 3014

36. Clementi, E. Andre, J.M., Andre, M.Cl., Klint, D. & Hahn, D. (1969) Acta.Phys.Hung. 27, 493

37. Silver, D.M., Mehler, E.L. & Ruedenberg, K. (1970) J.Chem.Phys. 52, 1174

38. Bender, C.F. & Davidson, E.R. (1969) Phys.Rev. 183, 23

39. Kaufman, J.J. & Sachs, L.M. (1970) J.Chem.Phys. 52, 3534

40. Wilkinson, P.G. (1965) Astrophys.J. 138, 778

41. Pipano, A., Gilman R.R., Bender, C.F. & Shavitt, I. (1970) Chem.Phys.Letters 4, 583

42. Rauk, A., Allen, L.C. & Clementi, E. (1970) J.Chem.Phys. 52, 4133

43. Geluo, M., Staemmler, V., Ahlrichs, R. & Kutzelnigg, W. (1970) Chem.Phys.Letters 7, 503

44. Feinberg, M.J., Ruedenberg, K. & Mehler, E.L. (1970) "Advances in Quantum Chemistry" Volume 5, p. 27, Academic Press

45. Linnett, J.W. (1964) "Electronic Structure of Molecules - a New Approach" Methuen

46. Das, G. & Wahl, A.C. (1966) J.Chem.Phys. 44, 87

47. Wahl, A.C. & Das, G. (1970) "Advances in Quantum Chemistry" Volume 5, p. 261, Academic Press

48. Murrell, J.N., Kettle, S.F.A. & Tedder, J.M. (1965) "Valence Theory" p. 194, Wiley

49. Tang, K.C. & Edmiston, C. (1970) J.Chem.Phys. 52, 997

50. Robb, M.A. & Csizmadia, I.G. (1970) Int.J.Quant.Chem. 4, 365

51. McWeeny, R. & Ohno, L. (1960) Proc.Roy.Soc.London A253, 367

52. Klessinger, M. & McWeeny, R. (1965) J.Chem.Phys. 42, 3343

53. Ahlrichs, R. & Kutzelnigg, W. (1968) Chem.Phys.Letters 1, 651

54. Ladrier, R.C. & Goddard III, W.A. (1969) J.Chem.Phys. 51, 1073

55. See reference 3 and King, H.F., Stanton, R.E., Kim, H., Wyatt, R.E. & Parr, R.G. (1967) J.Chem.Phys. 47, 1936

56. Clementi, E. (1969) Int.J.Quant.Chem. 3S, 179

57. Pipano, A. & Shavitt, I. (1968) Int.J.Quant.Chem. 2, 741

58. Olympia, P.L. & Smith, D.W. (1970) J.Chem.Phys. 52, 67

59. See, for example, Klessinger, M. (1968) Symposia of the Faraday Society 2, 73. For methane, the percentage is less than 13%.

SECTION 2

THE THEORY OF THE NON-ORTHOGONAL PAIR FUNCTION METHOD AND ITS APPLICATION TO SOME FEW-ELECTRON SYSTEMS

In this section, explicit expressions for the one and two particle density functions of the non-orthogonal pair function (NPF) wave function,

$$\Psi = \hat{A}\,\Lambda^1(12)\Lambda^2(34)\,\ldots\,\Lambda^N(2N-1,2N)\ ,$$

where $\hat{A}$ is the antisymmetrisation operator $\sum_P^{N!}(-1)^P P$, will be derived for the case in which the spin pair functions, $\Lambda^I(12)$, are expanded in an arbitrary basis of one electron orbitals, χ_i^I,

$$\Lambda^I(12) = \sum_{ij}^{n_I} C_{ij}^I \chi_i^I(1)\chi_j^I(2)\alpha(1)\beta(2)$$

The variational equations, which determine the optimum pair functions by minimisation of the energy with respect to variation of the free parameters C_{ij}^I, will also be considered. And finally, applications of the theory to some four electron systems will be presented.

Although we shall restrict out attention to orbital expansions of the pair functions, this is by no means necessary. An expansion in which inter-electronic distances are explicitly introduced is an attractive alternative since electron correlation is known to be much better described when functions of r_{ij} are present in the wavefunction than by simple orbital products. This method of development of the pair function is only possible within a non-orthogonal formulation of the pair function wavefunction: orthogonality constraints, such as are imposed in the 'separated pair' method, exclude an expansion of this type. However, the integral evaluation problem, which such an approach entails, has yet to be solved satisfactorily for many-electron systems, so, for this practical reason, we shall adhere to an orbital expansion.

The only restriction imposed on the wavefunction is that each pair function, and hence the total wavefunction, has the spins of its electrons coupled to a singlet: the expansion coefficient matrix $\underline{C}^I$ must therefore be symmetric, i.e. $C^I_{ij} = C^I_{ji}$. This is a natural restriction for a wavefunction designed to describe molecules with localised two-electron bonds. In cases, where a dominant bonding scheme is not apparent, the addition of alternative spin couplings to the "perfect-pairing" one may be necessary. This may be achieved by allowing C^I_{ij} to be different from C^I_{ji}, and is equivalent to mixing some triplet character into the singlet pair function. The total wavefunction is then no longer an eigenfunction of the spin squared operator and a suitable spin projection operator must be applied.[1]

There have been few attempts to develop the theory of the pair function wavefunction without the imposition of orthogonality constraints. McWeeny and coworkers have considered in detail the situation where the wavefunction may be approximated by an antisymmetrised product of two non-orthogonal groups, each group being comprised of an arbitrary number of electrons, with special attention being paid to the spin coupling of the two groups.[2] The forms of the wavefunctions for the separate groups has not been elaborated so that an explicit general formulation of the method is not possible. The simple case of two pairfunctions has been studied by Silver.[3] Although the strong orthogonality constraint is relaxed, the one-electron orbital basis is assumed to be orthogonal: this is no constraint so long as each pair function is expanded in the complete set of orthogonal orbitals. However, in practical terms, the condition means that the approach which may be feasible for four-electrons is certainly not so for larger systems. Since we are particularly interested in the possibility of the application of the NPF method to molecules with localised two-electron bonds, it is vital that short pair function expansions, such as are used in the "Separated Bond Pair" method, may be applied. One-electron orbital orthogonality is therefore not compatible with this approach.

Before considering the density functions, we shall discuss in some detail the derivation of the explicit expression for the normalisation integral, as this will serve to illustrate the method of expansion of the density functions. The normalisation integral is

$$S = \int \Psi(1,2,\ldots 2N)\Psi(1,2,\ldots 2N)d\tau_{1\ldots 2N}$$

which, by applying the "turn-over" rule, reduces to

$$\begin{aligned} S &= (2N)!\int \Delta^1(12)\Delta^2(34)\ldots\Delta^N(2N-1,2N)\hat{A}\Delta^1(12)\ldots\Delta^N(2N-1,2N)d\tau_{1\ldots 2N} \\ &\equiv (2N)!\ S_R \end{aligned}$$

Expanding the pair functions in terms of spin orbitals,

$$\Delta^I(12) = \sum_{ij} C^I_{ij}\ \phi^I_i(1)\ \bar{\phi}_j{}^I(2)$$

where $\phi^I_i(1) = \chi^I_i(1)\alpha(1)$ and $\bar{\phi}_j{}^I(2) = \chi^I_j(2)\beta(2)$, we may write the "reduced" normalisation integral, S_R, as

$$S_R = \sum_{ijk\ell} \ldots \sum_{pqrs} \ldots C^1_{ij}\, C^2_{k\ell} \ldots C^1_{pq}\, C^2_{rs} \ldots$$

$$\times \int \phi^1_i(1)\bar{\phi}^1_j(2)\phi^2_k(3) \ldots \hat{A}\, \phi^1_p(1)\bar{\phi}^1_q(2)\phi^2_r(3) \ldots d\tau_{1\ldots 2N}$$

It may be noted that a non-antisymmetrised form is preferred for the individual pair functions: the use of the antisymmetric form gives a result which differs only by a constant. The second line of the above equation may be written as a determinant

$$\begin{vmatrix} S^{11}_{ip} & 0 & S^{12}_{ir} & 0 & \ldots & S^{1N} \\ 0 & S^{11}_{jq} & 0 & S^{12}_{js} & \ldots & \ldots \\ S^{21}_{kp} & \ldots & \ldots & \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ S^{N1} & \ldots & \ldots & \ldots & \ldots & S^{NN} \end{vmatrix},$$

where $S^{IJ}_{ij} = \int \chi^I_i(1)\chi^J_j(1)dV_1$. The zeros arise through the integration over the spin coordinates. Re-arrangement of the rows and columns of the determinant bring it to the block diagonal form,

$$\left|\begin{array}{c|c} A & 0 \\ \hline 0 & B \end{array}\right|$$

which is equivalent to the product of two determinants of order N, $|A|.|B|$. Thus

$$S_R = \sum_{ijk\ell..} \sum_{pqrs..} C^1_{ij}\, C^2_{k\ell} \cdots C^1_{pq}\, C^2_{rs} \cdots$$

$$\times \begin{vmatrix} S^{11}_{ip} & S^{12}_{ir} & \cdots \\ S^{21}_{kp} & S^{22}_{kr} & \cdots \\ \cdots & \cdots & \cdots \end{vmatrix} \begin{vmatrix} S^{11}_{jq} & S^{12}_{js} & \cdots \\ S^{21}_{\ell q} & S^{22}_{\ell s} & \cdots \\ \cdots & \cdots & \cdots \end{vmatrix}$$

Since a given summation subscript occurs exclusively in one row or column, we may multiply each element in that row or column by a coefficient which has the same subscript and perform the summation over this dummy index. By multiplying the rows of the first determinant by the first set of coefficients, i.e. C^1_{ij}, $C^2_{k\ell}$ ..., and the columns of the second by the set, C^1_{pq}, C^2_{rs}..., and rearranging the result using the relations $S^{IJ}_{ij} = S^{JI}_{ji}$ and $C^I_{ij} = C^I_{ji}$, we find that

$$S_R = \sum_{j\ell} \cdots \sum_{pr} \cdots \begin{vmatrix} CS(1,1)_{jp} & CS(1,2)_{jr} & \cdots \\ CS(2,1)_{\ell p} & CS(2,2)_{\ell r} & \cdots \\ \cdots & \cdots & \cdots \end{vmatrix}$$

$$\times \begin{vmatrix} CS(1,1)_{pj} & CS(2,1)_{rj} & \cdots \\ CS(1,2)_{p\ell} & CS(2,2)_{r\ell} & \cdots \\ \cdots & \cdots & \cdots \end{vmatrix}$$

where the notation $\mathrm{CS}(I,J)_{ij}$ is used for the matrix product $\underline{C}^{I}\underline{S}_{ij}^{IJ}$. (See Appendix 1). Now, introducing the standard determinant expansion

$$|D| = \sum_{x} \sum_{y\neq x} \sum_{z\neq x,y} \cdots (-1)^{P} D_{x1} D_{y2} D_{z3} \cdots$$

$$= \sum_{a} \sum_{b\neq a} \sum_{c\neq a,b} \cdots (-1)^{Q} D_{1a} D_{2b} D_{3c} \cdots ,$$

where $(-1)^{P}$ and $(-1)^{Q}$ are the parities of the transpositions, the 'reduced' normalisation integral becomes

$$S_{R} = \sum_{j\ell..} \sum_{pr..} \sum_{\{X,Y,Z..\}} \sum_{\{A,B,C..\}} (-1)^{P+Q}$$

$$\mathrm{CS}(X,1)_{xp}\ \mathrm{CS}(Y,2)_{yr}\ \mathrm{CS}(Z,3)_{zt} \cdots$$

$$\times\ \mathrm{CS}(1,A)_{pa}\ \mathrm{CS}(2,B)_{rb} \cdots$$

where $\sum_{\{X,Y,Z..\}}$ denotes the summations $\sum_{X} \sum_{Y\neq X} \sum_{Z\neq X,Y} \cdots$. It should be noted that each superscript has associated with it a summation subscript which is carried with it under all permutations. Thus the subscript x in the above equation stands for the summation index which corresponds to the superscript X. Once the value of X is fixed, then x is to be replaced by the requisite dummy index: for example, if X=1, then $x\equiv j$, and if X=2, then $x\equiv\ell$.

Summation over the set of subscripts p,r,... may now be performed, giving

$$S_R = \sum_{j\ell..}\ \sum_{\{XYZ..\}}\ \sum_{\{ABC..\}} (-1)^{P+Q}\ \mathrm{CSCS}(X,1,A)_{xa}$$

$$\times\ \mathrm{CSCS}(Y,2,B)_{yb}\ \mathrm{CSCS}(Z,3,C)_{zc}\cdots$$

This may be written in the more succinct form

$$\begin{aligned} S_R &= \sum_{j\ell..} P^N Q^N\ \mathrm{CSCS}(P_1,1,Q_1)_{p_1q_1}\ \mathrm{CSCS}(P_2,2,Q_2)_{p_2q_2} \\ &\times\ \mathrm{CSCS}(P_3,3,Q_3)_{p_3q_3}\ \cdots \\ &= \sum_{j\ell..} P^N Q^N \prod_{I=1}^{N} \mathrm{CSCS}(P_I,I,Q_I)_{p_Iq_I} \\ &= \mathrm{TR}[P^N Q^N \prod_{I=1}^{N} \mathrm{CSCS}(P_I,I,Q_I)_{p_Iq_I}] \end{aligned} \qquad (2.1)$$

where

$$\begin{aligned} \mathrm{CSCS}(I,J,K)_{ij} &\equiv \{\mathrm{CS}(I,J) \times \mathrm{CS}(J,K)\}_{ij} \\ &\equiv \underline{C^I S^{IJ} C^J S^{JK}}_{ij} \end{aligned}$$

$P^N = \sum_{P}^{N!} (-1)^P\ P$ is the permutation operator of order N which operates on the superscripts P_I and subscript summation indices p_I simultaneously, and TR is the set of summations over all subscript labels. The normalisation integral is thus a sum of traces of matrix products. For example, with N = 2,

$$S_R = \sum_{j\ell} \{\mathrm{CSCS}(1,1,1)_{jj}\ \mathrm{CSCS}(2,2,2)_{\ell\ell}$$

$$- \mathrm{CSCS}(1,1,2)_{j\ell}\ \mathrm{CSCS}(2,2,1)_{\ell j} - \mathrm{CSCS}(2,1,1)_{\ell j}\ \mathrm{CSCS}(1,2,2)_{j\ell}$$

$$+ \mathrm{CSCS}(2,1,2)_{\ell\ell}\ \mathrm{CSCS}(1,2,1)_{jj}\}$$

$$= [\mathrm{CSCS}(1,1,1)][\mathrm{CSCS}(2,2,2)] + [\mathrm{CSCS}(2,1,2)]^2$$

$$- 2[\mathrm{CSCS}(1,1,2) \times \mathrm{CSCS}(2,2,1)]$$

where [ ] means that the matrix trace is to be taken, and × stands for the matrix product. In standard matrix notation, S_R is

$$\mathrm{tr}(\underline{C^1S^{11}C^1S^{11}})\,\mathrm{tr}(\underline{C^2S^{22}C^2S^{22}}) + \{\mathrm{tr}(\underline{C^1S^{12}C^2S^{21}})\}^2$$

$$- 2\mathrm{tr}(\underline{C^1S^{11}C^1S^{12}C^2S^{22}C^2S^{21}}).$$

From now on, the alternative matrix notation of Appendix 1 will be used. The reduction of S_R so that it involves only matrix algebra is particularly convenient for practical applications:

the computational aspects, with simple rules for the expansion of S_R, are discussed in more detail in Section 3.

It is well known that the expectation value of any symmetric operator involving the coordinates of m electrons may be evaluated once the m'th order density function for the system is known. Thus, for an operator $\hat{X}(1,\ldots m)$, the expectation value, X, is given by

$$X = \int \hat{X}(1,\ldots m)\rho(1,2,\ldots m;\ 1'\ldots m')d\tau_{1'\ldots m' \to 1..m}$$

where the m'th order density function

$$\rho(1,\ldots m;\ 1'\ldots m') = \frac{(2N)!S^{-1}}{m!(2N-m)!}$$

$$\times \int \Psi(1,..2N)\Psi(1',\ldots m',m+1,\ldots 2N)d\tau_{m+1\ldots 2N}$$

for a 2N electron system. S is the normalisation integral, $\int \Psi(1,\ldots 2N)\Psi(1,\ldots 2N)d\tau_{1\ldots 2N}$, and the notation $\int \ldots d\tau_{x'\to x}$ is used to indicate that x' is made equal to x before the integration over x is performed. Specifically, the energy is the expectation value of the Hamiltonian which contains operators involving one and two particle coordinates only.

$$\hat{H} = \sum_{i=1}^{2N} \hat{f}(i) + \sum_{i<j}^{2N} \hat{g}(i,j)$$

where f(i) is the sum of the kinetic energy and nuclear attraction energy operators and g(i,j) the electron repulsion energy operators. The energy of the system, E, is therefore

$$E = \int \Psi \hat{H} \Psi \, d\tau \Big/ \int \Psi \Psi \, d\tau$$

$$\int \hat{f}(1) \, \rho(1;1') \, d\tau_{1'\to 1} + \int g(1,2)\rho(1,2;1',2')d\tau_{1'2'\to 12} \qquad (2.2)$$

Since, in practice, one is never concerned with operators involving more than two electrons, we shall limit our attention to the one-electron density function

$$\rho(1;1') = 2N\, S^{-1} \int \Psi(1,\ldots 2N)\Psi(1',2\ldots 2N)d\tau_{2\ldots 2N}$$

and the two-electron density function

$$\rho(1,2;1',2') = N(2N-1)S^{-1} \int \Psi(1,\ldots 2N)\Psi(1'2',3\ldots 2N)d\tau_{3\ldots 2N}$$

Most operators of interest are spin independent, for example, the Hamiltonian, dipole moment operators; it is then sufficient to determine the reduced density functions obtained by integrating over the remaining spin coordinates

$$P(1,\ldots m;1',\ldots m') = \int \rho(1,\ldots,;1'\ldots m')dS_{1'\ldots m'\to 1\ldots m}$$

Let us first consider the evaluation of the reduced one-electron density function

$$P(1;1') = 2N\, S^{-1} \int \Psi(1\ldots 2N)\Psi(1',2\ldots 2N)\,d\tau_{2\ldots 2N}\,dS_{1'\to 1}$$

By introducing the operator

$$\hat{O}_1 = \sum_{i=1}^{2N} \hat{O}(i)$$

where the one-electron operator $\hat{O}(i)$ removes the integration over the spatial coordinates of electron i and replaces, after integration over all other coordinates, the electron labels i by 1 in the terms originating from the lefthand wavefunction and 1' in those from the righthand wavefunction, the density function may be rewritten

$$P(1;1') = S_R^{-1} \int \psi(1\ldots 2N)\hat{O}_1\ \hat{A}\ \psi(1\ldots 2N)\,d\tau_{1\ldots 2N}$$

where

$$\psi(1\ldots 2N) = \prod_{I=1}^{N} \Delta^{I}(2I-1,2I)$$

Since the spin integration gives a zero result for any permutation which interchanges an even and an odd numbered electron, the total antisymmetriser $\hat{A}$ may be replaced by $\hat{A}_+\hat{A}_-$ where $\hat{A}_+$ permutes the even electrons only and $\hat{A}_-$ the odd. Thus, with $\Phi(1\ldots 2N)$ the product of spinless pair functions,

$$P(1;1') = S_R^{-1} \int \Phi(1\ldots 2N)\hat{O}_1 \; \hat{A}_+\hat{A}_- \; \Phi(1\ldots 2N)dV_{1\ldots 2N}$$

Writing $\hat{O}_1$ as $\hat{O}_+ + \hat{O}_-$, where

$$\hat{O}_+ = \sum_{i \text{ even}} \hat{O}_1(i)$$

and

$$\hat{O}_- = \sum_{i \text{ odd}} \hat{O}_1(i)$$

$$P(1;1') = S_R^{-1} \left\{ \int \Phi \; \hat{O}_+\hat{A}_+\hat{A}_- \; \Phi \; dV + \int \Phi \; \hat{O}_-\hat{A}_+\hat{A}_- \; \Phi \; dV \right\}$$

The spinless pair function, $\lambda^I(12) = \sum_{ij} C^I_{ij} \; \chi^I_i(1)\chi^I_j(2)$, is symmetric with respect to interchange of the electron labels: the label interchange, $1 \leftrightarrow 2$, $3 \leftrightarrow 4$, etc., thus leaves Φ unchanged. Hence the two terms on the right-hand side of the above equation are identical and we may write

$$P(1;1') = 2S_R^{-1} \int \Phi(1\ldots 2N)\hat{O}_-\hat{A}_+\hat{A}_- \; \Phi(1\ldots 2N)dV_{1\ldots 2N}$$

The expansion of the one-particle density function now follows the same path as the expansion of the normalisation integral.

$$P(1;1') = 2S_R^{-1} \sum_{I=1}^{N} \; \sum_{ijk\ell\ldots} \; \sum_{pqrs\ldots} \; C^1_{ij} \; C^2_{k\ell}\ldots \; C^1_{pq} \; C^2_{rs}\ldots$$

$$\begin{vmatrix} S^{11}_{ip} & S^{12}_{ir} & \cdots \\ \cdots & \cdots & \cdots \\ T^{I1}_{xp} & T^{I2}_{xr} & \cdots \\ \cdots & \cdots & \cdots \end{vmatrix} \begin{vmatrix} S^{11}_{jq} & S^{12}_{js} & \cdots \\ \cdots & \cdots & \cdots \\ S^{I1}_{yq} & S^{I2}_{ys} & \cdots \\ \cdots & \cdots & \cdots \end{vmatrix}$$

T^{IJ}_{ij} is an integral over the operator $\hat{O}_-$ and is hence an orbital product,

$$T^{IJ}_{ij} = \int \chi^I_i(2I-1)\ \hat{O}(2I-1)\ \chi^J_j(2I-1)dV_{2I-1}$$

$$= \chi^I_i(1)\chi^J_j(1')$$

Completing the expansion, we eventually arrive at the final expression for the reduced one-particle density function,

$$P(1;1') = 2S_R^{-1} \sum_{I=1}^{N} \mathrm{TR}[P^N Q^N\ \mathrm{CSCT}(P_I,I,Q_I)_{p_I q_I} \times \prod_{J\neq I}^{N} \mathrm{CSCS}(P_J,J,Q_J)_{p_J q_J}] \qquad (2.3)$$

where $\mathrm{CSCT}(I,J,K)_{ij} \equiv \underline{C^I S^{IJ} C^J T^{JK}}_{ij}$.

It is evident from inspection of the above equation that the one-particle density function normalisation condition is obeyed,

$$\int P(1;1')dV_{1'\to 1} = 2N$$

The two-electron density function is evaluated in an entirely analogous manner.

$$P(1,2;1',2') = S_R^{-1}\int \psi(1\ldots 2N)\hat{O}_2\hat{A}\,\psi(1\ldots 2N)d\tau_{1\ldots 2N}$$

where $\hat{O}_2 = \sum_{i<j}^{2N} \hat{O}(i,j)$.

$\hat{O}(i,j) = \hat{O}(i)\hat{O}(j)$ removes the integration over the spatial coordinates of electrons i and j. The operator is then expanded to give

$$\hat{O}_2 = \hat{O}_D + \hat{O}_{++} + \hat{O}_{--} + \hat{O}_{+-} + \hat{O}_{-+}$$

where $\hat{O}_D = \sum_i \hat{O}(i,i+1)$, i odd,

$\hat{O}_{++} = \sum_{i<j} \hat{O}(i,j)$, i and j even,

$\hat{O}_{--} = \sum_{i<j} \hat{O}(i,j)$, i and j odd,

$\hat{O}_{+-} = \sum_{i<j} \hat{O}(i,j)$, i even, j odd,

and $\hat{O}_{-+} = \sum_{ij} \hat{O}(i,j)$, i odd, j even and $i + 2 < j$.

Reduction of the antisymmetriser by spin integration, as before, and expansion of $\hat{O}_2$ allows the density function to be simplified to

$$\begin{aligned} P(1,2;1',2') = S_R^{-1} \{ & \int \Phi \, \hat{O}_D \hat{A}_+ \hat{A}_- \, \Phi \, dV_{1\ldots 2N} \\ & + 2 \int \Phi \, \hat{O}_{--} \hat{A}_+ \hat{A}_- \, \Phi \, dV_{1\ldots 2N} \\ & + 2 \int \Phi \, \hat{O}_{+-} \hat{A}_+ \hat{A}_- \, \Phi \, dV_{1\ldots 2N} \} \end{aligned}$$

Expansion of the pair functions into an orbital basis yields the final expression

$$\begin{aligned} P(1,2;1',2') = S_R^{-1} \{ & \sum_{I=1}^{N} \mathrm{TR}[P^N Q^N \; \mathrm{CTCT}(P_I,I,Q_I)_{p_I q_I} \\ & \times \prod_{J \neq I}^{N} \mathrm{CSCS}(P_J,J,Q_J)_{p_J q_J} \\ & + 2 \sum_{I<J}^{N} \mathrm{TR}[P^N Q^N \; \mathrm{CSCT}(P_I,I,Q_I)_{p_I q_I} \\ & \times \{\mathrm{CSCT}(P_J,J,Q_J)_{p_J q_J} + \mathrm{CTCS}(P_J,J,Q_J)_{p_J q_J}\} \\ & \prod_{K \neq IJ}^{N} \mathrm{CSCS}(P_K,K,Q_K)_{p_K q_K}]\} \end{aligned} \tag{2.4}$$

Further expansion of the density functions expressions with particular reference to practical computational methods for molecular systems will be postponed until Section 3.

We are now in a position to calculate, from equations (2.1) (2.2), (2.3) and (2.4), the electronic energy, given the integrals of the Hamiltonian operator over the chosen one-electron basis orbitals and the pair function coefficients C^I_{ij}. The conditions which determine the optimal values of the coefficients may be derived by appeal to the variation theorem, which states that an approximate wavefunction, containing variable parameters, always gives, under any variation of these parameters, an upper bound to the true energy. The objective is thus to find the conditions giving the values of the coefficients C^I_{ij} which make the energy a minimum, these values then defining the optimum wavefunction. The minimum in the electronic energy, E, is located by setting the variation of E with respect to the coefficients equal to zero,

$$dE/dC^I_{ij} = 0, \qquad \text{for all } I,\ i,\ j.$$

Since

$$E = \int \Psi\, \hat{H}\, \Psi\, d\tau \Big/ \int \Psi\, \Psi\, d\tau \equiv H/S$$

these conditions may be written

$$dH/dC^I_{ij} - E\ dS/dC^I_{ij} = 0$$

By expansion of H and S, which are quadratic in $\underline{C}^I$, in terms of these coefficients, a set of equations of the form

$$\sum_{k\ell} C^I_{k\ell}(H^I_{ijk\ell} - E\ S^I_{ijk\ell}) = 0, \qquad \text{all } ij$$

may be derived. These equations define a secular problem of order $(n_I)^2$, where n_I is the number of one-electron orbitals in expansion of pair function I, and which may be solved by standard techniques. Since C^I_{ij} must be equal to C^I_{ji}, the order of the secular equations may be reduced to $n_I(n_I+1)/2$. The coefficients C^J_{mn} $(J \neq I)$ must have fixed values for the construction of the matrix elements $H^I_{ijk\ell}$ and $S^I_{ijk\ell}$ required in the secular problem for the optimum C^I_{ij}: the set of N such secular problems must therefore be solved iteratively, starting from guessed trial coefficients, until the coefficients of all pair functions are self-consistent. The presence of the permutation operators in the expressions for H and S make the derivation of explicit general expressions for $H^I_{ijk\ell}$ and $S^I_{ijk\ell}$ a very complicated process, although, in particular cases, straightforward algebra will give the desired expressions.

Some preliminary considerations of the practical aspects of the NPF method merit some attention at this stage. The application of the NPF method to any 2N electron molecular system may be summarised as

1. The choice of a one-electron, generally atomic, orbital basis, the partition of the orbital basis into the pair functions, and the evaluation of the integrals over the one-electron orbitals.
2. The determination of the optimum coefficients by the iterative solution of the secular equations.
3. The evaluation of the density functions for the computation of molecular properties.

The choice of basis depends on the accuracy required: this, and the integral evaluation problem, will not be further considered here. The partitioning of the basis will depend on the system under study: in the case of molecules, which are chemically described in terms of two-electron bonds and e/ lon~~g~~ pairs, the use of a basis of directed or hybrid atomic orbitals will enable an unambiguous choice of partition to be made. In molecular systems, where a dominant bonding scheme is not apparent, some atomic orbitals may contribute to more than one pair function or 'bond': this possibility is allowed in a non-orthogonal formulation of the pair function wavefunction, but is prohibited in a strongly-orthogonal pair function formulation. Naturally, the more atomic orbitals that are allowed to contribute to each pair function, the more accurate will be the wavefunction. The effects of different basis orbital partitions will be considered further in some practical applications at a later stage. Step 3 proceeds straightforwardly, using equations (2.1), (2.3) and (2.4), once the pair function coefficients are fixed either by optimisation as in Step 2 or by the transfer of coefficients optimised in a similar but simpler molecule, for example, using the optimum coefficients for the C-H bond in methane, in a calculation on the ethane molecule. (See Section 4). However, it is apparent that the number of terms generated by the permutation operators in the expansions of the density functions increases rapidly with the number of electrons: for instance, the number of terms in the normalisation integral, equation (2.1), is $(N!)^2$ and each term is a product of 4N matrices. Thus, for N=2,4 and 6, the number of terms is 4, 576 and 518,400. And this is only for the evaluation of the normalisation integral once: the density function expansions involve many times more terms, and the solution of the secular problem, as outlined above, requires the repeated evaluation of even more complex expressions. Even with the aid of modern computers, it appears that the computation time necessary for a complete calculation will be such as to rule out the NPF method as a practical proposition for all but the smallest molecular systems.

The situation may be compared to that in the VB method. The matrix elements, $H_{ij} = \int \Psi_i \hat{H} \Psi_j \, d\tau$, where Ψ is a "structure" wavefunction, that is, a linear combination of determinants with coefficients determined by the molecular symmetry and the spin state, which are required for the determination of the optimum coefficients of the VB wavefunction, may be evaluated by two distinct methods. Firstly, H_{ij} may be decomposed into a sum of matrix elements $H^{ij}_{k\ell}$,

$$\text{i.e.} \quad H_{ij} = \sum_{k\ell}^{M} d_{ik} d_{j\ell} \int \Psi^i_k \hat{H} \Psi^j_\ell \, d\tau$$

$$= \sum_{k\ell}^{M} d_{ik} d_{j} \; H^{ij}_{k\ell}$$

where Ψ^i_k is a single determinant. By this method, the permutation problem - the "N! catastrophe" - is avoided, since $H^{ij}_{k\ell}$ is completely and simply defined in terms of the inverse of the overlap matrix of the orbitals which make up the determinants and the energy integrals.[4] Thus, the $(N!)^2$ permutation terms need not be considered explicitly. On the other hand, the number of terms $H^{ij}_{k\ell}$ per matrix element H_{ij} increases very rapidly with the number of Heitler-London "bonds", N_b. M is of the order 2^{2N_b}, so that the method soon becomes impractical, for example, ethane with seven bonds gives rise to about 16,000 $H^{ij}_{k\ell}$ terms in the "perfect-pairing" structure alone. The alternative method, due to Pauling and recently studied in detail by Shull, deals with the "structure" matrix element H_{ij} as a whole:[5] then, however, all permutations must be considered explicitly, and the problem is the same as that encountered in the NPF method.

An advantage of the NPF formalism, compared to the VB method, is that instead of having to compute each matrix element, H_{ij}, with its $(N!)^2$ permutation terms, separately, the contributions from each matrix element are computed simultaneously by the use of matrix algebra. A more important advantage is however associated with the NPF method: complete ab initio calculations, using non-orthogonal orbital based wavefunctions, especially iterative variational calculations, will not be possible for many-electron systems, but approximate methods may be available which will alleviate the computational problem without sacrificing the potential accuracy of such wavefunctions. The vague formulation of VB theory means that it is not well suited to such developments: the NPF method, on the other hand, is amenable to this type of development. An approach along these lines is dealt with in detail in Section 3.

For few-electron systems, the NPF method is computationally tractable without the invokation of any approximation. To investigate the potential of the NPF wavefunction, three widely different four-electrons systems have been studied - the lithium hydride molecule, the interaction of two helium atoms and the pi electrons of butadiene. The two requirements for the construction of an accurate wavefunction are a flexible one-electron orbital basis, capable of describing properly all relevant regions of space, and a flexible type of wavefunction which approaches closely the exact' wavefunction as the basis is improved. In an attempt, in the calculations on lithium hydride and the helium atom interaction, to steer a middle course between the twin demands, the combination of simple NPF wavefunctions with small extended orbital bases has been taken. The calculations are not aimed at simply obtaining a very accurate energy but rather at finding whether the method is capable of giving a satisfactory compromise between accuracy, simplicity and computational economy. The energy expression and variational equations for N=2 are readily derived and are given in detail in Appendix 2. Fully automatic programs have been written in Egdon Algol for the University of Glasgow KDF9 computer to perform the variational calculations and the evaluation of the density functions for any four-electron system with an arbitrary number of basis orbitals in each pair function.

The testing ground for novel quantum mechanical approximation methods is nowadays the lithium hydride molecule which combines the advantages of being small enough to be studied extensively, yet being large enough for valid extrapolations to larger systems to be made - a property now recognised not to be possessed by the hydrogen molecule - and for the investigation of factors met in more chemically interesting molecules i.e. a tightly bound core, a heteronuclear two-electron bond, with inter- and intra-pair correlation effects.

There are three ways open of choosing a flexible orbital basis - an n-extended basis, that is, one with orbitals with higher quantum numbers than those of the valence orbitals, an exponent-optimised valence orbital basis or an extended valence orbital basis, with more than of each of the valence orbitals. To avoid the neccessity of carrying out time consuming non-linear parameter optimisation, the third alternative was adopted, with the exponents chosen on a chemically intuitive basis. The seven orbital basis, first used by Matsen and coworkers[6], is made up of the following Slater type orbitals: for lithium, the optimum split-shell core orbitals, 1s and 1s', for Li^+, with exponents 3.30 and 2.065, and valence 2s and 2p orbitals, with a common exponent fixed by Slater's rules at 0.65; for hydrogen, the free atom 1s orbital, h, exponent 1.0, and, since the bond is highly polarised towards the proton, the optimum split-shell hydride ion 1s orbitals, h' and h", with exponents 1.38 and 0.62.

Four different partitions of the orbital basis into pair function sets were made. In the first three, the valence and core orbitals are separated into different pair functions; in the last, the core pair function is augmented by the lithium 2p orbital to allow for core polarisation. The results for the various partitions are given in Tables 1 and 2. Since the Hartree-Fock energy is not available for this basis, the exact amount of correlation energy gained by the NPF wavefunction

Table 1   Lithium Hydride R = 3.01 au.

	Calculation	Energy (au)
1*	NPF(1s,1s')(2s,2p,h)	-8.0070
2*	NPF(1s,1s')(2s,2p,h',h")	-8.0010
3*	NPF(1s,1s')(2s,2p,h,h',h")	-8.0109
4*	NPF(1s,1s',2p)(2s,2p,h)	-8.0071
5†	CI(4,13)	-7.9836
6†	VB(4,5)	-7.9845
7†	MCSCF(16,17)	-8.0177
8†	CI(7,41)	-8.019

* This work: the brackets indicate the partition of the orbitals into pair functions

† Comparative calculations: references

5. Fraga, S. & Ransil, B.J.: (1962) J.Chem.Phys. 36, 1127
6. Murrell, J.N. & Silk, C.L.: (1968) Symp.Farad.Soc. 2, 84
7. Mukherjee, N.G.: (1969) Ph.D. Thesis, University of Sheffield
8. Ebbing, D.: (1962) J.Chem.Phys. 36, 1361

The numbers in brackets are the size of the basis and the number of configurations.

cannot be calculated. However, the Hartree-Fock energy may be estimated to be -7.98 a.u. (For example, Ransil's optimised minimal basis gives -7.970 a.u. whereas Ruedenberg, and McWeeny and Mukherjee, with large extended bases, obtained energies of -7.987 a.u. and -7.979 a.u. respectively)[7]. About 0.03 a.u. of the correlation energy must therefore have been obtained compared to the exact value of 0.08 a.u. Although this amounts to less than 40% of the total correlation energy, it compares well with the results of many more complex wavefunctions - see Table 1. Only sigma-type orbitals are included in the basis: Bender and Davidson[8], from their extensive natural orbital CI calculations, have estimated that the minimum energy obtainable with sigma-type orbitals alone is -8.038 a.u. The NPF wavefunction thus accounts for about 60% of the sigma-type correlation energy, a sizeable proportion for a fairly simple basis and wavefunction. The introduction of core polarisation, by allowing the 2p orbital to contribute both to the bond and the core, has little effect on the electronic energy. So far only the energy has been computed: it would be of interest to compute the values of some molecular properties to obtain a more balanced picture of the NPF wavefunction.

Information regarding the nature of the bonding in the molecule is contained in the pair function coefficients, which are given in Table 2. Although the electron density and pair density distributions are the more fundamental quantities determining "what the electrons are doing", the coefficients of simple wavefunctions such as these can yield some insight into the nature of the two-electron bond. From the results of Table 2, we may firstly conclude that the core of the lithium atom is adequately described by a pure split-shell wavefunction, 1s1s', the weights of the $1s^2$ and $1s'^2$ configurations being negligible. Thus one electron of the core may be considered to be moving close to the nucleus and the other at a larger radius. The electron-pair bond is a much more complex entity: the normal conception of the bond in lithium hydride is of a bond strongly polarised in the sense Li^+H^-, viz. the large dipole moment of 5.9D. It might therefore be expected that Li^-H^+ type configurations would be of negligible importance and that a hydride ion split-shell description of the electrons in the region of the proton would give a good representation of the molecular electron distribution. It is apparent however that the weights of the Li^-H^+ configurations are by no means negligible.

Table 2 Pair Function Coefficients for LiH

Calculation*	Core pair function[#]		
	1s	1s'	2p
1, 2, 3	-0.007	1.1	
		0.009	
4	-0.01	1.1	0.01
		0.002	-0.01
			-0.003

Calculation*	Valence pair function				
	2s	2p	h	h'	h"
1, 4	-0.23	-0.14	1.49		
		-0.10	0.89		
			0.75		
2	-0.13	-0.02		0.49	0.26
		0.01		0.34	0.14
				0.06	0.09
					0.28
3	-0.37	-0.24	1.21	-0.05	0.84
		-0.16	0.49	0.06	0.66
			-0.16	0.75	1.34
				-0.93	-0.01
					-3.10

* See Table 1 for details of calculations

[#] The coefficients are tabulated in upper diagonal matrix form: for example,

$$\begin{array}{cc} 1s & 1s' \\ -0.007 & 1.1 \\ & .009 \end{array} \equiv -0.007(1s^2) + 1.1(1s1s'+1s'1s) + 0.009(1s'^2)$$

And, secondly, the split-shell, h'h", configuration is by no means adequate: for example, the calculation which omits the free atom hydrogen 1s orbital yields a markedly poorer energy than that obtained when it is included. Also, in the wavefunction which omits the orbital, the h''^2 configuration is dominant in the bond pair function, and, when it is included, the order of importance of Li^+H^- structures is h''^2, hh", h'^2, hh', with the expected dominant term h'h" making a negligible contribution. The more important terms appear to be the expanded orbital ones. Too much must not be made of the precise values of the coefficients, but they do indicate that the electron-pair bond is far more complex than the simple picture of an electrostatic bond between a hydride ion and a lithium ion would suggest.

The short to medium range interaction of two helium atoms is a system on which the relative merits of the NPF and SBP methods may profitably be tested: one expects intuitively that a wavefunction which allows the wavefunctions of the two atoms to overlap and interpenetrate will give a better representation of the interaction of the atoms than one in which they are forced to be exclusive at all internuclear separations. To examine this supposition, pair function wavefunctions have been calculated for He_2 for interatomic distances (R) from 0.5 to 5 a.u. A simple basis, the optimum split-shell orbitals, 1s and 1s', for the free helium atom, with exponents 1.19 and 2.18365 - the Eckart orbitals[9], was chosen, as in the previous calculations, as a compromise between accuracy and economy. The Eckart wavefunction for the free atom gives an energy of -2.8756 a.u., 0.0276 a.u. above the exact energy and 0.0281 a.u. better than a single 1s orbital wavefunction. Four series of calculations were performed with different pair function wavefunctions. The first is the simple NPF wavefunction in which each atom is described by a spherically symmetric, one-centre pair function; that is, for atom A,

$$\Delta^A(12) = \{C^A_{11} 1s_A(1) 1s_A(2) + C^A_{12}(1s_A(1) 1s'_A(2) + 1s'_A(1) 1s_A(2))$$

$$+ \; C^A_{22} 1s'_A(1) 1s'_A(2)\} \; (\alpha(1)\beta(2) - \beta(1)\alpha(2)),$$

and similarly for atom B. The Eckart wavefunction for He_2 is obtained by setting the coefficients of the diagonal terms, e.g. C^A_{11} and C^A_{22}, to zero in the above expansions. The symmetrically orthogonalised set of basis functions was used to construct the SBP wavefunction, the orthogonal counterpart of the simple NPF wavefunction. The final wavefunction is designed to take into account charge-transfer between the atoms: to the simple NPF pair function basis is added the more diffuse orbital centred on the other atom, e.g. Δ^A is constructed from $1s_A$, $1s'_A$ and $1s_B$. This is then a five parameter wavefunction, compared to two for the simple pair function wavefunctions and none for the Eckart wavefunction. The results are given in Table (3) and Figures (1) and (2). The following conclusions may be drawn. The split-shell or Eckart description of the helium atoms is adequate, compared to the simple NPF description, except at small R: that is, the weights of the configurations $1s_A^2$ and $1s'^2_A$ in the pair function Δ^A are negligible, and the energies given by the two wavefunctions are virtually identical. However, at R = 0.593 a.u., the weights increase dramatically, from less than 10% to about 75% of the split-shell configuration weight, with a correspondingly dramatic decrease in the energy of the simple NPF wavefunction of about 1.1 a.u. The introduction of the charge-transfer terms makes a pronounced improvement in the electronic energy and, more importantly, the interaction energy, at all R: the weights of the charge-transfer terms are important at all

Table 3 Helium-Helium Interaction - V(R)[+] (a.u.)

	Wavefunction[#]	Eckart	NPF	SBP	CT-NPF	Beck**	MBD
N*							
1	*	4.25	3.18	3.67	2.57	2.13[++]	2.31
2		0.998	0.997	0.755	0.689	0.631	0.614
3		0.258	0.258	0.205	0.207	0.182	0.163
4		0.0692	0.0693	0.0675	0.0592	0.0480	0.0433
5		0.0185	0.0185	0.0243	0.0162	0.0118	0.0115
6		0.00487	0.00488	0.00873	0.00422	0.00271	0.00304[++]
7		0.00125	0.00125	0.00304	0.00102	0.00053	0.00081[++]
8		0.00031	0.00031	0.00123	0.00022	0.00005	0.00021[++]

\+ $V(R) = E^{(R)}_{\text{electronic}} - 2E_{\text{He}} + 4R^{-1}$

$E_{\text{He}} = -2.87566$ a.u. for the Eckart wavefunction,

$E_{\text{He}} = -2.87572$ a.u. for the pair function wavefunctions

* The internuclear separation is R = 0.593N a.u.

\# See text for details: CT - charge transfer

** An empirical potential fitted to experimental and theoretical results

Beck, D.E. (1968) Mol.Phys. 14, 311

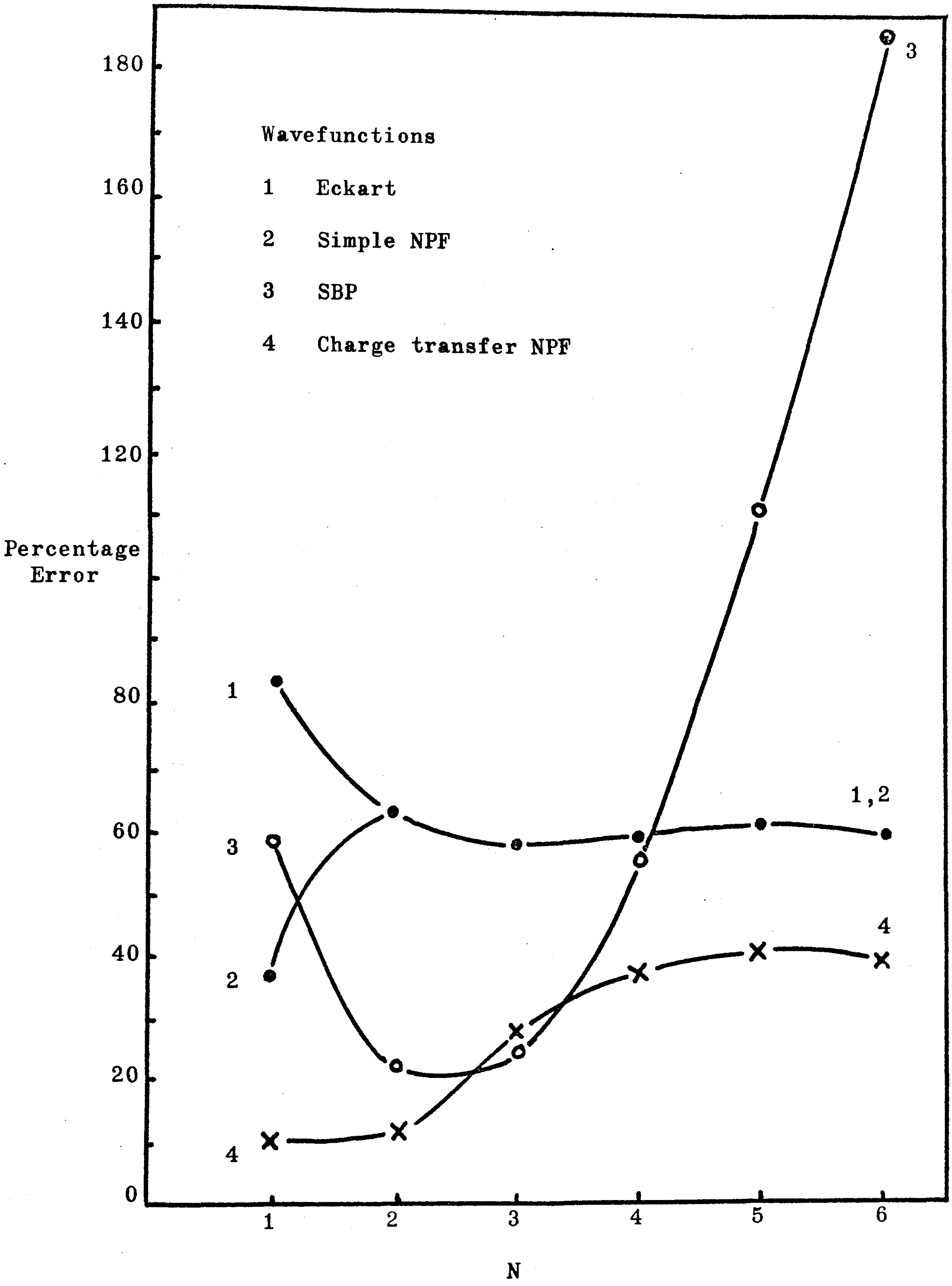
++ Outside the estimated range of applicability of the empirical potentials internuclear separations but are particularly large at small R. The addition of explicit polarisation orbitals, e.g. 2p, to the pair function bases would probably reduce the importance of the charge transfer configurations, as well as giving a better representation of the interaction energy at large separations.

Figure (2) shows a plot, as a function of R, of the percentage error in the interaction energies calculated with the different wavefunctions compared to the "exact" energy calculated from the potential function,

$$V(R) = 237\ e^{-4.23R}\ \text{eV (R in Ångstroms)},$$

as estimated by Matsumoto, Bender and Davidson in very accurate calculations.[10] The most interesting point that this illustrates is that, whilst the non-orthogonal orbital based wavefunctions give interaction energies which are in reasonable accord with Matsumoto, Bender and Davidson's results at all R, the accuracy of the energies calculated with the SBP wavefunction varies dramatically with internuclear separation. Thus, at small R, the interaction energy is poor; at large R, it is the wrong order of magnitude; and, at intermediate R, (2 to 3 a.u.), it is as accurate as that of the five-parameter charge-transfer NPF wavefunction. This

Figure 2. He-He Interaction : Accuracy of Calculated V(R)



Percentage Error = $100(V(R)_{calcd.} - V(R)_{MBD})/V(R)_{MBD}$

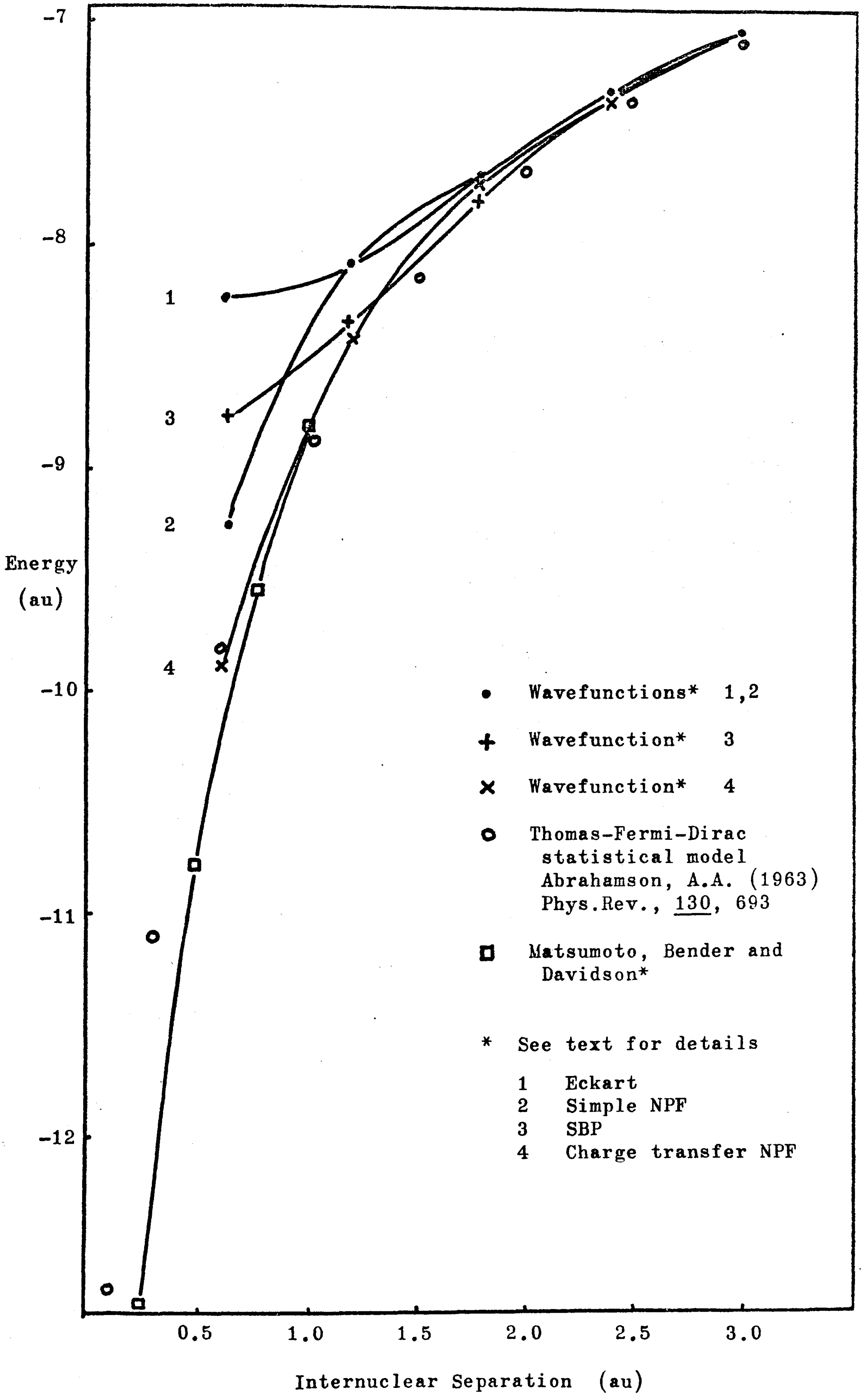
See text and table 3 for details of wavefunctions 1-4 and V(R)

Internuclear separation = 0.593N au.

means that the shape of the interaction potential is badly represented by the SBP wavefunction. The same conclusion has also been drawn from calcualtions on the interaction of two hydrogen molecules, where the barrier to the rotation of one molecule relative to the other was found to be much more accurately described by an NPF wavefunction than a SBP wavefunction.[11] In Figure (1), the electronic energy of the different wavefunctions is plotted as a function of R: it is evident that the electronic energies of the Eckart and SBP wavefunctions are tending to the wrong asymptotic limit as R goes to zero. In the case of the SBP wavefunction, this must be due to the orthogonality constraints which prevent the wavefunction going over to an approximation to the united atom, beryllium, wavefunction: it may not then be surprising that the interaction potential is so poorly represented by the SBP wavefunction.

Thus it must be concluded that the strong orthogonality constraint is not compatible with the correct description of the interaction between closed shell atoms and molecules. What the situation is regarding intra-molecular interactions is not clear: the SBP method appears to predict reasonably accurate molecular geometries as far as bonds to a central atom are

Figure 1. Helium-Helium Interaction : Electronic Energy



concerned - for water, a computed bond angle of about 100° is found compared to the experimental value of 105°.[12] Although the basic characteristics, for example, the positions of maxima and minima, of intramolecular potentials may be fairly accurately determined by SBP wavefunctions, it is likely that the details of the potentials will be better represented by NPF wavefunctions.

The pi-electron system of butadiene is of particular interest in the context of orthogonal versus non-orthogonal pair function wavefunctions since here we have a situation in which the localisation of the electrons pairs, in the terminal bonds, is generally considered to be weak. The question is how well does the non-orthogonality of completely localised pair functions account for the partial delocalisation of electrons into the central bond? The initial problem, in using a non-orthogonal basis in a system of this type where a core-valence separation is assumed, is the calculation of the energy integrals over the non-orthogonal orbitals. A basis of four pi-type orbitals centred on the carbon atoms is assumed: however the form these orbitals should take and the corrections to the integrals over these orbitals due to the presence of the core are by no means apparent. Two different approaches have been made to the problem. Either the orbitals are left entirely unspecified and the integrals found by fitting calculated results to experimental ones for test molecules, or the forms of the orbitals are guessed and the integrals evaluated theoretically under simplifying assumptions regarding the effects of the core. The former is satisfactory when the unspecified basis is assumed to be orthogonal, since it is known that the majority of the two-electron integrals then assume negligible values. When the basis is not orthogonal, the method is totally impractical and it is usual to resort to the latter approach. Thus, Parr and Mulliken, in calculations on butadiene, have used a basis of Slater-type orbitals with Slater's rules exponents, calculating some integrals theoretically, others by Mulliken's approximation.[13] Since the "best" empirical integrals, over the assumed orthogonal basis are far removed from the theoretical ones, over a strictly orthogonal basis, this method of calculating pi-electron integrals must be considered suspect. An alternative approach to the problem is to take the 'best' unspecified and assumed orthogonal basis, χ, transform it to the non-orthogonal basis ϕ,

$$\phi_i = \sum_{ij} S^{\frac{1}{2}}_{ij} \chi_j \qquad \text{where} \qquad S_{ij} = \int \phi_i \phi_j \, d\tau$$

and thence calculate the integrals over the non-orthogonal basis. For example, the one-electron integrals

$$f_{ij} = \int \phi_i(1)\ f(1)\ \phi_j(1)\ d\tau_1$$

$$= \sum_{k\ell} S^{\frac{1}{2}}_{ik}\ F_{k\ell}\ S^{\frac{1}{2}}_{j\ell}$$

where the $F_{k\ell}$ are the 'best' empirical integrals and are thus known. Of course, if one does not know the orbitals with which to calculate the integrals in the first place, one cannot either set up the overlap matrix, $\underline{S}$, to perform the transformation.* Viewed from a different standpoint, this difficulty may be avoided. Any basis transformation,

$$\phi'_i = \sum_{ij} U_{ij}\ \chi_j$$

is valid, if carried out properly, in the sense that the complete basis-set limit calculation, performed with any combination of a given set of orbitals, will yield the same wavefunction and energy. The objective may therefore be redefined as, not the evaluation of the integrals over a non-orthogonal basis, but as the finding of the basis $\underline{\phi}'$ which yields the best energy for a given type of wavefunction. Thus, the actual basis orbitals, $\underline{\chi}$, need never be known;

only the transformation matrix $\underline{U}$ linking bases need be specified. In the following the transformation $\underline{U} = \underline{S}^{\frac{1}{2}}$ is used, where the overlap integrals are obtained from STO's with exponent 1.625; the untransformed one and two electron energy integrals being calculated by standard semi-empirical methods. It must be emphasised that the resulting integrals are not to be regarded as integrals over a non-orthogonal basis. Our integrals are compared with those of Parr and Mulliken in Table 4. The correspondence between them is in fact quite remarkable: the reduction of all of Parr and Mulliken's integrals by a factor of 1.54 gives a set of integrals almost identical to those calculated by the transformation method.

Six different wavefunctions have been studied. These may be divided into three classes, single determinant, SBP and NPF wavefunctions. The two single determinant wavefunctions considered are the bond orbital (BO) wavefunction,

$$\hat{A}\, \phi_A\, \bar{\phi}_A\, \phi_B\, \bar{\phi}_B$$

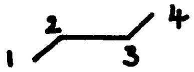
where

$$\phi_A = \chi_1 + \chi_2 \,, \quad \phi_B = \chi_3 + \chi_4$$

Table 4 Comparison of Semi-Empirical Integrals

Integral#	Parr and Mulliken*	Ours	Adjusted+
Overlap			
(12)	0.279	0.244	
(13)	0.0337	0.0339	
(14)	0.0021	0.0019	
(23)	0.233	0.244	
One Electron			
(11)	-30.87	-24.28	-24.34
(12)	-12.45	-10.04	-9.99
(13)	-1.68	-1.80	-1.38
(14)	-0.13	-0.19	-0.11
(22)	-36.41	-27.87	-28.18
(23)	-10.58	-10.04	-8.39
Two Electron++			
(1111)	16.93	10.97	10.99
(1112)	3.61	2.01	2.34
(1122)	9.24	5.50	6.00
(1212)	1.09	0.49	0.71
(1133)	5.52	3.90	3.58
(1113)	0.30	0.22	0.19
(1144)	3.70	2.91	2.40
(1234)	0.46	0.24	0.30
(1324)	0.009	0.007	0.006

\# Numbering convention



* Parr, R.G. and Mulliken, R.S. (1950) J.Chem.Phys. 18, 1338

\+ Parr and Mulliken's integrals with all two-electron integrals reduced by a factor of 1.54

++ Charge-cloud notation is used for the two-electron integrals and χ_i are the "non-orthogonal" atomic orbitals, and the MO wavefunction. The NPF wavefunctions are the simple NPF wavefunction $\hat{A}\,\Delta_A\,\Delta_B$ where Δ_A is built from χ_1 and χ_2, and Δ_B from χ_3, χ_4, and the 'delocalised' NPF wavefunction in which Δ^A is constructed from orbitals χ_1, χ_2, χ_3 and Δ^B from χ_2, χ_3, χ_4. Thus each pair function is delocalised over three atoms. The last two wavefunctions are of the SBP type: firstly, the simple SBP wavefunction $\hat{A}\,\hat{\Delta}_A\,\hat{\Delta}_B$ where $\hat{\Delta}_A$ and $\hat{\Delta}_B$ are constructed as for the simple NPF wavefunction but from orthogonal atomic orbitals, and secondly, the SBP wavefunction in which the pair functions are constructed from localised molecular orbitals (LMO): in this way partial delocalisation of the bonds is allowed. The results are given in Table 5. It is immediately apparent that the NPF wavefunctions perform better than their SBP counterparts. Particularly encouraging is the completely localised NPF wavefunction, which, although it only accounts for a small percentage of the correlation energy, performs very satisfactorily for a completely localised wavefunction in a situation where considerable delocalisation of the electrons out of the terminal bonds is expected. In fact, this is not the best energy that may be obtained with this

Table 5 Butadiene - Results

	Wavefunction*	Number of Free Parameters	Energy (eV)
1	BO	0	-82.287
2	SBP	2	-82.531
3	MO	1	-83.182
4	NPF	2	-83.297
5	SBP-LMO	3	-83.962
6	"Delocalised" NPF	5	-84.084

* See text for details

wavefunction: with the same basis transformation matrix, $\underline{S}^{\frac{1}{2}}$, but using an orbital exponent of 0.9 in the calculation of the elements of $\underline{S}$ an energy of -83.890 eV may be obtained. Since the SBP-localised MO wavefunction is known to be capable of giving 90% of the correlation energy, the partially delocalised NPF wavefunction must be, to all intents and purposes, the exact wavefunction within the basis. This is further encouraging since the basis-set limit wavefunction is a 20 configuration function, and the delocalised NPF wavefunction contains only 5 free parameters. The question, can a localised non-orthogonal orbital based wavefunction describe accurately the electron density in a weakly localised system, may be answered in the affirmative. Since such systems are not well described by localised SBP wavefunction, the NPF wavefunction should provide a useful tool in such situations.

Some general conclusions regarding the practical aspects of the ab initio NPF method may be drawn from these calculations. But, first, it should be said that although these calculations do indicate that the NPF wavefunction is capable of giving fairly accurate wavefunctions, they have been concentrated on the energy alone: a more balanced assessment would require the study of molecular properties as well. In connection with the calculations on lithium hydride, most of the correlation energy obtained must come from core correlation which, it is assumed, does not contribute much to the improvement of other molecular properties. A correlated core is however necessary to obtain any core-valence correlation which may be more important to molecular properties.

It has been found that the computing time for the variational process depends strongly on the number of orbitals in each pair function. For example, if each pair function is constructed from two orbitals, the total computation time is less than thirty seconds, but, in the lithium hydride calculations, with two core orbitals and five valence orbitals, each cycle takes about 2.5 minutes and, for four orbitals in each pair function, a computing time per cycle of more than five minutes is found. This rules out large basis calculations by the NPF method. As the number of pair functions increases, the main problem will not necessarily be the time factor, if small bases are used, but the programming of the variational equations. The explicit expressions are fairly complex for $N = 2$: for N greater than four, they will certainly be prohibitively complicated. The final point is the convergence of the iterative variational process: for well separated pairs, for example, those in lithium hydride and He_2 at long range, no difficulty has been encountered, but, when the interpair overlap is large, the convergence is often slow and very dependent on the trial coefficients. For example, at an internuclear separation of 0.593 a.u. in the He_2 calculations with the charge transfer NPF wavefunction, some thirty cycles were needed to reach convergence. Also, although the energy had converged to eight decimal places, the coefficients, in some cases, particularly the charge transfer wavefunctions, had still not settled down. On the other hand, the SBP wavefunctions for He_2 converged rapidly for all internuclear separations. These practical problems, although rather discouraging, should not be insurmountable, but do provide all the more reason for looking at approximation methods which are the subject of the next section. Overall, one might estimate that the practical upper limit of the ab initio NPF wavefunction is four pair functions with a small accurate basis, say three orbitals. If, for example, the 1s core of methane is removed from the problem by orthogonalisation of the valence orbitals to the fixed, possibly correlated, core, the scope of the NPF method may to a certain extent be enlarged.

REFERENCES

1. Lowdin, P.-O. "Advances in Chemical Physics" Volume 2, (1959) p. 307, Wiley - Interscience.
2. (a) Dacre, P.D. & McWeeny, R. (1970) Proc.Roy.Soc. London A317, 435
   (b) McWeeny, R. & Sutcliffe, B.T. (1963) Proc.Roy.Soc. London A273, 103
3. Silver, D.M. (1969) J.Chem.Phys. 50, 5108
4. Prosser, F. & Hagstrom, S. (1968) Int.J.Quant.Chem. 2, 89
5. Shull, H. (1969) Int.J.Quant.Chem. 3, 523
6. Robinson, J.M., Stuart, J.D. & Matsen, F.A. (1960) J.Chem.Phys. 32, 988
7. (a) Ransil, B.J. (1960) Rev.Mod.Phys. 32, 239
   (b) Miller, K.J. & Ruedenberg, K. (1968) 48, 3414
   (c) Mukherjee, N.G. & McWeeny, R. (1970) Int.J.Quant.Chem. 4, 97
8. Bender, C.F. & Davidson, E.R. (1966) J.Phys.Chem. 70, 2675
9. Linnett, J.W. (1960) "Wave Mechanics and Valencey" p. 68 Methuen
10. Matsumoto, G.H., Bender, C.F. & Davidson, E.R. (1967) J.Chem.Phys. 46, 402

11. Magnasco, V., Musso, G.F. & McWeeny, R. (1967) J.Chem.Phys. 47, 4617

12. Klessinger, M. (1969) Chem.Phys.Letters 4, 144

13. Parr, R.G. & Mulliken, R.S. (1950) J.Chem.Phys. 18, 1338

SECTION 3

APPROXIMATE SOLUTIONS OF THE NON-ORTHOGONALITY PROBLEM

The conclusions of the previous section, that the NPF method, whilst yielding accurate wavefunctions from small well-chosen bases, rapidly becomes intractable through the increasing complexity of the matrix elements required in the variational procedure and the time-consuming problem of computing the $(N!)^2$ terms arising from the anti-symmetry principle, indicates that the search for justifiable approximation methods will be necessary if the NPF method is to be applicable to molecules with a larger number of electrons. It is unfortunately the case that, in many calculations of molecular wavefunctions and properties, a large number of terms are computed which have a negligible effect on expectation values and on the path of the variational determination of optimum parameters. Because of the difficulty in general of specifying a priori which terms these are and of defining the effect of omission of terms from the computation, little work appears to have been directed towards the goal of approximate solutions in which such terms are neglected. An example where some headway has been made in this direction is the 'neglect-of-differential-overlap' approximation in MO theory, in which many integrals, over an orthogonal hybrid basis, involving the charge distribution of two orbitals on different centres, may, to good accuracy, be omitted from the computation. The NPF method should be particularly amenable to this type of treatment since the practical problems arise directly from the sheer number of terms that must be calculated. The search is therefore for an expansion of the density functions which possesses a well defined partition into a small number of dominant terms and a large number that are practically negligible. The approach that we are pursuing is approximate in the sense that not all terms in a given expression are computed, but is ab initio in the sense that it is not a semi-empirical approach where difficult-to-compute quantities are estimated by appeal to experiment or where quantities are similarly "adjusted" to correct for the inadequacies of the approximations. The task is thus more severe than in a semi-empirical scheme since the neglect of any term must be rigorously justified. However, as the approximations are relaxed, the approximate solution must converge in the limit to the true one, a property not possessed by semi-empirical theories.

The degree of accuracy demanded of an approximation scheme is dependent on several considerations. The computation of the basic integrals is generally achieved through numerical or analytical infinite series methods: a residual error must be present in these quantities. Also the actual process of evaluating optimum parameters and the values of observables is subject to computational round-off errors. It is unnecessary to demand an accuracy greater than can be obtained in practice in the exact calculation. More importantly, there is the fact that the wavefunctions used are only approximations to the true wavefunctions: they are eigenfunctions of an approximate (and not always defined) Hamiltonian. There is no virtue in asking for an exact solution to an approximate problem. The only requirements are that the errors induced by the approximate method of solution are smaller than those intrinsic to the approximate wavefunction, and that it should be possible to make some estimate of the partition of the errors between those inherent in the wavefunction and those in the method of solution. In the short run, it is a simpler task, where possible, to compute exact approximate solutions since this obviates the neccessity of determining which terms may be omitted and the errors induced in expectation values by their omission. In the long run, however, if reasonably accurate solutions are to be found for large molecules, and if computations on the small chemically important groups within large molecules are to be feasible, this approach must be examined in detail. We shall, in this section, devote our attention to two approaches conforming to this philosophy.

Valence bond wavefunctions have not had the attention they deserve due to the computational problems introduced by the antisymmetriser. The conventional method of avoiding the "non-orthogonality catastrophe" was to invoke the "neglect-of-overlap" approximation. This was justified by the observation that overlap integrals are necessarily less than unity so that powers of overlap integrals, brought in by the antisymmetriser, tend to zero as the powers and the distances between centres increase. To arrive at expressions of sufficient simplicity to be of practical use, it was necessary at that time to neglect <u>all</u> powers of all overlap integrals. Unfortunately, the wholesale omission of such terms cannot be justified: the resulting errors are too large. Thus it is found that there is no binding in the hydrogen molecule[1], and that exchange integrals, which are indisputably positive, have in this approximation negative values[2]. Approximate VB theory was therefore forced into a semi-empirical formalism, with, for example, exchange integrals becoming disposable parameters[3].

But the basic justification remains: it is possible to expand the expression for any observable, X, derived from a wavefunction based on non-orthogonal orbitals, in terms of powers of overlap integrals. Thus

$$X = N^{-1}(X_0 + \sum_{i<j} X_{ij}S_{ij} + \sum_{i<j} \sum_{k<\ell} X_{ijk\ell}S_{ij}S_{k\ell} + \ldots)$$

where X_{ij}, $X_{ijk\ell}$,... are constants, S_{ij} is the overlap integral $\int \phi_i(1)\phi_j(1)d\tau_1$ and the normalisation constant $N = N_0 + \sum_{i<j} \sum_{k<\ell} N_{ijk\ell} S_{ij} S_{k\ell} + \ldots$. The expansions consist of a finite number of terms, at most n terms, where n is the number of electrons. There is of course no guarantee that the series converge; certainly, the powers-of-overlap, e.g. $S_{ij} S_{k\ell} S_{mn}$, must become smaller as the number of integrals in the product increases, but no such condition binds the coefficients, X_{ij}, etc., or the summed terms. In the following, the expansion will be made in terms of inter-pair overlap only, no intra-pair overlap being neglected, contrary to the usual approximate VB theories. The convergence of the series is expected to be considerably improved over a general overlap integral expansion since interpair overlaps are generally much smaller than those within a pair. The inclusion of intra-pair overlap is quite arbitrary, since we may, without loss of generality, orthogonalise the orbitals within each pair: the orthogonalisation unfortunately introduces no simplification of the problem, so that it is consistent to retain intra-pair non-orthogonality throughout.

The problem, that has hitherto been avoided, is the development of a more realistic approach to the "neglect-of-powers-of-overlap" approximation, to determine in what situations the series is sufficiently convergent to justify a truncation which will appreciably reduce the computational difficulties, to find a systematic a priori truncation scheme and to assess the errors resulting from the truncation. The first requirement is the development of the density function expansion in a form which is sufficiently simple to allow the convergence properties of the overlap series to be studied in detail. Explicit expressions for the coefficients of the overlap integral products, for example, the X_{ij} in equation (3.1), are not readily derived. However, it is possible to arrive at expressions for the NPF wavefunction density functions which allow the convergence of the series to by systematically studied.

Let us first consider the one particle density function which is, from equation (2.3),

$$P(1,1') = 2 \sum_{I=1}^{N} P^{N} Q^{N} \, \mathrm{TR}[(\mathrm{CSCT}(P_I,I,Q_I) . \prod_{J \neq I}^{N} \mathrm{CSCS}(P_J,J,Q_J)] \qquad (3.2)$$

where the notation is that of Appendix 1. The summation subscripts of equation (2.3), p_I and q_I, have been suppressed since they may be inferred from the superscripts (cf. the expansion of the normalisation integral in Section 2): it should always be remembered that each term in (3.2) e.g. $CSCS(P_J, J, Q_J)$ is a matrix. The density function is not in a particularly convenient form as the orbital product $T_{ij}^{IJ} = \chi_i^I(1)\chi_j^J(1')$ occurs with an index determined by the permutation operator: a more useful expression may be obtained by expanding the permutation operators.

$$\begin{aligned} P^N &= P_I^{N-1}(1 - \sum_{K \neq I} P_{IK}) \\ Q^N &= Q_I^{N-1}(1 - \sum_{L \neq I} Q_{IL}) \end{aligned} \qquad (3.3)$$

where P_{IK} interchanges the left-hand superscripts (LSS) I and K (with their corresponding subscript summation indices), and P_I^{N-1} is the (N-1)th order permutation operator which permutes all left super- and subscripts except those in the I'th term. Q_{IL} and Q_I^{N-1} operate similarly for right-hand superscripts (RSS). Introducing the expansions (3.3) into equation (3.2)

$$P(1,1') = 2 \sum_{I=1}^{N} TR[P_I^{N-1} Q_I^{N-1}(1 - \sum_{K \neq I} P_{IK})(1 - \sum_{L \neq I} Q_{IL})$$

$$\mathrm{CSCT}(P_I,I,Q_I) \prod_{J\neq I}^{N} \mathrm{CSCS}(P_J,J,Q_J)$$

Expanding P_{IK} as $P_{I\to K}\ P_{K\to I}$ where $P_{I\to K}$ replaces the LSS I by K, and similarly for Q_{IL}, and operating with $P_{I\to K}$, $Q_{I\to L}$ we find

$$\begin{aligned} P(1,1') = 2 \sum_{I=1}^{N} \mathrm{TR}[&\mathrm{CSCT}(I,I,I)\ P_I^{N-1}\ Q_I^{N-1} \prod_{J\neq I}^{N} \mathrm{CSCS}(P_J,J,Q_J) \\ &- \sum_{K\neq I} \mathrm{CSCT}(K,I,I)\ P_I^{N-1}\ Q_I^{N-1}\ P_{K\to I} \prod_{J\neq I}^{N} \mathrm{CSCS}(P_J,J,Q_J) \\ &- \sum_{L\neq I} \mathrm{CSCT}(I,I,L)\ P_I^{N-1}\ Q_I^{N-1}\ Q_{L\to I} \prod_{J\neq I}^{N} \mathrm{CSCS}(P_J,J,Q_J) \\ &+ \sum_{K,L\neq I} \mathrm{CSCT}(K,I,L)\ P_I^{N-1}\ Q_I^{N-1}\ P_{K\to I}Q_{L\to I} \prod_{J\neq I}^{N} \mathrm{CSCS}(P_J,J,Q_J)] \end{aligned} \tag{3.4}$$

We may define the coefficient

$$D(K,L,I) = \sum_{T} P_I^{N-1} Q_I^{N-1}\ P_{K\to I}\ Q_{L\to I} \prod_{J\neq I}^{N} \mathrm{CSCS}(P_J,J,Q_J)$$

where T is the set of summation indices which contains all subscripts occuring twice in the coefficient expansion. If $K\neq L$ then D(K,L,I) is a matrix of dimension $n_L \times n_K$ where n_K is the number of orbitals in pair function K; otherwise it is a scalar. Thus equation (3.4) becomes

$$\begin{aligned} P(1,1') \quad = \quad 2 \sum_{I=1}^{N} \{&\mathrm{tr}(\mathrm{CSCT}(I,I,I)).D(I,I,I) \\ &- \sum_{K\neq I} \mathrm{tr}(\mathrm{CSCT}(K,I,I) \times D(K,I,I)) \end{aligned}$$

$$- \sum_{K \neq I} \mathrm{tr}(\mathrm{CSCT}(I,I,K) \times D(I,K,I))$$

$$+ \sum_{K \neq I} \mathrm{tr}(\mathrm{CSCT}(K,I,K) . D(K,K,I) \qquad (3.5)$$

$$+ \sum_{K,L \neq I} \sum{}' \mathrm{tr}(\mathrm{CSCT}(K,I,L) \times D(K,L,I))\} .$$

where tr stands for the normal matrix trace, and x for the matrix product. By separating out the orbital terms T^{IJ}_{ij}, we finally arrive at

$$P(1,1') = 2 \sum_{I=1}^{N} \mathrm{tr}(\underline{T^{II} D^{II}}) + 2 \sum_{I,J}^{N} \sum{}' \mathrm{tr}(\underline{T^{IJ} D^{IJ}}) .$$

where

$$D^{II}_{ij} = \mathrm{CSC}(I,I)_{ij} . D(I,I,I)$$

$$- \sum_{K \neq I} D(K,I,I) X \ \mathrm{CSC}(K,I)_{ij}$$

and

$$D^{IJ}_{ij} = - D(I,J,I) X \ \mathrm{CSC}(I,I)_{ij} + \mathrm{CSC}(J,I)_{ij} . D(J,J,I)$$

$$+ \sum_{K \neq I,J} D(K,J,I) X \ \mathrm{CSC}(K,I)_{ij} ,$$

and where we have used the notation

$$\mathrm{CSC}(I,J) = \underline{C^{I} S^{IJ} C^{J}}$$

We have achieved two advantages by this rearrangement of the density function expression. Firstly, the matrices $\underline{D}^{IJ}$ define the one particle density matrix: the expectation value of any one electron property is given by

$$X = \int P(1,1')\hat{X}(1)d\tau_{1\to 1'}$$

$$= \sum_I \sum_J \mathrm{tr}\,(\underline{X^{IJ}\ D^{IJ}})$$

Thus, once the $\underline{D}^{IJ}$ matrices are known, and they need only be computed once, the evaluation of the observable becomes trivial, requiring only the integrals for the property over the atomic orbital basis, $X^{IJ}_{ij} = \int \chi^I_i(1)\hat{X}(1)\chi^J_j(1)d\tau_1$. Secondly, the calculation of the density matrix is separated into two distinct processes: the evaluation of the coefficients D(K,L,I) and their combination with the CSC(K,L) matrices. The second step is essentially simple; the rate-determining step is the calculation of the coefficients. This involves an $((N-1)!)^2$ term sum of products of the matrices CSCS(K,L,M). Thus the full weight of the non-orthogonality problem is thrown into the computation of the coefficients. The essential simplicity of their definition makes them ideally suited to machine computation, unlike the unexpanded form of the density function, equation (3.1). The question of the rate of convergence of the powers-of-overlap series is now more or less a matter of the rate of convergence of the expansion of the coefficients, which, because of their simple nature, is a much more easily studied problem. We shall return to this in more detail later.

An analogous but considerably more tedious expansion is possible for the two particle density function. The permutation operator is expanded as

$$P^N + P_{IJ}^{N-2} (1 - \sum_{K \neq IJ} (P_{IK} + P_{JK}) + \sum_{K<L \neq IJ} P_{IK} P_{JL})(1 - P_{IJ})$$

where P_{IJ}^{N-2} is the (N-2)th order permutation operator working on all LSS except the ones in the I'th and J'th terms. This enables us to expand out the second and third terms of the two particle density function of equation (2.4): the first term is dealt with exactly as for the one particle density function. Taking, for example, the second term, it is readily seen that the expansions of P^N and Q^N will yield terms of the type

$$\mathrm{TR}[P_{IJ}^{N-2} Q_{IJ}^{N-2} P_{IK} P_{JL} Q_{IM} Q_{JN} \quad \mathrm{CSCT}(P_I, I, Q_I)$$

$$\mathrm{CSCT}(P_J, J, Q_J) \prod_{R \neq IJ}^{N} \mathrm{CSCS}(P_R, R, Q_R)$$

and hence by analogy with the one particle density function, the coefficient

$$D(K,L,M,N,I,J) = \sum_T P_{IJ}^{N-2} Q_{IJ}^{N-2} P_{K \to I} P_{L \to J} Q_{M \to I} Q_{N \to J} \prod_{R \neq IJ}^{N} \mathrm{CSCS}(P_R, R, Q_R)$$

where P^{N-2}_{KL} permutes all LSS, with their summation subscripts, in the product that follows, the K'th and L'th terms having I and J as their LSS in the identity permutation. T again denotes the set of summation indices that occur twice in the coefficient expansion. (Although the label N is used for the number of electron pairs and as a pair function name variable, no confusion should arise, the context making plain which is intended). An expression analogous to (3.5) for the one particle density matrix is straightforwardly derived: it is set out in detail in Appendix 3. The density matrix is separated, as before, into simple matrix products and coefficients involving the remaining permutations. Some properties which show the usefulness of the coefficients are set out below.

1. The normalisation integral

$$S_R = \mathrm{TR}[P^N Q^N \prod_{I=1}^{N} \mathrm{CSCS}(P_I, I, Q_I)]$$

is a special case of the general coefficient

$$D_N(A,B,C,D,E,F) = \sum_T P^{N-2}_{EF} Q^{N-2}_{EF} P_{A\to E} P_{B\to F} Q_{C\to E} Q_{D\to F} \prod_{G\neq EF}^{N} \mathrm{CSCS}(P_G, G, Q_G)$$

Thus

$$S_R = D_{N+2}(N+1, N+2, N+1, N+2, N+1, N+2)$$

where N+1 and N+2 are undefined dummy pair functions.

Similarly the coefficients defined for the one particle density function

$$D(K,L,I) \equiv D_{N+1}(K,N+1,L,N+1,I,N+1)$$

and finally

$$D(K,L,M,N,I,J) \equiv D_N(K,L,M,N,I,J)$$

Thus all one and two electron properties are completely, and relatively simply, defined in terms of the general coefficient $D_N(A,B,C,D,E,F)$.

2. $D_N(K,L,K,L,I,J)$ is a scalar; $D_N(K,L,M,L,I,J)$ is a matrix, A^{MK}_{ij}, of dimension $n_m \times n_k$; and $D_N(K,L,M,N,I,J)$ is a supermatrix - the difference of two matrix scalar products, that is, $D_N(K,L,M,N,I,J) \equiv B^{MK}_{ij}\,B^{NL}_{k\ell} - B^{ML}_{i\ell}\,B^{NK}_{kj}$ where, for example, $\underline{B}^{MK}$ is a matrix of dimension $n_m \times n_k$.

3. The following parameter interchange relations hold

a) $D_N(K,L,M,N,I,J) = -D_N(L,K,M,N,I,J)$

$= -D_N(K,L,N,M,I,J) = D_N(L,K,N,M,I,J)$

b) The coefficients $D_N(K,L,M,L,I,J)$ and $D_N(M,L,K,L,I,J)$ are related, in the simplified notation of 2, by

$$\underline{A^{MK}\;C^{K}}_{ij} = \underline{A^{KM}\;C^{M}}_{ji}$$

and $D_N(K,L,M,N,I,J)$ and $D_N(M,N,K,L,I,J)$ by

$$\underline{B^{MK}C^{K}}_{ij}\;\underline{B^{NL}C^{L}}_{k\ell}\;-\;\underline{B^{NK}C^{K}}_{kj}\;\underline{B^{ML}C^{L}}_{i\ell}$$

$$=\;\underline{B^{KM}C^{M}}_{ji}\;\underline{B^{LN}C^{N}}_{\ell k}\;-\;\underline{B^{KN}C^{N}}_{jk}\;\underline{B^{LM}C^{M}}_{\ell i}$$

These relations are useful in reducing the number of coefficients to be calculated since they always occur in conjunction with the $\underline{C}^K$ matrices in the manner of the above relations. The relations are readily proved by examination of the definition of the coefficients. They do not hold when K or M = I, or L or N = J, but, in these cases, the related coefficients are never required.

Simple rules may be derived which allow easy evaluation of the coefficients. These are, for the general coefficient $D_N(A,B,C,D,E,F)$, set out below.

1. Write down the centre superscript (CSS), that is, all numbers from 1 to N except E and F.
2. If $A \neq E,F$, write E on left hand side of A;

   if $B \neq E,F$, write F on left hand side of B;

   if $C \neq E,F$, write E on right hand side of C;

   if $D \neq E,F$, write F on right hand side of D.

   Fill in the remaining left and right superscripts with the same number as the CSS.

3. Each of the sets of three numbers, i.e. (LSS, CSS, RSS) or (X, Y, Z), correspond to the matrix product $(\underline{C^{X} S^{XY} C^{Y} S^{YZ}})$ or, in the previous notation, CSCS(X,Y,Z). If, among the sets, there are any in which the LSS equals the RSS, then take the trace of the matrix. If the LSS of one set equals the RSS of another, take the matrix product. Continue combining sets until the LSS equals the RSS of the combined set. Take the trace. Continue until all sets are combined as far as possible. The product of the combined sets is the leading term and is a scalar, matrix or supermatrix depending on the parameters.

4. Permute all LSS and all RSS, i.e. $((N-2)!)^2$ terms, combining the sets for each new term as in 3. The sum of the terms, with the parity of the permutations gives the final result.

Example. D(34, 54, 12), N = 6

Step 1 3 4 5 6

Step 2 $({}^{1}{}_{3}{}^{3})$ $({}^{2}{}_{4}{}^{2})$ $({}^{5}{}_{5}{}^{1})$ $({}^{6}{}_{6}{}^{6})$

Step 3 $({}^{5}{}_{5}{}^{1} \mid {}^{1}{}_{3}{}^{3})$ $({}^{6}{}_{6}{}^{6})$ $({}^{2}{}_{4}{}^{2})$

$$\equiv \text{CSCS}(5,5,1) \times \text{CSCS}(1,3,3).\text{tr CSCS}(6,6,6).$$
$$\text{tr CSCS}(2,4,2)$$

Step 4 Let us take the sample permutations P_{56} on LSS and P_{26} on RSS with parity +1, giving

$$({}^{1}{}_{3}{}^{3}) \quad ({}^{2}{}_{4}{}^{6}) \quad ({}^{6}{}_{5}{}^{1}) \quad ({}^{5}{}_{6}{}^{2})$$

and $({}^{5}{}_{6}{}^{2} \mid {}^{2}{}_{4}{}^{6} \mid {}^{6}{}_{5}{}^{1} \mid {}^{1}{}_{3}{}^{3})$

$$\equiv \mathrm{CSCS}(5,6,2) \times \mathrm{CSCS}(2,4,6) \times \mathrm{CSCS}(6,5,1) \times \mathrm{CSCS}(1,3,3)$$

The evaluation of the coefficients can be seen to be well suited to being performed on a computer, since the procedure is composed of only a few simple rules and consisting solely of matrix manipulations. It should also be noted that, if the molecule possesses any elements of symmetry, this is reflected in the coefficients and may greatly reduce the number to be computed in instances of high symmetry. For example, a calculation by the NPF method on the methane molecule (Td symmetry), a relatively large molecule by present standards for an accurate calculation, is well within present computing capabilities. Assuming a core orthogonal to all other orbitals, there are only 18 distinct coefficients to be calculated. And the number of coefficients does not increase as the basis is extended - only the matrix dimensions.

The problem of the convergence of the powers-of-overlap series is readily studied in terms of the coefficients which contain the major part of the non-orthogonality problem. The expansion which is most appropriate in this case is in terms of the interpair overlap matrices $\underline{S}^{IJ}$. Each of the $((N-2)!)^2$ terms which make up the coefficient may be readily classified according to its interpair overlap power. Every set (X,Y,Z) which has either the LSS or RSS different from the CSS introduces a power of overlap: thus $(X,Y,Z) \equiv \underline{C^X\ S^{XY}\ C^Y\ S^{YZ}}$ brings in two powers of overlap. The total power of a term is simply the sum of the number of noncoincidences of LSS and RSS's with the CSS's. Thus, in the example above, the leading term is a fourth power term, and the sample permutation gives a seventh power term. Each interpair overlap matrix is not of the same order of magnitude: the farther apart a pair, the smaller the elements of the overlap matrix. Thus it is consistent to neglect terms which involve overlap matrices between well spearated pair functions. This may be achieved by adding a distance weighting - the greater the separation, the greater the weight - for each interpair overlap matrix to the basic overlap power. The neglect-of-overlap approximation is therefore readily introduced into the general coefficient evaluation routine. The maximum weighted power of overlap to be included and a suitable distance weighting scheme are chosen. The distance weighted overlap order of each term in the coefficient expansion is evaluated: if it is less than, or equal to, the preset limit, then the value of the term is computed, otherwise it is omitted.

No intrapair overlap is neglected in the above approximation scheme, and, since physically separated pairs are likely to be most amenable to the NPF treatment, rapid convergence of the overlap series is expected. Because the coefficients and the approximation scheme is very easily introduced into the calculation, the possibility is open of performing extensive rigorous tests on the convergence of the series in different situations, rather than, as hitherto, rejecting all terms that are computationally or analytically too cumbersome.

In an attempt to determine the applicability of this approximation method, we have investigated a particularly simple model system - the linear chain of four hydrogen molecules with the intramolecular separation R varied between 2 and 4 a.u. Although the model system does not relate directly to the problems likely to occur in actual molecular systems, the range of R studied is in the short to medium range - the Van der Waal's radius of hydrogen is 2.3 a.u.; and coupled interactions, i.e. those between two molecules via intervening molecules, are present. The nearest neighbour inter pair overlap is of the same order as that expected in saturated hydrocarbons: at R = 3 a.u., the interpair overlap matrix is $\begin{pmatrix} 0.08 & 0.02 \\ 0.25 & 0.08 \end{pmatrix}$ compared to the interbond overlap matrix for methane, in a hybrid basis, $\begin{pmatrix} 0.00 & 0.15 \\ 0.15 & 0.28 \end{pmatrix}$.

The pair functions for the individual H_2 molecules were taken to be the optimum Heitler-London plus ionic wavefunction for the free molecule.

$$\Delta(12) = c_{11}(h_A^2 + h_B^2) + c_{12}(h_A h_B + h_B h_A)$$

where $h_A = \exp(-1.193 r_A)$, $c_{11} = 1.0$ and $c_{12} = 3.9$.

Only coefficients of the type D(I,J,I,J,K,L) were considered (although the general coefficient introduces no extra difficulty) because these reduce to scalars which makes the comparison of different approximations more easy. The coefficients studied are listed in Table 1 and were chosen to highlight the following problems:

a) the rate of convergence of the order of overlap expansion within a given coefficient

b) the relative magnitudes of the coefficients as parameters

I,J,K,L change - some types may be safely neglected.

c) the importance of environmental effects - it may be possible to approximate coefficients by equating those with the same relative interpair distances e.g. D(1,1,2) and D(2,2,3).

Tables 2 to 6 give the computed values of the coefficients at various levels of approximation. In each of the tables, the results of the simple powers-of-overlap and the distance-weighted powers-of-overlap approximations are presented. In the former, the number of interpair overlap matrices occuring in each term of the expansion of the coefficient is evaluated. This is the simple overlap power: if a term has a power greater than the preset limit, Z, then its value is not computed. For example, if Z = 3, a term such as tr CSCS(1,1,3) × CSCS(3,3,1) is computed having an order of 2 whilst tr CSCS(3,2,3).tr CSCS(4,3,4) is not, having an order of 4. In the latter approximation, a distance weighting scheme is added. The most satisfactory scheme, and the one used in all the Tables, was to give nearest neighbours zero weight, next nearest neighbours a weight of unity and third nearest neighbours a weight of four. The simple overlap order is evaluated as above, but to the order is added the distance weight for each interpair overlap matrix. Thus if the overlap matrix $\underline{S}^{13}$ is present, one is added to the power, pair functions 1 and 3 being next nearest neighbours. This weighted order is then compared against the limit Z as before. Thus the first example above has a weighted order of 2 + 2 = 4 and is not computed: the second example has an order 4 + 0 = 4. The following general conclusions may be drawn.

The rate of convergence within a given coefficient depends vitally on the interpair separation. At R = 2 a.u., the convergence of the normalisation integral is extremely poor and truncation is unlikely to be valid, whilst at R = 4 a.u., the convergence is very rapid. At the intermediate separation, where the interpair overlap is comparable to that in saturated hydrocarbons, the rate of convergence is good: thus, with Z = 4, about 15% of the terms give an accuracy of 4 parts in 10^5. The use of the distance weighting scheme reduces even further the number of terms required - 5% give the same accuracy whilst 13% give an accuracy of 3 parts in 10^8. It should be

Table 1

Type	Coefficient*
Normalisation	$D_6(5,6,5,6,5,6) \equiv D$
One electron	$D_5(1,5,1,5,1,5) \equiv D(1,1,1)$
	$D_5(1,5,1,5,2,5) \equiv D(1,1,2)$
	$D_5(1,5,1,5,3,5) \equiv D(1,1,3)$
	$D_5(1,5,1,5,4,5) \equiv D(1,1,4)$
	$D_5(2,5,2,5,3,5) \equiv D(2,2,3)$
Two electron	$D_4(1,2,1,2,1,2)$
	$D_4(1,2,1,2,3,4)$
	$D_4(1,4,1,4,2,3)$
	$D_4(1,3,1,3,2,4)$

* 1 to 4 label the H_2 molecules sequentially. 5 and 6 are dummy pair functions used to define the coefficient: they do not enter the calculation.

Table 2 Normalisation Coefficient R=3 a.u.

Approximation*	No. of terms#	D($\times 10^{-7}$)	% Error
Z=0	1(0)	2.14150442	10.7
Z=2	13(0)	1.92749676	0.36
Z=3	29(0)	1.92816992	0.32
Z=4	83(0)	1.93450560	0.0043
Z=5	179(55)	1.93447589	0.0027
Z=6	351(177)	1.93442326	1.10^{-5}
Z=7	495(313)	1.93442342	2.10^{-6}
Z=8	576(393)	1.93442346	0
Z=2/DWT	7(0)	1.92754819	0.36
Z=4/DWT	30(0)	1.93450442	0.0042
Z=6/DWT	76(0)	1.93442351	3.10^{-6}
All terms > 10^2	32	1.93447022	0.0024
All terms > 10^1	44	1.93443668	7.10^{-4}

* Z is the maximum overlap power included.

DWT - the distance weighting scheme described in the text is used.

\# The numbers in round brackets are the number of terms less than 10^{-2} included in the calculation.

Table 3 Normalisation Coefficient

Approximation*	No. of terms	$D(10^{-7})$#			
		R = 2.0 a.u.	R = 2.5 a.u.	R = 3.0 a.u.	R = 4.0 a.u.
Z=0	1	2.1415044(0)	2.14150442(0)	2.14150442(0)	2.14150442(0)
Z=2	13	1.20362064(0)	1.68245140(0)	1.92749676(0)	2.09997073(2)
Z=3	29	1.22317797(0)	1.68621860(0)	1.92816992(0)	2.10003047(2)
Z=4	83	1.34445803(1)	1.71533792(5)	1.93450560(23)	2.10026960(25)
Z=5	179	1.34060628(9)	1.71497849(37)	1.93447589(55)	2.10026911(97)
Z=2/DWT	7	1.20575574(0)	1.68279970(0)	1.92754819(0)	2.09998172(0)
Z=4/DWT	30	1.34422384(0)	1.71532093(0)	1.93450442(0)	2.10026954(0)

* See Table

\# The numbers in brackets are the number of terms less than 10^{-2}.

Table 4 One Electron Coefficient D(1,1,1) R = 3.0 a.u.

Approximation*	No. of Terms	D(1,1,1)($\times 10^{-5}$)	% Error	D(1,1,1)/D($\times 10^{-2}$)**	% Error
Z=0	1	3.14803268	7.0	1.47000989	3.3
Z=2	7	2.93831637	0.12	1.52442091	0.24
Z=3	11	2.93881104	0.11	1.52414526	0.22
Z=4	20	2.94192625	4.10^{-4}	1.52076388	4.10^{-3}
Z=5	32	2.94191547	2.10^{-6}	1.52078167	3.10^{-3}
Z=6	36	2.94191548	0	1.52082305	1.10^{-5}
Z=2/DWT	5	2.93835417	0.12	1.52439985	0.24
Z=4/DWT	15	2.94192544	3.10^{-4}	1.52076439	4.10^{-3}
Z=6/DWT	27	2.94191547	2.10^{-6}	1.52082285	3.10^{-6}

* See Table

** D(1,1,1) and D are computed to the same level of approximation.

Table 5 One Electron Coefficients

	R=2.0 a.u.	R=2.5 a.u.	R=3.0 a.u.	R=4.0 a.u.
D(1,1,1)	2.30258057(5)*	2.71525219(5)	2.94191548(5)	3.10749278(5)
	2.30388156(5)**	2.71537334(5)	2.94192544(5)	3.10749292(5)
	0.057(36/15)#	5.10^{-3}(36/15)	3.10^{-4}(36/15)	5.10^{-6}(36/15)
D(1,1,2)	1.37808306(4)	7.23021999(3)	3.44322731(3)	6.55684436(2)
	1.36505971(4)	7.21756303(3)	3.44214114(3)	6.55670325(2)
	0.95(36/5)	0.18(36/5)	0.032(36/5)	2.10^{-3}(36/5)
D(1,1,3)	4.35626595(2)	1.13574491(2)	2.59087317(1)	4.30631255(-1)
	4.56277465(2)	1.15664463(2)	2.60966738(1)	4.3237226(-1)
	4.7(36/4)	1.8(36/4)	0.73(36/4)	0.40(36/4)
D(1,1,4)	1.12666398(1)	1.59492304(0)	1.80228359(-1)	3.19565161(-5)
	0	0	0	0
	100(36/0)	100(36/0)	100(36/0)	100(36/0)
D(2,2,3)	1.54356706(4)	7.65238668(3)	3.53804272(3)	6.58085808(2)
	1.54053132(4)	7.65002552(3)	3.53788657(3)	6.58081071(2)
	0.20(36/5)	0.031(36/5)	4.10^{-3}(36/5)	7.10^{-4}(36/5)

* Exact value of coefficient, times power of ten in brackets

** Value of coefficient in Z=4/DWT approximation

Percentage error with, in brackets, the number of terms in the exact and approximate coefficient

Table 6 Two Electron Coefficients

	R=2.0 a.u.	R=2.5 a.u.	R=3.0 a.u.	R=4.0 a.u.
	3.96683818(3)*	4.30015539(3)	4.47412463(3)	4.59775041(3)
D(1,2,1,2,1,2)	3.96683818(3)**	4.30015539(3)	4.47412463(3)	4.59775041(3)
	0 (4/4)#	0 (4/4)	0 (4/4)	0 (4/4)
	1.43028390(-5)	3.12639310(-7)	1.16115580(-8)	5.520300383(-10)
D(1,2,1,2,3,4)	0	0	0	0
	100(4/0)	100(4/0)	100(4/0)	100(4/0)
	1.29963072(1)	2.91366796(0)	5.99397142(-1)	2.04239723(-1)
D(1,4,1,4,2,3)	1.31558290(1)	2.92568476(0)	6.00213554(-1)	2.04528328(-2)
	1.2(4/1)	0.41(4/1)	0.14(4/1)	0.14(4/1)
	1.30649415(1)	2.91871685(0)	5.99813350(-1)	2.04511907(-2)
	1.31558290(1)	2.92568476(0)	6.00213554(-1)	2.04528328(-2)
	0.70(4/1)	0.24(4/1)	0.067(4/1)	8.10^{-3}(4/1)

*, **, #. See Table 5

noted that the distance weighting scheme eliminates the third nearest neighbour terms, at all but the highest levels of approximation. Thus, in favourable circumstances and given an effective distance weighting scheme, very considerable reductions in the number of terms to be computed may be achieved at the cost of very little loss of accuracy. The convergence of the one electron coefficient D(1,1,1) is less marked since we are effectively dealing with a three molecule system - 42% of the terms are needed for 3 parts in 10^6 accuracy. The results in column 5 of Table 4 show that there is little or no cancellation of errors in the ratio of the one electron to the normalisation coefficient.

As the parameters refer to more and more distant pair functions, the magnitude of the coefficient decreases much more rapidly at large R than at small R. The error increases as the magnitude of the coefficient decreases, and as R becomes smaller. That the error increases as the magnitude decreases does not mean that the overall accuracy of, say, the energy will suffer since it is unlikely that the factor by which the coefficient is multiplied in evaluating an expectation value will increase as the magnitude of the coefficient decreases - rather the reverse since, if the coefficient involves overlap in the leading term, then so will the factor. The overall pattern of the results indicate that, if the situation is favourable, a consistent approximation scheme, for example, the neglect of powers of overlap greater than four with the distance weighting, allows the calculation to the relevant accuracy of a wide variety of coefficients.

The magnitude of environmental effects are evident in the values of D(1,1,2) and D(2,2,3): although of the same order, they differ by a significant amount, about 3% at $R = 3$ a.u. This difference is well reproduced by the approximation scheme. On the other hand, D(1,4,1,4,2,3) and D(1,3,1,3,2,4) are much smaller and differ by less: here the difference is not detected.

This very preliminary study of a model system is encouraging in that it indicates that, in circumstances which may not be atypical of molecular systems, a large proportion of the work may be avoided by a consistent overlap approximation scheme. It is also evident that the dividing line between a convergent and an unconvergent expansion is a fine one. The completion of the energy calculation for the model system is necessary to give the comprehensive data from which concrete conclusions may be drawn, regarding this aspect of the non-orthogonality problem.

A similar approach to the non-orthogonality problem has been made by Dacre and McWeeny[4] who considered the overlap expansion for the interaction of two groups, atoms or molecules in arbitrary spin states, retaining all intragroup non-orthogonality and using intergroup overlap in the expansion. They found that the energy converges rapidly - the interaction energy, including only terms up to fourth power in overlap, is in error, for two neon atoms, by only 0.2% at R = 2.5 a.u., and, for two nitrogen atoms, by 0.1% at R = 3.0 a.u. An advantage of their formalism is that the energy and normalisation integrals are readily computed to the same level of approximation: cancellation of errors increases the convergence quite markedly. Also non-orthogonal formalisms for solids have been considered, directed towards the evaluation of the inverse - of the overlap matrix. Only very simple systems have been tackled, for example, the ferromagnetic hydrogen lattice, where advantage may be taken of the symmetry[5]

In extending the NPF approximation to molecular systems, the question of the convergence of the overlap series arises. Because of the many nearest neighbours that are present, and hence the larger number of terms of the same order of magnitude, the convergence of the coefficients cannot be as marked as that in a model with only single nearest neighbours: possibly the net of hydrogen molecules would serve as a more realistic model. On the other hand, the use of directed atomic orbitals should lead to smaller interpair overlap integrals than in the model system. In fact, the approach requires, and deserves, much more investigation before it will be possible to say whether or not it constitutes a solution to the non-orthogonality problem.

Although the approximation method considered above is quite general, it suffers from the lack of a variational method for the pair function coefficients: the necessary matrix elements are far too complex to be of any practical use. Thus the pair function coefficients must be fixed at the outset of the computation. This should not be too great a disadvantage in many types of large molecule since the transferability of non-orthogonal pair functions for commonly occuring groups, for example, the CH bond, is expected to be good (see Section 4), and the coefficients could be found by optimisation in small molecules, e.g. methane. Also, if the neglect-of-overlap approximation can be used to significantly reduce the computational labour, numerical optimisation methods may, for a limited number of parameters, be practically possible.

However, as noted in Section 2, the situation regarding complete variational calculations is far from optimistic for more than six electrons: this would rule out even the possibility of finding the optimum coefficients in methane. And the neglect-of-overlap approximation is neither well enough understood nor likely to be of much advantage at this level. An approximate variational method would thus be highly desirable for medium sized molecules for determining the optimum basic pair functions for larger molecules.

The HF LCAO wavefunction, although neglecting opposite-spin electron correlation, is a good approximation to the exact molecular wavefunction: and the localised or bond orbital approximation, in which completely localised MO's replace the delocalised ones of HF LCAO theory, in many instances, gives results extremely close to those of the complete treatment[6]. We might therefore consider the expansion of the localised pair functions, $\Delta^I(12)$, in terms of a bond orbital product, $\phi_0^I(1)\phi_0^I(2)\alpha(1)\beta(2)$, plus a correlation correction.

$$\Delta^I(12) = \sum_{ij} d_{ij}^I \; \phi_i^I(1)\phi_j^I(2)\alpha(1)\beta(2)$$

$$= \phi_0^I(1)\phi_0^I(2)\alpha(1)\beta(2) + \delta\Delta^I(12)$$

where $$\delta\Delta^{I}(12) = \sum_{(ij)\neq 0} d_{ij}^{I}\phi_{i}^{I}(1)\phi_{j}^{I}(2)\alpha(1)\beta(2) ,$$

$(ij)\neq 0$ indicating that the $i=j=0$ term is excluded from the summation. The NPF wavefunction may then be written

$$\begin{aligned}\Psi &= \hat{A}\,\Delta^{1}(12)\Delta^{2}(34)\ \ldots\ \Delta^{N}(2N-1,2N) \\ &= \hat{A}\,[\psi_0 + \sum_{I=1}^{N}\psi_I + \sum_{I<J}^{N}\psi_{IJ} + \ldots + \psi_{12\ldots N}]\end{aligned} \tag{3.6}$$

where $$\psi_0 = \prod_{I}^{N} \phi_0^{I}(2I-1)\phi_0^{I}(2I)\alpha(2I-1)\beta(2I)$$

$$\psi_I = \delta\Delta^{I}(2I-1,2I)\prod_{J\neq I}^{N}\phi_0^{J}(2J-1)\phi_0^{J}(2J)\alpha(2J-1)\beta(2J)$$

and so on.

The energy expression is, via the above wavefunction expansion, given by

$$E = \int \Psi\,\hat{H}\,\Psi\,d\tau \Big/ \int \Psi\,\Psi\,d\tau = H/S$$

where the energy integral

$$H = H_{0,0} + 2\sum_{I} H_{0,I} + \sum_{IJ} H_{I,J} + \sum_{IJ}{}' H_{0,IJ} + \ldots + H_{12..N,12..N} \tag{3.7}$$

in which

$$H_{0,0} = \int \psi_0 \hat{H} \psi_0 \, d\tau$$

$$H_{0,I} = \int \psi_0 \hat{H} \psi_I \, d\tau$$

and, in general, $H_{IJ..P,KL..Q} = \int \psi_{IJ..P} \hat{H} \psi_{KL..Q} \, d\tau$

With analogous definitions, the normalisation integral

$$S = S_{0,0} + 2 \sum_I S_{0,I} + \sum_{IJ} S_{I,J} + \sum_{IJ}{}' S_{0,IJ} + \ldots \qquad (3.8)$$

Since the localised orbital description, ψ_0, provides a good approximate representation of the electron density and energy of the molecule, the correlation correction terms will only represent a small modification of the essentially correct bond orbital description: hence the correlation coefficient expansion coefficients may be expected to be small, that is, $d^I_{ij} << 1$, $(ij)\neq 0$, and the correlation terms, ψ_I, $\psi_{IJ,...}$, to make progressively smaller contributions to the total wave function. The expansions of the wavefunction and the energy should therefore converge rapidly and make valid a truncation of equations (3.7) and (3.8). If, in practice, such a truncation may be made to yield a large part of the energy of the NPF wavefunction, the variational determination of optimum coefficients, d^I_{ij}, using the approximate energy expression should give pair function coefficients close to the optimum ones for the exact NPF wavefunction. Such a method has been considered for the general cluster expansion of the wavefunction, where, because of the simple expressions for the matrix elements, the exact nature of the approximation may be studied in detail. It has been shown that, under certain conditions, it is even possible to decouple the variational equations, each pair function being determined by an independent equation[7]. These formal simplifications are paid for by the practical difficulties imposed by the orthogonality constraints that have to be satisfied. Although no decoupling of the variational equations is possible in a truncated NPF expansion, the lack of orthogonality problems allows the straightforward derivation of widely applicable variational equations.

The initial problem is to decide at what stage the expansion may be truncated. This can only be determined with certainty by practical experience: however it is also necessary, for any advantage to be gained, that the expansion is sufficiently short for the variational procedure to be a practical one. We shall here prejudge the issue by adopting a hopefully adequate approximation which permits of the derivation of relatively simple variational methods. The simplest approximation of equation (3.6) is to neglect all but the first correlation terms in the wavefunction

$$\Psi' = \psi_0 + \sum_I \psi'_I \equiv \sum_I \Psi_I$$

where

$$\Psi_I = \sum_{ij=0} d^I_{ij} \phi^I_i(2I-1)\phi^I_j(2I)\alpha(2I-1)\beta(2I)$$

$$\times \prod_{J\neq I} \phi^J_0(2J-1)\phi^J_0(2J)\alpha(2J-1)\beta(2J).$$

A dashed symbol is used to indicate that the quantity is an approximation to the exact NPF quantity. The normalisation convention that the coefficient of ψ_0 is unity is no longer imposed. The approximate energy expression is thus

$$E' = H'/S' = \sum_{IJ} H_{IJ}/\sum_{IJ} S_{IJ} \tag{3.9}$$

where

$$H_{IJ} = \sum_{ijk\ell} d'^I_{ij}\, d'^J_{k\ell}\, H^{IJ}_{ijk\ell}$$

and

$$H^{IJ}_{ijk\ell} = \int \psi^I_{ij}\, \hat{H}\, \psi^J_{k\ell}\, d\tau$$

ψ^I_{ij} is the determinant ψ_0 with $\phi^I_0\phi^I_0$ replaced by $\phi^I_i\phi^I_j$

But $H^{IJ}_{0000} = H^{KL}_{0000} \equiv H_0$, for all I, J, K and L

and $H^{IJ}_{00k\ell} = H^{KJ}_{00k\ell} \equiv H^{J}_{k\ell}$, for all I and K.

Hence,
$$H_{IJ} = d'^{I}_{00}\, d'^{J}_{00}\, H_0 + \sum_{(ij)\neq 0} d'^{I}_{ij}\, d'^{J}_{00}\, H^{I}_{ij} + \sum_{(k\ell)\neq 0} d'^{J}_{k\ell}\, d'^{I}_{00}\, H^{J}_{k\ell} + \sum_{(ij),(k\ell)\neq 0} d'^{I}_{ij}\, d'^{J}_{k\ell}\, H^{I}_{ij}$$

so that
$$H' = d^2{}_0\, H_0 + 2\, d_0 \sum_{I} \sum_{(ij)\neq 0} d'^{I}_{ij}\, H^{I}_{ij} + \sum_{IJ} \sum_{(ij)(k\ell)\neq 0} d'^{I}_{ij}\, d'^{J}_{k\ell}\, H^{IJ}_{ijk\ell}$$

with
$$d_0 = \sum_{I} d'^{I}_{00}$$

Analogous expressions are found for S'. Differentiating E' with respect to d_0 and d'^{K}_{pq} and equating the resulting expressions for $\partial E'/\partial d_0$ and $\partial E'/\partial d'^{K}_{pq}$ to zero leads to the following equations defining the optimum coefficients

$$d_0(H_0 - E'S_0) + \sum_{I}^{N} \sum_{(ij)\neq 0} d'^{I}_{ij}(H^{I}_{ij} - E'S^{I}_{ij}) = 0,$$
$$d_0(H^{K}_{pq} - E'S^{K}_{pq}) + \sum_{I}^{N} \sum_{(ij)\neq 0} d'^{I}_{ij}(H^{IK}_{ijpq} - E'S^{IK}_{ijpq}) = 0 \qquad (3.10)$$

for all K and (pq) $\neq$ 0. Using the condition that d'^{I}_{ij} must be equal to d'^{I}_{ji} to preserve a singlet spin eigenfunction, the above secular equations may be reduced to one of order M where

$$M = \sum_{I}^{N} n_I \; (n_I+1)/2 - (N-1)$$

and n_I is the number of orbitals in pair function I. Thus the optimum coefficients may be determined non-iteratively, by the solution of a single secular problem. The dimension of the secular equations increases rapidly with the number of bonds and will make the procedure impractical for large molecules, especially when larger than minimal bases are used: on the other hand, in situations of high symmetry, where there are relatively few distinct coefficients, the method is eminently practical.

It is worth considering the situation when it is not possible to solve the complete secular equation and we wish to determine the coefficients of each pair function separately. Assuming that the coefficients of all pair functions except K are known or have been guessed, the optimisation of the d'^{K}_{ij} leads to the equations

$$\sum_{ij} d'^{K}_{ij}(H^{KK}_{ijpq} - E'S^{KK}_{ijpq}) = - \sum_{I \neq K}^{N} (H^{KI}_{pq} - E'S^{KI}_{pq}) \qquad (3.11)$$

where $$H^{KI}_{pq} = \sum_{(ij)\neq 0} H^{KI}_{pqij}\, d'^{\,I}_{ij}$$

This gives a set of $n_K(n_K+1)/2$ equations for pair function K, but, because of the term on the right hand side of equation (3.11), the coefficients are not uniquely defined. A unique solution may however be found by reconsidering the complete eigenvalue equations (3.10): by partitioning the coefficients into the set $\underline{C}_A$ corresponding to those of pair function K and the set $\underline{C}_B$ corresponding to the coefficients of all pair functions except K, we may write the secular equation

$$(\underline{H} - E\,\underline{S})\underline{C} = \underline{0}$$

as

$$\begin{pmatrix} \underline{H}_{AA} - E\,\underline{S}_{AA} & \underline{H}_{AB} - E\,\underline{S}_{AB} \\ \underline{H}_{BA} - E\,\underline{S}_{BA} & \underline{H}_{BB} - E\,\underline{S}_{BB} \end{pmatrix} \begin{pmatrix} \underline{C}_A \\ \underline{C}_B \end{pmatrix} = \begin{pmatrix} \underline{0} \\ \underline{0} \end{pmatrix}$$

which, on expansion, yields

$$(\underline{H}_{AA} - E\,\underline{S}_{AA})\underline{C}_A + (\underline{H}_{AB} - E\,\underline{S}_{AB})\underline{C}_B = \underline{0} \tag{3.13a}$$

$$(\underline{H}_{BA} - E\,\underline{S}_{BA})\underline{C}_A + (\underline{H}_{BB} - E\,\underline{S}_{BB})\underline{C}_B = \underline{0} \tag{3.13b}$$

From (3.13b) we have

$$\underline{C}_B = -(\underline{H}_{BB} - E\,\underline{S}_{BB})^{-1}(\underline{H}_{BA} - E\,\underline{S}_{BA})\underline{C}_A$$

which gives, on insertion in (3.13a),

$$(\underline{H}^{F}_{AA} - E\,\underline{S}_{AA})\underline{C}_{A} = \underline{0} \tag{3.14}$$

where $\underline{H}^{F}_{AA} = \underline{H}_{AA} - (\underline{H}_{AB} - E\,\underline{S}_{AB})(\underline{H}_{BB} - E\,\underline{S}_{BB})^{-1}(\underline{H}_{BA} - E\,\underline{S}_{BA})$
The coefficients of pair function K may thus be found by an iterative process in which the energy E is guessed, $\underline{H}^{F}_{AA}$ is set up and the secular equation (3.14) solved to give a new E with which to revise $\underline{H}^{F}_{AA}$, and so on until $\underline{C}_{A}$ and E converge. The advantage of this approach lies in the fact that an effective Hamiltonian, of which the pair function is an eigenfunction, is defined for each pair of bonded electrons in the field of all other pairs. In practice, the approach requires the repeated inversion of a matrix of order $M - n_K(n_K+1)/2$ and is, in effect, equivalent to solving the complete secular problem.

Since, in the above approximation, it is the wavefunction that has been truncated, the part of the energy that is optimised always remains an upper bound to the exact energy. But the convergence of expectation values, especially the energy, is of greater importance than that of the wavefunction. It is readily seen from expansions (3.6) and (3.7) that the wavefunction truncated to n'th order does not give the energy to order 2n in the correlation correction: for example, the fourth term in equation (3.7) is a second order energy term, derived from a second order wavefunction term. Since the applicability of the approach depends vitally on the convergence rate of the energy expression - the larger the part of the energy that is optimised, the closer will be the coefficients so determined to the true optimum coefficients, it is preferable to focus attention on the energy expansion of equation (3.7) and to truncate it, rather than the wavefunction expansion, at a given order.

The most tractable approximation in this approach is the neglect of all terms greater than second order in correlation corrections. Thus

$$E' = H'/S' \tag{3.12}$$

where

$$H' = H_{0,0} + 2\sum_{I} H_{0,I} + \sum_{IJ} H_{I,J} + \sum_{IJ} H_{0,IJ}$$

and $$S' = S_{0,0} + 2\sum_{I} S_{0,I} + \sum_{IJ} S_{I,J} + \sum_{IJ} S_{0,IJ}$$

This approximation differs from the previous one, equation (3.9), only through the last term in the above expressions. It should be noted that $H_{0,IJ}$ is non-zero even in an orthogonal basis expansion due to the presence of single excitations in the pair functions. The influence of this term is difficult to assess: it is probably small in an orthogonal basis, but whether this is also true for a non-orthogonal basis, it is not possible to say.

The optimum coefficients, in this approximation, may be determined by exactly the same method as given above for the wavefunction truncation approximation, since it is readily shown, by expansion of the NPF wavefunction, that the energy integral, H', may be written

$$
\begin{aligned}
H' &= d_0^2 H_{0,0} + 2d_0 \sum_I \sum_{(ij)\neq 0} d'^{I}_{ij} H^{I}_{ij} \\
&+ \sum_I \sum_{(ij)\neq 0} \sum_{(k\ell)\neq 0} d'^{I}_{ij} d'^{I}_{k\ell} H^{I,I}_{ijk\ell} \\
&+ \sum_I \sum_J{}' \sum_{(ij)\neq 0} \sum_{(k\ell)\neq 0} d'^{I}_{ij} d'^{J}_{k\ell} (H^{I,J}_{ijk\ell} + H^{0,IJ}_{ijk\ell})
\end{aligned}
$$

An M'th order secular equation, as before, therefore defines the optimum coefficients.

As we have seen, the attempt to determine each set of pair function coefficients individually leads to the use of the partition method which, although defining an effective

Hamiltonian for each pair, is equivalent in practice to solving the complete problem. It would be advantageous to have a method of obtaining the optimum coefficients for a given pair function within the fixed space of all other pair functions. For example, one might wish, as a first approximation, in a calculation on methyl fluoride, to assume that the C-H bond pair functions are the same as those in methane, thus leaving the C-F bond pair function as the only undetermined quantity. Although the following solution to this problem is given explicitly for the energy truncation approximation, it is equally valid, by the omission of the $H^{0,IJ}$ terms, for the wavefunction truncation approximation.

An alternative prescription has been devised, based on the steepest descent method. Expanding the terms in the energy expression into the orbital basis, we find

$$H_{0,I} = \sum_{ij} d'^{I}_{ij} H^{0I}_{ij} \equiv [d'^{I} H^{0I}]$$

$$H_{I,J} = \sum_{ijk\ell} d'^{I}_{ij} d'^{J}_{k\ell} H^{IJ}_{ijk\ell} \equiv [d'^{I} H^{IJ} d'^{J}]$$

$$H_{0,IJ} = \sum_{ijk\ell} d'^{I}_{ij} d'^{J}_{k\ell} H^{0IJ}_{ijk\ell} = [d'^{I} H^{0IJ} d'^{J}]$$

where $H^{0IJ}_{ijk\ell} = \int \psi_0 \hat{H} \psi^{IJ}_{ijk\ell} \, d\tau$ and $\psi^{IJ}_{ijk\ell}$ is the determinant ψ_0 with $\phi^{I}_{0}\phi^{-I}_{0}$ replaced by $\phi^{I}_{i}\phi^{-I}_{j}$ and $\phi^{J}_{0}\phi^{-J}_{0}$ by $\phi^{J}_{k}\phi^{-J}_{\ell}$.

$$H^{0I}_{ij} = \int \psi_0 \hat{H} \psi^{I}_{ij} \, d\tau \qquad \text{and}$$

$$H^{IJ}_{ijk\ell} = \int \psi^{I}_{ij} \hat{H} \psi^{J}_{k\ell} \, d\tau$$

Analogous expansions exist for the overlap terms, $S_{0,I}$ etc.

Thus the energy integral H' may be written

$$H' = H_{0,0} + 2 \sum_{I} [d'^{I} H^{0I}] + \sum_{IJ}{}' [d'^{I} H^{0IJ} d'^{J}] + \sum_{IJ} [d'^{I} H^{IJ} d'^{J}]$$

$$= H_{0,0} + \sum_{I} \{[d'^{I} H^{I}] + [d'^{I} H^{II} d'^{I}]\}$$

where

$$\underline{H}^{I} = 2\underline{H}^{0I} + \sum_{J\neq I} \{\underline{H}^{IJ}\underline{d}'^{J} + \underline{H}^{0IJ}\underline{d}'^{J}\}$$

and $$(\underline{H}^{IJ}\underline{d}'^{J})_{ij} = \sum_{k\ell} H^{IJ}_{ijk\ell}\, d'^{J}_{k\ell}$$

Finally we have

$$H' = H_{0,0} + \sum_{I} [d'^{I}\, \bar{H}^{I}]$$

with $$\bar{\underline{H}}^{I} = \underline{H}^{I} + \underline{H}^{II}\, \underline{d}'^{I}$$

Similarly, $$S' = S_{0,0} + \sum_{I} [d'^{I}\, \bar{S}^{I}]$$

The energy change, $\delta E'$, due to a variation of the correlation correction coefficients, $d'^{I}_{ij} \rightarrow d'^{I}_{ij} + \delta d^{I}_{ij}$, is given by, to first order,

$$\begin{aligned}
\delta E' &= \delta(H'/S') = (\delta H' - E'\delta S')/S' \\
\delta H' &= [\delta d^{I}\, \bar{H}^{I}] + [d'^{I}\, \delta\bar{H}^{I}] \\
&= [\delta d^{I}\, H^{I}_{F}]
\end{aligned}$$

where $$\underline{H}^{I}_{F} = \underline{H}^{I} + 2\underline{H}^{II}\, \underline{d}'^{I},$$

and $$\delta S' = [\delta d^{I}\, S^{I}_{F}]$$

We may, without loss of generality, assume that the wave function is, to the level of the approximation, normalised, S' = 1, and that it remains normalised under the variation $\underline{\delta d}^I$, that is, $\delta S' = 0$. We are therefore seeking a variation which maximises $\delta E'$, and hence $\delta H'$, subject to the conditions that $\delta S' = 0$ and $\underline{\delta d}^I$ is symmetric, so that the wavefunction remains an eigenfunction of $\hat{S}^2$. These constraints are contained, to first order, in the variation

$$\underline{\delta d}^I = \underline{M}^I \underline{S}_F^I - \underline{S}_F^I \underline{M}^I$$

where $\underline{M}^I = \underline{D}^I - (\underline{D}^I)^+$ and $\underline{D}^I$ is an arbitrary matrix. The normalisation condition is obeyed to first order since

$$\delta S' = [\delta d^I S_F^I] = [M^I S_F^I S_F^I] - [S_F^I M^I S_F^I]$$

= 0, from the cyclic commutation property of the matrix trace. The symmetry condition is also satisfied since

$$\begin{aligned} (\underline{\delta d}^I)^+ &= (\underline{M}^I \underline{S}_F^I)^+ - (\underline{S}_F^I \underline{M}^I)^+ \\ &= \underline{S}_F^I (\underline{M}^I)^+ - (\underline{M}^I)^+ \underline{S}_F^I \\ &= \underline{M}^I \underline{S}_F^I - \underline{S}_F^I \underline{M}^I = \underline{\delta d}^I \end{aligned}$$

since $\underline{M}^I$ is antisymmetric and $\underline{S}_F^I$ symmetric. Thus,

$$
\begin{aligned}
\delta E' &= [\delta d^{I} \; H^{I}_{F}] \\
&= [(M^{I} \; S^{I}_{F} - S^{I}_{F} \; M^{I}) H^{I}_{F}] \\
&= 2[M^{I} \; S^{I}_{F} \; H^{I}_{F}] \\
&= 2[D^{I} \; S^{I}_{F} \; H^{I}_{F}] - 2[(D^{I})^{+} \; S^{I}_{F} \; H^{I}_{F}] \\
&= [(D^{I})^{+} \; A^{I}]
\end{aligned}
$$

where $\underline{A}^{I} = 2(\underline{H}^{I}_{F} \; \underline{S}^{I}_{F} - S^{I}_{F} \; \underline{H}^{I}_{F})$

Now, $\delta E'$ has its largest negative value for

$$\underline{D}^{I} = -\lambda \; \underline{A}^{I}$$

where λ is a positive scalar which fixes the length of the step taken down the direction of steepest descent determined by $\underline{D}^{I}$. The optimum value of λ is found by expanding E' to second order and maximising it with respect to λ.

$$\delta E' = E' - (E' - \lambda H_{1} + \lambda^{2} H_{2})/(1 + \lambda^{2} S_{2})$$

where $H_{1} = [B^{I} \; H^{I}]$, $H_{2} = [B^{I} \; H^{II} \; B^{I}]$, $S_{2} = [B^{I} \; S^{II} \; B^{I}]$ and $\underline{B}^{I} = 4(\underline{H}^{I}_{F} \; \underline{S}^{I}_{F} \; \underline{S}^{I}_{F} + \underline{S}^{I}_{F} \; \underline{S}^{I}_{F} \; \underline{H}^{I}_{F} - 2\underline{S}^{I}_{F} \; \underline{H}^{I}_{F} \; \underline{S}^{I}_{F})$.

Expanding $(1 + \lambda^{2} S_{2})^{-1}$ in a binomial series, $1 - \lambda^{2} S_{2} + \ldots$, and neglecting terms of order higher than λ^{2} in $\delta E'$, we find

$$
\begin{aligned}
\delta E' &= \lambda H_{1} - \lambda^{2}(H_{2} - E' S_{2}) \\
d(\delta E')/d\lambda &= H_{1} - 2\lambda(H_{2} - E' S_{2}) = 0
\end{aligned}
$$

Therefore, $\lambda_{opt} = H_1/2(H_2 - E'S_2)$

It may be noted that, at the minimum, $\underset{\sim}{S}_F^I$ and $\underset{\sim}{H}_F^I$ commute.

The solution by this method requires a double iteration process. The 'inner' steepest descent iteration locates the local minimum for pair function I in the fixed space of all other pair functions: the 'outer' iteration - application of the steepest descent procedure to each pair function in turn - is continued until they are self-consistent and the overall minimum reached.

So far we have not considered how the basis functions, ϕ_i^I, are to be chosen. As we have noted, the bond orbital approximation will yield a good starting point for the expansion, that is, the ϕ_0^I: these localised orbitals have been determined for a variety of molecules by standard SCF techniques and are readily available in the literature. Even if, in certain instances, they are not, they may be easily determined by well known methods. There is a wide choice open for the higher orbitals, ϕ_i^I, $i \neq 0$, due to the lack of orthogonality constraints - the localised virtual orbitals, which are the redundant solutions of the SCF equations, extended basis orbitals with higher principal or angular quantum numbers; orbitals from neighbouring bonds which introduce charge delocalisation;

and any other orbitals not vital to the basic description of the bond. How vital the choice of the bond orbitals, ϕ_0^I, is, depends entirely on how sensitive the rate of convergence of the expansion is to the ϕ_0^I, and can only be determined by practical experience by evaluation of the higher neglected terms in the series.

The matrix elements, for example, $H_{ijk\ell}^{IJ}$, required in the optimisation procedures are complicated by the non-orthogonality of the basic orbitals: they may be evaluated by standard techniques which are given in detail in Appendix 4. Since the matrix elements are fixed once the choice of basis functions has been made, they need only be computed once and should not constitute a grave computational problem.

The approximate variational methods offer a real hope of obtaining NPF wavefunctions for medium sized molecules such as ethane which are outside the scope of a complete treatment. The importance of the NPF method, as compared to the SBP method, lies in its potential accuracy and its extended scope. Because of the orthogonality restrictions, the SBP method is confined to molecules where the bonds are well separated: the NPF method is not so restricted. For example, molecular reactions where bonds are being broken and formed e.g. Walden inversion; orbital deficient molecules such as SF_6, without the need to invoke d orbitals and yet retaining the equivalence of all bonds; molecular interactions which are poorly represented by the SBP method; hydrogen bonded systems where a hydrogen atom is bonded to two other atoms; these situations all fall within the scope of the NPF method.

For large molecular systems without any symmetry and large molecules where only one small group is chemically important, the majority of current wavefunction methods will be out of the question because the molecule must be treated as a whole, e.g. the MO method, or basis orthogonalisation and integral transformations are required. The NPF method, on the other hand, contains the possibility of division of the molecule into "spheres of influence" outside which contributions may be neglected, and gives a basis for rigorous approximation techniques.

REFERENCES

1. Slater, J.C. (1951) J.Chem.Phys. 19, 220
2. McWeeny, R. (1954) Proc.Roy.Soc.London A223, 63, 306
3. Coulson, C.A. (1961) "Valence" p. 174, Oxford University Press
4. Dacre, P.D. & McWeeny, R. (1970) Proc.Roy.Soc.London A317, 435
5. (a) Herring, C. (1966) in "Magnetism IIB" Ed. Rado, G.T. & Sulh, H. Academic Press

   (b) Lowdin, P.-O. (1951) J.Chem.Phys. 18, 365
6. Petke, J.D. & Whitten, J.L. (1969) J.Chem.Phys. 51, 3166
7. (a) Sinanoglu, P. (1964) "Advances in Chemical Physics" Volume 6, p. 315, Wiley-Interscience

   (b) McWeeny, R. & Steiner, E. (1965) "Advances in Quantum Chemistry" Volume 2, p. 93, Academic Press

SECTION 4

Valence Theory and the Non-Orthogonal Pairfunction Wavefunction

One of the major problems of molecular quantum mechanics is to relate the concepts of empirical valence theory to quantum theory, to examine their theoretical justification and to extend and quantify them where this is possible. Coulson has defined the questions that valence theory must answer as: why are molecules formed at all; why do atoms exhibit particular valencies; and why do molecules take up the specific geometries that they possess[1]? Not only must valence theory be explanatory, it must also be predictive: from the answers to the above questions must follow the details of molecular electronic structure, the reason why a certain molecule has a particular value for a given property and why similar molecules possess different values, and the ability to predict the changes from molecule to molecule.

The advances that have so far been made are not particularly impressive[2]. For example, the question of why ammonia is pyramidal instead of planar with a single (lone pair) electron on either side of the plane, which might appear on electrostatic grounds to be energetically more favourable, cannot be satisfactorily answered. The problems are many. There is the difficulty that, electrons cannot be treated as static yet we cannot enquire into their motions except in terms of probabilities. Secondly, partitioning of the value of a given property into quantities referring to the constituent parts of the molecule, atoms or bonds, and their interactions is necessary. For example, the total energy must be partitioned into bond and bond interaction energies to say whether methane is tetrahedral because of the increase in bond repulsions or the decrease in the bond energy as the geometry changes from tetrahedral. Since none of these quantities can be an observable, no unique partition exists and the merits of a particular partition rests solely on the practical results obtained. Also, the practical problem of obtaining wavefunctions for molecules of interest is still not completely solved. Most practical work is, at present, directed towards molecules that are observed: to answer the questions that are the province of valence theory, it is surely necessary to also study those molecules that do not exist so that the reasons for the stability of particular molecules can be found. This has so far not been attempted. Although little advance has been made in terms of fundamental explanations, many details of molecular properties have been rationalised on the basis of quantum mechanical calculations, simple physical pictures and molecular parameters such as net atomic charges and bond orders. An example is the correlation of the proton chemical shifts of the hydrocarbon series C_nH_{2n+2} with the net charge carried by the proton.[3]

To explain the existence, geometry and properties of molecules it is vital to achieve a partition of the quantum mechanical expression for the property of interest, e.g. energy, dipole moment. Since there is no God-given partition, a degree of subjectivity is unavoidable. It might be objected that subjectivity has no part in scientific research and should be avoided at all costs; if science were simply a search for facts, this would be acceptable, but if it is a search for explanations, then partitioning is inescapable. This is a new dimension introduced by quantum mechanics: at the macroscopic level, the subdivision of a system is always apparent - we may talk about the motion of one billiard ball independently of all others that may be in motion on the table. At the microscopic level it is not so simple, but it is only a matter of degree. There are forces acting on the system of billiard balls that are (quite correctly) neglected: this is a partitioning of the total system. In a molecule, the forces between, say, bonds are certainly larger, but it is an undisputable experimental fact that the bonds in molecules are to a great extent independent entities.

How may we achieve a partition of the expression for a particular property? Three distinct approaches, for an orbital wavefunction, are conceivable - to divide up the expression directly into quantities referring to bonds and/or atoms using the atomic orbitals as the basis for the partition; to divide up the space that the molecule occupies into bond or atomic regions and to compute the contribution of each region to the property; or to divide up the wavefunction into bond or atomic functions and to follow this separation through to the expression for the property. As an example, take the x-component of the dipole moment $\bar{\mu}$ of a molecule: this may be expressed in terms of the one-electron density function P(1) i.e.

$$\bar{\mu} = \int x_1 \, P(1) \, dV_1$$

For an arbitrary orbital wavefunction,

$$P(1) = \sum_{ij} P_{ij} \, \phi_i(1)\phi_j(1), \text{ where the } \phi_i \text{ are}$$

here assumed to be atomic orbitals. Thus

$$\bar{\mu} = \sum_{ij} P_{ij} \int \phi_i(1)x_1\phi_j(1)dV_1 = \sum_{ij} P_{ij} \, \mu_{ij}$$

For the first partitioning method, we might take $\mu_A = \sum_{ij\varepsilon A} P_{ij}\mu_{ij}$

as the contribution of atom A to the dipole moment (where the sum is over all orbitals on atom A), and $\mu_{AB} = 2 \sum_{i\varepsilon A} \sum_{j\varepsilon B} P_{ij}\mu_{ij}$ as the contribution of the bond between atoms A and B (where the sum is over all orbitals that may be considered to contribute to the bond). In the second method, the three dimensional space might be divided up into spherical regions of a given radius round each atom plus interbond regions. Then the quantity

$$\mu_A = \int_0^{R_A} x_1 \, P(1) \, dV_{1A}$$

where the integration is over the sphere of radius R_A occupied by atom A, would be the contribution of atom A to the dipole moment. And thirdly, if the wavefunction is a product of functions describing the individual bonds

$$\Psi = \hat{A} \, \Delta^1(12)\Delta^2(34) \, \ldots \, \Delta^N(2N-1,2N)$$

then it is possible, under certain conditions, to write the density function in terms of bond densities[3]

$$P(1) = \sum_I P^I(1) \quad \text{where} \quad P^I(1) = \int \Delta^I(12)\Delta^I(12) d\tau_2 ds_1$$

and hence the dipole moment is

$$\bar{\mu} = \sum_I \mu_I \, , \qquad \mu_I = \int P^I(1) \, x_1 \, dV_1 \, ,$$

a sum of bond dipole moments.

The second partitioning method must at present be ruled out, despite its attractiveness[5], on the grounds of the practical difficulties it entails. The remaining two methods suffer from a common problem: how to decide which orbitals to associate with which bond, in the former at the partitioning stage, and in the latter at the stage of setting up the wavefunction. The latter method does have the advantage that, if there is any doubt, an orbital may be allowed to contribute to more than one bond function, the variational process sorting out the degree to which it contributes to each. This is especially important when an extended basis is used. Also, because of this difficulty, it is easier, when partitioning the expression for the property, to define quantities referring to atoms than to bonds. But, since molecules are thought of more in terms of bonds than of atoms, the method of partitioning the wavefunction, which defines bond quantities directly, is preferable. Although theoretically we might think the third method best, it must be borne in mind that the ultimate choice depends on the usefulness of the results achieved.

Although most observables are single numbers (with appropriate units) and hence require a method of partition, an exception is the molecular electron density distribution. Might it not be possible to obtain all the information we need and avoid the neccessity of partitioning either wavefunction or observable expression into atom or bond quantities by studying the electron distribution? This appealing idea has been of great interest recently in terms not only of the electron density but also of density difference maps in which the electron density of the free atoms are subtracted from that of the molecule, showing how the electrons are redistributed on molecule formation.[6] Although the general picture is useful, the results cannot be made quantitative and the subjectivity of interpretation one wishes to avoid is still present.

Many of the important contributions of quantum mechanics to valence theory in the rationalisation of the differences of molecular properties between molecules have been based on the concept of net atomic charges and bond orders. In empirical valence theory, it has long been the practice to denote atoms which are supposedly deficient, or have an excess, of charge, relative to the free atom, by the symbols $\delta+$ and $\delta-$ e.g. $A^{\delta+} - B^{\delta-}$, and the concept of bond order is evident at its crudest level in the molecular valence diagram in which each electron pair bond, denoted by a single line between the atoms, contributes one unit to the bond order. Quantification of these concepts is achieved by the partition of the total number of electrons, n, via the density function.

$$n = \int P(1)\, dV_1 = \sum_{ij} P_{ij} S_{ij}$$

where $$S_{ij} = \int \phi_i(1)\phi_j(1)\, dV_1.$$

Thus, we have the Mulliken population analysis[7], corresponding to the first method of partition, giving the gross atom population

$$n_A = \sum_{ij\varepsilon A} P_{ij} S_{ij} + \sum_{B} \sum_{i\varepsilon A} \sum_{j\varepsilon B} P_{ij} S_{ij}$$

and the bond order between atoms A and B, n_{AB},

$$n_{AB} = 2 \sum_{i\varepsilon A} \sum_{j\varepsilon B} P_{ij} S_{ij}$$

The addition of the interatomic terms to n_A is necessary so that the sum of the atom populations equals the number of electrons, n. An equipartition of the bond charge n_{AB} between the two atoms is assumed in the Mulliken method. Various alternative methods have been proposed which achieve a more realistic partition[8].

All orbital wavefunctions are based ultimately on non-orthogonal atomic orbitals, so that the Mulliken population analysis is always possible. But, in most semi-empirical and some ab initio methods the most convenient basis for the expansion of the density function is an orthogonal one. Then all n_{AB} are zero since $S_{ij} = \delta_{ij}$. It is customary in these cases to define the bond order as simply

$$n_{AB} = \sum_{i\varepsilon A} \sum_{j\varepsilon B} P_{ij}$$

It might be considered unrealistic to base a partition on orthogonal orbitals since orthogonalisation of the atomic orbital bases mixes orbitals on different centres. In practice, it appears that the spatial extent of the orbitals is little affected and that the resulting parameters are as meaningful as the ones based on the non-orthogonal basis.[9] An advantage is that no partition problem is then encountered in the definition of atom populations since the sum of the diagonal elements P_{ii} sum to the number of electrons directly.

It is difficult, as we have discussed, in the Mulliken type of analysis to define uniquely which orbitals contribute to which each bond, a problem avoided by using a wavefunction partitioned such that it yields directly bond density functions. When the density is so divided, it is an easy matter to construct unambiguous bond orders and bond polarities by performing a Mulliken analysis on each bond density function individually. This is an important step since bond polarities are likely to be more informative than total atom populations and the bond orders more reliable.

Let us look at the partitioning of the number of electrons for the various wavefunctions considered in the introduction. The first and most striking point is that the VB wavefunction, despite the chemical basis on which it is set up, is, in practice, not very informative since a formal bond partition is not built into the wavefunction. Thus only a Mulliken partition, with all its deficiencies, may be used. It is interesting that the MO wavefunction is an improvement in this respect on the VB wavefunction, since the density function may be cast in a form which brings out the bond picture of the molecule. Thus

$$P(1) = \sum_i \phi_i(1)\phi_i(1)$$

where the ϕ_i are the delocalised MO's. But, by a unitary transformation, the MO's may be localised without destroying the diagonal form of the density function.

$$P(1) = \sum_I \chi_I(1)\chi_I(1) = \sum_I P^I(1)$$

where the χ_I are localised MO's. Thus the density is in a sum-of-bond-densities with, one might think, the ability to define bond orders and bond polarities. The situation is unfortunately not so clear-cut: although the χ_I are localised, they do contain small secondary contributions from other atoms. These secondary contributions represent the interactions between bonds, and, as in the Mulliken analysis, are not readily separated out or unambiguously assigned to other bonds. Bond quantities are therefore not directly definable in the MO method.

The simple SBP wavefunction is of particular interest in this connection since the density function may also be written in a sum-of-bond-densities form[4]. However, the method of construction of the pair functions ensures that they are completely localised within the bond (disregarding "orthogonality" delocalisation). Thus bond orders and polarities are readily defined. A fundamental disadvantage of the SBP method is that the orthogonality constraint precludes the allocation of an orbital to more than one bond function. But one of the main advantages of the wavefunction partition method was precisely that the variational process could sort out the contribution of one orbital to different bond functions. That this is not allowed reduces the usefulness of the SBP method to valence theory. There is a further disadvantage of the SBP method: although the bonds are more accurately described than in the MO method, the one-electron interactions between bonds is completely neglected. The NPF wavefunction, on the other hand, due to the non-orthogonality of the basis, includes these interactions and separates out the bond and interbond interaction densities. From Equation (23) of Section 2,

$$P(1) = \sum_{I} P^{I}(1) + \sum_{IJ}{}' P^{IJ}(1)$$

where

$$P^{I}(1) = \sum_{ij} \phi_i^{I}(1)\phi_j^{J}(1)\,[\mathrm{CSC}(I,I)_{ij}D(I,I,I) - \sum_{K\neq I} D(K,I,I)\times \mathrm{CSC}(K,I)_{ij}]$$

and

$$P^{IJ}(1) = \sum_{ij} \phi_i^{I}(1)\phi_j^{J}(1)\,[\mathrm{CSC}(J,I)_{ji}D(J,J,I) - D(I,J,I)\times \mathrm{CSC}(I,I)_{ji}$$

$$+ \sum_{K\neq I,J} D(K,J,I)\times \mathrm{CSC}(K,I)_{ji}]$$

By integration over the coordinates of electron 1, the electron count partition is achieved. Thus

$$\int P(1)dV_1 = n = \sum_{I} n_I + \sum_{IJ}{}' n_{IJ}$$

where

$$n_I = \int P^{I}(1)dV_1 \ , \qquad n_{IJ} = \int P^{IJ}(1)dV_1$$

Now $$P^I(1) = \sum_{ij} P^I_{ij} \phi^I_i(1) \phi^I_j(1)$$

and $$P^{IJ}(1) = \sum_{ij} P^{IJ}_{ij} \phi^I_i(1) \phi^J_j(1)$$

so that

$$n_I = \sum_{ij} P^I_{ij} S^{II}_{ij}$$

and $$n_{IJ} = \sum_{ij} P^{IJ}_{ij} S^{IJ}_{ij}$$

A Mulliken population analysis of the n_I, n_{IJ} will yield bond orders and bond polarities. The bond interactions or long range bond orders of MO theory have so far not provided any useful results: the explicit interbond 'bond' orders and interaction functions of the NPF wavefunction may be more informative.

It is interesting to look at the orders of magnitude and signs of the contributions to the bond and interbond density functions. The coefficients of the atomic orbitals in the bond density may be expanded in powers of interpair overlap giving terms of order $1 - S^2 + S^3$ ... whilst, for the interbond densities, the expansion series is $-S + S^2 + S^3$ ... . Thus, for small interbond overlap, the bond density is positive and lays down a basic electron density, and the interbond density is negative, in general, subtracting out electron density from the regions between the bonds. With this partition, since the one-electron density is negative, the one-electron energy associated with it will, except possibly for long range interactions, contribute to the repulsion between bonds.

So far we have only considered the partitioning of the electron number, n, but the NPF wavefunction partition also provides a separation of the electron density and hence gives the opportunity to plot out the densities of the different parts of the molecule, i.e. bond and interbond, to "see what the molecule looks like". Although valuable for forming a mental picture of the molecular electronic structure, they cannot, like the density difference maps, be more than a guide to our construction of hypothesis about the nature of the bonding which must be tested by numerical analysis of the partition of the expression for the particular property. For example, for the geometry of the ammonia molecule, the energy is the property of interest. This may be partitioned, for an NPF wavefunction as

$$
\begin{aligned}
E = & \sum_{I} (f_{II} + g_{IIII} + n_{II}) \\
& + \sum_{IJ}{}' (f_{IJ} + g_{IIJJ} + g_{IJIJ} + g_{IIIJ} + g_{IJJJ} + n_{IJ}) \\
& + \sum_{IJK}{}' (g_{IJKK} + g_{IKJJ} + g_{JKII}) \\
& + \sum_{IJKL}{}' g_{IJKL}
\end{aligned}
$$

where the one-electron terms f_{II}, f_{IJ} are derived from the partition of the one-electron density

$$
\begin{aligned}
\int P(1)\hat{f}(1)dV_1 &= \sum_{I} \int P^{I}(1)\hat{f}(1)dV_1 + \sum_{IJ}{}' \int P^{IJ}(1)\hat{f}(1)dV_1 \\
&= \sum_{I} f_{II} + \sum_{IJ}{}' f_{IJ}
\end{aligned}
$$

where f(1) is the sum of the kinetic energy and nuclear attraction operators, and the g_{IJKL} terms from an analogous partition of the two-electron density function with

$$g_{IJKL} = \int P^{IJKL}(12)\ r_{12}^{-1}\ dV_{12}$$

$\sum_{IJ} n_{IJ}$ is the nuclear repulsion energy. If such an analysis is carried out on the energy of the planar and pyramidal configurations of ammonia, it should be possible to explain why ammonia is, in fact, pyramidal. A similar type of analysis has been proposed by Clementi for HF and MC SCF wavefunctions[10] but, because of reasons already discussed, it is necessary to reduce the analysis to an 'atomic' base rather than the more acceptable 'bond' base.

It is thus suggested that, if any real advance is to be made in quantum mechanical valence theory and explanations as well as rationalisations of molecular properties is to be forthcoming, the study of the partition of observables for existent as well as non-existent molecules must be made by the methods outlined above.

We have considered the molecule as a whole to be broken down into fixed entities - bonds, lone pairs, etc. So far we have not considered the question of why bonds form at all. The testing ground for theories about the reason for the chemical bond is the hydrogen molecule. The Schrödinger equation for this two-electron system - the prototype of the two-electron bond - is

$$\hat{H}(12)\ \Psi(12) = E\ \Psi(12)$$

where

$$\hat{H}(12) = -\tfrac{1}{2}\nabla^2(1) - \tfrac{1}{2}\nabla^2(2) - \frac{1}{r_{1A}} - \frac{1}{r_{2A}} - \frac{1}{r_{1B}} - \frac{1}{r_{2B}} + \frac{1}{r_{12}} + \frac{1}{r_{AB}}$$

is the molecular Hamiltonian. Binding may be explained in terms of the more rapid increase in the attractive nuclear attraction energy compared to that of the repulsive kinetic, and electron and nuclear repulsion energies, as the equilibrium internuclear separation is approached. Once it is passed, the increase in the kinetic energy is dominant so that energy curve becomes repulsive.[11] This is the reason for the hydrogen molecule bond, but is it true for the bonds in all molecules? Generally, it is not possible to set up a Schrödinger equation for a bond analogous to that for the hydrogen molecule, but, in the NPF method, it is feasible because the wavefunction is based on a product of two-electron functions, with no orthogonality constraints. If all pair functions except one are known, then an effective Hamiltonian operator may be constructed such that

$$H^{I}(12)\Lambda^{I}(12) = E_{I}\,\Lambda^{I}(12)$$

where $\Lambda^{I}(12)$ is the unknown antisymmetrised pair function, and is subject to no restrictions. The derivation of the expression for $H^{I}(12)$ is straightforward, proceeding as for the derivation[12] of the Hartree-Fock effective Hamiltonian, but, as it is rather complex due to the permutation operators, we shall not give it in detail. This pseudo-Schrödinger equation for bond I may be solved by the usual expansion techniques - this is in effect what has been done for two pair functions in Appendix 2. The nature of any chemical bond may thus be studied in detail as for the hydrogen molecule in the NPF approximation.

In the SBP and HF methods, the functions that are to be found cannot be written as solutions of unconstrained pseudo-Schrödinger equations, because of the orthogonality constraint. Pseudopotential methods must be used for determining the valence HF orbitals to prevent variational collapse into the core space. The unconstrained form of the pseudo-Schrödinger equation in the NPF method has considerable potential. It may be possible to replace the exact effective Hamiltonian by a model Hamiltonian with approximately the same potential but much easier to solve. It might then even be possible to construct the potential for a bond in a large molecule knowing only the molecular geometry and hence for example find the wavefunction for an O-H bond in a steroid molecule. The excellent results from absurdly simple model potentials for the Rydberg series of molecules and atoms are encouraging,[14] although the two problems are orders of magnitude apart in physical and mathematical complexity.

Since the same type of bond, e.g. CH, in different molecules, usually exhibit extremely similar properties, it might be expected that the bond wavefunctions - pairfunctions - similarly change little from molecule to molecule. Should this be so, a valuable approximation method is added to the NPF method; the transfer of pairfunction coefficients optimised in a small molecule to a larger one where the optimisation process is not feasible. Certainly, this technique has been shown to be fairly accurate in completely localised MO (bond orbital) calculations.[15] Also the SBP method rests in the orthogonality of the atomic orbital basis: thus each hybrid contains contributions - "tails" - from orbitals on all other atoms in the molecule. The 'tails' of the hybrids are obviously not transferable from one molecule to another because of the different geometries. Are then the pair function coefficients transferable? Surprisingly enough, the answer does seem, to reasonable accuracy, to be yes: for example the OH bond of the water molecule may be transfered to hydrogen peroxide without undue loss of accuracy.[16] These facts and the intuitively plausible idea that non-orthogonal bond functions should be more transferable than orthogonal ones suggest that quite a high degree of accuracy may be expected in the transfer of NPF pairfunctions.

The NPF wavefunction has been shown to be, in theory, a potentially powerful tool for the explanation of the basic questions of valence theory and the rationalisation of the differences in the properties of molecules, and to possess clear advantages in this field over most current wavefunctions. Its usefulness will however be limited by the ability to solve the non-orthogonality problem: the methods of Section 3 and the possibility of the transferability of pairfunctions are signs that solutions may not be far away.

REFERENCES

1. Coulson, C.A. (1961) "Valence", p.1, Oxford University Press
2. Although see the detailed studies of Ruedenberg and coworkers
   (a) Rue, R. & Ruedenberg, K. (1964) J.Phys.Chem. 68, 1676
   (b) Feinberg, M.J., Ruedenberg, K. & Mehler, E.L. (1970) "Advances in Quantum Chemistry" Vol. 5, p.27, Academic Press.
3. Baird, N.C. & Whitehead, M.A. (1966) Theoret.chim.Acta 6, 167
4. (a) Hurley, A.C., Lennard-Jones, J.E. & Pople, J.A. (1953) Proc.Roy.Soc. A220, 446
   (b) Klessinger, M. & McWeeny, R. (1965) J.Chem.Phys. 42, 3343
5. MacLagan, R.G.A.R. (1971) Chem.Phys.Letters 8, 114
6. Bader, R.F.W., Keeveny, I. & Cade, P.E. (1967) J.Chem.Phys. 47, 3381
7. Mulliken, R.S. (1955) J.Chem.Phys. 23, 1833
8. Doggett, G. (1969) J.Chem.Soc.A. 229
9. Cook, D.B., Hollis, P.C. & McWeeny, R. (1967) Mol.Phys. 13, 553
10. Clementi, E. (1967) J.Chem.Phys. 46, 3842
11. Slater, J.C. (1963) "Quantum Theory of Molecules and Solids I", p.34, McGraw-Hill.
12. Reference 11, p. 256
13. Weeks, J.D., Hazi, A. & Rice, S.A. (1969) "Advances in Chemical Physics" Vol. 16, p.283, Wiley, Interscience.

14. Reference 13, p.307

15. Hoyland, J.R. (1969) J.Chem.Phys. 50, 473

16. Levy, M., Strauss, W.J., Shull, H. & Hagstron, S. (1970) J.Chem.Phys. 52, 5483

CHAPTER 2

SPIN THEORY FOR ELECTRON PAIR BONDS

1. Introduction

The inaccessibility of closed form solutions of Schrodinger's equation for systems with more than one electron has meant that approximation methods must be used in the construction of many electron wavefunctions. The basis of these techniques is the superposition method in which the wavefunction is constructed from a complete but arbitrary set of functions via the variational theorem. The wavefunction for an N electron system, in the Born-Oppenheimer approximation, is a function of 4N variables - 3N electron position coordinates $(x_i y_i z_i)$ and N electron spin coordinates (s_i). Thus the method of superposition of configurations gives, for the total electronic wavefunction,

$$\Psi = \sum_{i}^{n} a_i \hat{A} \Psi_i (x_1 y_1 z_1 x_2 \ldots z_n s_1 s_2 \ldots s_n)$$

where Ψ_i is one of the configurations from the basis set, $\hat{A}$ is the permutational operator that ensures that Ψ_i has the required antisymmetry with respect to the interchange of electrons, and a_i is a coefficient to be determined variationally. The space and spin variables may be separated to the extent that

$$\Psi = \sum_{i}^{n}\sum_{j}^{m} a_{ij}\, \hat{A}\, \psi_i(x_1y_1z_1\ldots z_n)\theta_j(s_1\ldots s_n) \qquad 1$$

As n tends to infinity so Ψ tends to the exact wavefunction: on the other hand, for practical caculations, m is always finite.

Most of the work that has gone into the construction of approximations to the exact wavefunction has been focussed on the problems of obtaining accurate spatial wavefunctions ψ_i, the summation over spin functions θ_j usually being truncated to a single term. This is either because the form of the approximate spatial wavefunction is such that only one spin function can be constructed - the case with closed shell Hartree-Fock theory - or because chemical 'evidence' suggests that one term dominates the expansion - for example, the 'perfect pairing' approximation in Valence Bond (VB) theory, in which each electron pair bond is associated with a singlet spin function. Also, the properties of general interest - energy, electron distribution, etc., - appear to be rather insensitive to the refinement of the spin part of the wave function for the even electron molecules in singlet ground states most frequently studied: however, spin dependent properties and the electronic structure of molecules with odd electrons or even electron molecules in states of higher spin multiplicities will depend more critically on the approximations made in the spin wavefunction. This has been shown by calculations on the spin density at the nucleus of small atoms[1] and in the theory of the proton hyperfine interaction in pi electron radicals[2].

Determination of the optimum wavefunction with respect to the spin variables is, in theory, possible since the sum in Equation 1 over θ_j is finite for finite N. However, it rapidly becomes an unmanageable problem as the number of electrons increases. To give some idea of the extent of the problem, the number of independent spin functions that may be constructed, for a singlet state, increases by approximately a factor of four as a further pair of electrons is added. Thus, just as in the construction of the spatial wavefunction, practical considerations dictate that a truncation of the sum over spin wavefunctions is made.
The basis functions θ_j, from which the spin wavefunction is constructed, may be chosen in infinitely many different ways. This arbitrariness may be taken advantage of by choosing a set of θ_j such that the sum over j of Equation 1 converges as quickly as possible so that only a few terms

need be included. We take 'convergent' to mean, in this context, that, in the case where the spatial wavefunction is approximated by a single term, if the spin wavefunctions can be arranged in decreasing order of importance i.e. $c_{11} > c_{12} > \ldots c_{ij} > \ldots c_{1m}$, then the sum is convergent. 'Convergence' cannot however be understood in the rigorous mathematical sense with well defined intervals between subsequent terms since the θ_j are unrelated. The greater the disparity between subsequent terms, in a general sense, the better is the convergence, and hence the better is the approximation of the spin wavefunction truncated to $m'<m$ terms to the exact spin wavefunction. That certain approximations to the spatial wavefunction fall into this category is implicitly assumed, e.g. VB and configuration interaction (CI) theories, and is found in practice to be true.

Now, the question of whether a particular choice of basis functions leads to a convergent series depends entirely on the system being studied and the approximate spatial wavefunction used: a basis which is convergent in one case may not be so in another. In this chapter, we shall suggest a solution to this problem for molecules which are described chemically as consisting of independent, localised two-electron bonds, for example H_2O, C_2H_6, SF_6. In recent years, there has been a great deal of interest in this class of molecule, based on the use of pair or group function methods[3], whereby the spatial electron distribution of each bond is described in terms of a localised two-electron (pair) function. We shall, in the second part of the chapter, combine this with the spin theory presented in the first part to produce a method capable of describing accurately both the space and spin electron distribution of pair bonded molecules: the theory is then extended to cover the positive ions of these molecules.

2. Current Spin Function Bases

Two different approaches are at present made to the problem of constructing and using N electron spin eigenfunctions. The first approach has lead to two different formulations due to Kotani[4] and Löwdin[5]. The vector coupling method of Kotani builds up spin eigenfunctions for the N electron system with total spin S and component M, θ^N_{SM} from the appropriate N-1 electron spin eigenfunctions according to formulae of the type

$$\theta^N_{SM} = K_1 \, \theta^{N-1}_{S+\frac{1}{2}\;M-\frac{1}{2}} \, \alpha + K_2 \, \theta^{N-1}_{S-\frac{1}{2}\;M+\frac{1}{2}} \, \beta$$

Starting from the one electron states, α and β, electrons are added one by one using the vector coupling formulae until the required N electron spin eigenfunction is reached. On the other hand, Löwdin has proposed a method in which the spin eigenfunctions are obtained by a projection operator technique i.e.

$$\theta^N_{SM} = \hat{O}_{spin}\, \theta^N_M$$

where

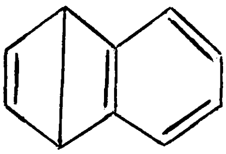
$$\hat{O}_{spin} = \prod_{T \neq S} \frac{\hat{S}^2 - T(T+1)}{S(S+1)-T(T+1)}$$

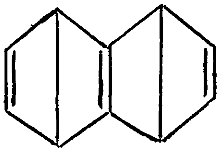
is the projection operator and $\hat{S}^2$ is the square of the spin angular momentum operator. θ^N_M is an arbitrary N electron spin function with the required component M.

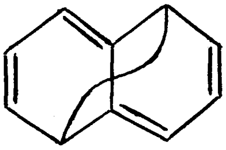
These methods are closely related and can in fact be shown to be equivalent on orthogonalisation of the Löwdin spin functions[5]. This approach is concerned only with the generation of the complete set of spin functions, and is not designed for use with any particular form of the spatial wavefunction. It is not possible, without a calculation, to pick out the most important terms in the spin wavefunction expansion: the approach is therefore impractical except in the case of a very small number of electrons where the whole basis may be used without prejudice.

We have not yet indicated by what procedure one is to classify spin functions as important or unimportant. The selection of the dominant terms from a complete set is a perennial problem in practical quantum chemistry and not soluble in exact terms. The normal procedure is to choose a basis, space and spin, which mirrors in some sense the physical conception of the molecular structure: those terms which represent most closely this picture are assumed to be the important ones, and the more a basis function deviates from the chemical picture, the less important it is.

The second approach to spin function construction - the bond eigenfunction method due to Rumer[6] - falls into this category and is hence well suited to practical application. Combined with a spatial wavefunction of singly occupied directed orbitals, Rumer's method yields the classic Valence Bond theory of quantum chemistry. A singlet spin function, $\alpha(1)\beta(2)-\beta(1)\alpha(2)$, is associated with a pair of atomic orbitals: these orbitals are then said to be 'bonded'. Different partitions of the atomic orbitals of the molecule into pairs give rise to different spin functions and different molecular 'bond' patterns. That spin function in which all the paired orbitals correspond to the chemical bonds is the major term in the spin wavefunction expansion i.e. the 'perfect pairing' term (where spatial symmetry does not require a combination of the spin functions to be taken). The more paired orbitals that are associated with chemical bonds, the more important is that spin function, and the more remote they are - viz., a 'long' bond, the less important is that spin function. For example, in naphthalene, the bonding pattern (i) is considered more important than either (ii) or (iii). The setting up of the spin functions by the Rumer method is essentially simple: the rules for the generation of a linearly independent set are well known. The disadvantages are that the method is really only well suited to the study of planar systems such as the treatment of the pi electronic structure of aromatic hydrocarbons. For non-planar molecules such as saturated hydrocarbons, the construction of symmetry adapted functions is a complex problem. Secondly, the bond eigenfunction


(i)


(ii)


(iii)

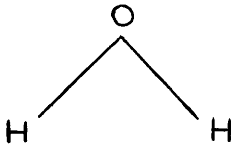
method is designed for use with VB wavefunctions: the combination of this spin function basis with pair function wavefunctions leads to considerable difficulties. In its usual form, the method is restricted to the construction of singlet spin functions only, and the spin eigenfunctions are non-orthogonal which may be a drawback in developing a general theory.

The spin functions introduced in the next section are specifically designed to overcome these problems especially when combined with a pair function approximation to the spatial wavefunction. The basis is intended for the study of molecules with localised two-electron bonds, and is as unsuited on spatial symmetry grounds, to the study of planar pi electron systems as the Rumer set is to sigma bonded molecules.

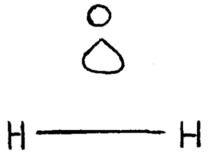
3. Construction of the Spin Functions

The necessary properties of a practical approach to the problem of spin wavefunction optimisation are that the basis of spin functions, as well as being easily constructed for an arbitrary number of electrons and arbitrary spin, is capable, in conjunction with a suitable spatial wavefunction, of a physical interpretation, in a manner analogous to the 'short' bond - 'long' bond classification of the Rumer-VB method, allowing the more important terms to be selected. Also, the convergence of the series should be fast. This last property cannot be guaranteed nor a precise definition given to 'fast' except by calculation in a given situation: however, the nature of the physical interpretation does suggest that the basis possesses, in a general sense, the property of rapid convergence. It should be emphasised that, whilst we must consider the form of the spatial wavefunction when we wish to choose a truncated expansion, the method of spin eigenfunction construction is quite general: the basis may be used with any spatial wavefunction although, of course, the interpretation may then no longer be valid.

In the development of the spatial wavefunction for sigma bonded molecules such as methane, by the VB and generalised methods, the existence of localised bonds is recognised and incorporated from the outset. It would be an advantage if, when developing the spin wavefunction, the spin functions were also constrained to this arrangement. The Rumer method does not allow this possibility since alternative spin couplings to the one of 'perfect pairing' correspond to alternative bonding schemes: for example, the water molecule can be represented by a mixture of a structure (i) with two O-H bonds (the dominant term) with a structure (ii) with a hydrogen-hydrogen bond and an oxygen intra-atom bond.



(i)



(ii)

To derive an alternative basis with the required properties, we look first at the two electron problem. There are four combinations of the two possible one-electron spin wavefunctions, α and β, which are eigenfunctions of the operators $\hat{S}^2$, the square of the total spin angular momentum operator, and $\hat{S}_z$, the z component of angular momentum operator. These are

$$\begin{array}{llcl} S \equiv & P(0,0) & = & \alpha(1)\beta(2)-\beta(1)\alpha(2) \equiv \alpha\beta-\beta\alpha \\ T_1 & P(1,1) & = & \alpha\alpha \\ T_0 & P(1,0) & = & \alpha\beta+\beta\alpha \\ T_{-1} & P(1,-1) & = & \beta\beta \end{array}$$

where, in $P(a,b)$, a is the total spin

i.e. $\hat{S}^2 P(a,b) = a(a+1)P(a,b)$

and b the z component of the spin

i.e. $\hat{S}_z P(a,b) = b\, P(a,b)$

S is the singlet function and the set T the three components of the triplet.

(Here, and in the following, where electron spin coordinates are suppressed, the labelling is assumed to follow sequentially. We shall use the unnormalised convention for P(0,0) and P(1,0) to simplify numerical coefficients).

The set of functions, for a 2N electron system,

$$\theta^N_{M,i} = \prod_{I=1}^{N} P(x^i_I, y^i_I) \qquad i=1,\ldots f_{NM}$$

obtained by taking all possible products of N pair spin functions P(a,b) such that $\sum_J y^i_J = M$, form a basis for the construction of spin eigenfunctions θ^N_{SM} with $S = M, M+1, \ldots N$. For example, for N=2, M=0, the basis set is

$$\theta^2_{0,1} = SS \qquad \theta^2_{0,2} = T_0 T_0$$

$$\theta^2_{0,3} = T_1 T_{-1} \qquad \theta^2_{0,4} = T_{-1} T_1$$

That these basis functions possess the property of retaining the chemical bond pattern is obvious if one expresses the spatial wavefunction in the form

$$\psi = \prod_{I=1}^{N} \Delta^I (2I-1, 2I)$$

where $I \equiv x_I y_I z_I$.

Instead of associating different spin couplings with different bond patterns, one describes them in terms of different assignments of the singlet and triplet pair spin functions (PSF) to the fixed bond pattern. We have now to show how the spin functions of a given spin are constructed, and to prove that they are linearly independent.

The problem is to find those combinations of the basis functions $\theta^N_{M,i}$ which are eigenfunctions of the spin angular momentum operator $\hat{S}^2$ with eigenvalue S(S+1)

$$\text{i.e.} \qquad \hat{S}^2\,\theta^N_{SM} = S(S+1)\,\theta^N_{SM}$$

$$\text{where} \qquad \theta^N_{SM} = \sum_i a^S_i\,\theta^N_{M,i}$$

It can readily be shown that

$$\hat{S}^2 = \hat{S}_-\hat{S}_+ + \hat{S}^2_z + \hat{S}_z$$

$$\text{where} \qquad \hat{S}_+ = \hat{S}_x - i\hat{S}_y$$

$$\text{and} \qquad \hat{S}_- = \hat{S}_x + i\hat{S}_y$$

are the usual step-up and step-down operators. Each of these 2N electron operators is a sum of one-electron operators: thus, for example,

$$\hat{S}_z = \sum_{i=1}^{2N} \hat{S}_z(i)$$

where $\hat{S}_z(i)$ operates on the spin function of electron i only. Straightforward expansion gives the following results

$$\hat{S}_-\hat{S}_+ = \sum_i^N \sum_j^N \hat{T}_{ij} \quad , \quad \text{where}$$

$$\hat{T}_{ij} = \hat{S}_-(2i-1)\hat{S}_+(2j-1) + \hat{S}_-(2i-1)\hat{S}_+(2j) + \hat{S}_-(2i)\hat{S}_+(2j-1) + \hat{S}_-(2i)\hat{S}_+(2j)$$

$$\hat{S}_z = \sum_i^N \hat{U}_i \quad , \quad \text{where}$$

$$\hat{U}_i = \hat{S}_z(2i-1) + \hat{S}_z(2i)$$

By denoting, for example, $\alpha(2i-1)\beta(2i) + \beta(2i-1)\alpha(2i)$ by $P_i(1,0)$, and using the well known rules for operating with $\hat{S}_+(i)$, $\hat{S}_-(i)$ and $\hat{S}_z(i)$, we find that

$$(\hat{T}_{ij}+\hat{T}_{ij})P_i(a,b)P_j(c,d) = 4(P_i(a,b-1)P_j(c,d+1) + P_i(a,b+1)P_j(c,d-1))/(1+|b|)(1+|d|) \qquad 2$$

where $P(x,y) = 0$ if $|y|>x$ or $|y|>1$,

$$\hat{T}_{ii}\, P_i(a,b) = (1+a+b)(a-b)P_i(a,b) \qquad 3$$

and

$$\hat{U}_i\, P_i(a,b) = b\, P_i(a,b) \qquad 4$$

We are now in a position to prove the following.

If θ^{N}_{SM} is an eigenfunction of $\hat{S}^2$ with spin S and z component M, then the spin function obtained by introducing the pair spin function P(0,0) in <u>any position</u> is also an eigenfunction of $\hat{S}^2$ with the same values of S and M. If the spin function θ^{N}_{SM} is a linear combination of terms, then P(0,0) is to be introduced in the same position in all terms.

The theorem follows directly from the expansion of $\hat{S}^2$ embodied in equations 2 to 4. If the inserted PSF is $P_k(0,0)$ and the resulting spin function $\theta^{N+1}_{S'M}$ then

$$\hat{S}^2\,\theta^{N+1}_{S'M} = \left[\sum_{i<j}^{N+1}\sum(\hat{T}_{ij}+\hat{T}_{ji}+2\hat{U}_i\hat{U}_j) + \sum_{i}^{N+1}(\hat{T}_{ii}+\hat{U}_i^2+\hat{U}_i)\right]\theta^{N+1}_{S'M}$$

$$= \left[\sum_{i<j\neq k}^{N+1}\sum(\hat{T}_{ij}+\hat{T}_{ji}+2\hat{U}_i\hat{U}_j) + \sum_{i\neq k}^{N+1}(\hat{T}_{ii}+\hat{U}_i^2+\hat{U}_i)\right]\theta^{N+1}_{S'M}$$

$$+ \left[\sum_{i\neq k}^{N+1}(\hat{T}_{ik}+\hat{T}_{ki}+2\hat{U}_i\hat{U}_k) + \hat{T}_{kk} + \hat{U}_k^2 + \hat{U}_k\right]\theta^{N+1}_{S'M}$$

The upper line gives $S(S+1)\theta^{N+1}_{S'M}$ since θ^{N}_{SM} is an eigenfunction of $\hat{S}^2$ with total spin S: the lower line gives zero by virtue of the definitions (2), (3) and (4). Thus $\theta^{N+1}_{S'M}$ is an eigenfunction of $\hat{S}^2$, with S' = S.

The rules for the construction of the 2N electron spin eigenfunctions of any S and M can now be given. Firstly, set up all θ^R_{SM} with R<N and involving only triplet PSFs. Interpose in all possible ways the N-R singlet PSFs. That these are eigenfunctions with the correct S and M follows from the above theorem. This generates $\sum_R {}^NC_R$ of the spin eigenfunctions. The remainder which involve only triplet PSFs - 'all-triplet' spin functions - are constructed from the vector coupling formulae given below. These are adapted from the formulae given by Kotani[4] and require knowledge of the 'all-triplet' spin functions for R<N.

$$\theta^N_{SM} = K_{11}\,\theta^{N-1}_{S+1\ M-1}\,T_1 - K_{12}\,\theta^{N-1}_{S+1\ M}\,T_0 + K_{13}\,\theta^{N-1}_{S+1\ M+1}\,T_{-1} \qquad 5$$

$$\theta^N_{SM} = K_{21}\,\theta^{N-1}_{S-1\ M-1}\,T_1 + K_{22}\,\theta^{N-1}_{S-1\ M}\,T_0 + K_{23}\,\theta^{N-1}_{S-1\ M+1}\,T_{-1} \qquad 6$$

$$\theta^N_{SM} = -K_{31}\,\theta^{N-1}_{S\ M-1}\,T_1 + K_{32}\,\theta^{N-1}_{SM}\,T_0 + K_{33}\,\theta^{N-1}_{S\ M+1}\,T_{-1} \qquad 7$$

where

$$K^2_{11} = \frac{(S-M+1)(S-M+2)}{(2S+2)(2S+3)}$$

$$K^2_{12} = \frac{(S-M+1)(S+M+1)}{(2S+2)(2S+3)} \qquad K^2_{13} = \frac{(S+M+1)(S+M+2)}{(2S+2)(2S+3)}$$

$$K^2_{21} = \frac{(S+M)(S+M-1)}{2S(2S-1)} \qquad K^2_{22} = \frac{(S+M)(S-M)}{2S(2S-1)}$$

$$K_{23}^2 = \frac{(S-M)(S-M-1)}{2S(2S-1)} \qquad K_{31}^2 = \frac{(S-M+1)(S+M)}{2S(S+1)}$$

$$K_{32}^2 = \frac{M^2}{2S(S+1)} \qquad K_{33}^2 = \frac{(S-M)(S+M+1)}{2S(S+1)}$$

There is a slight error in the last formula on page 135 of reference 4.

That the spin functions so obtained are linearly independent follows from the orthogonality of the spin functions. Those with a different number of P(0,0) PSFs or with P(0,0) PSFs interposed in different positions are orthogonal by virtue of the orthogonality of the PSFs. The 'all-triplet' spin functions generated by the vector coupling formulae (5) to (7) are also orthogonal.

Although the method of construction bears, on the surface, a close resemblance to that of Kotani, the number of spin functions which must be generated by the vector coupling method is, in fact, small, so that the construction of the complete set of spin eigenfunctions by this method is a very much easier task than it is by either the Kotani or Löwdin methods. The results in Table I, where the number of 'all-triplet' spin functions for S = 0 and various values of N is compared with the total number of spin functions, underlines the value of the simple theorem proved above.

Table I

N	Total number of spin functions	'All-triplet' spin functions
4	14	3
5	42	6
6	132	15
7	429	36
8	1430	91

It may be noted that the notation used results in compact expressions for the θ^N_{SM} which is an advantage for large N where the expressions in α, β form are long and complex. Equations 2 to 4 yield a ready means of checking the expressions for θ^N_{SM}.

As an example, we shall derive the set of spin functions for six electrons in a singlet state. The total number of independent spin functions, with a given S, is found from

$$f^{2N}_S = {}^{2N}C_{N-S} - {}^{2N}C_{N-S-1}$$

For N=3, S=0, there are five spin functions. The first is obviously the 'all-singlet' spin function

$$\theta_1 = \frac{1}{2\sqrt{2}}\,[SSS] \qquad 8a$$

From Equation 5, the 'all-triplet', S=0, spin function for four electrons is

$$\theta^{2}_{00} = \frac{1}{2\sqrt{3}}\,[2T_1T_{-1} + 2T_{-1}T_1 - T_0T_0]$$

which gives θ_2, θ_3 and θ_4 by the introduction of S in the three possible ways

$$\begin{aligned} \theta_2 &= \frac{1}{2\sqrt{6}}\,[2T_1T_{-1}S + 2T_{-1}T_1S - T_0T_0S] \\ \theta_3 &= \frac{1}{2\sqrt{6}}\,[2T_1ST_{-1} + 2T_{-1}ST_1 - T_0ST_0] \\ \theta_4 &= \frac{1}{2\sqrt{6}}\,[2ST_1T_{-1} + 2ST_{-1}T_1 - ST_0T_0] \end{aligned} \tag{8b}$$

The final spin function is an 'all-triplet' one and is generated via θ^{2}_{11}, θ^{2}_{10} and θ^{2}_{1-1} and Equation 5, giving

$$\begin{aligned} \theta_5 = \frac{1}{2\sqrt{3}}\,[&T_0T_{-1}T_1 - T_0T_1T_{-1} + T_{-1}T_1T_0 \\ &- T_1T_{-1}T_0 + T_1T_0T_{-1} - T_{-1}T_0T_1] \end{aligned} \tag{8c}$$

Comparison with the set of spin functions obtained by Kotani's method shows that θ_1, θ_2 and θ_4 are identical to those associated with the branching diagrams ,  and  but that θ_3 and θ_5 are a linear combination of the functions characterised by the diagrams  and .

Extension of the method to non-integer values of S is easily achieved by the introduction of the single electron vector coupling formulae[4],

$$\theta^{N+\frac{1}{2}}_{S+\frac{1}{2}\ M+\frac{1}{2}} = -\sqrt{\frac{S-M+1}{2S+3}}\ \theta^{N}_{S+1\ M}\ \alpha+\sqrt{\frac{S+M+2}{2S+3}}\ \theta^{N}_{S+1\ M+1}\ \beta$$

$$\theta^{N+\frac{1}{2}}_{S+\frac{1}{2}\ M+\frac{1}{2}} = -\sqrt{\frac{S+M+1}{2S+1}}\ \theta^{N}_{SM}\ \alpha + \sqrt{\frac{S-M}{2S+1}}\ \theta^{N}_{S\ M+1}\ \beta$$

The 2N electron spin functions are generated as described above.

4. Interpretation and Classification of Spin Functions

A set of spin functions only provides a practical basis for calculations if the important terms in the expansion of Equation 1 may be selected on intuitive grounds, allowing the series to be truncated after a few terms. For molecules with localised two-electron bonds, an interpretation and selection criterion is readily available for the spin functions based on PSFs. Assuming a pair function wavefunction,

$$\Psi = \hat{A}\ \Delta^1(12)\Delta^2(34)\ldots\Delta^N(2N-1,2N)\sum_J C_J\theta^N_J$$

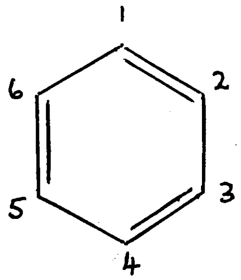
where $\Delta^I(2I-1,2I)$ is a localised two-electron spatial function describing bond I, for example, a Heitler-London function, $a_I(1)b_I(2) + b_I(1)a_I(2)$. Expanding θ^N_I as a product of PSFs,

$$\Psi = \hat{A} \sum_{I} C_I \, [\Delta^1(12)P(x_I^1 y_I^1)\Delta^2(34)P(x_I^2 y_I^2)\ldots]$$

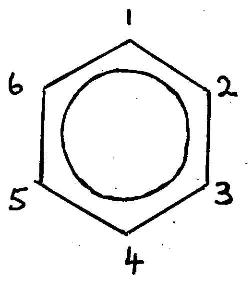
Since a pair function with electron spins coupled to a singlet leads to a build up of electron density in the bond, whilst triplet coupling leads to removal of charge from between the bonded atoms, the spin function in which bonds are triplet coupled and essentially 'broken' will be of high energy and hence only occur with small weights in the total wavefunction. In sigma bonded systems, where the bond singlet coupling is expected to be strong, only terms which involve a few triplet coupled bonds will be important. A natural selection criterion is thus established. It might also be expected that triplet coupling between a pair of adjacent bonds will be more significant than between a well separated pair. As the triplet coupling of a pair of bonds leads to the delocalisation of electron density from the bond regions into the regions behind the atoms and between the bonds, the weight with which such a term enters the wave function may well give a good indication of the degree of interdependence of a pair of bonds. This description of the bonding in terms of the type and amount of coupling within and between bonds may have more physical significance than that afforded by the Rumer method which compares spin functions in terms of the relative importance of hypothetical 'long' bonds, cf. Shull's criterion for electron pair bonds[7].

Although the setting up of the spin functions is not as trivial as in the case of the Rumer method, the possibility of truncation of the basis in a practical situation means that only the simpler spin functions with a few triplet couplings are required: these are readily generated once and for all and can be used for any value of N.

Naturally, the situation as described above only applies to sigma systems: in delocalised systems, such as the pi electrons of aromatic hydrocarbons, it is impossible to specify one dominating bonding scheme, and the spin functions involving many triplet PSFs may be as important as the 'perfect pairing' spin function. The reason for this is easily seen. Taking benzene, in the VB approximation, as an example, the bonding scheme represented by θ_1, Equation 8a, is given by Figure (i).



(i)

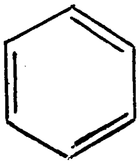


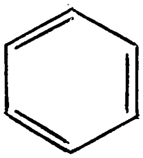
(ii)

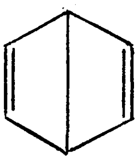
This structure leads to increased electron density between atoms 1 and 2, 3 and 4, and 5 and 6. But it is well known that the true situation is that of Figure (ii) with equal electron densities between atoms 1 and 2, and 2 and 3, and so on. The terms in the wavefunction involving triplet coupled spin functions give rise to an electron density distribution which is decreased between the 'bonded' atoms, e.g. 1 and 2, and increased between the 'non-bonded' atoms, e.g. 2 and 3. To achieve the required equal distribution, large weights of the triplet functions will be needed. A quantitative measure of this may be found by expressing the Rumer functions for benzene in terms of the pair based functions. The Rumer wavefunction is

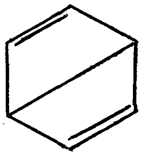
$$\Psi = \hat{A}\,\psi_{\text{space}}\,[(\theta_1^R + \theta_2^R) + \lambda(\theta_3^R + \theta_4^R + \theta_5^R)]$$

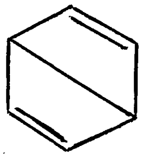
where the θ_i^R represent the bonding schemes


1


2


3


4


5

and λ is a mixing parameter to be determined. Straightforward expansion of the spin functions gives the alternative expression,

$$\Psi = \hat{A}\,\psi_{\text{space}}\left[\left(\frac{5-6\lambda}{4}\right)\theta_1 + \frac{\sqrt{3}}{2}(2\lambda-1)(\theta_2+\theta_3+\theta_4) + \frac{\sqrt{6}}{4}\theta_5\right]$$

where $\theta_1,\ldots\theta_5$ have the same meaning as in Equation 8. Coulson[8] gives 0.434 for the optimum value of λ, which on substitution yields

$$\Psi = \hat{A}\,\psi_{\text{space}}\,[0.60\,\theta_1 - 0.06(\theta_2+\theta_3+\theta_4) + 0.61\,\theta_5]$$

Thus the 'all-triplet' spin function is as important as the 'all-singlet' one. In delocalised systems, we have therefore, besides the problem of projecting out spatial symmetry eigenfunctions, no valid criterion for the truncation of the series in Equation 1.

5. Spatial Symmetry

The problem of setting up molecular wavefunctions involves not only the construction of spin eigenfunctions but also of wavefunctions having the necessary spatial symmetry properties. The complexity of this task depends in large measure on the spin function basis chosen. Consider the symmetry operator $\hat{R}$ working on the wavefunction Ψ

$$\hat{R}\,\Psi = \hat{R}(x)\psi(x)\theta(y)$$

where x stands for the spatial and y the spin variables. We write $\hat{R}(x)$ to show that $\hat{R}$ affects only the spatial coordinates. In the case of systems described by wavefunctions with permutational symmetry i.e. each one or two electron function goes completely into another under all symmetry operations,

$$\hat{R}(x)\psi(x)\theta(y) = \psi(x')\theta(y)$$

where $\psi(x')$ differs from $\psi(x)$ only in the labelling of the electron spatial coordinates. We may think of $\hat{R}(x)$ as a permutational operator, and it follows that, by relabelling the electrons (space and spin),

$$\hat{R}\,\Psi = \psi(x')\theta(y) = \psi(x)\hat{R}^{-1}(y)\theta(y)$$

where $\hat{R}^{-1}(y)$ is the inverse of the permutation operator $\hat{R}(x)$ but now working on the spin variables y. We can therefore, in this case, speak of 'operating on the spin wavefunction with symmetry operators'. The symmetry adapted VB wavefunctions for planar pi systems are readily set up if the Rumer basis is used, e.g. in benzene, one 'Kekulé' spin function transforms into the other under a C_6 rotation. However, if this is not so, we have a much more complicated problem: the bond eigenfunctions for methane do not transform in a permutational manner, and the tedious 'uncrossing' rules must be invoked. It is therefore worthwhile choosing a basis of spin functions which minimises the amount of work required. The spin functions introduced in this Chapter, in conjunction with a bond function spatial wavefunction, have ideal symmetry properties for sigma bonded molecules since each bond of the fixed bonding pattern is maintained under all symmetry operations. Thus the PSFs are never mixed, and the projection of symmetry adapted wavefunctions is a relatively easy task. To show this, and to give an indication of the power and conciseness of the method for cases with relatively large N, we shall consider the problem of spin optimisation for the methane molecule, point group Td, assuming that the spatial wavefunction is approximated by an unspecified single term bond function wavefunction. The fourteen singlet spin functions for an eight electron system are characterised as follows.

(a) $\theta_1 = SSSS$

(b) $\theta_2, \ldots \theta_7$ are the ${}^4C_2 = 6$ ways of inserting two singlet PSFs into the four electron 'all-triplet' function,

$$2T_1T_{-1} + 2T_{-1}T_1 - T_0T_0$$

(c) $\theta_8, \ldots \theta_{11}$ are the ${}^4C_1 = 4$ ways of placing one singlet PSF in the six electron 'all-triplet' function,

$$T_0T_{-1}T_1 - T_0T_1T_{-1} + T_{-1}T_1T_0 - T_1T_{-1}T_0 + T_1T_0T_{-1} - T_{-1}T_0T_1$$

(d) $\theta_{12}, \ldots \theta_{14}$ are 'all-triplet' functions constructed by the vector coupling formulae (5), (6) and (7).

$$\begin{aligned}\theta_{12} = {} & 12(T_1T_1T_{-1}T_{-1} + T_{-1}T_{-1}T_1T_1) \\ & -3(T_0T_{-1} + T_{-1}T_0)(T_0T_1 + T_1T_0) - 3(T_0T_1 + T_1T_0)(T_0T_{-1} + T_{-1}T_0) \\ & +2(T_{-1}T_1 + T_1T_{-1} + T_0T_0)(T_{-1}T_1 + T_1T_{-1} + T_0T_0)\end{aligned}$$

$$\theta_{13} = (2T_1T_{-1} + 2T_{-1}T_1 - T_0T_0)(2T_1T_{-1} + 2T_{-1}T_1 - T_0T_0)$$

$$\begin{aligned}\theta_{14} = {} & (T_1T_0 - T_0T_1)(T_0T_{-1} - T_{-1}T_0) \\ & +(T_0T_{-1} - T_{-1}T_0)(T_1T_0 - T_0T_1) \\ & +2(T_1T_{-1} - T_{-1}T_1)(T_{-1}T_1 - T_1T_{-1})\end{aligned}$$

The projection operator, which projects from an arbitrary wavefunction an eigenfunction of a given spatial symmetry, is

$$\hat{O}_R^{\lambda} = \sum_R \chi_R^{\lambda} \hat{R}(x)$$

where χ_R^{λ} is the character for the operation $\hat{R}$ of symmetry species λ. The sum is over all symmetry operations of the group. As discussed above, we may replace $\hat{R}(x)$ by $\hat{R}^{-1}(y)$. Since, for methane ${}^1A_1$, all χ_R^{λ} are unity,

$$\hat{O}_R^{A_1} = \sum_R \hat{R}(y)$$

There are three totally symmetric states to be found. Since all symmetry operations simply permute the PSFs, it is immediately seen that θ_1 is one of the totally symmetric states. Since the PSFs are not mixed under any of the symmetry operations, $\hat{R}\theta_i$ always remains within the subspace of the set (a),...(d) to which it belongs. Thus $\hat{O}_R\theta_2$ yields the second ${}^1A_1$ state which is the in-phase combination of the spin functions of set (b). It is readily seen that $\hat{O}_R$ operating on any member of set (c) gives zero i.e. there is no ${}^1A_1$ state with three bonds triplet coupled. Similar cancellation occurs with θ_{14}: the third state is a linear combination of θ_{12} and θ_{13}.

Thus

$$\begin{aligned}
\theta_1({}^1A_1) &= \theta_1 \\
\theta_2({}^1A_1) &= \theta_2+\theta_3+\theta_4+\theta_5+\theta_6+\theta_7 \\
\theta_3({}^1A_1) &= 2\theta_{12}+\sqrt{5}\,\theta_{13}
\end{aligned}$$

The actual work of obtaining the symmetry adapted functions is straightforward, if somewhat tedious, involving only permutations of the PSFs. The method could thus be readily implemented on a computer.

From the arguments we have advanced above, the order of importance of the symmetry adapted spin functions will be $\theta_1>\theta_2>\theta_3$: θ_3 could well make a negligible contribution to the total wavefunction.

It is interesting to note that the spin functions for ammonia (C_{3V}) and water (C_{2V}) are very similar to those for methane. $\theta_1(^1A_1)$ is the 'all-singlet' term and $\theta_2(^1A_1)$ is the 'in-phase' combination of spin functions with two bonds triplet coupled: there are no symmetrised spin functions with three bonds triplet coupled.

6. Spin Theory and the Pair Function Method

In the previous sections, we have proposed a new basis of spin functions which is ideally suited for describing the spin properties of molecules with localised two-electron bonds. We have not so far considered in detail the particular form of the spatial wavefunction to be used with these spin functions: any spatial wavefunction may, in principle, be used, but one particular choice - the separated pair wavefunction - is of particular interest since it has been developed with localised two-electron bond systems specifically in mind. It also leads to simple expressions for the matrix elements required in the optimisation of the spin coupling coefficients.

Spatial wavefunctions which recognise, from the outset, the existence of electron pair bonds have long been of interest as a step towards the translation of the chemical picture into quantum mechanical terms. The prototype of this class of wavefunction is the VB wavefunction in which the two electron bond is described by a Heitler-London pair function

$$\Delta(12) = a(1)b(2) + b(1)a(2)$$

It has unfortunately been found that the problems posed by this type of wavefunction are not easily solved. The use of non-orthogonal atomic orbitals presents severe computational difficulties while attempts to orthogonalise the basis orbitals lead to very poor wavefunctions, requiring many ionic structures to correct the imbalance[9]. Recent results[11] have shown that the orthogonality conditions may be validated if the form of the pair function is generalised.

Thus the pair function,

$$\begin{aligned}\Delta(12) = {} & c_{11}a(1)a(2) + c_{22}b(1)b(2) \\ & + c_{12}a(1)b(2) + c_{21}b(1)a(2)\end{aligned} \qquad 9$$

where a and b are now orthogonal atomic orbitals, includes the intra-bond ionic structures and gives a wavefunction of high accuracy - better than the comparable molecular orbital (MO) wavefunction - whilst being able to retain the orthogonality of the atomic orbitals to simplify the practical aspects of the calculation. The pair function wavefunction is

$$\psi = \hat{A}\, \Delta^1(12)\Delta^2(34)\ldots\Delta^N(2N-1,2N)$$

for a molecule with N electron pair bonds, with

$$\Delta^I(12) = \sum_{ij} c^I_{ij}\, \chi^I_i(1)\, \chi^I_j(2)$$

The orthogonality conditions are

$$\int \Delta^I(12)\, \Delta^J(13)\, dv_1 = 0\ , \qquad I \neq J$$

that is, strong or one-electron orthogonality, which means that the pair functions must be constructed from separate subsets of the atomic orbital basis, and in practice usually means that a minimal basis is chosen. The standard procedure is to set up hybrid atomic orbitals, symmetrically orthogonalise them - this has been shown to have little effect on their localisation properties[10] - and then to assign them in pairs to the pair functions.

Current pair function theory only allows for a single spin function - the 'all-singlet' or 'perfect pairing' one, and can only be used for singlet states. By combining pair function theory with the spin functions of Section 3, the method may be extended to allow the treatment of states with $S \neq 0$, whilst retaining the chemical description afforded by pair theory, as well as giving a better description of S=0 states. The optimisation of the spin wavefunction of even electron saturated molecules with S≠0 will, of course, improve the electron spin distribution. This may be of great importance in the investigation of magnetic effects on which there is a wealth of experimental data. Also the introduction of alternative spin couplings will improve the charge distribution, by delocalising electron density out of the bond regions. The combination of PSF based spin functions with pair theory is of particular value since it leads to simple expressions for the matrix elements, making calculations, which would be extremely complex if the Rumer, Löwdin or Kotani spin bases were used, computationally tractable. It must be borne in mind that the spin function optimisation is necessarily relative to the spatial wavefunction, due to the incomplete separation of space and spin in Equation 1. The form of the spatial wavefunction determines exactly how good the description of spin properties is. The same is true of the improvement of the charge distribution. It is found that the introduction of alternative spin couplings does not introduce any of the inter-bond terms not included in the simple pair theory density function.

Should any of these terms be vital to a good description of the charge distribution, an alternative method will be required.

Since the spin functions are based on singlet and triplet PSFs, symmetric and antisymmetric space pair functions are needed to give the total wavefunction the obligatory antisymmetry. In an arbitrary basis, this would make it necessary to perform a combined triplet pair function - structure coefficient optimisation. However, in the simple case of a minimal basis, the triplet pair function contains <u>no</u> variational parameter i.e.

$$\Delta^{I}(12) = a(1)b(2) - b(1)a(2)$$

The tacit assumption that the pair functions optimised for the single 'perfect pairing' spin function may be transferred without change has been made above: it does not seem likely that the improvements gained by reoptimisation of the pair functions would justify the vastly increased work this would entail.

The general spin optimised separated pair wavefunction for a 2N electron molecule is

$$= \sum_{i} a_i \hat{A}_p \Psi_i$$

$$= \sum_{i} a_i \hat{A}_p \left[\prod_{I=1}^{N} \Delta_i^{I}(2I-1,2I)\theta_i^{I}(2I-1,2I)\right]$$

The coefficients a_i are parameters to be determined variationally, subject to the spin and space symmetry requirements. $\hat{A}_p$ is the partial antisymmetriser which includes only permutations between pair functions.

$$E = \int \Psi \hat{H} \Psi \, d\tau \Big/ \int \Psi \Psi \, d\tau$$

$$\sum_i \sum_j a_i a_j H_{ij} \Big/ \sum_i \sum_j a_i a_j S_{ij}$$

Optimisation of the a_i leads to the standard secular equation, where the matrix elements are fixed linear combinations of the

$$H_{ij} = \int \Psi_i \hat{H} \hat{A}_p \Psi_j \, d\tau$$

and
$$S_{ij} = \int \Psi_i \hat{A}_p \Psi_j \, d\tau$$

Assuming that the space pair functions are normalised, we have

$$H_{ij} = \int \prod_{I=1}^{N} \Delta_i^I(2I-1,2I)\,\theta_i^I(2I-1,2I)\,\hat{H}\,\hat{A}_p$$

$$\times \prod_{J=1}^{N} \Delta_j^J(2J-1,2J)\,\theta_j^J(2J-1,2J)\,d\tau$$

$$S_{ij} = \int \prod_{I=1}^{N} \Delta_i^I(2I-1,2I)\,\theta_i^I(2I-1,2I)\,\hat{A}_p$$

$$\times \prod_{J=1}^{N} \Delta_j^J(2J-1,2J)\,\theta_j^J(2J-1,2J)\,d\tau$$

Taking first the diagonal elements, one readily finds the standard expression

$$H_{ii} = \sum_{I=1}^{N} F_I^i + \tfrac{1}{2} \sum_{I} \sum_{J \neq I} (J_{IJ}^i - K_{IJ}^i)$$

$$S_{ii} = \prod_{I=1}^{N} S_I^i$$

where

$$F_I^i = X_I^i \int \Delta_i^I(12)\theta_i^I(12)\hat{H}(12)\Delta_i^I(12)\theta_i^I(12)d\tau_{12}$$

$$J_{IJ}^i = X_{IJ}^i \int \Delta_i^I(12)\theta_i^I(12)\Delta_i^J(34)\theta_i^J(34)\hat{H}(1234)$$

$$\times \; \Delta_i^I(12)\theta_i^I(12)\Delta_i^J(34)\theta_i^J(34)d\tau_{1234}$$

$$K_{IJ}^i = X_{IJ}^i \int \Delta_i^I(12)\theta_i^I(12)\Delta_i^J(34)\theta_i^J(34)\hat{HP}(1234)$$

$$\times \; \Delta_i^I(12)\theta_i^I(12)\Delta_i^J(34)\theta_i^J(34)d\tau_{1234}$$

and

$$\hat{H}(12) = -\tfrac{1}{2}\nabla^2(1) - \tfrac{1}{2}\nabla^2(2) - \sum_A \left(\frac{Z_A}{r_{1A}} + \frac{Z_A}{r_{2A}}\right) + \frac{1}{r_{12}}$$

$$\hat{H}(1234) = \frac{1}{r_{13}} + \frac{1}{r_{14}} + \frac{1}{r_{23}} + \frac{1}{r_{24}}$$

$$\hat{HP}(1234) = \frac{P_{13}}{r_{13}} + \frac{P_{14}}{r_{14}} + \frac{P_{23}}{r_{23}} + \frac{P_{24}}{r_{24}}$$

The coefficients are defined by

$$X_I^i = \prod_{J \neq I}^{N} S_J^i$$

$$X_{IJ}^i = \prod_{K \neq IJ}^{N} S_K^i$$

$$S_I^i = 2, \text{ if } \theta_i^I = S, T_0, 1 \text{ otherwise}$$

The expressions may be brought into a more useful form by the expansion of the pair function into an orbital basis. Thus

$$F_I^i = X_I^i \int \Delta_i^I(12)\theta_i^I(12)\hat{H}(12)\Delta_i^I(12)\theta_i^I(12)d\tau_{12}$$

$$= X_I^i \sum_{pqrs} c_{pq}^I(i)\ c_{rs}^I(i) \int \chi_p^I(1)\chi_q^I(2)\,[\hat{f}(1)+\hat{f}(2)+\hat{g}(12)]$$

$$\chi_r^I(1)\chi_s^I(2)dv_{12} \times \int \theta_i^I(12)\theta_i^I(12)ds_{12}$$

$$= X_I^i Y_i^I\{2(\underline{c^I(i)f^{II}c^I(i)}) + [c_{pq}^I(i)\ g_{pqrs}^{IIII}\ c_{sr}^I(i)]\}$$

where

$$F_{ij}^{IJ} = \int \chi_i^I(1)\,[-\tfrac{1}{2}\nabla^2(1) - \sum_A \frac{Z_A}{r_{1A}}]\chi_j^J(1)dv_1 \quad ,$$

$$g_{ijk\ell}^{IJKL} = \int \chi_i^I(1)\chi_j^J(1)\ \frac{1}{r_{12}}\ \chi_k^K(2)\chi_\ell^L(2)dv_{12}$$

and

$$Y_i^I = - S_I^i \text{ if } \theta_i^I \, \varepsilon T, \; S_I^i \text{ otherwise}$$

The notation (_____) means that the trace of the matrix product is to be taken and [ ] the summation over all dummy indices. The interpair integrals are expanded in a similar manner.

$$J_{IJ} = 4\chi_{IJ}^i Y_i^I Y_i^J \left[\underline{c^I(i) c^I(i)}_{pq} \, g_{pqrs}^{IIJJ} \, \underline{c^J(i) c^J(i)}_{rs} \right]$$

$$K_{IJ} = 4\chi_{IJ}^i Z_{IJ}^{ii} \left[\underline{c^I(i) c^I(i)}_{pq} \, g_{psrq}^{IJJI} \, \underline{c^J(i) c^J(i)}_{rs} \right]$$

where

$$Z_{IJ}^{ij} = (-1)^{P_j^I + P_j^J} \int \theta_i^I(12) \theta_i^J(34) P_{13} \theta_j^I(12) \theta_j^J(34) ds_{1234}$$

and

$$P_i^I = 1, \text{ if } \theta_i^I \, \varepsilon T, \; 0 \text{ otherwise}$$

The coefficients Z_{IJ}^{ij} may be rewritten as

$\int A(23) B(23) ds_{23}$ where

$$A(23) = \int \theta_i^I(12) \theta_j^I(13) ds_1 \times (-1)^{P_j^I}$$

$$B(23) = \int \theta_i^J(34) \theta_j^J(24) ds_4 \times (-1)^{P_j^J}$$

Table 2

B(23) A(23)	SS	ST_0	ST_1	ST_{-1}	T_0S	T_0T_0	T_0T_1	T_0T_{-1}	T_1S	T_1T_0	T_1T_1	T_1T_{-1}	$T_{-1}S$	$T_{-1}T_0$	$T_{-1}T_1$	$T_{-1}T_{-1}$
SS	2	.	.	.	.	-2	.	.	.	.	-1	.	.	.	.	-1
ST_0	.	-2	.	.	2	.	.	.	.	.	-1	.	.	.	.	1
ST_1	.	.	.	1	.	.	.	1	1	1	.	.	.	.	.	.
ST_{-1}	.	.	1	.	.	.	-1	.	.	.	.	.	1	-1	.	.
T_0S	.	2	.	.	-2	.	.	.	.	.	1	.	.	.	.	-1
T_0T_0	-2	.	.	.	.	2	.	.	.	.	1	.	.	.	.	1
T_0T_1	.	.	.	1	.	.	.	1	1	1	.	.	.	.	.	.
T_0T_{-1}	.	.	-1	.	.	.	1	.	.	.	.	.	-1	1	.	.
T_1S	.	.	1	.	.	.	-1	.	.	.	.	.	1	-1	.	.
T_1T_0	.	.	-1	.	.	.	1	.	.	.	.	.	-1	1	.	.
T_1T_1	-1	1	.	.	-1	1	.	.	.	.	1	.	.	.	.	.
T_1T_{-1}	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
$T_{-1}S$	.	.	.	1	.	.	.	1	1	1	.	.	.	.	.	.
$T_{-1}T_0$	.	.	.	1	.	.	.	1	1	1	.	.	.	.	.	.
$T_{-1}T_1$	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
$T_{-1}T_{-1}$	-1	-1	.	.	1	1	.	.	.	.	.	.	.	.	.	1

If there is no entry in any position in the Table, then $Z_{IJ}^{ij} = 0$.

and are given for all possible combinations of A and B in Table 2.

Thus we have finally for the diagonal matrix element

$$H_{ii} = \sum_{I}^{N} X_I^i Y_i^I \{2(\underline{c^I(i) f^{II} c^I(i)}) + [c^I_{pq}(i)\; g^{IIII}_{prqs}\; c^I_{sr}(i)]\}$$

$$+ \sum_{I<J}^{N}\sum 4X^i_{IJ}\,[\underline{c^I(i)c^I(i)}_{pq}\; \{Y_i^I Y_i^J\; g^{IIJJ}_{pqrs} - Z^{ii}_{IJ}\; g^{IJJI}_{psrq}\}$$

$$\underline{c^J(i)c^J(i)}_{rs}\,] \qquad 10a$$

$$S_{ii} = \prod_{I}^{N} S_I^i$$

The off-diagonal elements are greatly simplified by the orthogonality conditions. No more than single interchanges between pair functions need be considered as more will produce mismatched overlap integrals. There are also no Coulomb terms since a permutation of electrons is required to destroy the orthogonality of the spin functions. For one pair spin function non-coincidence, in pair I,

$$H_{ij} = \sum_{J\neq I} K^{ij}_{IJ} \qquad 10b$$

$$S_{ij} = 0$$

where

$$K^{ij}_{IJ} = 4X^{i}_{IJ}\ Z^{ij}_{IJ}\ [\underline{c^{I}(i)c^{I}(j)}_{pq}\ g^{IJJI}_{psrq}\ \underline{c^{J}(i)c^{J}(j)}_{rs}]$$

For two non-coincidence, in I and J,

$$H_{ij} = -\ K^{ij}_{IJ} \qquad 10c$$

$$S_{ij} = 0$$

and, for more than two non-coincidences, the matrix element vanishes.

Since the spin functions are readily set up (assuming that a limited set of the more important ones is to be used), and the matrix elements between the functions Ψ_i are easily calculated according to Equation 10, the method is well suited to machine computation.

The effect of optimising the spin part of the wavefunction is, as we have said, to redistribute the electrons, alleviating the excessive localisation of the electrons in the bonds by delocalisation into the interbond regions. We can see to what extent this is achieved by studying the changes in the one-electron density function, $\rho(1,1')$.

$$\rho(1,1') = \int \Psi\ \Psi\ d\tau_{2...2N}$$

$$= \sum_i \sum_j a_i\, a_j \int \Psi_i\, \hat{A}_p\, \Psi_j\, d\tau_{2\ldots 2N}$$

$$= \sum_i \sum_j a_i\, a_j\, \delta_{ij}\, [\sum_K^N P_K^i(1,1')]$$

Since, if $i \neq j$, then spin integration for no interchanges gives zero contribution whilst, for more interchanges, the space integration gives zero.

$$P_K^i(1,1') = X_K^i \int \Delta_i^K(12)\theta_i^K(12)\Delta_i^K(1'2)\theta_i^K(1'2)d\tau_2$$

$$\rho(1,1') = \sum_i a_i^2\, [\sum_K^N P_K^i(1,1')]$$

$$= \sum_K^N P_K(1,1')$$

where

$$P_K(1,1') = \sum_i a_i^2\, P_K^i(1,1')$$

The density is thus still in a 'sum-of-bond-densities' form. Expansion of $P_K^i(1,1')$ into an orbital basis yields the following expressions

Singlet S $\quad P_K^i(1,1') = X_K^i \sum_p \sum_q \underline{c^K c^K}_{pq}\, \chi_p^K(1)\chi_q^K(1')\,[\alpha(1)\alpha(1')+\beta(1)\beta(1')]$

$$\text{Triplet } T_0 \qquad = -X^i_K \sum_p \sum_q \underline{c^K c^K}_{pq} \chi^K_p(1)\chi^K_q(1')\,[\alpha(1)\alpha(1')+\beta(1)\beta(1')]$$

$$T_1 \qquad = -X^i_K \sum_p \sum_q \underline{c^K c^K}_{pq} \chi^K_p(1)\chi^K_q(1')\,[\alpha(1)\alpha(1')]$$

$$T_{-1} \qquad = -X^i_K \sum_p \sum_q \underline{c^K c^K}_{pq} \chi^K_p(1)\chi^K_q(1')\,[\beta(1)\beta(1')]$$

The approximation of a minimal basis gives a particularly simple form for the bond density $P_K(1,1')$.

$$\begin{aligned}
P_K(1,1') = \tfrac{1}{2}\{ & \chi^K_1(1)\chi^K_1(1')\,[x^K a^K_S + \tfrac{1}{2}a^K_{T_0} + a^K_{T_1}]\alpha\alpha \\
& + \chi^K_2(1)\chi^K_2(1')\,[(1-x^K)a^K_S + \tfrac{1}{2}a^K_{T_0} + a^K_{T_1}]\alpha\alpha \\
& + \chi^K_1(1)\chi^K_1(1')\,[x^K a^K_S + \tfrac{1}{2}a^K_{T_0} + a^K_{T_{-1}}]\beta\beta \\
& + \chi^K_2(1)\chi^K_2(1')\,[(1-x^K)a^K_S + \tfrac{1}{2}a^K_{T_0} + a^K_{T_{-1}}]\beta\beta \\
& + (\chi^K_1(1)\chi^K_2(1') + \chi^K_2(1)\chi^K_1(1'))\,[y^K a^K_S](\alpha\alpha+\beta\beta)\}
\end{aligned}$$

where $x^K = (c^K_{11})^2 + (c^K_{12})^2 \;,\quad 1-x^K = (c^K_{22})^2 + (c^K_{12})^2 \;,$

$y^K = c^K_{12}(c^K_{11} + c^K_{22}),$

and $\quad a^K_S = \sum\limits_{i,K=S} a^2_i \, S_{ii}$, and so on, where the summation is over all Ψ_i that have the K'th bond singlet coupled.

Since $a_S^K < 1$ ($a_S^K = 1$ when the bond K is coupled to a singlet in all terms) and $a_{T_0}^K$, $a_{T_1}^K$, $a_{T_{-1}}^K > 0$, the effect of the inclusion of alternative spin couplings is to decrease the electron density in the bonding regions i.e. the bond order $y^K \rightarrow y^K a_S^K$, and to increase the electron density round the atoms. The bond polarity is reduced by the same factor as the bond order $(1-2x^K) \rightarrow a_S^K(1-2x^K)$. Although it is not obvious that electron density is redirected into the interbond regions, this is the likely result of the optimisation since it is the atomic terms χ_i^2 which will make the major contributions to the interatomic regions. The spin density function $Q_K(1,1')$, for bond K,

$$Q_K(1,1') = \int \hat{S}_z(1) P_K(1,1') dS$$

$$= \tfrac{1}{2}(\chi_1^K(1)\chi_1^K(1') + \chi_2^K(1)\chi_2^K(1'))[a_{T_1}^K - a_{T_{-1}}^K]$$

The spin density in bond K is thus equally shared by the two orbitals, and there is no 'bond order' spin density term. No spin density is found in any bond that is not, in at least one term, triplet coupled.

The magnitude of the spin coupling in different molecular types is of considerable importance: the combination of pair function theory and spin theory allows, through the matrix element expressions (10), some estimate to be made of the effect of spin optimisation. It also lets us substantiate our earlier statements regarding the convergence of the spin function series. For molecules, whose symmetry operations simply permute the bonds, the projection of a state of the correct spatial symmetry conserves the number of triplet PSFs in each term. Hence, for even electron singlet states, the lowest term in the wavefunction contains no triplet PSFs while higher terms will have 2,3,...N triplet PSFs. As we have shown, the matrix elements connecting states with one or two PSF non-coincidences consist solely of exchange integrals. Since the interbond exchange integrals will be small - we are assuming an orthogonalised hybrid basis - the mixing will be small. Since there is no matrix element between states in which the number of triplet PSFs differs by more than two, the majority of states must interact with the 'all-singlet' term via intermediate states. The more triplet PSFs, the more intermediate states required for the coupling and the smaller the resulting interaction. Also, the matrix element between two terms differing by two triplet PSFs will be larger, the closer are the bonds involved, since the exchange integrals will fall off rapidly with increasing separation. In fact, at certain levels of semi-empirical schemes, e.g. CNDO, the interaction terms are zero, and the states are, in this approximation, uncombining: only in more accurate theories, e.g. INDO, in which one centre exchange integrals are retained, will the different spin couplings mix. Although the effect on the energy is thus expected to be small, this is not to say that other properties may not be affected to a much greater extent.

Pi electron systems, with symmetries such that a single partitioning of the molecule into bonds is not possible, e.g. benzene, are only tractable within this scheme in the simplified situation where the pair function is represented by a Heitler-London function. The symmetry operators mix spin functions with different numbers of triplet PSFs and hence the magnitudes of the interaction terms depend on the symmetry determined coefficients and may in fact be large. The class of pi electron systems which do not possess symmetries, such that the assigned bonds are mixed, provide an interesting subject. For example, butadiene, represented by a structure with two localised pair functions, produces two uncombining spin states, in the pi electron approximation, where exchange integrals are neglected. This would imply that butadiene is closer to the sigma bonded type of molecule than to the traditional conception of it as a delocalised pi system. The same principle predicts that the infinite linear polyene chain is alternant, as is the benzene analogue, fulvene.

So far we have considered only the chemically oriented form of pair theory where the pair function is built from a pair of directed orbitals within the bond. A more accurate wavefunction is obtained with a localised MO basis. Not only are charge transfer terms introduced directly, but the above conclusions on the magnitude of the effect of spin optimisation must also be slightly modified. Since the MO's contain contributions from all atomic orbitals, the interaction matrix elements will now be composed of, as well as exchange integrals, Coulomb integrals. However the weight of these integrals will be small, so that the conclusions are largely unaltered.

7. Molecular Positive Ions

The analysis of the preceding sections showed that the optimisation of the spin coupling in saturated molecules is unlikely to alter radically the energy or charge distribution of an even electron, low spin molecule. However, the optimisation is expected to be of much greater significance when one is dealing with spin properties or systems with odd electrons. In this section, we shall be concerned with the generalisation of the theory to the positive ion radicals, derived from even electron saturated molecule parents. The model for the positive ion is a direct extension of the localised bond model for the parent: the electronic structure is considered as a resonance mixture between states with localised bond ionisations. Thus

$$\Psi = \sum_K \sum_i a^K_i \Psi^K_i$$

where K labels the bond ionised, i the spin function.

$$\Psi^K_i = \hat{A}_p \Delta^1(12)\Delta^2(34)\ldots\Delta^N(2N-3,2N-2)\phi^K(2N-1)\theta_i(1,2,\ldots 2N-1)$$

where ϕ_K is a one electron orbital constructed from the basis orbitals of which the parent pair function Δ^K is built. The orthogonality conditions are

$$\int \Delta^I(12)\Delta^J(13)d\tau_1 = 0 \qquad I \neq J$$

$$\int \Delta^I(12)\phi^J(1)d\tau_1 = 0 \qquad I \neq J$$

The expansions into an orbital basis are

$$\Delta^I(12) = \sum_{ij} c^I_{ij}\, \chi^I_i(1)\chi^I_j(2)$$

$$\phi^I(1) = \sum_i d^I_i \chi^I_i(1)$$

The spin functions θ_i are products of N-1 PSFs multiplied by an α or β function for the odd electron: spin eigenfunctions may be constructed by use of the vector coupling formulae of Section 3. The energy is given by

$$E = \int \Psi \hat{H} \Psi \, d\tau \Big/ \int \Psi \Psi \, d\tau$$

$$= \sum_K \sum_L \sum_i \sum_j a^K_i a^L_j H^{KL}_{ij} \Big/ \sum_K \sum_L \sum_i \sum_j a^K_i a^L_j S^{KL}_{ij}$$

Optimisation of the a^K_i leads to the secular equation

$$(\underline{H} - E\,\underline{S})\underline{a} = \underline{0}$$

In practice, the condition that Ψ is constructed from spin eigenfunctions allows the order of the secular equation to be reduced. To derive expressions for the matrix elements H^{KL}_{ij}, we expand the Hamiltonian $\hat{H}$ as

$$\hat{H} = \sum_{i=1}^{N-1} \hat{H}(2i-1,2i) + \tfrac{1}{2}\sum_i^{N-1} \sum_j{}' \hat{H}(2i-1,2i,2j-1,2j)$$

$$+ \hat{H}(2N-1) + \sum_{i=1}^{N-1} \hat{H}(2i-1,\, 2i,\, 2N-1)$$

where

$$\hat{H}(12) = -\tfrac{1}{2}\nabla^2(1) - \tfrac{1}{2}\nabla^2(2) - \sum_A Z_A\left(\frac{1}{r_{1A}} + \frac{1}{r_{2A}}\right) + \frac{1}{r_{12}}$$

$$\hat{H}(1) = -\tfrac{1}{2}\nabla^2(1) - \sum_A \frac{Z_A}{r_{1A}}$$

$$\hat{H}(1234) = \frac{1}{r_{13}} + \frac{1}{r_{23}} + \frac{1}{r_{14}} + \frac{1}{r_{24}}$$

$$\hat{H}(123) = \frac{1}{r_{13}} + \frac{1}{r_{23}}$$

The matrix elements are given below, using the symbols and conventions introduced for the 2N electron matrix elements.

(1) Diagonal element H^{KK}_{ii}

$$H^{KK}_{ii} = \int \Psi^K_i \, \hat{H} \, \Psi^K_i \, d\tau \tag{11a}$$

$$= \sum_{I\neq K}^{N} F^i_I + \sum_{I<J\neq K}^{N}\sum (2J^i_{IJ} - K^i_{IJ})$$

$$+ \bar{F}^i_K + \sum_{I\neq K} (\bar{J}^i_{IK} - \bar{K}^i_{IK})$$

F^i_{IJ}, J^i_{IJ}, K^i_{IJ} are defined in Equation 10.

$$\bar{F}^i_K = S^{KK}_{ii} \int \phi^K(1) \, \hat{H}(1) \, \phi^K(1) \, d\tau_1$$

$$= S^{KK}_{ii} \, \underline{(d^{+K} \, f^{KK} \, d^K)}$$

where

$$S^{KK}_{ii} = \int \Psi^K_i \Psi^K_i \, d\tau = \prod_{I \neq K}^{N} S^I_i$$

$$\bar{J}^i_{IK} = X^i_I \int \Delta^I_i(12)\phi^K(3) \, \hat{H}(123) \, \Delta^I_i(12)\phi^K(3) dv_{123}$$

$$\times \int \theta^I_i(12)\theta^K_i(3)\theta^I_i(12)\theta^K_i(3) \, ds_{123}$$

$$= 2X^i_I \, Y^I_i \, [\underline{c^I(i)c^I(i)}_{pq} \, g^{IIKK}_{pqrs} \, d^K_r \, d^K_s \,]$$

where $\theta^K_i(1)$ is the one electron spin function, α or β.

$$\bar{K}^i_{IK} = X^i_I \int \Delta^I_i(12)\phi^K(3)\theta^I_i(12)\theta^K_i(3) \, [\, \frac{P_{13}}{r_{13}} + \frac{P_{23}}{r_{23}} \,]$$

$$\Delta^I_i(12)\phi^K(3)\theta^I_i(12)\theta^K_i(3)d\tau_{123}$$

$$= 2X^i_I \, \bar{Z}^{ii}_{IK} \, [\underline{c^I(i)c^I(i)}_{pq} \, g^{IKKI}_{psrq} \, d^K_r \, d^K_s \,]$$

where

$$\bar{Z}^{ij}_{IJ} = (-1)^{P^I_j} \int A_i(123)B_j(123)ds_{123}$$

where

$$A_i(123) = \theta^I_i(12)\theta^J_i(3) \text{ and } B_j(123) = \theta^I_j(32)\theta^J_j(1)$$

The values of $\bar{Z}^{ij}_{IJ}$ for all combinations of A_i and B_j are given in Table 3.

Table 3

A_j	$S\alpha$	$S\beta$	$T_0\alpha$	$T_0\beta$	$T_1\alpha$	$T_1\beta$	$T_{-1}\alpha$	$T_{-1}\beta$
A_i								
$S\alpha$	1	.	-1	.	.	-1	.	.
$S\beta$	.	1	.	1	.	.	-1	.
$T_0\alpha$	1	.	-1	.	.	-1	.	.
$T_0\beta$	.	-1	.	-1	.	.	-1	.
$T_1\alpha$	.	.	.	.	-1	.	.	.
$T_1\beta$	-1	.	-1	.	.	.	.	.
$T_{-1}\alpha$	.	1	.	-1	.	.	.	.
$T_{-1}\beta$	.	.	.	.	.	.	.	-1

For positions in the Table without entries, $\bar{Z}^{ij}_{IJ} = 0$

(2) Off-diagonal element - different bonds ionised, matched PSFs.

$$H^{KL}_{ii} = \int \Psi^K_i \hat{H} \Psi^L_i \, d\tau \qquad \text{11b}$$

Space integration gives a zero result for all terms except those involving a permutation of electrons between K and L to produce a matched overlap integral.

$$H_{ii}^{KL} = X_L^i \int \Delta_i^L(12)\phi^K(3)\theta_i^L(12)\theta_i^K(3)\ \hat{HP}(123)$$

$$\times\ \Delta_i^K(12)\phi^L(3)\theta_i^K(12)\theta_i^L(3)\ d\tau_{123}$$

where

$$\hat{HP}(123) = -\frac{P_{13}}{r_{12}} - \frac{P_{13}}{r_{23}} - \frac{P_{23}}{r_{13}} - \frac{P_{23}}{r_{12}} - P_{13}f_2 - P_{23}f_1$$

which gives an expansion

$$H_{ii}^{KL} = -\ 2X_L^i\ \bar{Z}_{LK}^{ii}\ \{[\underline{c^K(i)d^K}_p\ g_{prsq}^{KLLL}\ c_{qr}^L(i)d_s^L\ ]$$

$$+\ [\underline{d^Lc^L(i)}_p\ g_{pqsr}^{LKKK}\ c_{qr}^K(i)d_s^K] + [\underline{c^K(i)d^K}_p\ f_{pq}^{KL}\ \underline{d^Lc^L(i)}_q]\}$$

(3) Off-diagonal element - same bond ionised, different spin functions.

$$H_{ij}^{KK} = \int \Psi_i^K\ \hat{H}\ \Psi_j^K\ d\tau \qquad 11c$$

(a) Single PSF non-coincidence in I

$$H_{ij}^{KK} = -\sum_{J\neq IK} K_{IJ} - \bar{K}_{IK}$$

where

$$K_{IJ} = 4X_{IJ}^i\ Z_{IJ}^{ij}\ [\underline{c^I(i)c^I(j)}_{pq}\ g_{psrq}^{IJJI}\ \underline{c^J(i)c^J(j)}_{rs}\ ]$$

and

$$\bar{K}_{IK} = 2X_I^i \, \bar{Z}_{IK}^{ij} \, [\underline{c^I(i)c^I(j)}_{pq} \, g_{psrq}^{IKKI} \, d_r^K \, d_s^K \,]$$

(b) Two PSF non-coincidences in I, J

$$H_{ij}^{KK} = - K_{IJ}$$

(c) Two PSF non-coincidences in I, K

$$H_{ij}^{KK} = - \bar{K}_{IK}$$

(d) More than two PSF non-coincidences

$$H_{ij}^{KK} = 0$$

(4) Off-diagonal element - different bonds ionised, different spin functions.

$$H_{ij}^{KL} = \int \Psi_i^K \, \hat{H} \, \Psi_j^L \, d\tau \qquad \text{11d}$$

(a) One or two PSF non-coincidences in K, L.

$$H_{ij}^{KL} = X_L^i \int \Delta_i^L(12)\phi^K(3)\theta_i^L(12)\theta_i^K(3) \;\; H\hat{P}(123)$$

$$\Delta_j^K(12)\phi^L(3)\theta_j^K(12)\theta_j^L(3) \;\; d\tau_{123}$$

$$= -2X_L^i \, \bar{Z}_{LK}^{ij} \, \{[\underline{c^K(j)d^K}_p \, g_{pqrs}^{KLLL} \, c_{sq}^L(i) \, d_r^L]$$

$$+ \, [\underline{d^L c^L(i)}_p \, g_{pqrs}^{LKKK} \, c_{qs}^K(j) \, d_r^K]$$

$$+ \, [\underline{c^K(j)d^K}_p \; f^{KL}_{pq} \; \underline{d^L c^L(i)}_q]\}$$

(b) More than two PSF non-coincidences or non-coincidences not involving K, L.

$$H^{KL}_{ij} = 0$$

The existence of the matrix element between functions with different ionised bonds means that ions of different symmetries, with coefficients differing only in sign, have different energies. This is contrary to the conclusion drawn previously[11] that these states are degenerate within the model.

The charge and spin density functions are straightforwardly derived: however the results are too complex to warrant inclusion. We can see that there will be an interbond term in the density connecting the ionised pair functions: the density is thus no longer in a 'sum-of-bond-densities' form. It is expected that the effect of this term will be to remove electron density from between the bonds and localise it in the bonds, thus restoring some of the lost bond strength.

The degree of mixing of different spin functions is expected to be much greater for the positive ion than for the parent due to the interaction term (11d) which involves interbond one-electron integrals, non-negligible even in the pi-electron approximation, and much larger than the other matrix elements which are composed of exchange integrals. To investigate the importance of spin optimisation, and to study the degree of splitting of ionised states of different symmetry, we have performed some simple calculations on the butadiene positive ion. Butadiene gives rise on ionisation to two states, one of gerade (g) and the other of ungerade (u) symmetry. The doublet spin functions that may be constructed for the three electron system are

$$\theta_1 = S\alpha \qquad \theta_2 = 2T_1\beta - T_0\alpha$$

giving the symmetrised wavefunctions

$$\psi_1^{\pm} = \tfrac{1}{2}(\Delta^1\,\phi^2 \pm \Delta^2\,\phi^1)S\alpha$$

$$\psi_2^{\pm} = \frac{1}{\sqrt{12}}(\Delta^1\,\phi^2 \pm \Delta^2\,\phi^1)(2T_1\beta - T_0\alpha)$$

where the upper sign is for g, the lower for u symmetry.

The matrix elements were evaluated according to Equation 11: standard pi electron theory approximations were invoked in calculating the integrals, the values being, in the usual notation, in units of electron volts, $f_{11} = -23.363$, $f_{22} = -25.638$, $f_{12} = -2.618$, $f_{23} = -2.113$, $\gamma_{11} = \gamma_{22} = 11.130$, $\gamma_{12} = 5.472$,

$\gamma_{23} = 5.181$, $\gamma_{13} = 3.845$, $\gamma_{14} = 2.906$. Various levels of sophistication of the pair function were employed: the VB approximation gives

$$\Delta(12) = \frac{1}{\sqrt{2}}(\chi_1(1)\chi_2(2) + \chi_2(1)\chi_1(2))$$

and the bond orbital (BO) approximation

$$\Delta(12) = \psi(1)\psi(2) = \tfrac{1}{2}(\chi_1(1)+\chi_2(1))(\chi_1(2)+\chi_2(2)).$$

The coefficients of the pair function (PF)

$$\Delta(12) = c_{11}\chi_1(1)\chi_1(2) + c_{12}(\chi_1(1)\chi_2(2) + \chi_2(1)\chi_1(2))$$

$$+ c_{22}\,\chi_2(1)\chi_2(2)$$

where obtained by minimising the energy of the parent molecule with a single spin coupling: the values used here are $c_{11} = c_{22} = 0.4$, $c_{12} = 0.6$. The one-electron orbitals were in all cases chosen as

$$\phi(1) = \frac{1}{\sqrt{2}}(\chi_1(1) + \chi_2(1))$$

No optimisation of the weights of χ_1 and χ_2 was attempted in view of the lack of polarity of the pair function of the parent molecule. The results appear to be fairly insensitive to small changes in the coefficients, so that any optimisation is unlikely to be worth the extra labour. The antisymmetric pair function, for combination with triplet PSFs, was taken, in all cases, to be

$$\Delta_{as}(12) = \frac{1}{\sqrt{2}}(\chi_1(1)\chi_2(2) - \chi_2(1)\chi_1(2))$$

The results are listed in Table 4. The optimisation of the spin coupling has little effect on the g-u energy gap in either the BO or PF approximations, although considerable improvement is found for the VB wavefunction. In all cases, the energy lowering accompanying spin optimisation is significant, and the mixing coefficient large. In fact, the alternative spin coupling in the VB approximation to the u state has a greater weight than that of the standard coupling θ_1. Also the energy lowering is much greater in this approximation, indicating the importance of alternative spin couplings in orthogonal VB theory. However, the energy gap is still too small and the poor energy obtained reflects the neccessity of including ionic structures in orthogonal VB theory.

The localised or exciton model of ionisation, in which the electron is removed from a localised bond, the charge hole being delocalised by resonance amongst structures with different bonds ionised, is far removed from the rather simpler (in mathematical

Table 4. Butadiene Positive Ion

Wavefunction		BO	VB	PF	MO
Δ(g-u)	BSO	2.11	1.06	2.03	
	ASO	2.06	1.41	2.00	2.46
ΔE_{opt}	g	0.59	0.76	0.47	
	u	0.54	1.11	0.43	
Cg	1	0.909	0.771	0.938	
	2	-0.416	-0.637	-0.347	
Cu	1	0.924	0.637	0.945	
	2	0.383	0.771	0.327	
Ip u	BSO	9.88	8.00	10.55	
	ASO	9.34	6.89	10.11	9.85
g	BSO	11.99	9.06	12.58	
	ASO	11.40	8.30	12.11	12.31

All results in electron-volts.

Δ(g-u)	Energy difference between g and u ionisation potentials
ΔE_{opt}	Energy lowering on optimisation of spin coupling
Cg, Cu	Coefficients of optimised wavefunctions
Ip	Ionisation potential
BSO	Before spin optimisation
ASO	After spin optimisation

terms) picture presented by MO theory in which one electron is removed from a delocalised molecular orbital, the ionisation energy being given by Koopman's theorem as the negative of the orbital energy. The energy gap between g and u states predicted by MO theory is 2.46eV, in good agreement with the results in Table 4. The MO approximation is an adiabatic one in which there are no free parameters to allow the system to respond to the ionisation: the spin optimised exciton model does allow for some relaxation and should hence be more accurate. The exciton model may give a greater insight into the physical changes resulting on ionisation. The weights of different resonance structures, (determined in the above case by symmetry, but no so in, for example, ethane where relative weights of C-C and C-H bond ionisations are not fixed), will give some indication of from which bond the electron is being removed.

The ionisation potentials calculated by the BO and PF versions of the exciton model compare well with the MO results and, as far as can be determined, with experiment which gives values of 9.08, 11.25, 12.14, 13.23ev, symmetry undetermined[7]. Although the g-u gap is little changed by optimisation of the spin wavefunction, the ionisation potential because, in the parent molecule, there is no matrix element between the alternative spin couplings and, consequently, no lowering of the energy. The two models, which give entirely different physical pictures, can thus be seen to give quantitative results of comparable accuracy, providing the pair function is well represented.

One might hope that, as optimised minimal basis pair functions for even electron species are becoming readily available, the study of the electronic structure of the positive ions may be tackled along the same chemical lines of investigation by the spin optimised pair function method presented above. Molecular positive ions have received some attention recently through the technique of photo-electron spectroscopy, primarily because the use of Koopman's theorem in MO theory allows the prediction of a series of ionisation potentials, now measurable. The questions considered are entirely energetic: tackling the problem via a 'chemical' wavefunction may throw more light on the physical changes in the electronic structure on ionisation. The spin properties of odd electron molecules are not accurately or easily calculated by the MO method; for example, it is impossible to explain the negative spin densities that are experimentally observed within restricted Hartree-Fock theory. Since the spin optimisation procedure is entirely practical, accurate spin distributions may well be made accessible by this method: it also opens up the possibility of performing accurate calculations on molecular ions to test the approximations made in deriving the relation between the carbon atom spin density and the proton hyperfine coupling constant in pi electron systems[2].

Molecular ions exist and are closely related to their parents. But, with so much known about the bonds and bonding of the parents, it is surprising that so little work has been directed towards the investigation of the validity and details of the physical models - the exciton and delocalised models - for the positive ions. The theory presented in this section provides a framework within which to perform this investigation of the exciton model.

7. Conclusions

In this Chapter we have presented a new basis of spin eigenfunctions which has advantages in construction and application over the current methods. It has been shown that the basis, when used in conjunction with the strongly orthogonal pair function spatial wavefunction, provides a tractable method for the optimisation of the molecular spin coupling. The theory was then extended to molecular positive ions.

The scope of the pair function method, as applied to saturated molecules, now covers the wide range from the ground states of even electron molecules to excited states[11], states

CONCLUSION

CONCLUSION

By comparison of current methods of molecular wavefunction construction it has been shown that the VB method, for small molecules, possesses advantageous features that outweigh the practical disadvantages, and therefore its neglect as a practical procedure is unjustified. The NPF method was shown to be a more useful form of VB theory for larger molecular systems where the non-orthogonality problem becomes extremely severe. The theory of the NPF method, the most general pair function method and hitherto not studied in any detail, has been worked out; the one- and two-particle density functions were evaluated explicitly, and the variational equations considered. Accurate practical applications of the theory to lithium hydride, the helium-helium interaction and the pi-electrons of butadiene have been presented.

Approximate methods of solution of the non-orthogonality problem which at present limits the practical use of the VB and NPF theories have been proposed, within the framework of the NPF method, and the theory and practical aspects worked out in detail. These are ab initio methods in that no semi-empirical parameters are introduced, and the accuracy of the approximation may be rigorously tested at all stages. Some preliminary results of an investigation of a model system by the "neglect-of-powers-of-overlap" method were also presented. The advantages of the NPF formalism, compared to current alternatives, as regards their contribution to valence theory have been stressed.

A new method of construction of a spin function basis, which is particularly suited to use in conjunction with electron pair theories has been derived. Using this spin function basis, the SBP theory has been extended to allow spin optimisation of singlet states and the construction of SBP wavefunctions for even-electron molecules of arbitrary multiplicity and for molecular positive ions. The results of a simple application to the butadiene positive ion were presented.

APPENDIX 1

Appendix 1 Notation

Although all new notation used in Chapter 1 is explained as it is encountered, the complexity of the notation makes an appendix bringing together the different abbreviations necessary. The analysis of the NPF wavefunction makes extensive use of matrix algebra: where possible standard notation is used, that is, $\underline{A}$ for the matrix A, A_{ij} for the element in the i'th row and j'th column, $\underline{AB}$ for the product of matrices $\underline{A}$ and $\underline{B}$ and $\underline{AB}_{ij} = \sum_k A_{ik} B_{kl}$ for the i,j'th element of the product matrix. $\underline{\underline{A}}$ represents the supermatrix with elements $A_{ijk\ell}$ and tr $\underline{A}$ is the trace, or sum of the diagonal elements, of $\underline{A}$. However, a major part of the theory involves superscripted matrices, e.g. $\underline{A}^I$, $\underline{B}^{IJ}$, and products of such matrices. To ease the typographical congestion of super- and subscripts, a simplified notation has been resorted to in which the matrix superscripts are placed in brackets to the right of the matrix name and the underline omitted; for example, $\underline{A}^I \equiv A(I)$, $B^{IJ}_{ij} \equiv B(I,J)_{ij}$. The superscripts are always written in upper case, the subscripts in lower. The notation is particularly convenient for matrix products, as we may then group all the matrix names on the left in order and the superscripts within the bracket; thus, $\underline{A}^{I}\underline{B}^{JK} \equiv \mathrm{AB(I,J,K)}$. It is not immediately obvious which superscripts belong to which matrix name: however the notation is only used where the number of different matrix names is small so that the context makes plain the allocation of superscript to name, reading both superscript and name from left to right. Further typographical economy is achieved with this notation since, in all situations where use is made of it, the right-hand superscript of a matrix is always equal to the left-hand superscript of the next matrix in the product. If a matrix has a single superscript the left- and right-hand superscripts coincide, and the above rule still holds. It is therefore unneccessary to repeat the superscript in the bracket: $\underline{A}^{IJ}\underline{B}^{JK}$ is written AB(I,J,K) and ABCD(I,J,K), where $\underline{A}$ and $\underline{C}$ are singly superscripted and $\underline{B}$ and $\underline{D}$ doubly, is the abbreviation for $\underline{A^{I}B^{IJ}C^{J}D^{JK}}$.

It should also be noted that, in sections 2 and 3 of Chapter 1, the subscripts of the matrices are usually linked to the superscripts, i.e. the dummy summation subscript of a matrix is determined by the superscript. For example, if the superscripts I, J, K are associated with the subscripts i, j, k respectively, the representative element of the matrix $\underline{A^{IJ}B^{JK}}$ is $\underline{A^{IJ}B^{JK}}_{ik}$ and of $\underline{A^{IK}B^{KJ}}$, $\underline{A^{IK}B^{KJ}}_{ij}$. It is then not necessary to include explicitly the summation subscripts, and these may be dropped e.g. $\underline{A^{IK}B^{KJ}}_{ij} \equiv AB(I,K,J)$.

The main matrix names used are $\underline{C}$ for the pair function coefficient matrix which is singly superscripted and $\underline{S}$ for the pair function overlap matrices which are doubly superscripted. The charge distribution matrix, with typical element $\chi_i^I(1)\chi_j^J(1')$ where $\chi_i^I(1)$ is a one-electron orbital is denoted by $T(1|I,J)_{ij}$ or $T(I,J)_{ij}$ where the electron coordinate label is readily understood from the context. Coefficients D(I,J,K,L,M,N) and D(I,J,K) are defined and explained in Section 3, but, to avoid confusion, it should be stated that these are not necessarily simple matrices with six and three superscripts respectively.

When the bracket notation is used, the product of two matrices is denoted by a cross, e.g. $A(I,J)\times B(K,L) \equiv \underline{A^{IJ}B^{KL}}$, and the matrix trace by the usual notation or by square brackets, e.g. tr $\underline{A}^{IJ} \equiv \mathrm{TR}\, A(I,J) \equiv [A(I,J)]$.

APPENDIX 2

where the notation is explained in detail in Appendix 1. The one-electron part of the energy, $H_1 = \int \hat{H}(1)\ P(1;1')dV_{1'\to 1}$, is expanded via equation (2.3) to give

$$\begin{aligned} H_1 = 2\,S_R^{-1}\{ & [\mathrm{CSCF}(1,1,1)][\mathrm{CSCS}(2,2,2)] \\ & + 2[\mathrm{CSCF}(2,1,2)][\mathrm{CSCS}(1,2,1)] \\ & + [\mathrm{CSCF}(2,2,2)][\mathrm{CSCS}(1,1,1)] \\ & - 2[\mathrm{CSCS}(1,1,2)\times \mathrm{CSCF}(2,2,1)] \\ & - [\mathrm{CSCS}(1,2,2)\times \mathrm{CSCF}(2,1,1)] \\ & - [\mathrm{CSCS}(2,1,1)\times \mathrm{CSCF}(1,2,2)]\} \end{aligned} \tag{A.2}$$

where

$$\begin{aligned} \mathrm{CSCF}(I,J,K)_{ik} &= \int \hat{H}(1)\ \mathrm{CSCT}(1|I,J,K)_{ik}\ dV_1 \\ &= \sum_j \mathrm{CSC}(I,J)_{ij} \int \chi_j^J(1)\chi_k^K(1)\ \hat{H}(1)\ dV_1 \\ &= \sum_j \mathrm{CSC}(I,J)_{ij}\ F(J,K)_{jk} \end{aligned}$$

Expanding $P(1,2;1',2')$ from equation (2.4), the two-electron part of the energy in $H_2 = \int P(1,2;1',2')\hat{H}(12)dV_{1'2'\to 12}$, is

$$
\begin{aligned}
H_2 = S_R^{-1}\{ & [[C(1)\cdot G(1,1,1,1)\cdot C(1)]][CSCS(2,2,2)] \\
+ & [[C(2)\cdot G(2,2,2,2)\cdot C(2)]][CSCS(1,1,1)] \\
+ & 4[CSC(1,1)\cdot G(1,1,2,2)\cdot CSC(2,2)] \\
- & 2[[CSC(1,2)\cdot G(1,1,2,2)\cdot CSC(1,2)]] \\
+ & 4[CSC(1,2)\cdot G(1,2,1,2)\cdot CSC(1,2)] \\
- & 2[[CSC(1,1)\cdot G(1,2,1,2)\cdot CSC(2,2)]] \\
+ & 2[[C(1)\cdot G(1,2,1,2)\cdot C(2)]][CSCS(1,2,1)] \\
- & 4[[C(1)\cdot G(1,1,1,2)\cdot CSCSC(1,2,2)]] \\
- & 4[[CSCSC(1,1,2)\cdot G(1,2,2,2)\cdot C(2)]]\}
\end{aligned} \tag{A.3}
$$

where G(I,J,K,L) is the super matrix of two electron integrals

$$
\begin{aligned}
G(I,J,K,L)_{ijk\ell} &= \int \chi_i^I(1)\chi_j^J(1)\; r_{12}^{-1}\; \chi_k^K(2)\chi_\ell^L(2)\; dV_{12} \\
&= \int \hat{H}(12)\; T(1|I,J)_{ij}\; T(2|K,L)_{k\ell}\; dV_{12}
\end{aligned}
$$

and the supermatrix summation notation,

$$\sum_{ijk\ell} A_{ij}\, B_{ijk\ell}\, C_{k\ell} \equiv [A\cdot B\cdot C]$$

and $$\sum_{ijk\ell} A_{ik}\, B_{ijk\ell}\, C_{j\ell} \equiv [[A\cdot B\cdot C]] \text{ , is used.}$$

The optimum wavefunction is determined by demanding that the energy is stationary with respect to the variation of all coefficients. Differentiating the energy expression,

$$E = H/S_R \quad \text{where} \quad H = S_R(H_1 + H_2) \text{ ,}$$

with respect to C^I_{ij} and setting $\partial E/\partial C^I_{ij}$ equal to zero gives a set of equations

$$\frac{\partial H}{\partial C^I_{ij}} - E\,\frac{\partial S_R}{\partial C^I_{ij}} = 0, \quad \text{all } i, j.$$

Since H and S_R are quadratic in $\underline{C}^I$, the equations are linear and by separating out $C^I_{k\ell}$ from the expressions for $\partial H/\partial C^I_{ij}$ and $\partial S_R/\partial C^I_{ij}$, we arrive at a set of equations of the form

$$\sum_{k\ell} C^I_{k\ell}\, [H^{I'}_{ijk\ell} - E\, S^{I'}_{ijk\ell}] = 0, \quad \text{all } i, j.$$

This is a secular problem of order n_I^2, where n_I is the number of orbitals in pair function I, which determines the optimum coefficients $C^I_{k\ell}$ within the fixed space of all other pair functions.

Since $C^{I}_{k\ell}$ must equal $C^{I}_{\ell k}$, the order may be reduced to $n_I(n_I+1)/2$ with

$$\sum_{k\leqslant\ell} C^{I}_{k\ell}\ [H^{I}_{ijk\ell} - E\ S^{I}_{ijk\ell}] \ = \ 0, \qquad \text{all } i \leqslant j$$

where

$$H^{I}_{ijk\ell} = (H^{I'}_{ijk\ell} + H^{I'}_{ij\ell k} + H^{I'}_{jik\ell} + H^{I'}_{ji\ell k}) \times (1+\delta_{ij})^{-1}(1+\delta_{k\ell})^{-1}$$

and similarly for $S^{I}_{ijk\ell}$. The evaluation of the explicit expressions for the matrix elements $H^{I}_{ijk\ell}$, $S^{I}_{ijk\ell}$ is straightforward: it is most readily achieved by noting that there are only four types of term in the energy expression, equations (A.1), (A.2) and (A.3). These are, for the matrix elements for pair function 1,

$$W_1(A,B) \ = \ \mathrm{tr}(\underline{C^{1}AC^{1}B})$$

$$X_1(A,B) \ = \ \mathrm{tr}(\underline{C^{1}A}).\mathrm{tr}(\underline{C^{1}B})$$

$$Y_1(A;B,D,E,F) \ = \ \sum_{ijk\ell} A_{ijk\ell}\ (\underline{BC^{1}D})_{ij}\ (\underline{EC^{1}F})_{k\ell}$$

$$Z_1(A;B,D,E,F) \ = \ \sum_{ijk\ell} A_{ijk\ell}\ (\underline{BC^{1}D})_{ik}\ (\underline{EC^{1}F})_{j\ell}$$

The contributions from the terms of type $W_1(A,B)$ to the matrix elements $H^1_{ijk\ell}$, $S^1_{ijk\ell}$ are derived as follows. Separating out the coefficient matrices

$$W_1(A,B) = \sum_{k\ell mn} C^1_{k\ell}\, C^1_{mn}\, A_{\ell m}\, B_{nk}$$

$$= \sum_{k\leqslant\ell}\ \sum_{m\leqslant n} C^1_{k\ell}\, C^1_{mn}\, [A_{\ell m}B_{nk} + A_{\ell n}B_{mk}$$

$$+ \quad A_{km}B_{n\ell} + A_{kn}B_{m\ell}]t_{k\ell}t_{mn}$$

where

$$t_{k\ell} = (1 + \delta_{k\ell})^{-1}$$

Differentiating $W_1(A,B)$ with respect to C^1_{ij} gives

$$\partial W_1(A,B)/\partial C^1_{ij} = \sum_{k\leqslant\ell} C^1_{k\ell}\, [W_1'(A,B)_{ijk\ell}]$$

where

$$W_1'(A,B)_{ijk\ell} = [A_{\ell j}B_{ik} + A_{\ell i}B_{jk} + A_{ik}B_{\ell j}$$

$$+ \quad A_{jk}B_{\ell i} + A_{ki}B_{j\ell} + A_{kj}B_{i\ell}$$

$$+ \quad A_{j\ell}B_{ki} + A_{i\ell}B_{kj}]t_{ij}t_{k\ell}$$

is the contribution to the matrix element. Similarly the contributions of the other types of term are

$$X_1'(A,B)_{ijk\ell} = [(B_{k\ell}+B_{\ell k})(A_{ji}+A_{ij}) + (B_{ij}+B_{ji}) \times (A_{k\ell}+A_{\ell k})]t_{ij}t_{k\ell}$$

$$\begin{aligned} Y_1'(A;B,D,E,F)_{ijk\ell} &= \sum_{pqrs} A_{pqrs}\ [(B_{pi}D_{jq}+B_{pj}D_{iq}) \\ &\times (E_{rk}F_{\ell s} + E_{r\ell}F_{ks}) + (B_{rk}D_{\ell s}+B_{r\ell}D_{ks}) \\ &\times (E_{pi}F_{jq}+E_{pj}F_{iq})]t_{ij}t_{k\ell} \end{aligned}$$

$$\begin{aligned} Z_1'(A;B,D,E,F)_{ijk\ell} &= \sum_{pqrs} A_{pqrs}\ [(B_{pi}D_{jr}+B_{pj}D_{ir}) \\ &\times (E_{qk}F_{\ell s}+E_{q\ell}F_{ks}) + (B_{pk}D_{\ell r}+B_{p\ell}D_{kr}) \\ &\times (E_{qi}F_{js}+E_{qj}F_{is})]t_{ij}t_{k\ell} \end{aligned}$$

Using this notation, the matrix elements are

$$\begin{aligned} \underline{\underline{H}}^1 &= 2\{W_1'(S(1,1),F(1,1))\cdot[CSCS(2,2,2)] \\ &+ W_1'(S(1,1),S(1,1))\cdot[CSCF(2,2,2)] \\ &- 2W_1'(S(1,1),SCSCF(1,2,2,1)) \\ &- W_1'(SCS(1,2,2)\times CS(2,1),F(1,1)) \end{aligned}$$

$$- \; W_1'(\mathrm{SCF}(1,2,2)\times \mathrm{CS}(2,1),\mathrm{S}(1,1))$$

$$+ \; 2X_1'(\mathrm{SCF}(1,2,1),\mathrm{SCF}(1,2,1))\}$$

$$+ \; Y_1'(\mathrm{G}(1,1,1,1);\underline{1},\underline{1},\underline{1},\underline{1})[\mathrm{CSCS}(2,2,2)]$$

$$+ \; W_1'(\mathrm{S}(1,1),\mathrm{S}(1,1)) \; [[\mathrm{C}(2) \; \mathrm{G}(2,2,2,2)\cdot \mathrm{C}(2)]]$$

$$+ \; 2X_1'(((\mathrm{G}(1,2,1,2)\cdot \mathrm{C}(2))),\mathrm{SCS}(1,2,1))$$

$$- \; 4Y_1'(\mathrm{G}(1,1,1,2);\underline{1},\underline{1},\underline{1},\mathrm{SCSC}(1,2,2))$$

$$- \; 4W_1'(\mathrm{S}(1,1),\mathrm{S}(1,2)\times((\mathrm{G}(1,2,2,2)\cdot \mathrm{C}(2))))$$

$$+ \; 4W_1'(\mathrm{S}(1,1),(\mathrm{G}(1,1,2,2)\cdot \mathrm{CSC}(2,2)))$$

$$- \; 2Y_1'(\mathrm{G}(1,1,2,2);\underline{1},\mathrm{SC}(1,2),\underline{1},\mathrm{SC}(1,2))$$

$$+ \; 4Z_1'(\mathrm{G}(1,2,1,2);\underline{1},\mathrm{SC}(1,2),\underline{1},\mathrm{SC}(1,2))$$

$$- \; 2W_1'(\mathrm{S}(1,1),((\mathrm{G}(1,2,1,2)\cdot \mathrm{CSC}(2,2))))$$

where $\underline{1}$ is the unit matrix of dimension n_1, $((A.B))_{ik} = \sum_{j\ell} A_{ijk\ell} \, B_{j\ell}$ and $(A.B)_{ij} = \sum_{k\ell} A_{ijk\ell} \, B_{k\ell}$.

$$\underline{\underline{S}}^1 = W_1'(\mathrm{S}(1,1),\mathrm{S}(1,1))[\mathrm{CSCS}(2,2,2)]$$

$$- \; 2W_1'(\mathrm{S}(1,1),\mathrm{SCSCS}(1,2,2,1))$$

$$+ \quad X_1'(SCS(1,2,1),SCS(1,2,1))$$

The matrix elements for pair function 2 are obtained by interchanging 1 and 2 throughout the expressions for $\underline{\underline{H}}^1$ and $\underline{\underline{S}}^1$.

The solution of the variational problem is iterative since the matrix elements for each pair function depend on the coefficient matrix of the other. Starting from a guessed set of coefficients, $\underline{C}^2$, the optimum coefficients for pair function 1 are calculated. These are then used to get a new $\underline{C}^2$, the process being cycled until the coefficients for both pair functions are self-consistent.

APPENDIX 3

Appendix 3 Expansion of the Two Particle Density Function

It was stated that a coefficient D(K,L,M,N,I,J) whose properties and method of construction were dealt with in Section 3 of Chapter 1, could be extracted from the general expression for the two-particle density function of the NPF wavefunction

$$P(1,2;1',2') = P_D(1,2;1',2') + P_{OD}(1,2;1',2')$$

where

$$P_D(1,2;1',2') = \sum_{I=1}^{N} P^N Q^N TR[CTCT(P_I,I,Q_I) \prod_{J\neq I}^{N} CSCS(P_J,J,Q_J)]$$

and

$$\begin{aligned} P_{OD}(1,2;1',2') = {} & 2\sum_{I<J}^{N} P^N Q^N TR[CSCT(P_I,I,Q_I) \\ & CSCT(P_J,J,Q_J) \prod_{K\neq IJ}^{N} CSCS(P_K,K,Q_K)] \\ + {} & 2\sum_{I<J}^{N} P^N Q^N TR[CSCT(P_I,I,Q_I) \\ & CTCS(P_J,J,Q_J) \prod_{K\neq IJ}^{N} CSCS(P_K,K,Q_K)] \end{aligned}$$

In this Appendix we shall derive the complete expansion of

P(1,2;1',2') in terms of these coefficients. The notation of Appendix 1 is used throughout. The diagonal term of the two-particle density function $P_D(1,2;1',2')$ may be expanded in an analogous way to that for the one-particle density function, giving the result

$$
\begin{aligned}
P_D(1,2;1',2') = & \sum_I \{ \sum_{xyi} T(1|I,I)_{xy}\, T(2|I,I)_{zi} \\
& \times [C(I)_{ix}\, C(I)_{yz}\, D(I,I,I)] \\
& - 2 \sum_{K \neq I} T(1|I,I)_{xy}\, T(2|I,K)_{zk} \\
& \times [C(I)_{yz}\, (D(I,K,I) \times C(I))_{kx}] \\
& + 2 \sum_{K<L \neq I} T(1|K,I)_{xy}\, T(2|I,L)_{z\ell} \\
& \times [C(I)_{yz}\, (D(K,L,I) \times C(K))_{\ell x}] \\
& + \sum_{K \neq I} T(1|K,I)_{xy}\, T(2|I,K)_{zk} \\
& \times [C(K)_{kx}\, C(I)_{yz}\, D(K,K,I)] \}
\end{aligned}
$$

For the expansion of the off-diagonal term $P_{OD}(1,2;1',2')$, the N'th order permutation operator, P^N, must be expanded into three parts, an N-2 order operator, P^{N-2}, a second order operator, P^2, and one including all the cross terms, P^{CT}. The definition of the permutation operators, i.e. whether they operate on the labels or the positions, depends on the order in which we place them. P_{12} operating on (312) gives (321) if it is working on the labels, but (132) if on the positions. If the cross term is placed last $P^N = P^{CT} P^{N-2} P^2$, then all three operators work either on positions or on labels. It is necessary for us to use the alternative order $P^N = P^{N-2} P^{CT} P^2$ in which case P^{CT} and P^2 operate on labels and P^{N-2} on the positions (or vice versa). Using the stated definition, we write, for P^N,

$$P^N = P^{N-2}_{IJ} \left(1 - \sum_{K\neq IJ} (P_{IK} + P_{JK}) + \sum_{K<L\neq IJ} P_{IK}P_{JL}\right) \times (1 - P_{IJ})$$

$$= P^{N-2}_{IJ} \left(\sum_{K<L} X_{KL} \left(P_{IK}P_{JL} - P_{IK}P_{JL}P_{IJ}\right)\right)$$

$$= P^{N-2}_{IJ} \; P^2_R$$

where $X_{KL} = 1$ if $K\neq I$, $L\neq J$ or $K=I$, $L=J$

$= -1$ if $K\neq I$, $L=J$ or $K=I$, $L\neq J$

P^{N-2}_{IJ} is the N-2 order operator working on all positions except I and J and P_{KL} interchanges labels K and L irrespective of position, simultaneously with their corresponding summation subscripts. P_R^2 may be expanded as

$$P_R^2 = \sum_{K<L} X_{KL}[P_{I\to K}P_{J\to L}(1-P_{IJ})][P_{K\to I'}P_{L\to J'}]$$

where $P_{IK} = P_{I\to K}P_{K\to I'}$ and $P_{I\to K}$ replaces I by K and the subscript corresponding to K with that corresponding to I: the prime is removed after operation with the second operator. The operators in the second square brackets and P_{IJ}^{N-2} may now, since they operate on all terms except I and J, be taken through to the product part of the expansion of $P_{OD}(1,2;1',2')$, $\prod_{K\neq IJ} CSCS(P_K,K,Q_K)$. With a similar expansion for Q^N we find that

$$P_{OD}(1,2;1',2') = 2 \sum_{I<J} \sum_{K<L} \sum_{M<N} \sum_{U} X_{KL} X_{MN}$$

$$[P_{I\to K}P_{J\to L}(1-P_{IJ})][Q_{I\to M}Q_{J\to N}(1-Q_{IJ})]$$

$$CSCT(P_I,I,Q_I) \times (CSCT(P_J,J,Q_J) + CTCS(P_J,J,Q_J))$$

$$\sum_{T} P_{IJ}^{N-2} Q_{IJ}^{N-2} P_{K\to I}P_{L\to J}Q_{M\to I}Q_{N\to J} \prod_{R\neq IJ} CSCS(P_R,R,Q_R)$$

The primes are dropped by restricting the operation of the operator in the second line to the third line of the above equation.

The last line of the equation defines the coefficient D(K,L,M,N,I,J). To calculate the coefficient, the product $\prod_{R\neq IJ} \mathrm{CSCS}(R,R,R)_{rr}$ is set up, r being the subscript corresponding to R. The left superscripts K, L are replaced by I, J and right superscripts M, N by I, J; the summation subscripts k, m, corresponding to K, M are replaced by i corresponding to I and ℓ, n corresponding to L, N by j corresponding to J. The summation over all subscripts occuring twice in the product - the set T - is then made. This defines the leading term: the total coefficient is the signed sum of all terms generated by the action of the operators P_{IJ}^{N-2}, Q_{IJ}^{N-2}, operating on positions and permuting super and subscripts simultaneously. U stands for all subscripts not contained in T.

In this way we arrive at

$$P_{OD}(1,2;1',2') = 2 \sum_{I<J} [\sum_{K<L} \sum_{xy} \sum_{k\ell mn} (T(1|I,K)_{xk}$$

$$T(2|J,L)_{y\ell} - T(1|I,L)_{x\ell}\, T(2|J,K)_{yk})$$

$$\times \sum_{M<N} (\mathrm{CSC}(M,I)_{mx}\, \mathrm{CSC}(N,J)_{ny} - \mathrm{CSC}(N,I)_{nx}$$

$$\mathrm{CSC}(M,J)_{my})\; X_{MN} X_{KL}\; D(M,N,K,L,I,J)$$

$$
\begin{aligned}
+ & \sum_{M<L} \sum_{K<N} \sum_{xyz} \sum_{k\ell mn} (T(1|I,K)_{xk}\ T(2|J,L)_{zy} \\
& CSC(M,I)_{mx}\ C(L)_{\ell y}\ CS(J,N)_{zn} \\
- & T(1|I,K)_{xk}\ T(2|J,M)_{zy}\ CSC(L,I)_{\ell x}\ C(M)_{my}\ CS(J,N)_{zn} \\
- & T(1|I,N)_{xn}\ T(2|J,L)_{zy}\ CSC(M,I)_{mx}\ C(L)_{\ell y}\ CS(J,K)_{zk} \\
+ & T(1|I,N)_{xn}\ T(2|J,M)_{zy}\ CSC(L,I)_{\ell x}\ C(M)_{my}\ CS(J,K)_{zk}) \\
\times & X_{ML} X_{KN}\ D(M,L,K,N,I,J)]
\end{aligned}
$$

The summation subscripts x, y, z come from the expansion of the CSCT and CTCS terms whilst k, ℓ, m, n are the dummy indices associated with the superscripts K, L, M, N. As explained in Section 3, if, in D(K,L,M,N,I,J), K=M and L=N then the coefficient is a scalar, if K=M and L≠N we have a matrix $D(K,L,K,N,I,J)_{n\ell}$ and if K≠M and L≠N, a supermatrix $D(K,L,M,N,I,J)_{k\ell mn}$. To extract the K=M, L=N etc., terms so that the subscripts may be assigned explicitly results in a prohibitively lengthy expression: we have therefore omitted the subscripts in the coefficient in the above expansion, and the number and order of the subscripts is to be inferred from the superscripts K, L, M, N according to the rules of Section 3. Also, if K=M, then in the pre-coefficient terms the dummy index m is to be made into k and the summation over m disregarded, and similarly with ℓ and n if L=N. Thus, if K=M and L≠N, we get a contribution from the first term in the expansions of the above expression of

$$X_{KL}X_{KN} \sum_{xy} \sum_{k\ell n} T(1|I,K)_{xk}\ T(2|J,L)_{y\ell}\ [CSC(K,I)_{kx}\ CSC(N,J)_{ny}]\ D(K,N,K,L,I,J)_{\ell n}$$

The expansion of $P_{OD}(12,1'2')$ may be written in a more useful form by separating out the coefficient D(K,L,M,N,I,J), which is the most time consuming part in practice, and the orbital containing part, which is rapidly set up in, for example, the evaluation of the electron repulsion energy. Thus we find on dropping direct reference to the electron coordinates,

$$P_{OD}(12,1'2') = 2 \sum_{I<J} \sum_{K<L} \sum_{M<N} \sum_{xyk\ell mn} D(M,N,K,L,I,J)$$

$$\times\ X_{MN}X_{KL}[(T(I,K)_{xk}\ T(J,L)_{y\ell} - T(I,L)_{x\ell}$$

$$T(J,K)_{yk})(CSC(M,I)_{mx}\ CSC(N,J)_{ny} - CSC(N,I)_{nx}$$

$$CSC(M,J)_{my}) + \sum_{z} (T(I,K)_{xk}\ T(J,N)_{zy}\ CSC(M,I)_{mx}$$

$$C(N)_{ny}\ CS(J,L)_{z\ell} - T(I,K)_{xk}\ T(J,M)_{zy}\ CSC(N,I)_{nx}$$

$$C(M)_{my}\ CS(J,L)_{z\ell} - T(I,L)_{x\ell}\ T(J,N)_{zy}\ CSC(M,I)_{mx}$$

$$C(N)_{ny}\ CS(J,K)_{zk} + T(I,L)_{x\ell}\ T(J,M)_{zy}\ CSC(N,I)_{nx}$$

$$C(M)_{my}\ CS(J,K)_{zk})]$$

It would be useful, if instead of extracting the coefficients, the orbital terms could be separated out to give an expression of the form

$$P_{OD}(1,2,1'2') = \sum_{I,J} \sum_{K,L} \sum_{ijk\ell} T(I,K)_{ik}\ T(J,L)_{j\ell}$$

$$\times\ X(I,J,K,L)_{ijk\ell}$$

since this is the form required for the partition of two-electron expectation values, for example, in the separation of the electron repulsion energy into two, three and four bond interaction energies. Such an expression is not however available in a simple form: the partition must therefore be derived in particular cases from the explicit expansion of the above equation.

APPENDIX 4

Appendix 4 Matrix Elements for the Approximate Variational Methods

The matrix elements that were introduced in Section 3 in connection with the approximate variational methods are complicated by the non-orthogonality of the one-electron orbitals. Their evaluation therefore deserves some attention. The matrix elements required are

$$H_{OO} = \int \psi_0 \hat{H} \psi_0 \, d\tau$$

$$H_{ij}^{O,I} = \int \psi_0 \hat{H} \psi_{ij}^{I} \, d\tau$$

$$H_{ijk\ell}^{I,J} = \int \psi_{ij}^{I} \hat{H} \psi_{k\ell} \, d\tau$$

$$H_{ijk\ell}^{O,IJ} = \int \psi_0 \hat{H} \psi_{ijk\ell}^{IJ} \, d\tau$$

where ψ_0 is a determinant of doubly occupied non-orthogonal orbitals

$$\psi_0 = \hat{A} \, \phi_0^1(1) \bar{\phi}_0^1(2) \phi_0^2(3) \ldots \bar{\phi}_0^N(2N)$$

ψ_{ij}^{I} differs from ψ_0 by the substitution of the orbitals $\phi_i^I \bar{\phi}_j^I$ for $\phi_0^I \bar{\phi}_0^I$, and $\psi_{ijk\ell}^{IJ}$ by the substitution of $\phi_i^I \bar{\phi}_j^I \phi_k^J \bar{\phi}_\ell^J$ for

$\phi_0^{I-I}\phi_0^{J-J}\phi_0\phi_0$.

The matrix element between a pair of determinants constructed from non-orthogonal spin orbitals

$$\psi_a = \hat{A}\, a_1(1)a_2(2)\ \ldots\ a_n(n)$$

$$\psi_b = \hat{A}\, b_1(1)b_2(2)\ \ldots\ b_n(n)$$

may be expressed in terms of the first and second order cofactors of the spin orbital overlap matrix, $\underline{S}$, where $S_{ij} = \int a_i(1)b_j(1)d\tau_1$. Thus

$$H_{ab} = n!\,[\sum_{ij}^{n} f_{ij}\, D_{ij} + \sum_{i<k}^{n}\ \sum_{j<\ell}^{n} (g_{ijk\ell}-g_{i\ell kj}) \times D_{ijk\ell}]$$

where

$$f_{ij} = \int a_i(1)\,[-\tfrac{1}{2}\nabla^2(1) - \sum_{\alpha} \frac{Z_\alpha}{r_{12}}]\,b_j(1)d\tau_1$$

and $$g_{ijk\ell} = \int a_i(1)b_j(1)\,\frac{1}{r_{12}}\,a_k(2)b_\ell(2)d\tau_1 d\tau_2$$

D_{ij} and $D_{ijk\ell}$ are the cofactors of $\underline{S}$, obtained by striking out the i'th row and j'th column, and the i and k'th rows and j and ℓ'th columns respectively. If ψ_a and ψ_b are constructed from n_α alpha spin orbitals and n_β beta spin orbitals, and the spin orbitals are arranged so that all α spin orbitals precede the β spin orbitals, we may simplify the expressions by integration over spin coordinates. It is readily seen that the overlap matrix then has block diagonal form

$$\left(\begin{array}{c|c} \underline{S}^{\alpha} & \underline{0} \\ \hline \underline{0} & \underline{S}^{\beta} \end{array}\right)$$

where $\underline{S}^{\alpha}$ is the overlap matrix of α spin orbitals and $\underline{S}^{\beta}$ that of the β spin orbitals, and the determinant relation $|S| = |S^{\alpha}||S^{\beta}|$ holds. The first order cofactor matrix has the same form so that

$$\begin{aligned} D_{ij} &= D^{\alpha}_{ij} \cdot |S^{\beta}| && i,j < n \\ &= D^{\beta}_{ij} \cdot |S^{\alpha}| && i,j > n \\ &= 0 && \text{otherwise} \end{aligned}$$

where $\underline{D}^{\alpha}$ is the matrix of cofactors of $\underline{S}^{\alpha}$. The second order cofactors are related to the second order minors of the first order cofactor matrix, by

$$D_{ijk\ell} = (D_{ij}\, D_{k\ell} - D_{i\ell}\, D_{kj})/|S|$$

Spin integration gives the reduction of $D_{ijk\ell}$ to

$$
\begin{aligned}
D_{ijk\ell} &= D^{\alpha\alpha}_{ijk\ell}\,|S^{\beta}| \\
&= (D_{ij}\,D_{k\ell} - D_{i\ell}\,D_{kj})\,|S|\,|S^{\alpha}|^{2}, \quad ijk\ell < n_{\alpha} \\
&= D^{\beta\beta}_{ijk\ell}\,|S^{\alpha}| = (D_{ij}\,D_{k\ell} - D_{i\ell}\,D_{kj})\,|S|\,|S^{\beta}|^{2}, \quad ijk\ell > n_{\alpha} \\
&= D^{\alpha\beta}_{ijk\ell} = D^{\alpha}_{ij}\,D^{\beta}_{k\ell}, \quad ij < n_{\alpha} \text{ and } k\ell > n_{\alpha} \\
&= 0, \quad \text{otherwise}
\end{aligned}
$$

The remaining problem is to evaluate the cofactor matrices, $\underline{D}^{\alpha}$ and $\underline{D}^{\beta}$: this is achieved most economically by using the relation between cofactors and the elements of the inverse overlap matrix,

$$D_{ij} = S^{-1}_{ji}\,|S|$$

We thus have for the spin-integrated cofactor

$$D^{\alpha}_{ij} = (S^{\alpha})^{-1}_{ji}.|S^{\alpha}|$$

$$\text{and} \qquad D^{\beta}_{ij} = (S^{\beta})^{-1}_{ji}.|S^{\beta}|$$

This method is only valid for $|S| \neq 0$. Since the basis orbitals are deliberately non-orthogonal, the only situation in which $\underline{S}$ may be expected to be singular is when a term of a different symmetry is added to the description of a bond, for example, pi orbitals in the description of a sigma bond. The overlap matrix in such a case has at least two rows or columns zero: the number of non-zero cofactors is consequently much reduced and they may possibly be economically calculated directly. The biorthogonalisation technique of Prosser and Hagstrom[1] may alternatively be used when $|S| = 0$.

Although the matrix inversion-method is much faster than the computation of the cofactors directly, in view of the large number of matrix elements needed, any computational shortcuts, which can reduce the computing time per matrix element, will be of great value. Now, the matrices, $\underline{S}'^{\alpha}$ and $\underline{S}'^{\beta}$, that have to be inverted for the evaluation of the general matrix elements $H^{O,I}_{ij}$, $H^{I,J}_{ijk\ell}$ and $H^{O,IJ}_{ijk\ell}$, differ, at most, in two rows or columns from those, $\underline{S}^{\alpha}$ and $\underline{S}^{\beta}$, for the leading term, $H_{O,O}$. The standard matrix inversion methods make no use of any information gained by inversion of a similar matrix and hence duplicate much of the computation, especially when dealing with large numbers of electrons. This may be avoided by use of the following theorem.

If $\underline{A} = \underline{B} + \underline{u}.\underline{v}$ where $\underline{A}$ and $\underline{B}$ are square matrices, $\underline{u}$ is a column vector and $\underline{v}$ a row vector, then

$$\underline{A}^{-1} = \underline{B}^{-1} - \underline{B}^{-1} \underline{u}\, \underline{v}\, \underline{B}^{-1}/(1 + \underline{v}\, \underline{B}^{-1} \underline{u}).$$

The overlap matrices $\underline{S}'^{\alpha}$ and $\underline{S}'^{\beta}$ for the matrix element $H_{ij}^{O,I}$ differ from $\underline{S}^{\alpha}$ and $\underline{S}^{\beta}$ by one column. Thus, with $\underline{A} = \underline{S}'^{\alpha}$ and $\underline{B} = \underline{S}^{\alpha}$, we have

$$\underline{A} = \underline{B} + \underline{u}.\underline{v}$$

where $u_J = S_{i0}^{IJ} - S_{00}^{IJ}$, $S_{ij}^{IJ} = \int \phi_i^I(1)\phi_j^J(1)dV_1$ and $v_J = \delta_{JI}$. The inverse matrix, $(\underline{S}'^{\alpha})^{-1}$, is then given by

$$A_{ij}^{-1} = B_{ij}^{-1} - x_i\, B_{Ij}^{-1}/(1 + x_I)$$

where

$$x_i = \sum_k B_{ik}^{-1} u_k$$

$(\underline{S}'^{\beta})^{-1}$ is evaluated similarly. The matrix elements $H_{ij}^{O,I}$ may therefore be rapidly computed once the inverses of $\underline{S}^{\alpha}$ and $\underline{S}^{\beta}$ are known. By repeated application of the above theorem, expressions for the inverse matrices required in the computation of the matrix elements $H_{ijk\ell}^{I,K}$ and $H_{ijk\ell}^{O,IJ}$ are readily found: these are given below.

With $\underline{A} = \underline{S}'^{\alpha}$ for the matrix element $H_{ijk\ell}^{I,J}$ and $\underline{B} = \underline{S}^{\alpha}$, we have

$$\underline{A} = \underline{B} + \underline{u}\,\underline{v} + \underline{x}\,\underline{y} + \underline{w}\,\underline{z}$$

where $u_K = S^{JK}_{k0} - S^{JK}_{00}$, $x_K = \delta_{IK}$ and $w_K = \delta_{IK}$ are column vectors, and $v_K = \delta_{JK}$, $y_K = S^{IK}_{i0} - S^{IK}_{00}$ and $z_K = \delta_{JK}(S^{IJ}_{ik} + S^{IJ}_{00} - S^{IJ}_{i0} - S^{IJ}_{0k}) \equiv Z_{IJ}\,\delta_{JK}$ are row vectors. Application of the theorem three times and rearrangement of the final expression results in

$$A^{-1}_{k\ell} = B^{-1}_{k\ell} - \frac{\hat{u}_k}{1+\hat{u}_J}\,B^{-1}_{J\ell} - \frac{A^I_k\,A^J_\ell}{X_{IJ}(1 + \hat{u}_J)}$$

where

$$\hat{u}_k = \sum_m B^{-1}_{km}\,u_m,$$

$$\hat{y}_k = \sum_m y_m\,B^{-1}_{mk},$$

$$A^I_k = B^{-1}_{kI}(1 + \hat{u}_J) - \hat{u}_k\,B^{-1}_{JI},$$

$$A^J_k = \hat{y}_k(1 + \hat{u}_J) + B^{-1}_{Jk}(Z_{IJ} - t),$$

$$X_{IJ} = 1 + \hat{u}_J + A^J_I,$$

and

$$t = \sum_k y_k\,\hat{u}_k$$

Similarly, for $H^{0,IJ}_{ijk\ell}$, with $\underline{A} = \underline{S}'^{\alpha}$ and $\underline{B} = \underline{S}^{\alpha}$, we have

$$\underline{A} = \underline{B} + \underline{u}\,\underline{v} + \underline{w}\,\underline{x}$$

where $u_K = S^{IK}_{i0} - S^{IK}_{00}$ and $w_K = S^{JK}_{k0} - S^{JK}_{00}$ are column vectors and $v_K = \delta_{IK}$ and $x_K = \delta_{JK}$ are row vectors. The inversion theorem applied twice gives

$$A^{-1}_{k\ell} = B^{-1}_{k\ell} + ((a_k b_J - b_k(1 + a_J))B^{-1}_{I\ell}$$

$$- (a_k(1 + b_I) - b_k a_I)B^{-1}_J) \times c$$

where

$$a_k = \sum_{\ell} B^{-1}_{k\ell}\, w_{\ell}$$

$$b_k = \sum_{\ell} B^{-1}_{k\ell}\, u_{\ell}$$

and $$c = ((1 + b_I)(1 + a_J) - b_J\, a_I)^{-1}$$

Reference

1. Prosser, F. & Hagstrom, S. (1968) Int.J.Quant.Chem. 2, 89