Development of Methods in Spin-Coupled Theory

Thesis submitted in accordance with the requirements of the University of Liverpool for the degree of Doctor of Philosophy by Thorstein Thorsteinsson

May 1995

To Zuhal

Abstract

Apart from incidental involvement in optimisation of the spin-coupled wavefunction by the so-called super-cofactor approach, and development of pair population analysis, the work in this Ph.D. has taken the form of three major projects: valence bond interpretation of complete active space self-consistent field (CASSCF) wavefunctions, biorthogonal orbital optimisation, and calculation of inter-ionic potentials. These are all related to non-orthogonal orbital optimisation in particular, and non-orthogonal methods in general, and so a fair amount of introductory detail has been given to these subject areas in the context of the spin-coupled method.

A new highly efficient method for exact transformations of general CASSCF structure spaces has been developed, which may be of interest for other researchers in the field of multi-configuration orbital optimisation. In this work it has been used to obtain valence bond representations of 'N in N' CASSCF wavefunctions based on the form of the spin-coupled wavefunction. Four optimisation criteria (CASVB1–CASVB4) have been investigated that seek either to maximise the covalent component of the CASSCF wavefunction, or to minimise the energy of this component. The results have not only highlighted the striking similarities between the spin-coupled and CASSCF methods for most cases, but also suggested alternative descriptions in the cases of ozone and diborane.

A method for biorthogonal orbital optimisation based on the form of the spin-coupled wavefunction has been developed. A fully second-order treatment was found to be necessary to ensure satisfactory convergence, and such a scheme has been implemented. In this context, we appear to be the first to consider general optimisation of non-linear, non-symmetrical wavefunction parameters. The results show good agreement with those obtained from the variationally optimised spin-coupled wavefunction. At the very least this offers *qualitatively* correct results if the number of active electrons in the optimisation is larger than what may at present be treated using the spin-coupled method.

The advantages of valence bond approaches for the study of intermolecular forces have been recognised for some time. Such a scheme may be applied with only minor modifications to the problem of obtaining reliable potentials for modelling of ionic solids. The main difference lies in the realistic simulation of the crystalline environment, and this has also proven to be the main factor determining the accuracy of the potentials. Provided that a Madelung field of appropriate strength is applied very good agreement with existing potentials can be obtained. This suggests defining a suitable variable to be fitted to experimental data, or alternatively estimating the strength of the field in a simulation from the local surroundings of each individual ion. This will provide a consistent set of *transferable* potentials for the modelling of a variety of important systems and phenomena.

Acknowledgements

I would like to begin by thanking my previous supervisor in Copenhagen Dr Sten Rettrup who got me interested in quantum chemistry in the first place, and without whom I would not have had the opportunity of coming here to Liverpool.

I would also like to show my appreciation to those people in the Department of Chemistry here at Liverpool who have helped to make my stay here so enjoyable. Included in this list are my predecessors Dr Stephen Loades and Brian Petch; contemporaries Nick Clarke and Terry Cunningham; current students Andy Buttery and Kath Mort; and former departmental secretary Joyce Macintosh.

Special thanks go to my family and my fiancée who have never ceased supporting me, particularly in those hectic writing-up days.

Without the persistence and dedication of Dr Joe Gerratt (Bristol), Prof. Mario Raimondi (Milan) and my own supervisor Dr David Cooper the spin-coupled method would not have reached its current level of sophistication and as a consequence much of the work presented in this thesis would not have been possible. I thank the 'Gang of Three' for this, and also for countless helpful suggestions along the way.

Finally, I would like to express my sincere gratitude to my supervisor Dr David L. Cooper for taking me on with his grant from BP, and for being a truly excellent supervisor. He has taught me much about science, programming and writing, and his work ethics and dedication have been a real inspiration.

Contents

List of Figures

List of Tables

Chapter 1

General spin-coupled valence bond theory

1.0 The electronic Schrödinger equation

Before considering the more specialised quantum chemical ideas of modern valence bond theory, it is appropriate to recount briefly the relation to general physics and quantum mechanics.

The appropriate equation of motion for molecular systems was formulated by Schrödinger in 1926 [1]. In its non-relativistic form it may be written as

$$i\hbar\frac{\partial}{\partial t}\Psi = \hat{H}\Psi, \tag{1.1}$$

in which $\hat{H}$ is the Hamiltonian operator (see below) and Ψ is the wavefunction of the system—a function of all the particle coordinates. Among cases where relativistic corrections are significant belong the coupling between the spin and orbital angular momenta, and situations where electrons attain very high velocities close to the heavier nuclei (more pronounced for the innermost s electrons). See for example [2] or [3].

We will in this work consider only the time-independent form of this equation, applicable for the description of stationary states when the Hamiltonian contains

no time-related parts:

$$\hat{H}\Psi = E\Psi. \tag{1.2}$$

Again if relativistic effects are neglected the Hamiltonian may be written as

$$\hat{H} = -\sum_p \frac{\hbar^2}{2m_p}\nabla_p^2 + \sum_{p<q} \frac{e^2 Z_p Z_q}{4\pi\epsilon_0 r_{pq}}, \tag{1.3}$$

where p and q are particles of the system. m_p is the mass of particle p and Z_p its charge in units of the (positive) electronic charge e. The first term in (1.3) is thus the kinetic energy term, the second term takes into account the electrostatic interactions.

A major simplification may be achieved by separating the nuclear and electronic motion as proposed first by Born and Oppenheimer [4]. This is achieved by writing the wavefunction as a product of nuclear and electronic wavefunctions, $\Psi(\mathbf{r},\mathbf{R})=\Psi(\mathbf{r})\Psi(\mathbf{R})$ where $\mathbf{r}$ is a position vector for the electrons and $\mathbf{R}$ a position vector for the nuclei, and neglecting certain coupling terms. This yields the electronic Schrödinger equation for the motion of the electrons as

$$\hat{H}(\mathbf{R})\Psi(\mathbf{r}) = E\Psi(\mathbf{r}) \tag{1.4}$$

where the Hamiltonian depends parametrically on the nuclear coordinates according to

$$\hat{H}(\mathbf{R}) = -\sum_i \frac{\hbar^2}{2m_e}\nabla_i^2 + \sum_{i<j} \frac{e^2}{4\pi\epsilon_0 r_{ij}} - \sum_{i,A} \frac{e^2 Z_A}{4\pi\epsilon_0 r_{iA}} + \sum_{A<B} \frac{e^2 Z_A Z_B}{4\pi\epsilon_0 R_{AB}}. \tag{1.5}$$

This may be solved for the electronic motion while the nuclei are kept fixed in their positions. The approximation is likely to break down when electronic states calculated in this way become degenerate or near-degenerate (see for example [5]). Examples of such cases are the well known Jahn-Teller distortion and the 'avoided crossing' rule. The equation above is the basis of a large amount of work in quantum chemistry, and its solution will be the subject of the remainder of this thesis.

One inescapable relativistic notion for comparing theoretical predictions with experiment, is the concept of particle spin. The electron possess an intrinsic

angular momentum, its 'spin', amounting to $\hbar/2$. We may thus write

$$\hat{s}_z\alpha = +\frac{1}{2}\hbar\alpha$$

$$\hat{s}_z\beta = -\frac{1}{2}\hbar\beta, \tag{1.6}$$

where α and β are functions of the spin degree of freedom, σ, and both eigenfunctions of $\hat{s}^2$ and $\hat{s}_z$. It is often convenient to construct many electron wavefunction from spin orbitals, such that each electron occupies an orbital which is a function of four coordinates, for example $\phi(x_i, y_i, z_i)\alpha(\sigma_i)$.

The electron spin has two major implications. Firstly, having a half-integral spin the electron may be classified as a fermion, the crucial consequence of which is formulated by the Pauli principle stating that the wavefunction must be anti-symmetric under any interchange of two electronic coordinates:

$$\Psi(\mathbf{r}_i, \mathbf{r}_j) = -\Psi(\mathbf{r}_j, \mathbf{r}_i). \tag{1.7}$$

Alternatively, for a general permutation of electron coordinates we may write

$$\Psi(\mathbf{r}_1, \mathbf{r}_2, \cdots, \mathbf{r}_N) = \epsilon_P \Psi(\mathbf{r}_{p_1}, \mathbf{r}_{p_2}, \cdots, \mathbf{r}_{p_N}), \tag{1.8}$$

where ϵ_P is $+1$ for an even permutation, P, and -1 for an odd permutation. A wavefunction having this property may be realised very easily by taking simply a linear combination of determinants of spin orbitals (Slater determinants).

Secondly, if the Hamiltonian is independent of spin, as is the case in equation (1.5), the total wavefunction must also be an eigenfunction of the total spin operator (since the two operators commute: $[\hat{H}, \hat{S}^2]$=0) which may be defined according to

$$\hat{S}^2 = \sum_i \hat{s}_i^2. \tag{1.9}$$

The construction of spin eigenfunctions has been subject to a great deal of attention in the past. There are two basic strategies that may be employed. The simplest uses the theory of angular momentum to couple groups of electrons successively, according to a given scheme, until a final wavefunction with correct spin

quantum number is obtained. More elegant, perhaps, is the use of group theoretical methods, in which one for example can classify spin eigenfunctions as forming bases for irreducible representations of the symmetric group S_N. Normally the spin function will be constructed to ensure that the total wavefunction also adheres to (1.8). We will not consider the construction of spin eigenfunctions any further here—for a comprehensive treatment of this area the reader is referred to the book by Pauncz [6].

Having thus defined the form of acceptable wavefunctions, we may now consider how to approximate solutions to the time-independent Schrödinger equation, equation (1.2). There are several different ways of approaching this problem, but by far the most widely used strategy is the use of the variation theorem. For the optimisation of a ground state wavefunction it is straightforward to show that the energy expectation value of any trial wavefunction must always be above, or equal to, the exact ground state energy:

$$\langle E \rangle = \frac{\langle \Psi | \hat{H} | \Psi \rangle}{\langle \Psi | \Psi \rangle} \geq E_0. \tag{1.10}$$

Furthermore, when lowering the energy of the trial wavefunction one cannot obtain a *worse* approximation to the ground state wavefunction. So it is then very natural to assume that the best approximation to the ground state will be represented by a trial wavefunction with a minimum value of $\langle E \rangle$. This function may then be minimised with respect to the parameter set used to define the trial wavefunction, Ψ. This is indeed a very powerful statement, but it should be kept in mind that such a formulation not necessarily yields accurate estimates of molecular properties other than the energy.

As the reader may have become aware, several fundamental physical constants recur in most of the past equations. In order to ease the notation we shall therefore (unless otherwise stated) employ the so-called 'atomic units' in the remainder of this thesis. In this very common convention in the quantum chemical literature the ubiquitous constants m_e, e, $\hbar$ and κ_0 (κ_0=$4\pi\epsilon_0$) are set to unity. Physical quantities may then be given as dimensionless numbers in that the appropriate

powers of m_e, e, $\hbar$ and κ_0 are understood. Often the derived unit of length will be useful (bohr):

$$a_0 = \frac{\hbar^2 \kappa_0}{m_e e^2}, \tag{1.11}$$

or the corresponding energy unit (hartree):

$$E_h = \frac{m_e e^4}{\kappa_0^2 \hbar^2}. \tag{1.12}$$

1.1 Non-orthogonal orbital optimisation

The two main reasons for the take-off of molecular orbital (MO) theory and the relative stagnation of valence bond (VB) based methods beginning in the early 1950's, were the technical difficulties associated with the so-called '$N!$ problem' in practical VB calculations, as well as the quantitatively discouraging results these calculations tended to give. It was thought that sensible calculations could be constructed by direct use of atomic orbitals, but today it seems clear that an unprejudiced, initial orbital optimisation step is crucial.

In variational orbital optimisation schemes the VB wavefunction must in general take a multi-determinant form. Firstly, a large number of doubly occupied orbitals corresponds in the VB case to high-energy configurations, so the case of a closed shell Hartree-Fock wavefunction (or something resembling it) would not be appropriate. Secondly, and more simply, in a single-determinant wavefunction the orbitals are only determined within a linear transformation. Any criteria defining non-orthogonal orbitals from such a wavefunction would thus not be able to avoid a certain degree of arbitrariness. For this reason, valence bond orbital optimisation schemes are more complicated than simple Hartree-Fock theory, having more in common with standard, orthogonal multi-configuration self consistent field (MCSCF) methods.

The first general scheme of non-orthogonal orbital optimisation is probably that of Goddard [7]. It was originally referred to as the 'GI wavefunction' (GI referring to the combined spin-projection and antisymmetrisation operator G_i^μ), and

later the 'GVB (generalised valence bond) wavefunction'. Like the spin-coupled wavefunction, it is a single configuration of N electrons in N distinct orbitals, but, at least initially, optimisation of the spin coupling was not considered for this wavefunction. The spin function was chosen as one of the f_S^N linearly independent Kotani functions, most commonly of course the perfect-pairing function. In this way the optimisation problem could be solved by iterative diagonalisation of the N effective one-electron operators as defined in section 1.5. Besides this single spin function or 'perfect-pairing' (PP) approximation, the 'strong-orthogonality' (SO) approximation was also often employed in practical calculations. In this approach only overlaps between singlet-coupled orbitals are allowed to be non-zero. This simplifies the matrix element evaluation dramatically, but it is also justified by the notion of the *electron pair* as a single entity.

Complete optimisation of the spin-coupled wavefunction and its theoretical consequences was first explored by Gerratt and co-workers [8, 9, 10]. The development of theory and applications is still ongoing, see for example refs. [11, 12, 13]. Further details concerning aspects of the past theoretical development likely to be of particular practical utility will be given in the remainder of this chapter.

Besides the aforementioned work, mainly by Cooper, Gerratt and Raimondi, an early application of spin-coupled theory to atoms may be found in ref. [14]. More recently theoretical problems associated specifically with the spin-coupled wavefunction have been considered by Doggett, Fletcher and co-workers [15]. Their approach differs most significantly in the optimisation procedure adopted. A super-CI (configuration interaction) scheme is used whereby each of the N orbitals is updated in turn, after which the spin-coupling coefficients are re-optimised. This is then repeated until self-consistency. A very significant advantage of this is that excited states (of the same symmetry as the ground state) may also be described. Thus the relaxation of orbitals and spin coupling associated with very specific electronic excitations may be gauged. In many cases it is of paramount importance that ground and/or excited states are described at

equivalent levels of theory.

The complete non-orthogonal generalisation of MCSCF (which is presently referred to as 'VB-SCF') has been considered, first by van Lenthe and Balint-Kurti [16], and more recently by Verbeek and van Lenthe in their work on the programming package 'TURTLE' [17]. The work is based on a super-CI approach which allows for a relatively simple implementation of the orbital optimisation procedure for even the most complicated forms of wavefunction.

The development of methods for non-orthogonal orbital optimisation of multi-configurational wavefunctions is an exciting field, and likely to grow in importance in the future. As has been found both in the case of standard MCSCF as well as for the spin-coupled wavefunction, it is likely that 'direct' optimisation schemes will be superior to the super-CI procedure, both with regard to computational efficiency, as well as in the stability of the convergence. The efficient implementation of such an algorithm is likely to be a monumental task. Work in this field has been undertaken by Penotti [18] in his 'OBS-MCSC' (optimised basis set—multi-configuration spin-coupled) approach. In this, first and second derivatives with respect to orbital, spin-coupling and (spatial) configuration coefficients as well as even basis function parameters (of Slater functions) are evaluated analytically. An early conclusion of this work is that very accurate wavefunctions may be obtained using an impressively small number of variational parameters, with this combination of non-orthogonal orbitals and optimisation of basis function parameters.

Since the computational difficulties associated with the non-orthogonal orbital optimisation are so immense, there has been a greater emphasis on approximate methods here than compared with, for example, standard MCSCF methods. The discussion of such methods lies outside the scope of this chapter, but we shall return to two such approximations—the biorthogonal method and the concept of orbital localisation—in later chapters.

1.2 The spin-coupled wavefunction

For further information on the theory and applications of SC theory we refer the reader to a series of reviews on the subject, refs. [11, 19, 20, 21], and references therein.

The spin-coupled wavefunction for N electrons with spin S is defined as

$$\Psi = \hat{\mathcal{A}}\big[\phi_1(1)\phi_2(2)\cdots\phi_N(N)\Theta_S^N(1,\cdots,N)\big] \tag{1.13}$$

with

$$\Theta_S^N(1,\cdots,N) = \sum_{k=1}^{f_S^N} c_{Sk}\Theta_{S;k}^N(1,\cdots,N). \tag{1.14}$$

Here $\hat{\mathcal{A}}$ is the 'antisymmetriser', ensuring that the total wavefunction adheres to the Pauli principle. $\Theta_S^N(1,\cdots,N)$ is a general spin-eigenfunction for N electrons and spin S, given as the linear combination of f_S^N linearly independent spin functions, as in equation (1.14). These could be defined by the Rumer, Kotani or Serber scheme, or indeed in any way convenient for the problem at hand. The name 'spin-coupled' alludes to the fact that there is complete freedom in the coupling of the electron spins. The electron coordinates are often suppressed in these equations (and will be below). In addition to defining the *form* of the wavefunction, we also have to state the fact that the spin-coupled orbitals are expressed as completely general linear combinations of basis functions:

$$\phi_\mu = \sum_p c_{p\mu}\chi_p, \tag{1.15}$$

the important point in this context being the absence of any orthogonality constraints. The basis functions will in most cases be based on atomic orbitals centred on the nuclei of the molecular system in question. These could be either in the form of (in general contracted) Gaussian type orbitals (GTOs), or Slater orbitals,[1] or, alternatively, in the form of an intermediate basis formed from these, e.g. molecular orbitals found in a preceding Hartree-Fock calculation. The spin-coupled wavefunction is optimised with respect to both the $c_{p\mu}$ coefficients in

[1] For a discussion of the choice of basis sets see for example ref. [22] or [23].

equation (1.15) and the c_{Sk} coefficients in equation (1.14). In general this will be done without the imposition of any constraints besides the trivial normalisation conditions: for the orbitals

$$\langle \phi_\mu | \phi_\mu \rangle = 1, \tag{1.16}$$

and for the spin-coupling coefficients

$$\langle \Theta_S^N | \Theta_S^N \rangle = 1. \tag{1.17}$$

It seems useful at this point, having defined the spin-coupled wavefunction, to present some arguments regarding its appropriateness. The major point in favour of the spin-coupled wavefunction is likely to be its easy interpretation. This is related to the fact that there are not many orbitals, but also the spin coupling provides useful information in an easily digestible form. These statements are particularly true from the viewpoint of most chemists. Quantum chemistry may be said to be situated in the border region between chemistry and physics, and while MO based methods can be said to be highly physical in nature, the concepts of VB methods, especially as provided by spin-coupled theory, are much closer to the way of thinking of most non-theoretical chemists.

The fundamental guideline for all attempts at interpreting the spin-coupled wavefunction is the ansatz that the (first order) description of a covalent bond between two atomic centres is represented by two singlet-coupled electrons associated with two orbitals localised on each of these centres. This notion dates all the way back to the early calculations of Heitler and London [24]. From this perspective, it is more appealing to use the Rumer basis to span the spin space, because there will often be a one to one correspondence between the spin-coupled structures and the resonance structures of classical VB theory. These ideas are prevalent in classical valence bond theory as well, and in many ways the Rumer basis may be termed the traditional VB spin basis.

What distinguishes the spin-coupled wavefunction from alternative forms with non-orthogonal orbital optimisation? Enlarging the wavefunction will decrease

the energy, but at the loss of simplicity and interpretability. So the criterion for a suitable *first order* wavefunction, we argue, is the simplest possible wavefunction that gives qualitatively correct results for the systems of interest.

We shall first consider the possibility of further *simplifying* the spin-coupled wavefunction by way of restricting the spin space. In many cases the approximate spin coupling is obvious with hindsight. For example, often the perfect-pairing functions constitutes 95% or more of the total spin function, so that a PP approximation would be justified. There is no way of knowing this *a priori*, however, and on more than one occasion where the bonding has seemed 'obvious', employing the full spin space has yielded totally unexpected solutions. A commonly adopted approach is to perform an initial optimisation using only one spin function. Thereafter the remaining spin-parameters may be released, and if the one spin function approximation was valid the calculation will typically converge in 2–3 more iterations. The danger of such an approach is when the true solution is radically different from what was supposed, since it is then very easy to converge onto a local minimum on the energy hyper-surface. In actual fact, restricting the spin space can be viewed in much the same way as imposing constraints on the orbitals—both have similar problems. One should always take care not to unduly prejudice the final solution.

Retaining the full spin space is preferable in the general case, so that the spin-coupled wavefunction will normally be the *simplest* wavefunction we will wish to consider. We shall now consider a few general, chemically relevant situations, to give an idea of if, and when, it might be appropriate to augment the single configuration description:

- The fully covalent situation
- Hybridisation
- Charge transfer
- Inherent multi-configurational situations

The fully covalent situation—where the N spin-coupled orbitals are each centred on different atoms—is a situation to which a single configurational wavefunction can be said to be tailored. The most famous example of this is probably the description of the π-electron system of benzene [25]. The six π-orbitals are each centred on a carbon atom, meaning that the five Rumer functions correspond closely to the two Kekulé and the three Dewar structures. The two Kekulé structures dominate, as may be expected, contributing a total of approximately 80% to the wavefunction (the precise figure depending on the basis set used) [25]. Accordingly the orbitals are deformed slightly in the main bond-forming directions—of the neighbouring C atoms. This illustrates the fact that a large proportion of the energy lowering stems from the overlap between the pair-wise singlet-coupled orbitals.

The situation where a single atom forms bonds simultaneously to several other atoms, even in the same resonance structure, is only slightly more complicated. In this case more than one orbital will be associated with a given atom, and these will often reproduce the well-known hybridisation schemes of early quantum chemistry. A good example of this is provided by the example of methane [26]. As in the fully covalent case the different modes of spin coupling correspond to different valence bond structures, and including the full spin space is necessary to avoid any prejudice regarding the final solution. With the core orbital, $1s$ on C, excluded, the picture of the bonding corresponds closely to the classical sp^3 hybridisation scheme on C, with the remaining four electrons being accommodated by $1s$ orbitals on the hydrogens. The sp^x-type hybrids are slightly deformed towards the H atom to which they point, and all hydrogen $1s$ orbitals are slightly distorted towards the central carbon atom.

In both these examples, where the classical description is well known, the spin-coupled picture confirms the qualitative ideas from traditional VB, while augmenting (as, for example, in adding the concept of orbital deformation) and quantifying it. Not only may accurate energies and observables be obtained, but

furthermore questions regarding, for example, details of the spin coupling may be answered. Additionally, in a large number of systems for which no classical VB analogues exist, a consistent interpretation of the bonding is obtained by applying these same concepts to the analysis of the spin-coupled wavefunction.

Systems where charge transfer effects are important occur typically where electro-negative and electro-positive entities are found together. A simple example is provided by the formation of LiH (recently described in ref. [27], but see also references therein). In the long range, the wavefunction can be described as isolated fragments of the lithium and hydrogen atoms, whereas in the bond-forming region the wavefunction adopts significant Li^+H^- character. One might therefore *a priori* expect that a two configuration wavefunction would be necessary to adequately describe the transition between these two extremes—one configuration corresponding to the LiH situation, the other to the Li^+H^- situation. As it turns out, however, (although a two configuration description will always yield a lower energy than a one configuration one) the spin-coupled wavefunction also in this case gives a qualitatively correct description of the bond-forming process. It is a feature of the spin-coupled method that deforming the orbitals away from the purely one-centre case corresponds to the inclusion of so-called ionic (or charge-transfer) structures with purely one-centre orbitals. Analogously, in the spin-coupled description of LiH, a Li($2s$) orbital adopts increasingly more character of an $H^-(1s)$ orbital, in order to give the wavefunction its ionic, or charge transfer, character. So although the two configuration description cannot be ruled out in the case of LiH, the interpretational benefit in having only one set of orbitals seems to favour the single configuration description of this system.

The situations where the single-configuration description demonstrably fails fall into two types. Firstly, there are the cases where more than one configuration is needed in order to obtain an overall wavefunction of the correct symmetry. The simplest example of this is for the dissociation limit of H_2^+, where two configurations corresponding to H^+–H and H–H^+ are needed. The spin-coupled wavefunc-

tion will converge onto one of these, and while this may be perfectly acceptable on energy grounds and for most practical purposes, it is not aesthetically very pleasing. In a few theoretical investigations (as for example the calculation of interionic potentials treated in a later chapter) it may be desirable to have a symmetry-pure wavefunction. In such cases the spin-coupled wavefunction may be easily 'symmetrised' (for example by applying the symmetry operations of the system to the spin-coupled wavefunction and solving the generalised eigenvalue problem defined by the thus generated set of structures). However, in such cases the ability to create a fully variational wavefunction of the correct symmetry would probably be preferable. The second type of failing of the single-configuration description is more obscure, having to do with effects going beyond the purely 'non-dynamical' electron correlation provided by the spin-coupled method. Examples of such are extremely rare. One such is the case of ozone (O_3) where two conflicting spin-coupled descriptions lie very close in energy (see chapter 4). Typically one finds for such systems that the equivalent N in N CASSCF (complete active space self-consistent field) calculation is also inadequate.

We can conclude from this discussion that the cases where the spin-coupled method fails to get the *qualitative* picture right, although they certainly exist and are easily characterisable, are few and far between indeed. This is in direct opposition to the traditional (pre-CASSCF) MCSCF methods, where the included configurations must be chosen carefully to suit the chemical problem at hand. As such spin-coupled can be viewed as the natural first order method in modern valence bond methods, and may be said to be for VB what the Hartree-Fock method is for MO theory.

As for the Hartree-Fock method, the spin-coupled description may be improved, i.e., further electron correlation may be included, by way of a non-orthogonal CI step. The *initial* convergence of this expansion is in general rather better than for corresponding orthogonal CI calculations. However, due to the computational cost associated with setting up and solving the generalised

eigenvalue problem the number of CSFs that can be treated is correspondingly smaller—at least at the present stage of development. So for equivalent basis sets the orthogonal methods generally 'win out' in terms of the correlation energy that can be recovered.

If VB methods in this way can be said to lag behind MO theory in the quantitative accuracy attainable, it is in the interpretation of the wavefunction that valence bond methods come into their own (the interpretation of the spin-coupled wavefunction will be described further in section 1.6). The use of classical VB concepts, as well as the modern development of these ideas in the present formulation, are powerful interpretational tools indeed. This particularly holds true for the 'traditional' chemists for which the qualitative bonding picture, or the description of bond formation/breaking, is important. Traditionally quantum chemistry has had very little to offer this group of people, being more oriented towards physicists, physical chemists or spectroscopists. There is therefore the potential of a huge turn-around in the way of thinking of this section of the scientific community, if the ideas of spin-coupled theory can be made to gain wider acceptance.

1.3 Optimising the spin-coupled wavefunction

For both MCSCF and SC optimising with respect to the linear spin-coupling coefficients, or coefficients of configuration state functions (CSFs), is not problematic, this for example being achieved by solving a matrix eigenvalue problem. It is the optimisation of the non-linear orbital parameters which represents the major difficulty. For this purpose McWeeny identifies three main families of methods in the case of MCSCF wavefunctions: Fock operator based methods, methods based on the direct minimisation of the energy using first (and perhaps higher) derivatives of the energy, and finally methods aimed at satisfying a generalised Brillouin condition, i.e., ensuring that matrix elements between the

optimised state and singly excited configurations over the Hamiltonian (of the form $\langle\Psi_0|\hat{H}|\Psi_i^j\rangle$) become zero [23]. All these have also been applied to the optimisation of the spin-coupled wavefunction, the conclusion being that it is the direct optimisation schemes that provide the most rapid and stable convergence combined with maximum computational efficiency (Brillouin condition methods in the form of 'super-CI'-strategies may have other advantages as mentioned in section 1.2).

The expectation value for the (electronic) energy of the spin-coupled wavefunction may be written on the form

$$E = \frac{W}{\mathbf{D}} = \frac{1}{\mathbf{D}}\Big(\sum_{\mu,\nu}\mathbf{D}(\mu|\nu)\langle\phi_\mu|\hat{H}_1|\phi_\nu\rangle + \frac{1}{2}\sum_{\mu,\nu,\sigma,\tau}\mathbf{D}(\mu\nu|\sigma\tau)\langle\phi_\mu\phi_\nu|\hat{H}_2|\phi_\sigma\phi_\tau\rangle\Big), \tag{1.18}$$

where $\mathbf{D}$, $\mathbf{D}(\mu|\nu)$ and $\mathbf{D}(\mu\nu|\sigma\tau)$ are the zeroth, first and second order density matrices respectively. $\hat{H}_1$ and $\hat{H}_2$ are the one- and two-electron parts of the Hamiltonian. We can approach this as a standard optimisation problem of numerical analysis, optimising the function E with respect to all variational parameters—in this case the orbital coefficients, $c_{p\mu}$, and spin-coupling coefficients, c_{Sk}. It is generally considered necessary to employ a complete, second order optimisation procedure to ensure rapid and stable convergence. We shall therefore first consider the expressions for the first and second derivatives, the computational difficulties encountered in their evaluation, after which we proceed to describe the various second order methods currently in use.

1.3.1 First and second derivatives of the energy

The derivative of the fraction given in (1.18) with respect to a generic parameter, P, is given by

$$\frac{\partial}{\partial P}\frac{W}{\mathbf{D}} = \mathbf{D}^{-1}\frac{\partial W}{\partial P} - \frac{W}{\mathbf{D}^2}\frac{\partial \mathbf{D}}{\partial P}, \tag{1.19}$$

whereas the second derivative, with respect to P_1 and P_2, in a similar manner can be seen to be

$$\frac{\partial^2}{\partial P_1 \partial P_2}\frac{W}{\mathbf{D}} = \mathbf{D}^{-1}\frac{\partial^2 W}{\partial P_1 \partial P_2} - \mathbf{D}^{-2}\left(\frac{\partial W}{\partial P_1}\frac{\partial \mathbf{D}}{\partial P_2} + \frac{\partial W}{\partial P_2}\frac{\partial \mathbf{D}}{\partial P_1} + W\frac{\partial^2 \mathbf{D}}{\partial P_1 \partial P_2}\right)$$
$$+2\mathbf{D}^{-3}W\frac{\partial \mathbf{D}}{\partial P_1}\frac{\partial \mathbf{D}}{\partial P_2}. \tag{1.20}$$

These expressions are straightforward but rather cumbersome, so we shall restrict further efforts directly to the first and second derivatives of the overlap, $\mathbf{D}$, and un-normalised energy, W.

The simplest derivatives to consider are those with respect to the spin-coupling coefficients which may all be derived from the relation

$$\mathbf{D} = \sum_{k,l} c_{Sk} c_{Sl} \mathbf{D}_{kl}. \tag{1.21}$$

In this equation c_{Sk} and c_{Sl} are the spin-coupling coefficients encountered in (1.14) and $\mathbf{D}_{kl}$ is the transition density matrix defined from the two structures with spin functions $\Theta^N_{S;k}$ and $\Theta^N_{S;l}$ respectively. From this we get

$$\frac{\partial \mathbf{D}}{\partial c_{Sm}} = \sum_k c_{Sk}\left(\mathbf{D}_{km} + \mathbf{D}_{mk}\right) \equiv \mathbf{D}_m, \tag{1.22}$$

where we have conveniently defined the symmetrised, single-index density matrix, $\mathbf{D}_m$, and

$$\frac{\partial^2 \mathbf{D}}{\partial c_{Sm}\partial c_{Sn}} = \mathbf{D}_{mn} + \mathbf{D}_{nm}. \tag{1.23}$$

These results hold for any order of density matrix.

For the dependence of an orbital on the change of any given orbital parameter, we may use (1.15) to get

$$\frac{\partial \phi_\nu}{\partial c_{p\mu}} = \delta_{\mu\nu}\chi_p. \tag{1.24}$$

When considering expressions for the overlap or energy one must strictly speaking consider both updates to the bra and to the ket form of the orbital. This can be expressed

$$\frac{\partial}{\partial c_{p\mu}} = \frac{\partial}{\partial \langle c_{p\mu}|} + \frac{\partial}{\partial |c_{p\mu}\rangle}, \tag{1.25}$$

where $\langle c_{p\mu}|$ is the coefficient of the bra-orbital and $|c_{p\mu}\rangle$ the coefficient of the ket-orbital (these will in general be complex conjugates). While this consideration is only really important in the case of complex orbitals and coefficients, it nevertheless represents a useful short-cut to obtaining the derivative expressions. Not only is it simpler to consider the changes to bra- and ket-orbitals separately, but due to the bra-ket symmetry of the expressions we shall consider, the two terms in equation (1.25) will always be each others complex conjugate.

The orbitals will enter the energy expression in the form of overlap, one-electron or two-electron integrals. In all these cases the operations of differentiation and integration may be interchanged to give

$$\frac{\partial}{\partial\langle c_{p\mu}|}\langle\phi_{\mu_1}|\phi_{\nu_1}\rangle = \delta_{\mu\mu_1}\langle\chi_p|\phi_{\nu_1}\rangle, \tag{1.26}$$

$$\frac{\partial}{\partial\langle c_{p\mu}|}\langle\phi_{\mu_1}|\hat{H}_1|\phi_{\nu_1}\rangle = \delta_{\mu\mu_1}\langle\chi_p|\hat{H}_1|\phi_{\nu_1}\rangle \tag{1.27}$$

and

$$\begin{aligned}\frac{\partial}{\partial\langle c_{p\mu}|}\langle\phi_{\mu_1}\phi_{\mu_2}|\hat{H}_2|\phi_{\nu_1}\phi_{\nu_2}\rangle &= \delta_{\mu\mu_1}\langle\chi_p\phi_{\mu_2}|\hat{H}_2|\phi_{\nu_1}\phi_{\nu_2}\rangle \\ &+\delta_{\mu\mu_2}\langle\phi_{\mu_1}\chi_p|\hat{H}_2|\phi_{\nu_1}\phi_{\nu_2}\rangle\end{aligned} \tag{1.28}$$

for the three cases respectively.

We note first that an orbital derivative of a density matrix is independent of the form of a bra-orbital, say, if the corresponding orbital label occurs on the left-hand side of the vertical bar (similarly for a ket-orbital). In other words

$$\frac{\partial}{\partial\langle c_{p\mu}|}\mathbf{D}(\mu_1\cdots\mu_n|\nu_1\cdots\nu_n) = 0 \quad \text{if} \quad \mu\in\{\mu_1,\cdots,\mu_n\}. \tag{1.29}$$

The density matrix with a μ label on the left-hand side is constructed from determinants of matrices in which the corresponding row has been *removed* (cf. appendix 1.a). If μ does not occur on the left-hand side the simplest way of proceeding is by using the 'induction' relation between different orders of density matrix. The general form of this is

$$\mathbf{D}(\mu_1\cdots\mu_{n-1}|\nu_1\cdots\nu_{n-1}) = \sum_{\nu_n}\mathbf{D}(\mu_1\cdots\mu_{n-1}\mu_n|\nu_1\cdots\nu_{n-1}\nu_n)\langle\phi_{\mu_n}|\phi_{\nu_n}\rangle, \tag{1.30}$$

for the relationship between density matrices of orders $n-1$ and n. This holds for any μ_n not contained in the set $\{\mu_1, \cdots, \mu_{n-1}\}$. Applying this to the evaluation of the derivative gives

$$\frac{\partial}{\partial\langle c_{p\mu}|}\mathbf{D}(\mu_1\cdots\mu_n|\nu_1\cdots\nu_n) = \frac{\partial}{\partial\langle c_{p\mu}|}\sum_{\nu}\mathbf{D}(\mu_1\cdots\mu_n\mu|\nu_1\cdots\nu_n\nu)\langle\phi_\mu|\phi_\nu\rangle$$

$$= \sum_{\nu}\mathbf{D}(\mu_1\cdots\mu_n\mu|\nu_1\cdots\nu_n\nu)\langle\chi_p|\phi_\nu\rangle. \tag{1.31}$$

Indeed, since the value of the density matrix is zero if repetition occurs in either the μ or ν labels, equation (1.31) can be seen to encompass also the case of (1.29). This is the conventional form of the orbital derivative of a density matrix, given in terms of only the bra-orbital coefficient.[2]

Summing up we can see that since differentiating with respect to an orbital parameter requires a density matrix of one order higher, whereas the derivative with respect to a spin-coupling parameter adds a spin-coupling label to the density matrix, the following density matrices are needed:

- $\mathbf{D}_{kl}$ required up to second order for the spin-spin part of the Hessian
- $\mathbf{D}_k$ required up to third order for the spin-orbital part of the Hessian
- $\mathbf{D}$ required up to fourth order for the orbital-orbital part of the Hessian

In addition various non-, semi- and fully transformed integrals are used:

- Overlap integrals: $\langle\phi_\mu|\phi_\nu\rangle$, $\langle\chi_p|\phi_\mu\rangle$ and $\langle\chi_p|\chi_q\rangle$.
- One-electron integrals: $\langle\phi_\mu|\hat{H}_1|\phi_\nu\rangle$, $\langle\chi_p|\hat{H}_1|\phi_\mu\rangle$ and $\langle\chi_p|\hat{H}_1|\chi_q\rangle$.
- Two-electron integrals: $\langle\phi_\mu\phi_\nu|\hat{H}_2|\phi_\sigma\phi_\tau\rangle$, $\langle\chi_p\phi_\mu|\hat{H}_2|\phi_\nu\phi_\sigma\rangle$, $\langle\chi_p\chi_q|\hat{H}_2|\phi_\mu\phi_\nu\rangle$ and $\langle\chi_p\phi_\mu|\hat{H}_2|\chi_q\phi_\nu\rangle$.

[2]The reader may wish to verify that the complete change with respect to an orbital coefficient, $c_{p\mu}$, in the case of real orbitals indeed reduces to the sum of changes with respect to $\langle c_{p\mu}|$ and $|c_{p\mu}\rangle$. (This can for example be done by using (1.30) twice, expanding first over a bra-label and then over a ket-label. In this way μ can be made to occur on *both* sides in the density matrix.)

The semi-transformed integrals may be extracted from most transformation programs with very little additional effort, since these normally transform one index at a time.

These quantities are basic for *any* exact second order method for the optimisation of the spin-coupled wavefunction, although they need not necessarily occur explicitly in the above form. The time spent transforming integrals is likely to be significant only in cases with relatively large numbers of electrons and underlying basis functions. So for the majority of applications, evaluating the density matrices will be the most demanding task. At present what might be termed a 'brute force' approach is employed. The spin-coupled wavefunction is written in terms of Slater determinants (as are the f_S^N spin-coupled structures), which means the density matrices are expressed in terms of simple cofactors of the overlap matrix between (spin-) orbitals (see appendix 1.a). A detailed account of the algorithm in use may be found in refs. [12, 28]. Alternatives to this 'brute force' scheme could, for example, entail using a different orbital basis to express the density matrices (or even different bra- and ket- orbital sets as in the 'biorthogonal' method). Such schemes could serve to simplify the *structure* of the density matrices. Another possibility is the use of group theoretical methods to establish non-obvious relations between density matrices of different orders or in terms of different spin functions. Various such schemes have been tested in the past, but so far the 'super-cofactor approach' has proven to be clearly the most efficient.

1.3.2 Optimisation strategies

In choosing a given optimisation scheme the behaviour of the objective function must obviously be considered. As such the behaviour far from convergence is important to gauge the stability of a given strategy, whereas the behaviour close to convergence will determine the (maximum) convergence multiplicity (quadratic-, cubic- etc.) of a given algorithm. But apart from this, it is important to realise the relative cost of computing the various quantities required for a given procedure.

Such considerations may dictate that some variables be updated more often than others, or that other quantities be only approximated or estimated from previous iterations. Other considerations regarding the choice of a given optimisation procedure include the handling of constraints and the elimination of redundant variables. This has been discussed in more detail in ref. [13] where the optimisation of the core orbital set warranted a more sophisticated treatment of redundant variables and constraints than is probably otherwise necessary. A noteworthy observation from this investigation is that the incorporation of the constraints *into* the objective function (as for example in Lagrange multiplier methods) actually hinders the convergence by making the function not as well behaved. So in this case a proper non-linear transformation to the set of free variables is worthwhile. In all these considerations the non-orthogonality of the orbitals is, of course, of very little relevance, and it is therefore possible to take advantage of the traditional MCSCF literature. Here a vast amount of effort has been invested in order to obtain computationally efficient optimisation procedures—reviews may for example be found in refs. [29, 30]. It is likely that significant advances can be made in spin-coupled theory by in this way borrowing ideas from MCSCF.

If it is the *ground state* of the molecule which is of interest, an optimisation procedure must locate the global minimum on the energy hyper-surface. We may arrange all free variational parameters in a vector, $\mathbf{x}$, and it is then the objective function $E(\mathbf{x})$ that must be optimised. (Note in particular that redundancies connected with the normalisation conditions (1.16) and (1.17) must be eliminated from $\mathbf{x}$.) A minimum in the energy is characterised by having no first order change, and any second order change leads to an increase in energy, i.e., the gradient is zero

$$\mathbf{g} = \mathbf{0} \tag{1.32}$$

and the Hessian positive (semi) definite, which may be expressed

$$\tilde{\mathbf{s}}\mathbf{G}\mathbf{s} \geq 0, \qquad \forall \mathbf{s}. \tag{1.33}$$

Here $\mathbf{g}$ and $\mathbf{G}$ are respectively the gradient vector and Hessian matrix corre-

sponding to the variational parameter vector $\mathbf{x}$. These criteria may define a *local* minimum, but convergence onto such a point is normally quite rare, and can as a rule be easily diagnosed from the nature of the solution.

Conditions corresponding to (1.32) and (1.33) for the case of excited states may also be constructed (see for example ref. [29]). These will also be stationary points (i.e., (1.32) must be satisfied), but the Hessian will generally have one or more negative eigenvalues. Optimisation procedures for such cases will be a bit more involved than for the simple minimisation problem, and we shall forego the discussion of these.

If one assumes a second order behaviour of the objective function the condition (1.32) can be satisfied as in the Newton-Raphson procedure by solving

$$\mathbf{G}\boldsymbol{\delta} = -\mathbf{g} \tag{1.34}$$

for the update $\boldsymbol{\delta}=\mathbf{x}'-\mathbf{x}$. In realistic cases, i.e. cases that are not exactly second order, an iterative procedure will be required to obtain a solution, $\mathbf{x}$, where (1.32) is satisfied.

The Newton-Raphson scheme is rarely adopted without some form of trust region control. It is generally of paramount importance for the stability of optimisation procedures (for a discussion see for example ref. [33]). If the third and higher derivatives of the energy all take reasonable values, the second order model can be expected to hold with fair accuracy within a region around the current point in the parameter space (defined by the vector $\mathbf{x}$). The present strategy is to define a trust sphere size, H, above which no correction is accepted, i.e.,

$$\sqrt{\tilde{\boldsymbol{\delta}}\boldsymbol{\delta}} < H. \tag{1.35}$$

The trust sphere size is normally given an initial value of 0.1, and then adjusted up or down in subsequent iterations according to the accuracy of the second order method. This may be gauged by comparing the actual energy improvement with that predicted from the second order method:

$$r = \frac{E' - E}{\tilde{\mathbf{g}}\boldsymbol{\delta} + \frac{1}{2}\tilde{\boldsymbol{\delta}}\mathbf{G}\boldsymbol{\delta}}. \tag{1.36}$$

A scheme for choosing the next trust sphere size, H, based on this ratio, r, may be found in ref. [33]. Updates leading to actual increase in energy are rejected, followed by a severe decrease in the trust sphere size. Restricted step methods are subtly different from (1.34) in that in these the objective function is sought minimised, based on the second order model, subject to the constraint (1.35).

Provided that the Hessian matrix is positive definite from the outset, the Newton-Raphson algorithm will normally converge with reasonable stability and quadratic convergence close to the minimum. In most cases, of course, the Hessian will have one or more negative eigenvalues in which case the Newton-Raphson scheme is just as likely to converge onto an alternative stationary point. To amend this problem, the 'GQT method' may be employed [31]. In this approach equation (1.34) is replaced by

$$(\mathbf{G} + \alpha\mathbf{I})\boldsymbol{\delta} = -\mathbf{g}, \tag{1.37}$$

where $\mathbf{I}$ is the unit matrix, and the parameter α ensures that $(\mathbf{G} + \alpha\mathbf{I})$ is positive definite. In practical calculations a value for α is chosen such that the update-size is maximum subject to (1.35). The GQT method can be viewed as a linear combination of the Newton-Raphson and steepest decent methods [10]. When the Hessian in the course of the optimisation procedure becomes positive definite, the parameter α in equation (1.37) is set to zero whereby the GQT method reduces to the pure Newton-Raphson procedure.

In a variation on the basic GQT method, the *direction* of the previous update is taken into account [32]. The α parameter is determined as in the basic GQT method, but the unit matrix in equation (1.37) is replaced by the matrix $\mathbf{Q}$, according to

$$Q_{ij} = \delta_{ij} + \frac{(\beta^2 - 1)d_i d_j}{\mathbf{d}\tilde{\mathbf{d}}}. \tag{1.38}$$

Here $\mathbf{d}$ is the update-vector from the previous iteration, and β an adjustable parameter given an initial value of 0.9, but varied in subsequent iterations according to the quality of updates produced. This method is commonly referred to as the

'GQT2 method'. It has the effect of modifying the trust region from the purely spherical in the case of GQT, to an ellipsoid in the case of GQT2, elongated in the direction of the previous update.

For most applications, if a 'reasonable' starting guess is employed, convergence may be achieved in 10–15 iterations, the last 2–3 being taken up by the pure Newton-Raphson procedure. Convergence is defined by the satisfaction of the following three criteria

1. The Hessian is positive definite
2. The gradient is 'small'
3. The predicted update is 'small': $\delta_i < 10^{-6}$

Here 2. and 3. are of course interrelated, 3. chosen as the determining factor in the present strategy. The exact numerical values in the case of 2. and 3. may be varied according to the application. Note that obtaining the correct coefficients within 10^{-6} means that the spin-coupled energy is determined with an accuracy of 10^{-12}. In some cases the converged solution will have an effectively singular Hessian due to (near-) redundancies in the parameter space, in which cases the third criterion above may be dispensed with. Furthermore, in order to increase the stability of the optimisation procedure the Hessian may in such cases be modified slightly, according to equation (1.37)—even close to convergence. Setting α=1×10^{-4}, say, thus making the Hessian just positive definite will stabilise the size of the updates in the convergence limit.

1.4 Core-valence separation

A complete spin-coupled treatment is possible for only the smallest systems. At present much more than about 14 electrons is too computer intensive to be feasible on most available machines. While this number is likely to increase somewhat with increasing computational power and further development of the spin-coupled

method, there will clearly always be the need to include more electrons than the spin-coupled method itself can handle. Often a particular group of electrons will be particularly important for a given application, so these may be given a more refined description, while the remaining electrons are described more crudely in order to limit the computational difficulties. In the area of chemistry it is the valence electrons that are of particular interest, which is the justification for the core/valence terminology, but this should not be taken too literally.

In this section we shall concentrate our efforts on describing a core consisting only of doubly occupied orbitals. Alternatives, such as the usage of a SOPP-GVB core, have been proposed [13] but no applications are available at the present time. With such a core the spin-coupled wavefunction may be written

$$\Psi = \hat{\mathcal{A}}\left[\psi_1^2\psi_2^2\cdots\psi_{N_c/2}^2\phi_1\phi_2\cdots\phi_N\Theta_{00;f_0^{N_c}}^{N_c}\Theta_S^N\right] \tag{1.39}$$

each pair of core electrons by necessity being singlet-coupled, as signified by $\Theta_{00;f_0^{N_c}}^{N_c}$. With this definition, it can be shown that the wavefunction is invariant under orbital transformations of the form

$$\psi_j \rightarrow \psi_j + \lambda\psi_i, \qquad i \neq j \tag{1.40}$$

and

$$\phi_\mu \rightarrow \phi_\mu + \lambda\psi_i. \tag{1.41}$$

Writing the wavefunction as a linear combination of Slater determinants shows that such transformations correspond to 'column operations' on a subset of these determinants. That a determinant is unchanged by column operations is a basic result in linear algebra. For the purposes of this section, the consequences of this are that the core orbitals may be orthogonalised and the valence orbitals orthogonalised on the core orbital space. This is useful because it simplifies the expression for the energy and we shall in the following assume that such an orthogonalisation has taken place.

We shall take the form of the energy expression (1.18) as our starting point.

Letting ξ signify either a core or a valence orbital this becomes

$$E = \frac{W}{\mathbf{D}} = \frac{1}{\mathbf{D}}\Big(\sum_{a,b}\mathbf{D}(a|b)\langle\xi_a|\hat{H}_1|\xi_b\rangle + \frac{1}{2}\sum_{a,b,c,d}\mathbf{D}(ab|cd)\langle\xi_a\xi_b|\hat{H}_2|\xi_c\xi_d\rangle\Big), \quad (1.42)$$

where the a, b, c, d labels now assume the full range from 1 to $N + N_c$. This expression is general for any form of wavefunction with the proviso that the summations over orbitals be over all non-virtuals. In the present case, however, it is possible to simplify the evaluation of this expression considerably by identifying the many vanishing density matrix elements.

The first thing to note is that the overlap may be factorised into core and valence parts according to $\mathbf{D}$=$\mathbf{D}^c\times\mathbf{D}^v$. With orthonormal core orbitals we furthermore have $\mathbf{D}^c$=1.

The core one-electron density matrix is identical to the Hartree-Fock case, namely

$$\mathbf{D}^c(i|j) = 2\delta_{ij}, \quad (1.43)$$

whereas the core-valence block, due to the previously performed orthogonalisation, is zero

$$\mathbf{D}(i|\mu) = 0. \quad (1.44)$$

Similarly to equation (1.44) the two-electron density matrix will be zero if the numbers of spin-coupled orbital labels on the bra- and the ket-side are not the same. With at least one core orbital label on each side we can use the relation

$$\mathbf{D}(ab|cd) = 1/2\mathbf{D}(a|c)\mathbf{D}(b|d) - 1/4\mathbf{D}(a|d)\mathbf{D}(b|c). \quad (1.45)$$

This is a well known result in the case of a single determinant of doubly occupied orbitals, but using the formulae given in appendix 1.a and performing the spin integration will yield this result also in the present case. Direct application of this equation combined with (1.43) leads to the expressions for the core two-electron density matrix as

$$\mathbf{D}^c(ij|ij) = 2$$

$$\mathbf{D}^c(ij|ji) = -1, \quad (1.46)$$

which is a well known result in Hartree-Fock theory. For the core-valence block of the two-electron density matrix one similarly gets

$$\mathbf{D}(\mu i|\nu i) = \mathbf{D}(i\mu|i\nu) = \mathbf{D}(\mu|\nu)$$

$$\mathbf{D}(\mu i|i\nu) = \mathbf{D}(\mu i|i\nu) = -1/2\mathbf{D}(\mu|\nu). \tag{1.47}$$

Inserting all these expressions into (1.42) then gives for the energy

$$\begin{aligned} E = 2\sum_i \langle\psi_i|\hat{H}_1|\psi_i\rangle + \sum_{i,j}\Big(2\langle\psi_i\psi_j|\hat{H}_2|\psi_i\psi_j\rangle - \langle\psi_i\psi_j|\hat{H}_2|\psi_j\psi_i\rangle\Big) \\ + \frac{1}{\mathbf{D}^c}\Big\{\sum_{\mu,\nu}\mathbf{D}(\mu|\nu)\langle\phi_\mu|\hat{H}_1|\phi_\nu\rangle \\ + \sum_{i,\mu,\nu}\mathbf{D}(\mu|\nu)\Big(\langle\phi_\mu\psi_i|\hat{H}_2|\phi_\nu\psi_i\rangle - 1/2\langle\phi_\mu\psi_i|\hat{H}_2|\psi_i\phi_\nu\rangle\Big) \\ + \frac{1}{2}\sum_{\mu,\nu,\sigma,\tau}\mathbf{D}(\mu\nu|\sigma\tau)\langle\phi_\mu\phi_\nu|\hat{H}_2|\phi_\sigma\phi_\tau\rangle\Big\}. \end{aligned} \tag{1.48}$$

If one wishes to optimise also the core orbitals, the derivatives of this expression with respect to the respective orbital variations may be worked out and incorporated into a second order optimisation scheme in the same way as was done for the spin-coupled orbitals in section 1.3. This has been done in ref. [13]. Other than the obvious theoretical advantages of having a fully variational wavefunction, optimising the core is likely to be important when the distribution of the valence orbitals differs widely from the, say, SCF description which would otherwise determine the core.

We may define a Fock operator for the core orbitals, in a similar way as done in Hartree-Fock theory, according to

$$\hat{F} = \hat{H}_1 + \sum_i(2\hat{J}_i - \hat{K}_i). \tag{1.49}$$

In this way the energy expression (1.48) becomes

$$\begin{aligned} E = \sum_i\langle\psi_i|\hat{H}_1 + \hat{F}|\psi_i\rangle + \frac{1}{\mathbf{D}^c}\Big\{\sum_{\mu,\nu}\mathbf{D}(\mu|\nu)\langle\phi_\mu|\hat{F}|\phi_\nu\rangle \\ + \frac{1}{2}\sum_{\mu,\nu,\sigma,\tau}\mathbf{D}(\mu\nu|\sigma\tau)\langle\phi_\mu\phi_\nu|\hat{H}_2|\phi_\sigma\phi_\tau\rangle\Big\}. \end{aligned} \tag{1.50}$$

This formulation highlights the essential uncorrelated nature of the core electrons, and the lack of correlation between the motion of the core and valence electrons, the interaction being determined from simple Coulomb and exchange integrals.

Most applications in this thesis have been carried out using a frozen (i.e., not optimised) core, which in the majority of cases is more than satisfactory. It is clear from equation (1.50) that in this case optimising the spin-coupled orbitals can be done exactly as described in section 1.3 provided that the integrals over the Fock operator rather than the usual one-electron integrals are used (this was first shown in ref. [34]). Thus, as far as the computer intensive part of the calculation is concerned, the problem is reduced from an N_c+N to only an N electron problem.

The problem of choosing a core/valence separation is rarely of significance if it is possible to treat all *chemical* valence electrons by the spin-coupled method. In that case, if the SCF description is not wildly inadequate, the core may be chosen as the $N_c/2$ lowest lying (in orbital energy) molecular orbitals from the Hartree-Fock calculation (which in general may be open-shell). Although these will be delocalised over the entire molecule, they will be equivalent to the usual (but distorted) $1s^2$ etc. atomic orbitals (by a linear transformation). If the Hartree-Fock description fails to describe the core orbitals sensibly, these may be obtained from MCSCF or (more commonly) CASSCF calculations. If there is ambiguity the core orbitals may as before be identified from the orbital energies.

In the cases where the set of spin-coupled electrons does not coincide with the complete set of valence electrons, powerful localisation procedures may be used to separate the occupied SCF MOs (reviews may be found in ref. [35]). These serve to define a separation into orbitals associated with the core, individual chemical bonds and lone pairs. In most of these, a unitary transformation of the orbitals is sought (often iteratively) that satisfies a given localisation criterion. Two procedures, that have been used for a number of years, are based on, in some sense, minimising the average separation between two electrons occupying

a given MO. In Boys' scheme [36], the norm of the inter-electronic distance is minimised

$$\mathrm{B}\{\psi\} = \sum_i \langle \psi_i \psi_i | (\mathbf{r}_1 - \mathbf{r}_2)^2 | \psi_i \psi_i \rangle \tag{1.51}$$

with the sum including all m_{loc} MOs that are sought separated. In a procedure due to Edminston and Ruedenberg [37] the 'self-repulsion' energy is similarly maximised

$$\mathrm{ER}\{\psi\} = \sum_i \langle \psi_i \psi_i | \frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|} | \psi_i \psi_i \rangle. \tag{1.52}$$

Comparing these two schemes, the orbitals from the E-R schemes tend to be the more completely localised. Another advantage is that E-R preserves σ-π separation whereas the Boys scheme always gives so-called banana bonds [38]. Computationally, however, Boys's scheme is preferable, the effort scaling only as m_{loc}^3 rather than m_{loc}^5 for the E-R procedure.

A recent scheme giving orbitals with the same characteristics as in the Edminston-Ruedenberg scheme, but being computationally simpler is the Pipek-Mezey procedure [38]. This method is based on the Mulliken [39] populations for individual orbitals. For each orbital a measure of the delocalisation may be defined from

$$d_i = \Big\{ \frac{1}{4} \sum_A Q_{Ai}^2 \Big\}^{-1}, \tag{1.53}$$

where Q_{Ai} is the population of ψ_i on centre A, and the sum includes all atomic centres in the molecule. This quantity is a measure of the number of centres the orbital is delocalised over. For example for an orbital partitioned equally between P centres, we have

$$d_i = \Big\{ \frac{1}{4} P \Big(\frac{2}{P} \Big)^2 \Big\}^{-1} = P. \tag{1.54}$$

Having defined the delocalisation for a single orbital the mean delocalisation, defined from all MOs to be localised:

$$\mathrm{PM}\{\psi\} = m_{\mathrm{loc}} \Big\{ \sum_i d_i^{-1} \Big\}^{-1}, \tag{1.55}$$

may then be minimised as for the two other schemes. As an example of this, the valence MOs of diborane, B_2H_6, were localised by this scheme. The orbital popu-

MOs:	B_1	B_2	H_1^b	H_2^b	H_1^t	H_2^t	H_3^t	H_4^t
1	.5830	.5830	.3004	.3004	.0583	.0583	.0583	.0583
2	.5041	.5041	.0854	.0854	.2052	.2052	.2052	.2052
3	.3905	.3905	.6095	.6095	.0000	.0000	.0000	.0000
4	.5446	.5446	.0000	.0000	.2277	.2277	.2277	.2277
5	.5375	.5375	.0000	.0000	.2312	.2312	.2312	.2312
6	.4570	.4570	.0000	.0000	.2715	.2715	.2715	.2715
LMOs:	B_1	B_2	H_1^b	H_2^b	H_1^t	H_2^t	H_3^t	H_4^t
1	.4896	.4896	1.0322	.0004	-.0029	-.0029	-.0029	-.0029
2	1.0404	-.0216	-.0093	-.0093	1.0048	-.0076	.0017	.0010
3	.4896	.4896	.0004	1.0322	-.0029	-.0029	-.0029	-.0029
4	-.0216	1.0404	-.0093	-.0093	.0010	.0017	-.0076	1.0048
5	-.0216	1.0404	-.0093	-.0093	.0017	.0010	1.0048	-.0076
6	1.0404	-.0216	-.0093	-.0093	-.0076	1.0048	.0010	.0017

Table 1.1: Mulliken populations for the valence orbitals of B_2H_6.

lations before and after localisation are shown in table 1.1. After the localisation it is quite clear that orbitals 1 and 3 correspond to the three-centre two-electron bonds between the boron atoms and the bridging hydrogens, whereas the remaining 4 orbitals each correspond to a bond between a boron and terminal hydrogen atom. So if the spin-coupled description of the *bridging* region is of interest, the localised MOs 2,4,5 and 6 should be included in the core. For a comparison of the three localisation procedures in the case of B_2H_6, see ref. [38].

1.5 The Fock operators

In a similar way as may be shown in the case of MCSCF theory (see for example ref. [23]) the derivative of the energy expression with respect to an orbital parameter leads to a Fock-type eigenvalue equation. Due to the non-orthogonal orbitals the derivation is slightly more involved in this case since the density matrices here

are dependent on the orbital parameters. For the spin-coupled wavefunction the derivative of the energy expression (1.18) can be written

$$\frac{\partial}{\partial\langle c_{p\mu}|}\frac{1}{\mathbf{D}}\Big\{\sum_{\mu_1,\nu_1}\mathbf{D}(\mu_1|\nu_1)\langle\phi_{\mu_1}|\hat{H}_1|\phi_{\nu_1}\rangle$$

$$+\frac{1}{2}\sum_{\mu_1,\mu_2,\nu_1,\nu_2}\mathbf{D}(\mu_1\mu_2|\nu_1\nu_2)\langle\phi_{\mu_1}\phi_{\mu_2}|\hat{H}_2|\phi_{\nu_1}\phi_{\nu_2}\rangle\Big\}$$

$$=\frac{1}{\mathbf{D}}\Big\{\sum_{\nu}\mathbf{D}(\mu|\nu)\langle\chi_p|\hat{H}_1|\phi_\nu\rangle+\sum_{\mu_1,\nu_1,\nu}\mathbf{D}(\mu\mu_1|\nu\nu_1)\langle\chi_p|\phi_\nu\rangle\langle\phi_{\mu_1}|\hat{H}_1|\phi_{\nu_1}\rangle$$

$$+\sum_{\mu_1,\nu_1,\nu}\mathbf{D}(\mu\mu_1|\nu\nu_1)\langle\chi_p\phi_{\mu_1}|\hat{H}_2|\phi_\nu\phi_{\nu_1}\rangle$$

$$+\frac{1}{2}\sum_{\mu_1,\mu_2,\nu_1,\nu_2,\nu}\mathbf{D}(\mu\mu_1\mu_2|\nu\nu_1\nu_2)\langle\chi_p|\phi_\nu\rangle\langle\phi_{\mu_1}\phi_{\mu_2}|\hat{H}_2|\phi_{\nu_1}\phi_{\nu_2}\rangle$$

$$-\frac{W}{\mathbf{D}}\sum_{\nu}\mathbf{D}(\mu|\nu)\langle\chi_p|\phi_\nu\rangle\Big\}. \tag{1.56}$$

This may be shown by direct application of the formulae in section 1.3. The two first terms here relate to the derivative of the one-electron energy, the two next to the two-electron energy, and the last is the derivative of the normalisation integral $\mathbf{D}$. We can rewrite this as

$$\frac{\partial}{\partial\langle c_{p\mu}|}E=\mathbf{D}^{-1}\langle\chi_p|\sum_{\nu}\Big\{\hat{F}^{(\mu)}|\phi_\nu\rangle-|\phi_\nu\rangle\epsilon^{(\mu)}_{\nu\nu}\Big\}, \tag{1.57}$$

where we have isolated terms proportional to the overlap, $\langle\chi_p|\phi_\nu\rangle$, as[3]

$$\epsilon^{(\mu)}_{\nu\nu}=\sum_{\mu_1,\nu_1}\mathbf{D}(\mu\mu_1|\nu\nu_1)\langle\phi_{\mu_1}|\hat{H}_1|\phi_{\nu_1}\rangle$$

$$+\frac{1}{2}\sum_{\mu_1,\mu_2,\nu_1,\nu_2}\mathbf{D}(\mu\mu_1\mu_2|\nu\nu_1\nu_2)\langle\phi_{\mu_1}\phi_{\mu_2}|\hat{H}_2|\phi_{\nu_1}\phi_{\nu_2}\rangle-\frac{W}{\mathbf{D}}\mathbf{D}(\mu|\nu), \tag{1.58}$$

and other integrals between $\langle\chi_p|$ and $|\phi_\nu\rangle$ in the generalised Fock operator, $\hat{F}^{(\mu)}$, so that

$$\langle\chi_p|\hat{F}^{(\mu)}|\phi_\nu\rangle=\mathbf{D}(\mu|\nu)\langle\chi_p|\hat{H}_1|\phi_\nu\rangle+\sum_{\mu_1,\nu_1}\mathbf{D}(\mu\mu_1|\nu\nu_1)\langle\chi_p\phi_{\mu_1}|\hat{H}_2|\phi_\nu\phi_{\nu_1}\rangle. \tag{1.59}$$

The last term in this equation can be though of as the action of a generalised Coulomb operator, $\hat{J}_{\mu_1\nu_1}$, according to

$$\langle\chi_p|\hat{J}_{\mu_1\nu_1}|\phi_\nu\rangle=\langle\chi_p\phi_{\mu_1}|\hat{H}_2|\phi_\nu\phi_{\nu_1}\rangle. \tag{1.60}$$

[3]We have defined ϵ here as a two-index quantity for later convenience.

So, similarly to the Hartree-Fock case, the generalised Fock operator describes, besides the usual one-electron terms, the electron-electron interaction with the $N-1$ remaining electrons in the wavefunction.

The condition for the orbital derivative to be zero should be satisfied for any p in equation (1.57). This means that the vector $\sum_\nu\{\cdots\}$ therein should be the zero vector. Furthermore, as we may assume linear independence of the ϕ_ν orbital set,[4] each ν-term in the sum must vanish. So expanding the orbitals according to

$$|\phi_\nu^{(\mu)}\rangle = \sum_q |\chi_q\rangle c_{q\nu}^{(\mu)} \tag{1.61}$$

then leads to a generalised eigenvalue problem of the form

$$\sum_q \langle\chi_p|\hat{F}^{(\mu)}|\chi_q\rangle c_{q\nu}^{(\mu)} = \sum_q \langle\chi_p|\chi_q\rangle c_{q\nu}^{(\mu)} \epsilon_{\nu\nu}^{(\mu)} \tag{1.62}$$

or

$$\mathbf{F}^{(\mu)}\mathbf{c}^{(\mu)} = \mathbf{S}\mathbf{c}^{(\mu)}\boldsymbol{\epsilon}^{(\mu)}. \tag{1.63}$$

Solving this equation will give m possible eigensolutions, $\phi_\nu^{(\mu)}$, for the μth spin-coupled orbital. Since $\hat{F}$ is Hermitian these will be orthogonal, and the eigenvalues $\epsilon_{\nu\nu}^{(\mu)}$ may be interpreted as orbital energies. This would be the energy of an electron occupying the solution orbital $\phi_\nu^{(\mu)}$, experiencing the field of the remaining $N-1$ electrons. There are N generalised eigenvalue problems of the form (1.63)—one for each of the spin-coupled orbitals. Associated with each of these will be a set, or 'stack', of orbitals and orbital energies, and while orbitals in the same stack must be orthogonal, between stacks the orbitals will in general be non-orthogonal.

Getting expressions for the orbital energies, however interesting from a theoretical point of view, would be of little value on its own. It is the connection between the solutions of the one-electron problem with the description of excited states which is of paramount importance in valence bond theory. We begin by

[4]Even if more orbitals are used in the definition of the wavefunction, the sum in equation (1.18) can always be restricted to a linearly independent set.

noting that, as in simple Hartree-Fock theory (or indeed MCSCF theory) there is a strict correspondence between the Fock operator formulation and the Brillouin conditions. The generalised Brillouin theorem [40] (for MCSCF wavefunctions) has been adapted to VB wavefunctions by Grein and Chang [41]. Firstly we may write

$$\frac{\partial}{\partial|c_{p\mu}\rangle}|\Psi_0\rangle = |\Psi_\mu^p\rangle, \tag{1.64}$$

where Ψ_0 is the wavefunction being optimised and Ψ_μ^p is this wavefunction with the (single) orbital excitation $\phi_\mu \rightarrow \chi_p$ made. We then have

$$\frac{\partial}{\partial|c_{p\mu}\rangle}\frac{\langle\Psi_0|\hat{H}|\Psi_0\rangle}{\langle\Psi_0|\Psi_0\rangle} = \frac{1}{\langle\Psi_0|\Psi_0\rangle}\Big(\langle\Psi_0|\hat{H}|\Psi_\mu^p\rangle - \frac{\langle\Psi_0|\hat{H}|\Psi_0\rangle\langle\Psi_0|\Psi_\mu^p\rangle}{\langle\Psi_0|\Psi_0\rangle}\Big) \tag{1.65}$$

or

$$\frac{\partial E}{\partial|c_{p\mu}\rangle} = \frac{\langle\Psi_0|\hat{H}-E|\Psi_\mu^p\rangle}{\langle\Psi_0|\Psi_0\rangle}. \tag{1.66}$$

The condition for this to be zero is equivalent to (1.63), thus giving the same requirements for ϕ_μ. Setting χ_p=ϕ_μ it is easy to see that (1.66) vanishes, so the non-trivial case is when χ_p lies in the orthogonal complement of ϕ_μ. It should be clear that if the set $\{\chi_p\}$ is defined to coincide with the set $\{\phi_\nu^{(\mu)}\}$ as defined previously, not only will there be no direct interaction between the Brillouin state and the spin-coupled ground state, but matrix elements between different Brillouin states will also vanish. So to first order, solving the Fock equations will give orbitals that describe excited states of the molecule. The corresponding singly excited configurations therefore constitute an excellent basis for a non-orthogonal CI description of these states. (That they are not true eigenstates has to do with the fact that the $N-1$ other orbitals, as well as the spin-coupling coefficients, are likely to relax when the electron in ϕ_μ is excited.) A larger non-orthogonal CI may be performed by exciting into several virtuals, and also by performing excitations in more than a single stack. This is in a sense an over-determined problem, since in theory one is using an $m\times N$ dimensional basis set where an m dimensional one would be sufficient. However, as long as care is taken when choosing the excitations, linear dependence in the many-electron

configuration set is unlikely to present any problems.

1.6 Analysing the spin-coupled wavefunction

It has been mentioned in the previous sections that one of the major advantages of the spin-coupled wavefunction is its easy interpretation. We shall in this section elucidate this claim by taking a brief look at the tools available for extracting information from the spin-coupled wavefunction.

Central to the *interpretation* of the spin-coupled wavefunction is the notion, as mentioned section 1.2, that the first order description of a chemical bond is constituted by two orbitals with associated singlet-coupled orbitals. This idea, the very foundation of valence bond theory, may be attributed to Heitler and London's work on H_2 from 1927 [24]. In *modern* valence bond theory (as opposed to classical VB) we are interested in describing orbitals that are more complicated than the pure atomic form considered in ref. [24]. This was first considered by Coulson and Fischer [42]. In section 1.6.2 ways of representing the orbitals defined in this manner will be described. If the Heitler-London calculation furthermore is generalised to more than two electrons, the need for a more rigorous analysis of the coupling of the electron spins arises, see section 1.6.3. The spin-coupled wavefunction consists of a spatial and a spin part, and as such it is useful to separate the analysis of these two.

1.6.1 The symmetry conditions of the spin-coupled wavefunction

Before considering the detailed analysis of the orbitals and spin-coupling coefficients, it is useful to establish the basic *nature* of the solution when the molecule possesses some degree of symmetry. A thorough account of the case of a non-degenerate electronic wavefunction may be found in ref. [9], and we shall here just summarise the most important results. The case of a wavefunction belonging to

a degenerate irreducible representation has been considered in ref. [43].

We have a set of symmetry operations belonging to the point symmetry group, $\mathcal{G}$, defined by

$$\forall \hat{R} \in \mathcal{G} \qquad [\hat{H}, \hat{R}] = 0. \tag{1.67}$$

Assuming that the spin-coupled wavefunction belongs to a non-degenerate irreducible representation of $\mathcal{G}$, we furthermore have

$$\hat{R}\Psi = \zeta(R)\Psi, \qquad |\zeta(R)|^2 = 1. \tag{1.68}$$

This requirement for the spin-coupled wavefunction, with the assumption that the N spin-coupled orbitals are linearly independent, then gives the result

$$\hat{R}\phi_\mu = \zeta_{\mu\nu}(R)\phi_\nu, \qquad |\zeta_{\mu\nu}(R)|^2 = 1 \tag{1.69}$$

for the orbitals [9]. Each symmetry operation in this way leads to a permutation of the spin-coupled orbitals (the $\boldsymbol{\zeta}$-matrices are non-singular). Generally the orbitals may be divided into subsets such that *all* the symmetry operations either permute a subset cyclically or leave the order unchanged. A cycle of order higher than two will be associated with a corresponding rotation axis in the molecule.

The symmetry requirements for the spin function may be derived straightforwardly from the orbital permutation. So that if

$$\hat{R}\big(\phi_1(1)\cdots\phi_N(N)\big) = \zeta^{\text{orb}}(R)\big(\phi_1(p_1^R)\phi_2(p_2^R)\cdots\phi_N(p_N^R)\big) \tag{1.70}$$

the spin function must be unchanged, within a phase, under the same permutation of the electron labels:

$$\Theta(p_1^R, p_2^R, \cdots, p_N^R) = \zeta^{\text{spin}}(R)\Theta(1, 2, \cdots, N) \tag{1.71}$$

with

$$\zeta(R) = \zeta^{\text{orb}}(R)\zeta^{\text{spin}}(R). \tag{1.72}$$

To illustrate these ideas we take the singlet state of CH_2 as an example (for a more detailed discussion see chapter 4). The spin-coupled orbitals are shown in

figure 1.1, the bonding sp^x-type hybrids and the hydrogen orbitals being plotted in the molecular plane, whereas the lone-pair orbitals are plotted in the plane bisecting the HCH angle.

The lone-pair orbitals, LP1 and LP2, are the only orbitals containing components antisymmetric with respect to reflection in the molecular plane σ_v. Consequently these two orbitals are interchanged by this operation. The remaining orbitals contain antisymmetric components with respect to σ_v'—the plane bisecting the H–C–H angle—and by this operation SP1 and SP2 are interchanged, as are H1 and H2. This is in agreement with (1.69), indicating that the spin-coupled wavefunction is in fact symmetry-pure.

Setting $R{=}\sigma_v$ in (1.71), we can see that Θ must be either symmetric or antisymmetric with respect to the interchange of electron labels 1 and 2, i.e., electrons 1 and 2 must be perfectly singlet-coupled or perfectly triplet-coupled. As it happens the former is the case so that only two spin-coupling coefficients in the Rumer basis are non-zero (1 and 3).

1.6.2 Representing the orbitals

After a spin-coupled calculation N orbitals are available for analysis. These will as a rule be spatially well localised. This is an empirical finding based on a large number of actual calculations, but is also a likely form of solution from the point of view of maximising the distance between electrons in different orbitals, thus minimising the electron-electron repulsion energy.

For SC there is no simple correspondence between the orbitals and the total electron distribution as is the case for simple Hartree-Fock orbitals or for the natural representation of the CASSCF molecular orbitals. If a core has been taken out in the calculation one will normally insist that the spin-coupled orbitals be orthogonal to the core orbital space, in which case the total electron distribution will be the sum of contributions from the core and valence electrons respectively.[5]

[5]In some cases, however, as in the calculation of interionic potentials discussed in a later

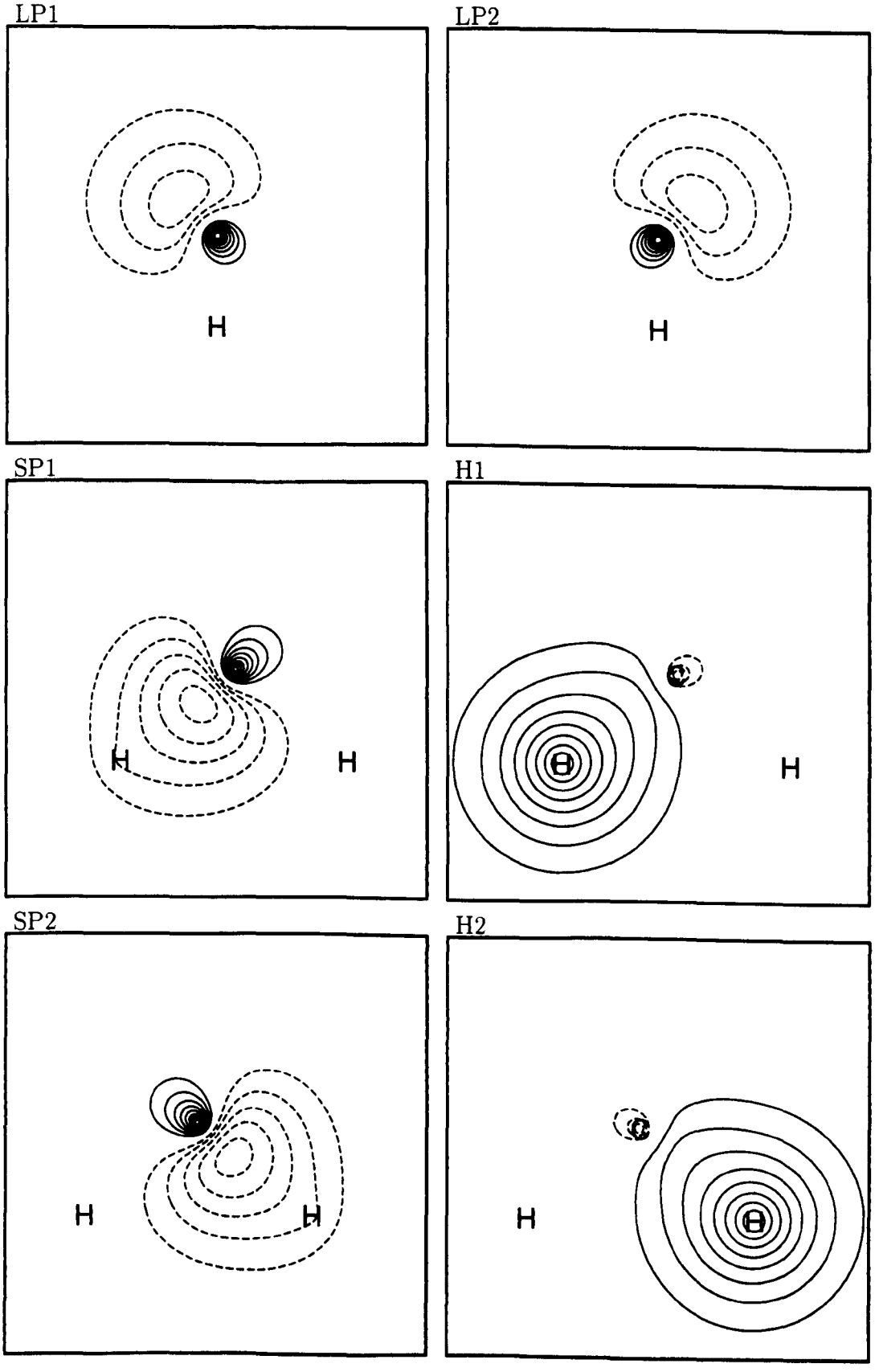

Figure 1.1: The different orbitals of CH_2 (singlet).

The simplest source of information regarding the relationship between the orbitals is represented by the orbital overlap matrix. In many cases the most striking feature of the overlap matrix is the large overlaps (typically in the range 0.6–0.9) between bond-forming orbitals (or orbitals for which the associated electron spins are close to being singlet-coupled—this will be explored further in section 1.6.3). It is well known (see for example ref. [23]) that the favourable energy contribution arising from the exchange interaction is due to terms of the form $\langle\phi_\mu|\phi_\nu\rangle\times\langle\phi_\mu|\hat{H}_1|\phi_\nu\rangle$, and so, to a certain extent, increasing the overlap between such orbitals will be favourable in energy terms. (The counteracting factor here is the inter-electronic repulsion, which will tend to keep the orbitals apart.) This simple notion is sufficient to rationalise a whole range of observations for the spin-coupled orbitals, such as deformation of atomic orbitals along bond-forming axes and hybridisation that optimises the *directional* properties of the atomic orbitals. It is generally more difficult to find patterns in the overlap between orbitals that are not associated with each other in the same way. There is of course a strong dependence on the spatial separation of the orbitals, and this should be take into account in the analysis.

In table 1.2 the overlaps for CH_2 (singlet) are given.[6] The values of 0.67262, 0.80601 and 0.80601 suggest, without further knowledge of the form of the orbitals, that the bond-forming (or: nearly singlet-coupled) pairs are LP1 and LP2 (the lone-pair orbitals), SP1 and H1 (one of the sp^x-type hybrid on carbon and one of the hydrogen 1s-like orbitals), and SP2 and H2 (the other sp^x-type hybrid and hydrogen orbital). (This is verified by the spin-analysis in the next section.) It is a noticeable fact that the overlap between the lone-pair orbitals is smallest in spite of the fact that these orbitals are on the same centre. Apart from this, the remaining overlap integrals follow a pattern consistent with the aforementioned general distance dependence.

chapter, it may be preferable not to insist on this property.

[6]Note that it is customary to choose the relative phases of the orbitals such that overlaps between bond-forming orbitals are positive and symmetry related integrals have the same sign.

	LP1	LP2	SP1	H1	SP2	H2
LP1	1	0.67262	0.19725	0.13170	0.19725	0.13170
LP2	0.67262	1	0.19725	0.13170	0.19725	0.13170
SP1	0.19725	0.19725	1	0.80601	0.31433	0.12042
H1	0.13170	0.13170	0.80601	1	0.12042	-0.05517
SP2	0.19725	0.19725	0.31433	0.12042	1	0.80601
H2	0.13170	0.13170	0.12042	-0.05517	0.80601	1

Table 1.2: Overlap matrix for CH_2.

	LP1	LP2	SP1	H1	SP2	H2
LP1	—	0.00000	1.50000	1.50000	1.50000	1.50000
LP2	0.00000	—	1.50000	1.50000	1.50000	1.50000
SP1	1.50000	1.50000	—	-0.00650	1.57278	1.43372
H1	1.50000	1.50000	0.02374	—	1.43372	1.57278
SP2	1.50000	1.50000	1.67571	1.30055	—	-0.00650
H2	1.50000	1.50000	1.30055	1.67571	0.02374	—

Table 1.3: Expectation values of $(\hat{s}_\mu + \hat{s}_\nu)^2$ for CH_2. Numbers above the diagonal are those based on Ψ, below the diagonal on Θ.

1.6.3 The spin-space

The first problem one encounters when setting out to analyse the spin-space is assigning weights to non-orthogonal functions. This problem occurs when analysing *any* non-orthogonal CI-expansion, and the various schemes available are outlined in appendix 1.b.

In analysing the coupling of the electron spin, there is a choice in which weights to use. One can base the weights on the f_S^N individual structures defined as

$$\Psi_k = \hat{\mathcal{A}}\left[\phi_1\phi_2\cdots\phi_N\Theta_{S;k}^N\right] \tag{1.73}$$

in the spin-coupled wavefunction Ψ or alternatively the weight can be obtained from the individual spin functions, $\Theta_{S;k}^N$ in the total spin function Θ_S^N. This may to a certain degree be a matter of taste, but there are conceptual advantages in separating completely the analysis of the space- and spin-parts.

Recently powerful analytical tools in the form of 'spin correlation analysis' [44] have been made available. We consider first the expectation value

$$\langle\Theta_S^N|(\hat{s}_\mu+\hat{s}_\nu)^2|\Theta_S^N\rangle. \tag{1.74}$$

This value will describe the spin coupling between the electrons μ and ν. It is useful to consider the value for this quantity in the following 'ideal' cases:

0 for perfectly singlet-coupled electrons

2 for perfectly triplet-coupled electrons

1.5 for completely un-coupled electrons

One will rarely obtain these ideal values, except when dictated by symmetry, but generally a small value for $\langle(\hat{s}_\mu+\hat{s}_\nu)^2\rangle$ will signify near-singlet-coupling of the associated electrons, a number close to 2 will signify near-triplet-coupling etc.

Instead of using the total spin function in the expectation value we can use the complete wavefunction. In this way we may *define* the expectation value according to

$$\langle\Psi_S^N|(\hat{s}_\mu+\hat{s}_\nu)^2|\Psi_S^N\rangle \equiv \langle\Psi_S^N|\hat{\mathcal{A}}\left[\phi_1\cdots\phi_N\left((\hat{s}_\mu+\hat{s}_\nu)^2|\Theta_S^N\right)\right]\rangle. \tag{1.75}$$

Defined in this way the expectation value gives a measure of the coupling of the electrons associated with the spin-coupled orbitals ϕ_μ and ϕ_ν. Similarly to the fact that the Chirgwin-Coulson weights based on the full wavefunction may sometimes give 'unphysical' negative weights, the spin expectation values occasionally lie outside the meaningful range of $[0;2]$. The choice between (1.74) and (1.75) is analogous to the choice between $\Theta^N_{S;k}$ and Ψ_k in the evaluation of weights for the various modes of spin coupling.

The above is not restricted to evaluating the total spin associated with pairs of orbitals, any subgroup of orbitals may be treated. It is often sufficient to consider all *pairs* of orbitals, however, in order to establish most patterns in the spin coupling.

As an example the expectation values associated with all the pairs of orbitals for CH_2 are given in table 1.3. For bond-forming orbital pairs the corresponding value will typically be in the range 0.0–0.2, and this can be seen to be the case for CH_2 also. As mentioned in section 1.6.1, the lone-pair electrons are perfectly singlet-coupled for symmetry reasons. This also means that there can be no coupling between a lone-pair electron to any other electron, and this leads to the exact values of 1.5 in the first two columns (and rows). The value of 0.024 for the coupling between the electrons associated with a hydrogen orbital and sp^x-type hybrid is in agreement with the strong dominance of the perfect-pairing function of 98.8%.

The notion of pair populations is closely related to spin correlation analysis using orbital pairs [27]. In traditional Mulliken population analysis a summation is restricted in the normalisation condition for the one-electron density matrix, leading to an electron population associated with the given atomic centres in the molecule. In the same way it is possible to restrict summations in the normalisation condition for the *two*-electron density matrix

$$\binom{N}{2} = \sum_{abcd} \mathbf{D}(ab|cd)\langle ab|cd\rangle, \tag{1.76}$$

this leading to *pair populations* associated with pairs of atomic centres. It is

useful for pair populations to distinguish between symmetric and antisymmetric pairs, so the two basic pair population quantities are defined according to

$$M_{AB}^{\pm} = \frac{1}{2} \sum_{a \in A} \sum_{b \in B} \sum_{cd} \mathbf{D}(ab|cd)(\langle a|c\rangle\langle b|d\rangle \pm \langle a|d\rangle\langle b|c\rangle), \tag{1.77}$$

associated with centres A and B. Of course to pairs symmetric in the spatial part of the wavefunction will be associated singlet-coupling of the electron spins, and to pairs spatially antisymmetric will be associated triplet-coupling. So we denote M^+, as defined above, as a singlet-like pair population, and M^- a triplet-like pair population. This may be made slightly more precise by noting the Dirac spin-exchange identity [45]

$$4\hat{s}_i \cdot \hat{s}_j = 2P_{ij}^{\sigma} - 1. \tag{1.78}$$

Using this relation it is for example straightforward to demonstrate that, provided that the density matrix is expressed in terms of spin-coupled orbitals, and that there to each centre is associated one spin-coupled orbital (or if the summations in (1.77) are restricted accordingly), the values for $2M_{AB}^{-}$ (which in this case may be denoted $2M_{\mu\nu}^{-}$) correspond exactly to the expectation values obtained from (1.75).

The pair population scheme thus has traits in common with the spin correlation analysis, but provides several advantages in that it is independent of the orbital basis used to express the density matrix. As such the analysis is not by any means restricted to the spin-coupled wavefunction. Another very interesting notion is the relationship between pair populations and bond order, as well as to the concept of valency. The quantity

$$M^{\text{eff}} = M^{+} - \eta M^{-} \tag{1.79}$$

where η=1/3 for systems with S=0 or S=1/2, will give a measure of the number of 'bond-forming' pairs. This would for a perfect-pairing spin function be the number of *exactly* singlet-coupled pairs, or $\frac{1}{2}N-S$. M_{AB}^{eff} may be interpreted as

the bond order between centre A and B, and the quantity

$$\mathcal{V}_A = \sum_{B \neq A} M_{AB}^{\text{eff}} \tag{1.80}$$

provides a value for the valency of the A atom. Further details regarding the application of these quantities may be found in ref. [27].

Spin-correlation and pair population analysis are powerful tools for discovering patterns in the spin coupling associated with the spin-coupled wavefunction. The values for $\langle(\hat{s}_\mu+\hat{s}_\nu)^2\rangle$ will not in general contain all the information inherent in the f_S^N spin-coupling coefficients, however. It is therefore often useful to transform the coefficients to a spin-basis, with an ordering of the orbitals, so as to highlight the observed trends established in the spin analysis. Highly efficient programs for doing this now exist [46, 47]. The most frequently used schemes are Rumer, Kotani and Serber, but for a thorough discussion of the choice of spin bases see ref. [6].

1.a Appendix: The Löwdin formula and density matrices

One of the main technical problems in the practical implementation of valence bond theory is the calculation of matrix elements between many-electron wavefunctions defined in terms of non-orthogonal orbitals. The Löwdin formula [48], which gives expressions for matrix elements between Slater determinants, is the foundation for any attempt at addressing this problem. Since any antisymmetric wavefunction can be written as a linear combination of Slater determinants, any exact rule for evaluating matrix elements may be derived from the Löwdin formula. The famous 'Slater-Condon' rules for the case of orthogonal orbitals are as such just a special case of this formula.

We define two (un-normalised) Slater determinants in terms of spin-orbitals

according to

$$\Phi_I = (N!)^{-\frac{1}{2}}\left|\varphi_1^I\varphi_2^I\cdots\varphi_N^I\right| = (N!)^{-\frac{1}{2}}\hat{\mathcal{A}}\left[\varphi_1^I(1)\varphi_2^I(2)\cdots\varphi_N^I(N)\right]$$

$$\Phi_J = (N!)^{-\frac{1}{2}}\left|\varphi_1^J\varphi_2^J\cdots\varphi_N^J\right| = (N!)^{-\frac{1}{2}}\hat{\mathcal{A}}\left[\varphi_1^J(1)\varphi_2^J(2)\cdots\varphi_N^J(N)\right]. \qquad (1.81)$$

The overlap between these can, by interchanging the operations of integration and antisymmetrisation, be seen to reduce to

$$\begin{aligned}\langle\Phi_I|\Phi_J\rangle &= \frac{1}{N!}\int\hat{\mathcal{A}}\left[\varphi_1^I(1)\varphi_2^I(2)\cdots\varphi_N^I(N)\right]\hat{\mathcal{A}}\left[\varphi_1^J(1)\varphi_2^J(2)\cdots\varphi_N^J(N)\right]\mathrm{d}\tau \\ &= \int\varphi_1^I(1)\varphi_2^I(2)\cdots\varphi_N^I(N)\hat{\mathcal{A}}\left[\varphi_1^J(1)\varphi_2^J(2)\cdots\varphi_N^J(N)\right]\mathrm{d}\tau \\ &= \left|\langle\varphi_1^I|\varphi_1^J\rangle\langle\varphi_2^I|\varphi_2^J\rangle\cdots\langle\varphi_N^I|\varphi_N^J\rangle\right| = |\mathbf{S}^{IJ}|, \qquad (1.82)\end{aligned}$$

where $\mathbf{S}^{IJ}$ is the matrix of orbital overlaps. From this the first order density matrix can be found by using the relation[7] (the general form of which is given in (1.30)):

$$\mathbf{D} = \sum_a \mathbf{D}(a|b)\langle\varphi_a|\varphi_b\rangle, \qquad (1.83)$$

since the overlap between two wavefunctions is nothing but the zeroth order (transition) density matrix. In this way we can identify

$$\mathbf{D}_{IJ}(a|b) = |\mathbf{S}^{IJ}|_{(a,b)}, \qquad (1.84)$$

where $|\mathbf{S}^{IJ}|_{(a,b)}$ is a first order cofactor. The first order cofactor may be defined compactly as

$$|\mathbf{S}^{IJ}|_{(a,b)} = \frac{\partial|\mathbf{S}^{IJ}|}{\partial\langle\varphi_a^I|\varphi_b^J\rangle} \qquad (1.85)$$

(as may indeed the first order density matrix from the zeroth order—cf. equation (1.83)), from which, by considering the definition of a determinant, it may be seen to equal

$$|\mathbf{S}^{IJ}|_{(a,b)} = (-1)^{a+b}|(\mathbf{S}^{IJ})_{(a,b)}|. \qquad (1.86)$$

[7]Note the exact analogy between this general relation for density matrices and the expression for the Laplace expansion of a determinant.

Here $(\mathbf{S}^{IJ})_{(a,b)}$ is the sub-matrix of $\mathbf{S}^{IJ}$ with the ath row and bth column removed. In a similar fashion the second order density matrix for determinants may be shown to be just the second order cofactor:

$$\mathbf{D}_{IJ}(ab|cd) = |\mathbf{S}^{IJ}|_{(ab,cd)}. \tag{1.87}$$

This reduces to the determinant of the overlap matrix with *two* columns and rows removed, plus a phase factor:

$$|\mathbf{S}^{IJ}|_{(ab,cd)} = \epsilon_{ab}\epsilon_{cd}(-1)^{a+b+c+d}|\mathbf{S}^{IJ}_{(ab,cd)}|, \tag{1.88}$$

with ϵ_{ab} being $+/-1$ according to whether (ab) form an even $(a < b)$ or odd $(a > b)$ permutation. Similar expressions may be derived for the higher order density matrices and cofactors.

Partial orthogonality between the spin orbitals will lead to a block diagonal overlap matrix $\mathbf{S}^{IJ}$, and thus simplify the evaluation of the determinant and cofactors. The partitioning into alpha and beta spin orbitals is the most commonly encountered, but point group symmetry of the molecule may lead to additional simplifications. A rigorous account of the consequences of a block diagonal overlap matrix may be found in ref. [49].

The spin-less density matrices may be evaluated from the density matrices in terms of spin-orbitals simply by integrating over the spin-variable:

$$\mathbf{D}(\mu|\nu) = \mathbf{D}(\overset{+}{\mu} \mid \overset{+}{\nu}) + \mathbf{D}(\bar{\mu} \mid \bar{\nu}). \tag{1.89}$$

Working with the spin-less density matrix is normally preferable, unless one is working with different spatial orbital sets for the alpha and beta spin orbitals (so-called unrestricted calculations).

Once the relevant density matrix has been evaluated, by considering all many-electron wavefunctions and all orbitals defining the wavefunction under consideration, it may be transformed to a preferred representation. The many electron space may be transformed, such that if

$$\Psi_k = \sum_I C_{Ik}\Phi_I, \tag{1.90}$$

we get a partially transformed density matrix as

$$\mathbf{D}_{kJ} = \sum_{I} C_{Ik}\mathbf{D}_{IJ}, \tag{1.91}$$

and a fully transformed density matrix as

$$\mathbf{D}_{kl} = \sum_{I,J} C_{Ik}C_{Jl}\mathbf{D}_{IJ}. \tag{1.92}$$

Similarly the orbital representation may be transformed. Transforming the density matrix, however, requires the inverse transformation (it transforms contravariant—this may be shown straightforwardly from (1.83)) such that if

$$\phi'_{\nu} = \sum_{\mu} c_{\mu\nu}\phi_{\mu}, \tag{1.93}$$

the first order density matrix, in terms of the ϕ' orbital set, can be written as

$$\mathbf{D}'(\sigma|\tau) = \sum_{\mu,\nu}(\mathbf{c}^{-1})_{\sigma\mu}(\mathbf{c}^{-1})_{\tau\nu}\mathbf{D}(\mu|\nu). \tag{1.94}$$

The higher order density matrices may be transformed in a similar fashion.

1.b Appendix: Non-orthogonal weights

Assigning weights when functions are non-orthogonal requires some consideration, as the choice of scheme is not as unambiguous as in the standard case. Whereas the squares of the coefficients is the only sensible choice of weight in the orthogonal case, there are an infinite number of ways of generalising this to the non-orthogonal case.

In general if the function F is given as the linear combination

$$F = \sum_{i} c_i f_i, \tag{1.95}$$

where the set of functions $\{f\}$ has the overlap matrix

$$S_{ij} = \langle f_i|f_j\rangle, \tag{1.96}$$

we wish to assign a number $o(f_i)$, assessing the occupation, or weight, of f_i in F. For a sensible scheme we may reasonably require the fulfillment of the following criteria:

1. With orthonormal functions the weights reduce to the usual $o(f_i) = c_i^2$

2. The weights add to one: $\sum_i o(f_i) = 1$

3. Each weight should be between 0 and 1

where 2. and 3. may be said to be requirements for any weight, or probability measure. In the case of orthogonal functions we have also the following property:

4. 'Linearity' so that $o(f_i + f_j) = o(f_i) + o(f_j)$

but it would probably be unreasonable to apply this also to non-orthogonal weights.

A brute force methods of defining weights is to simply orthogonalise the functions, and then use the new coefficients c_i' to define weights in the usual manner. It has been shown that Löwdin orthonormalisation, i.e., using $\mathbf{S}^{-\frac{1}{2}}$ as the transformation matrix, gives a set of functions resembling most closely the original (in a root mean square sense) [50]. It furthermore has the advantage of preserving any already existing partial orthogonality, thus, for example, not allowing any mixing between symmetries. The main drawbacks of this procedure is associated with determining the inverse of the overlap matrix. This can be quite a demanding task when the number of functions is large, and is also characterised by numerical instability in cases with near linear dependence. Especially in such cases the weights obtained by this procedure may be rather counterintuitive.

The most commonly used approach in the spin-coupled group, which is after Chirgwin and Coulson [51], takes advantage of the normalisation condition (where real functions are assumed),

$$\sum_{i,j} c_i S_{ij} c_j = 1, \tag{1.97}$$

for the function F. By restricting one summation in this expression occupation numbers that automatically add to one may be defined:

$$o(f_i) = \sum_j c_i S_{ij} c_j. \tag{1.98}$$

The main problem with this definition is that occasionally the weights defined in this manner lie outside the range [0; 1]. However, the three other criteria defined above may be seen to hold trivially for this scheme. In addition it is computationally extremely simple to obtain these weights, with quite a high tolerance against linear dependence.

A third approach worth mentioning is represented by the so-called inverse overlap weights [52, 53]. The main argument in this scheme is that since it is the *unique* part of f_i which is important for the energy or other properties of F, this should be the determining factor for the weight. Accordingly f_i may be orthogonalised to the other functions, i.e., $\langle f_i^{\perp}|f_j\rangle$=0 if $j\neq i$, and a weight defined as

$$o(f_i) = |\langle f_i^{\perp}|F\rangle|^2. \tag{1.99}$$

This may be shown to reduce to $o(f_i)=|c_i|^2/(\mathbf{S}^{-1})_{ii}$ [53]. The weights defined in this way do not automatically add to one, but may be scaled accordingly.

The choice of weights is largely a matter of taste. Familiarity with a particular scheme, its strong and weak points etc., is likely to be as important as choosing the theoretically 'best' procedure.

For readers familiar with the various population analysis schemes in existence, particularly the Mulliken method [39], it may be worth mentioning the correspondence between these and the assigning of non-orthogonal weights. The normalisation condition for the one-electron density matrix can be written

$$\sum_{p,q} \mathbf{D}(p|q)\langle\chi_p|\chi_q\rangle = N. \tag{1.100}$$

(The analogy with equation (1.97) might be clearer if we were there to define $\mathbf{D}(i|j)=c_ic_j$.) Similarly to what is done in the Chirgwin-Coulson scheme, this summation may be restricted to give

$$o(\chi_p) = \sum_{q} \mathbf{D}(p|q)\langle\chi_p|\chi_q\rangle, \tag{1.101}$$

essentially an occupation of the basis function χ_p. By adding up occupation numbers for basis function associated with a given atomic centre, the population

on that centre may be gauged as in the Mulliken population analysis scheme. Alternatively the Löwdin population scheme may be used, where weights of the form

$$o(\chi_p) = (\mathbf{S}^{\frac{1}{2}}\mathbf{D}\mathbf{S}^{\frac{1}{2}})_{pp} \tag{1.102}$$

are added up according to the specific centres. It is easily shown that the Löwdin population analysis scheme is equivalent to the Mulliken scheme with symmetrically orthonormalised basis functions [22]. Thus in the same way as the Mulliken population analysis corresponds to the Chirgwin-Coulson scheme, the Löwdin population analysis corresponds to Löwdin orthonormalisation and use of standard weights.

Bibliography

[1] E. Schrödinger, Ann. Phys. **79**, 169 (1926).

[2] P. Pyykkö, Adv. Quantum Chem. **11**, 353 (1978).

[3] I. P. Grant, in: "*Lecture Notes to the Advanced NATO Study Institute on Methods in Computational Molecular Physics*", Max-Planck-Institut für Astrophysik, Bad Windsheim (1991).

[4] M. Born and R. Oppenheimer, Ann. Phys. **84**, 457 (1927).

[5] B. T. Sutcliffe, in: "*Computational Techniques in Quantum Chemistry*" (ed. G. H. F. Diercksen, B. T. Sutcliffe and A. Veillard), Reidel, Boston (1975). Or: B. T. Sutcliffe, in: "*Lecture Notes to the Advanced NATO Study Institute on Methods in Computational Molecular Physics*", Max-Planck-Institut für Astrophysik, Bad Windsheim (1991).

[6] R. Pauncz, "*Spin Eigenfunctions*", Plenum Press, New York (1979).

[7] W. A. Goddard, III, Phys. Rev. **157**, 73 (1967).

W. A. Goddard, III, Phys. Rev. **157**, 81 (1967).

W. A. Goddard, III, J. Chem. Phys. **48**, 450 (1968).

W. A. Goddard, III, J. Chem. Phys. **48**, 5337 (1968).

[8] J. Gerratt and W. N. Lipscomb, Proc. Natl. Acad. Sci. U.S.A. **59**, 332 (1968).

[9] J. Gerratt, J. Adv. Atom. Mol. Phys. **7**, 141 (1971).

[10] N. C. Pyper and J. Gerratt, Proc. Roy. Soc. Lond. **A355**, 407 (1977).

[11] D. L. Cooper, J. Gerratt and M. Raimondi, Adv. Chem. Phys. **69**, 319 (1987).

[12] D. L. Cooper, J. Gerratt, M. Raimondi, M. Sironi and T. Thorsteinsson, Theor. Chim. Acta **85**, 261 (1993).

[13] P. B. Karadakov, J. Gerratt, M. Raimondi and D. L. Cooper, J. Chem. Phys. **97**, 7635 (1992).

[14] U. Kaldor and F. E. Harris, Phys. Rev. **183**, 1 (1969).

U. Kaldor, Phys. Rev. A **1**, 1586 (1970).

U. Kaldor, Phys. Rev. A **2**, 1267 (1970).

[15] G. D. Fletcher, G. Doggett and A. S. Howard, Phys. Rev. A **46**, 5459 (1992).

G. Doggett and G. D. Fletcher, J. Mol. Struc. **260**, 313 (1992).

G. Doggett, G. D. Fletcher and F. R. Manby, J. Mol. Struc. **300**, 191 (1993).

[16] J. H. van Lenthe and G. G. Balint-Kurti, Chem. Phys. Lett. **76**, 138 (1980).

J. H. van Lenthe and G. G. Balint-Kurti, J. Chem. Phys. **78**, 5699 (1983).

[17] J. Verbeek, PhD Thesis, University of Utrecht, Utrecht, (1990).

[18] F. E. Penotti, Int. J. Quantum Chem. **46**, 535 (1993).

[19] D. L. Cooper, J. Gerratt and M. Raimondi, Int. Rev. Phys. Chem. **7**, 59 (1988).

[20] D. L. Cooper, J. Gerratt and M. Raimondi, in: "*Valence bond theory and chemical structure*", eds. D. J. Klein and N. Trinastić, Elsevier, 287 (1990).

[21] D. L. Cooper, J. Gerratt and M. Raimondi, Chem. Rev. **91**, 929 (1991).

[22] A. Szabo and N. S. Ostlund, "*Modern Quantum Chemistry*", 1. ed., McGraw-Hill (1989).

[23] R. McWeeny, "*Methods of Molecular Quantum Mechanics*", 2. ed., Academic Press, London (1989).

Or: R. McWeeny and B. T. Sutcliffe, "*Methods of Molecular Quantum Mechanics*", Academic Press, London (1969).

[24] W. Heitler and F. London, Z. Phys. **44**, 455 (1927).

[25] D. L. Cooper, J. Gerratt and M. Raimondi, Nature **323**, 699 (1986).

[26] F. E. Penotti, D. L. Cooper, J. Gerratt and M. Raimondi, J. Mol. Struc. **169**, 421 (1988).

[27] D. L. Cooper, R. Ponec, T. Thorsteinsson and G. Raos, Int. J. Quantum Chem. (1994) (in press).

[28] M. Sironi, Dottorato di ricerca in science chimice, Università di Milano (1989).

[29] J. Olsen, D. L. Yeager and P. Jørgensen, Adv. Chem. Phys. **54**, 1 (1983).

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

[31] S. M. Goldfeld, R. E. Quant and H. F. Trotter, Econometrica **34**, 541 (1966).

[32] S. M. Goldfeld, R. E. Quant and H. F. Trotter, Research Memorandum No. 95, Economic Research Program, Princeton University (1968).

[33] R. Fletcher, "*Practical Methods of Optimization*", 2. ed., Wiley, New York (1987).

[34] R. McWeeny, Proc. R. Soc. Lond. **A223**, 306 (1954).

[35] D. A. Kleier, T. A. Halgren, J. H. Hall, Jr. and W. N. Lipscomb, J. Chem. Phys. **61**, 3905 (1974).

R. Kari, Int. J. Quantum Chem. **25**, 321 (1984).

D. Caldwell and H. Eyring, Adv. Quantum Chem. **2**, 93 (1978).

W. L. Luken and D. N. Beratan, Theoret. Chim. Acta. **61**, 265 (1982).

[36] S. F. Boys, Rev. Mod. Phys. **32**, 296 (1960).

J. M. Foster and S. F. Boys, Rev. Mod. Phys. **32**, 300 (1960).

S. F. Boys, in: "*Quantum Theory of Atoms Molecules and the Solid State*", ed. P.-O. Löwdin, Academic, New York, 253 (1966).

[37] C. Edminston and K. Ruedenberg, Rev. Mod. Phys. **35**, 457 (1963).

C. Edminston and K. Ruedenberg, J. Chem. Phys. **43**, S97 (1963).

C. Edminston and K. Ruedenberg, in: "*Quantum Theory of Atoms Molecules and the Solid State*", ed. P.-O. Löwdin, Academic, New York, 263 (1966).

[38] J. Pipek and P. G. Mezey, J. Chem. Phys. **90**, 4916 (1989).

[39] R. Mulliken, J. Chem. Phys. **23**, 1833 (1955).

R. Mulliken, J. Chem. Phys. **23**, 2343 (1955).

[40] B. Levy and G. Berthier, Int. J. Quantum Chem. **2**, 307 (1968).

[41] F. Grein and T. C. Chang, Chem. Phys. Lett. **12**, 44 (1971).

[42] C. A. Coulson and I. Fischer, Phil. Mag. **40**, 386 (1949).

[43] S. C. Wright, D. L. Cooper, J. Gerratt, and M. Raimondi, J. Phys. Chem. **96**, 7943 (1992).

[44] G. Raos, J. Gerratt, D. L. Cooper and M. Raimondi, Chem. Phys. **186**, 233 (1994).

G. Raos, J. Gerratt, D. L. Cooper and M. Raimondi, Chem. Phys. **186**, 251 (1994).

[45] P. A. M. Dirac, "The Principles of Quantum Mechanics", 4. ed., Oxford University Press, London (1958).

[46] G. Raos, J. Gerratt, D. L. Cooper and M. Raimondi, Mol. Phys. **79**, 197 (1993).

[47] P. B. Karadakov, J. Gerratt, D. L. Cooper and M. Raimondi, Theor. Chim. Acta. **90**, 51 (1995).

[48] P.-O. Löwdin, Phys. Rev. **97**, 1474 (1955).

[49] J. Verbeek and J. H. van Lenthe, Int. J. Quantum Chem. **40**, 210 (1991).

[50] P.-O. Löwdin, Arkiv Mat. Astr. Fysik **35A**, 9 (1947).

[51] B. H. Chirgwin and C. A. Coulson, Proc. Roy. Soc. Lond. **A201**, 196 (1950).

[52] G. A. Gallup and J. M. Norbeck, Chem. Phys. Lett. **21**, 495 (1973).

[53] G. A. Gallup, R. L. Vance, J. R. Collins and J. M. Norbeck, Adv. Quantum Chem. **16**, 229 (1982).

Chapter 2

Transformation of full-CI structure spaces

The following three chapters are all concerned with the interpretation of CAS wavefunctions by use of valence bond methods. The ideas described in this chapter are the foundation for the different spin-coupled-based criteria described in the following chapters. They are of considerable more general utility, but we shall here introduce the ideas with the transformation of a CAS wavefunction in mind. In view of the later interpretations, we have tried to look at the CAS wavefunction from a perspective that highlights the similarity with the spin-coupled method. Since a full discussion of the scope for further use of these ideas will only become possible once the various orbital criteria have been introduced and preliminary results presented, a collective discussion for all three chapters will be given in section 4.6.

2.1 Motivation for this research

Many of the existing *ab initio* packages have the ability to perform complete active space self-consistent field (CASSCF, or just CAS) calculations, so that these at present are essentially routine—much more so than for example spin-coupled calculations (for a review of this method, see for example ref. [1]). It is widely

agreed that the CAS method takes excellent account of *non-dynamical* electron correlation. Although a rather vague term, this may be defined as the correlation necessary to describe the forming or dissociation of chemical bonds. (In this way, the method is well suited for the needs of many chemists, including researchers not purely interested in quantum chemistry.) An alternative, more theoretical, definition is represented by the statement that non-dynamical correlation is that obtained by 'internal' orbital excitations, i.e., excitations between already occupied orbitals, so that the CAS wavefunction in this way is the most completely non-dynamically correlated wavefunction for a given basis set. As such a subsequent CI calculation based on excitations into the CAS virtual space will lower the energy, but not significantly alter the qualitative features of the energy surface (or that of most properties).

The initial motivation for undertaking this study was based on a series of findings relating to the CAS method. Several of these suggest that there is a strong link between the CAS and spin-coupled methods. The inherent conceptual simplicity, its lack of bias and also the easy availability of CAS has made it a popular method, and also a convenient one for comparison with the spin-coupled method. Furthermore, as mentioned in section 1.4, using the CAS method to define the core orbital set for the spin-coupled wavefunction is a commonly employed procedure. Traditionally the spin-coupled method uses one orbital to describe each of the N active electrons, so it is natural that the association with the CAS method is strongest with the corresponding 'N in N' calculations (N electrons distributed in N active orbitals—this will be defined in more detail in the next section). Since spin-coupled is, like CAS, a method aimed particularly at obtaining non-dynamical electron correlation, the following similarities with the 'N in N' CAS procedure are perhaps not so surprising:

1. Spin-coupled potential energy surfaces, as well as the molecular property surfaces, are generally parallel to, and close to, the equivalent CAS ones.

2. Particularly for smaller numbers of electrons, N, the spin-coupled energy

tends to be only a few millihartree above the CAS energy.

3. Performing a full CI based on all distributions of N electrons in the N spin-coupled orbitals yields a solution where the spin-coupled configuration is strongly dominant.

The qualitative and, to a certain extent, quantitative agreement in the energy and properties are indirect indications that the two wavefunctions are relatively similar. The third statement above, in which the spin-coupled structure space is expanded to be of N in N CAS form, combined with the first two, suggests that a representation might be found for the CAS wavefunction in which the spin-coupled configuration would dominate.

It is central for the further development of this theory, that the CAS wavefunction is invariant under any linear transformation of the defining active orbitals. In this way the basic concept is the same as, for example, the various localisation schemes in existence for SCF (see section 1.4). We wish to define a transformation of the orbitals that increases the interpretability of the wavefunction. Direct interpretation of the CAS wavefunction, as represented for example by the natural orbitals and their (fractional) occupation numbers, is from a chemist's viewpoint far from ideal. As argued in chapter 1, it is in the interpretation that valence bond based methods have one of their main advantages. The aim of this work is thus to define a transformation of the CAS orbitals to a set of easily interpreted valence bond orbitals. In this way linking a powerful interpretational tool to an already well established MO based method might persuade more people to employ valence bond related methods.

2.2 The CAS wavefunction

We shall in this section look more closely at the CAS wavefunction, with special emphasis on similarities with properties of the spin-coupled wavefunction.

In the CAS method one begins by partitioning the orbital set into four subspaces according to:

1. Frozen core
2. Optimised core
3. Active orbitals
4. Virtual orbitals

As for the spin-coupled wavefunction, the core orbitals are also for the CAS wavefunction doubly occupied in all the structures. The frozen core is the part which is not optimised in the calculation, and this may be chosen and effectively removed from the problem in the way discussed in section 1.4. The CAS method is more general than the traditional spin-coupled approach in that the number of active orbitals, m, does not necessarily coincide with the number of active electrons, N. (As such one often talks about 'N in m' CAS calculations.)

One of the great advantages of CAS, which it shares with the spin-coupled method, is in this way the lack of bias towards individual chemical systems. Once the partitioning of the orbital space has been chosen (particularly if all core orbitals are optimised), the wavefunction is fully defined. In this way the method has considerably more 'black box' character than, say, other MCSCF methods.

We can write the CAS wavefunction on the form

$$\Psi_{\mathrm{CAS}} = \sum_{I,k} \Phi_{Ik} = \sum_{I,k} c_{Ik} \hat{\mathcal{A}}[\Omega_I \Theta_k], \tag{2.1}$$

where Ω_I here is an orbital configuration of the form

$$\Omega_I = \phi_{i_1} \phi_{i_2} \cdots \phi_{i_N}. \tag{2.2}$$

These go through all possible occupancies of the N electrons in the m active orbitals. It is useful to classify each configuration according to the number of doubly occupied orbitals, p. In this way the total number of spatial configurations

can be calculated according to

$$N_{\text{conf}} = \sum_{p=0}^{N/2} \binom{m}{p} \times \binom{m-p}{N-2p}. \tag{2.3}$$

With each configuration there is associated a set of spin functions, the Θ_k's in equation (2.1). These are often chosen to be eigenfunctions for the spin operators $\hat{S}^2$ and $\hat{S}_z$ leading to $f_S^{N-2p_I}$ different functions for the Ith configuration. We shall consider using Rumer or Kotani functions for these in section 2.6. Combining the expression for $f_S^{N-2p_I}$ with each term of (2.3), the total dimension of the CAS CI space can be shown to be

$$N_{\text{CAS}}(N,m,S) = \frac{2S+1}{m+1}\binom{m+1}{N/2-S}\binom{m+1}{N/2+S+1}, \tag{2.4}$$

commonly referred to as the Weyl formula [2]. An alternative to using spin eigenfunctions is to use simple spin strings for the Θ_k's, in other words to expand the CAS wavefunction in terms of Slater determinants. This will be discussed further in section 2.6.

It is the sheer size of the CAS structure space, as signified by N_{CAS}, that limits the size of systems that can be treated by the CAS method. This number increases extremely rapidly, particularly with the number of active orbitals, m, so that N and m are normally chosen to be of comparable magnitudes in practical calculations.

It is instructive to compare the numbers of variational parameters to be optimised in typical CAS and spin-coupled calculations. We shall assume that the partitioning of the orbital space has been done in identical manners, so that the 'N in N' CAS and spin-coupled calculations in this respect are equivalent. This means that, for rotations between the core/active, core/virtual and active/virtual orbital spaces, there are the same number of free parameters, namely $m_{core}\times m$, $m_{core}\times m_{vir}$ and $m\times m_{vir}$ respectively. (The core/core and virtual/virtual orbital rotations of course leave both wavefunctions unchanged.) The active/active orbital rotations are redundant in the CAS procedure since these can always be expressed as variations of the CI expansion coefficients. For the spin-coupled

method there are in general $N(N-1)$ free, active/active orbital parameters. Only in special cases, such as maximum spin ($S=\frac{1}{2}N$), may some (or all) of these become redundant. The number of free linear parameters is in the CAS method $N_{\mathrm{CAS}}-1$ associated with the CI expansion coefficients, and for spin-coupled f_S^N-1 associated with the spin-coupling coefficients.

It is clear that variations in both active/active orbital parameters as well as in spin-coupling coefficients can be expressed in terms of the CI expansion coefficients in equation (2.1). So the parameter-space defined by variation of the two former types of coefficients is a sub-space of that spanned by the CI expansion coefficients. Furthermore, we may take the difference in the *dimensions* of these two spaces as an *a priori* indication of the improved quality of the CAS wavefunction compared with spin-coupled. The numbers of free parameters for various numbers of active electrons are shown in table 2.1. Since this is the *only* difference between the two methods—the remaining orbital rotation parameters are the same in the two cases—we can see why $E_{\mathrm{CAS}} \leq E_{\mathrm{SC}}$ must hold. Particularly for larger numbers of electrons it is clear from table 2.1 that the spin-coupled wavefunction is significantly more compact than the corresponding CAS. The fact that this difference in size is not normally reflected in a big difference in energy suggests that the spin-coupled wavefunction from a physical point of view takes a sensible form for describing the non-dynamical correlation effects.

2.3 Previous related work

As mentioned in section 2.2 the CAS wavefunction is invariant under linear transformations of the active orbital set. This has inspired a number of localisation schemes, analogous to those used for single-determinant SCF wavefunctions (see section 1.4). There has also been some interest in the transformation of the CAS orbitals to sets of non-orthogonal valence bond orbitals (we here define a localisation procedure as an optimisation scheme that *preserves* the orthogonality of

$N;S$	CAS	SC	$N;S$	CAS	SC
2 ; 0	2	2	8 ; 1	2352	83
2 ; 1	0	0	9 ; $\frac{1}{2}$	8819	113
3 ; $\frac{1}{2}$	7	7	10 ; 0	19403	131
4 ; 0	19	13	10 ; 1	29699	179
4 ; 1	14	14	11 ; $\frac{1}{2}$	104543	241
5 ; $\frac{1}{2}$	74	34	12 ; 0	226511	263
6 ; 0	174	34	12 ; 1	382238	428
6 ; 1	188	38	13 ; $\frac{1}{2}$	1288286	584
7 ; $\frac{1}{2}$	783	55	14 ; 0	2760614	610
8 ; 0	1763	69	14 ; 1	5010004	1182

Table 2.1: Number of free variational parameters for N in N CAS and spin-coupled calculations not including orbital rotations involving core or virtual orbitals.

the orbitals, otherwise it will be referred to as a valence bond based method).

An important distinction between schemes for reinterpreting the CAS orbitals, is whether these take into account the very complicated *structure* of the CAS wavefunction (in principle one could of course use SCF localisation methods for CAS orbitals and vice versa). We are currently aware of two methods particular to CAS that are independent of the exact form of the wavefunction. One localisation procedure involves projecting the optimised orbitals onto the atomic basis functions and then performing a symmetrical orthogonalisation [3]. Another, specific to the 'N in N' CAS wavefunction, involves minimising the energy of the perfect-pairing function [4].

For methods that *do* take the specific nature of the CAS wavefunction into account, the problem of transforming the structure space must be addressed. (This will be discussed in detail, in the next section.) For a general linear transforma-

tion of the active orbital set

$$\{\phi'\} = \{\phi\}\mathbf{O}, \tag{2.5}$$

there exists a corresponding transformation of the structure space

$$\{\Phi'\} = \{\Phi\}\mathbf{T}(\mathbf{O}) \tag{2.6}$$

in which $\{\Phi\}$ is the row-vector of the N_{CAS} CAS structures defined in terms of the orbitals $\{\phi\}$ as in (2.1), and $\{\Phi'\}$ is the corresponding row-vector defined in terms of $\{\phi'\}$.

This problem is also addressed in standard MCSCF calculations, but with slightly different emphasis. The conventional approach is to equate

$$\mathbf{O} = \exp(\mathbf{\Lambda}) \tag{2.7}$$

where $\mathbf{\Lambda}$ is anti-Hermitian (i.e., skew-symmetric for real orbital transformations). This is sufficient to ensure that the orbital transformation $\mathbf{O}$ is unitary [5]. Working directly in terms of the $m(m-1)/2$ parameters needed to specify this matrix then circumvents the need for a constrained optimisation procedure. A further important simplification [6] is that the transformation of the structure space can be written in the form

$$\Phi_i' = \hat{\Lambda}\Phi_i, \tag{2.8}$$

with

$$\hat{\Lambda} = \exp\Big(\sum_{r,s} \Lambda_{rs} a_r^\dagger a_s\Big). \tag{2.9}$$

The $a_r^\dagger$ and a_s are here the conventional second-quantisation creation and annihilation operators. This expression is not usually evaluated exactly, but it is a convenient way of identifying first, second, and possibly higher, order changes to the structure space. In most approaches (2.8) is used only as part of the evaluation of the various derivatives of the energy expression. For sizeable orbital updates the integrals are transformed, and the matrix elements between structures reevaluated. This type of approach could without doubt be generalised to non-orthogonal orbital transformations. However, when the full change to the

structure space is of interest, expressing the structure transformation using the exponential as in equation (2.9) is not necessarily very helpful.

A method for generating a spin-coupled-like representation of the CAS wavefunction has been developed by McDouall and Robb [7]. It is crucial for their scheme that the CAS orbitals have already been transformed to a set of orbitals relatively close to the spin-coupled representation—in their published work, the localisation procedure described in ref. [4] was used to achieve this. Using this orbital basis, the f_S^N CAS eigenvectors with largest projections onto the covalent space were chosen. The underlying assumption is that there is a one-to-one correspondence between these eigenvectors and the f_S^N spin-coupled eigenvectors. As such the eigenvalues are assumed to be the same, and the spin-coupling coefficients in each case are all assumed to coincide with the corresponding CI expansion coefficients. We can express that as

$$\Psi_{\mathrm{SC}}^{\lambda} = \sum_{i}^{f_S^N} c_i^{\lambda} \Psi_i = \sum_{i}^{f_S^N} c_i^{\lambda} \Phi_i + \sum_{i=f_S^N+1}^{N_{\mathrm{CAS}}} c_i^{\lambda} \Phi_i \tag{2.10}$$

where λ designates the different eigenvalues of the CI vectors. Transforming from the orbital guess to the resulting spin-coupled-like orbitals using (2.8), then gives the condition

$$\Psi_i = \hat{\Lambda} \Phi_i \qquad i = 1, \cdots, f_S^N \tag{2.11}$$

for the orbitals. This expression cannot in general be solved exactly because the functions Ψ_i are of a more general form than $\hat{\Lambda}\Phi_i$, but a solution may be approximated by multiplying both sides of (2.11) by $\langle \Psi_i |$, and truncating the exponential after the linear term.

When incorporating this number of assumptions and approximations into a method two questions should be addressed: a) the effect of the approximations on the end results; and b) the justification of the approximations in terms of the computational effort saved. The calculation on the π electron system of nitrone (CH_2NHO) seemed to give qualitatively similar results to spin-coupled, but it is of course difficult to base any conclusions on just one example. Particularly

testing a system with more than four active electrons seems imperative. Also the justification for the approximations may not be so clear-cut. As we shall demonstrate in the next and following sections, obtaining the *complete* transformation of the structure space is not nearly as computationally demanding as it might seem to the uninitiated. So clearly pursuing a more 'intuitively obvious' scheme ought to be worthwhile.

Recently Messmer and co-workers have taken advantage of the full-CI nature of the CAS space as a means of overcoming, or circumventing, the non-orthogonality problem [8]. Any set of valence bond expansion functions, $\{\Psi\}$, which is based on N active electrons expanded in no more than m orbitals, may be written as linear combinations of the corresponding 'N in m' CAS structures, i.e.,

$$\{\Psi\} = \{\Phi\}\mathbf{C}. \tag{2.12}$$

Once the matrix of coefficients, $\mathbf{C}$, has been determined, it is straightforward to transform the overlap and Hamiltonian matrices, $\mathbf{S}$ and $\mathbf{H}$, and any vector in the CAS space $\mathbf{c}$:

$$\begin{aligned} \mathbf{S}' &= \mathbf{C}^\dagger\mathbf{S}\mathbf{C} \\ \mathbf{H}' &= \mathbf{C}^\dagger\mathbf{H}\mathbf{C} \\ \mathbf{c}' &= \mathbf{C}^{-1}\mathbf{c}, \end{aligned} \tag{2.13}$$

to the basis of valence bond functions. In this way $\mathbf{S}$ and $\mathbf{H}$ may be constructed using orthogonal orbitals, and then transformed to any appropriate non-orthogonal orbital basis, so that any explicit evaluation of non-orthogonal matrix elements is avoided. This may not be computationally the most efficient procedure, but it has the advantage of being very simple to implement. In ref. [8] a hierarchy of increasingly correlated valence bond wavefunctions were investigated, with emphasis on recovering a large proportion of the correlation energy produced by the CAS wavefunction. The so-called 'orbital relaxed' GVB wavefunction (GVB/R) was suggested as the best compromise between compactness

and quality (in terms of correlation energy). This wavefunction consists, like the spin-coupled wavefunction, of f_S^N structures, but with each mode of spin coupling is used a different orbital product. The intrinsic over-completeness of such a description was not a problem in any of the examples examined, because of symmetry relations between the orbitals in different structures. It is an interesting idea in itself, but in my opinion it is not worth sacrificing the clarity of the spin-coupled model in order to gain such a relatively slight amount of, mainly non-dynamical, correlation energy.

The specific technical problem of obtaining the structure transformation corresponding to an orbital transformation has been considered by Malmqvist [9] with application to the so-called 'CASSCF state interaction method' [10]. The problem addressed here is how to calculate efficiently the transition density matrix between two separately optimised CAS wavefunctions (e.g., between wavefunctions describing the ground and excited states of the molecule). The core and active orbital sets defining a CAS will be individually orthonormal, but if the two CAS wavefunctions are freely optimised there will be non-vanishing overlaps between orbitals belonging to different CAS functions. In what seems to be a very general solution to this type of problem, the two orbital sets are biorthogonalised, rendering the calculation of any transition quantity quite straightforward (cf. chapter 5).

This leaves the question of how to re-express the CAS CI coefficients following a general, non-unitary transformation of the orbital set (in general the orthonormality of the orbital sets must be sacrificed in order to ensure biorthogonality). This is done by expressing the orbital transformation as a sequence of single orbital updates. Similarly to what is shown in section 2.5, this can be achieved by an LU decomposition of the orbital transformation matrix, so that, if $\mathbf{O}=\mathbf{LU}$, then

$$\{\phi'\} = \{\phi'\}\mathbf{L} + \{\phi\}\mathbf{U}, \tag{2.14}$$

where it is understood that the ϕ' orbitals are formed one at a time in ascending

order. It is useful to compare, as Malmqvist does, the computational effort with that of obtaining the one-electron energy from the expression

$$E_1 = \langle \Psi_{\mathrm{CAS}} | \sum_{\mu,\nu} \langle \phi_\mu | \hat{H}_1 | \phi_\nu \rangle \hat{E}_{\mu\nu} | \Psi_{\mathrm{CAS}} \rangle. \tag{2.15}$$

As this can be achieved by $N(N-1)/2$ applications of $\hat{E}_{\mu\nu}$ with $\mu \neq \nu$ (the one-electron term is symmetric), the effort must be roughly half that of a complete orbital transformation.

2.4 The structure transformation

When this research was initiated, it was assumed that the structure transformation would have to be approximated. Nevertheless it was thought worthwhile in the initial exploration to carry out the exact transformation as a means of gauging the particular computational characteristics of the problem. The conclusion of this is, as will be clear in the following sections, that determining the structure transformation is *not* sufficiently computationally demanding as to warrant any approximations of this sort.

To restate the problem defined in section 2.3, we wish to find the *complete* structure transformation matrix corresponding to a general, non-unitary orbital transformation:

$$\{\phi'\} = \{\phi\}\mathbf{O} \quad \rightarrow \quad \{\Phi'\} = \{\Phi\}\mathbf{T}(\mathbf{O}). \tag{2.16}$$

We consider first a 'brute force' approach to obtain this transformation—both to gain a better understanding of the essence of the problem, and to confirm why such a scheme is not very practical. A given 'new' structure is defined according to

$$\Phi'_{Ik} = \hat{\mathcal{A}}[\phi'_{i_1} \phi'_{i_2} \cdots \phi'_{i_N} \Theta_k]. \tag{2.17}$$

Inserting into this expression the expansion of the orbitals,

$$\phi'_i = \sum_j o_{ji} \phi_j, \tag{2.18}$$

gives

$$\Phi'_{Ik} = \sum_{j_1} \cdots \sum_{j_N} o_{j_1 i_1} \cdots o_{j_N i_N} \hat{\mathcal{A}}[\phi_{j_1} \phi_{j_2} \cdots \phi_{j_N} \Theta_k]. \tag{2.19}$$

In this expression the orbitals $\phi_{j_1} \cdots \phi_{j_N}$ may in general be in a non-standard order compared with the given ordering scheme adopted for the Φ_{Ik}'s. The problem of expressing the structures with reordered orbitals in terms of the 'old' structure set, $\{\Phi\}$, will be considered further in section 2.6. However, even if the computational cost associated with this could be considered negligible, the fact remains that each 'new' structure will have at least one contribution to every 'old' spatial configuration. As a minimum estimate, the computational effort will thus scale as $N_{\mathrm{CAS}} \times N_{\mathrm{conf}}$. It was verified by our initial calculations that this is not acceptable.

2.5 Sequential orbital updates

The approach outlined in this section takes advantage of some simple properties of the structure transformation. $\{\mathbf{T(O)}\}$ must form a *true* representation of $\{\mathbf{O}\}$ such that

$$\mathbf{O} = \mathbf{O}_1\mathbf{O}_2 \Leftrightarrow \mathbf{T(O)} = \mathbf{T(O}_1)\mathbf{T(O}_2). \tag{2.20}$$

A useful result deriving from this, since $\mathbf{T(I)} = \mathbf{I}$, is

$$\mathbf{T(O}^{-1}) = \mathbf{T}^{-1}\mathbf{(O)}. \tag{2.21}$$

The basic idea is to describe the orbital transformation as a sequence of the simplest possible updates of the forms

$$\mathbf{O}_{\mu\nu}(\lambda) : \phi_\nu \rightarrow \phi_\nu + \lambda\phi_\mu. \tag{2.22}$$

For the diagonal case of μ=ν this just represents a scaling of the orbital by (1+λ). The problem can thus be stated as

$$\mathbf{O}_{11}(\lambda_1) \cdots \mathbf{O}_{mm}(\lambda_{m^2}) = \mathbf{O}, \tag{2.23}$$

where we wish to find an ordering of the m^2 matrices and values for the λ-variables so that the equality holds. Once a solution is found, the corresponding structure

transformation matrix can be constructed by using (2.20). As will be clear in section 2.6, the structure transformation matrices corresponding to updates of the form (2.22) have a particularly simple structure.

The matrix corresponding to the update (2.22) is just

$$\mathbf{O}_{\mu\nu}(\lambda) = \mathbf{I} + \lambda\boldsymbol{\delta}_{\mu\nu}, \tag{2.24}$$

where $\mathbf{I}$ is the identity, and the matrix $\boldsymbol{\delta}_{\mu\nu}$ is null with the exception $\boldsymbol{\delta}_{\mu\nu}(\mu,\nu)=1$. The inverse of this has the same generic form, namely

$$\mathbf{O}_{\mu\nu}^{-1}(\lambda) = \mathbf{I} - \lambda\boldsymbol{\delta}_{\mu\nu} \qquad \text{for } \mu \neq \nu, \tag{2.25}$$

and

$$\mathbf{O}_{\mu\mu}^{-1}(\lambda) = \mathbf{I} - \frac{\lambda}{1+\lambda}\boldsymbol{\delta}_{\mu\mu}. \tag{2.26}$$

Multiplying an expression of this form from the left on a general matrix will have the effect of adding a multiple of the νth row to the μth row. Recasting (2.23) slightly, as

$$\left(\mathbf{O}_{mm}^{-1}(\lambda_{m^2}) \cdots \mathbf{O}_{\mu\nu}^{-1}(\lambda_i) \cdots \mathbf{O}_{11}^{-1}(\lambda_1)\right)\mathbf{O} = \mathbf{I}, \tag{2.27}$$

the problem can be seen to be equivalent to reducing $\mathbf{O}$ to the identity matrix using a sequence of such 'row-operations'. This problem is addressed in elementary numerical analysis, where several different methods exist for solving the general linear problem

$$\mathbf{A}\mathbf{x} = \mathbf{b}, \tag{2.28}$$

where $\mathbf{A}$ is square ($m\times m$, say) and $\mathbf{b}$ and $\mathbf{x}$ are column vectors, for $\mathbf{x}$, using this basic strategy. Examples are Gaussian elimination with back-substitution, Jordan's method and Crout's factorisation algorithm, all of which essentially reduce the coefficient matrix, $\mathbf{A}$, to the identity using $m\times m$ row-operations. In this work a strategy very similar to Jordan's method was used. Each orbital is updated in turn (using m updates), and within each orbital the 'diagonal' update ($\mu=\nu$) is carried out first. A property of such a strategy is that the ith λ-value,

appropriate to $\mathbf{O}_{\mu\nu}$ in (2.27), may be taken directly as the (μ,ν) element of the matrix

$$\mathbf{O}_{\sigma\tau}^{-1}(\lambda_i - 1)\cdots\mathbf{O}_{11}^{-1}(\lambda_1)\mathbf{O}, \tag{2.29}$$

where σ and τ are orbital labels appropriate to matrix $\lambda-1$ in equation (2.27). For diagonal updates, (μ=ν) λ may be determined as this value minus one.

Generally, the main consideration in the choice of algorithm should be the numerical stability rather than computational efficiency. Most schemes will as such allow for permutations of rows or columns, for example, in order to minimise the effect of round-off errors. The dimension of $\mathbf{O}$ will always be reasonable, so the computational cost associated with any of the mentioned schemes may safely be ignored. On the other hand, any inaccuracy in the factorisation (2.23) is likely to accumulate when the corresponding structure transformation is constructed.

2.6 Transforming the CI vector

In the previous section it was demonstrated that a sequence of orbital updates on the form (2.22) is sufficient to describe any non-singular transformation of m orbitals. In this section we will show how the corresponding change of the many-electron functions may be evaluated very efficiently. We shall assume that the spin orbital set is *restricted* so that the alpha and beta spin orbitals have identical spatial parts (this is of course not necessary for the general formalism described in the following). The transformation of the spatial orbital set, as defined in (2.5), will then apply equally to the alpha and beta spin orbital sets, i.e., $\mathbf{O}^\alpha$=$\mathbf{O}^\beta$=$\mathbf{O}$.

The structure transformation corresponding to the update (2.22) of an alpha spin orbital can be described by the action of the operator

$$\hat{\mathbf{I}} + \lambda\hat{\mathbf{E}}_{\mu\nu}^{\alpha} \tag{2.30}$$

on the structure set. The excitation operator $\hat{\mathbf{E}}$ is defined as discussed in appendix 5.a using an annihilation operator constructed from the *dual* orbital basis. With restricted orbitals the α- and β-updates are of course independent (the $\hat{\mathbf{E}}^{\alpha}$

and $\hat{\mathrm{E}}^{\beta}$ operators commute), which means that the order of the respective transformations may be chosen purely on the basis of computational considerations.

2.6.1 Determinants

Expanding the CAS CI space in Slater determinants yields the simplest possible algorithm for obtaining the transformation of the CI space (mainly because no evaluation of coupling coefficients is needed), and so this case will be considered first. Since an excitation of an alpha orbital will leave the beta-part of the determinant unchanged (and vice versa), it is convenient to define the determinants with the alpha orbitals arranged first followed by the orbitals with beta spin, e.g.,

$$\Phi_{1527} = | \overset{+}{\phi}_1 \overset{+}{\phi}_3 \overset{+}{\phi}_7 \overset{+}{\phi}_8 \overset{-}{\phi}_2 \overset{-}{\phi}_3 \overset{-}{\phi}_5 \overset{-}{\phi}_7 |. \tag{2.31}$$

This means, firstly, that any phase factor associated with the excitation of an alpha orbital, say, is determined by the alpha part of the determinants only. Secondly, the evaluation of indexing information is simplified. The determinant index can be defined from independent alpha and beta indices (independent since the structure space is of full CI form), as

$$I_{\mathrm{det}} = I^{\alpha} + (I^{\beta} - 1)N^{\alpha}_{\mathrm{det}}, \tag{2.32}$$

where I^{α} and I^{β} index the alpha and beta strings respectively. With this definition an alpha excitation will affect the alpha index only, making the change to I_{det} trivial to calculate.

The minimum number of determinants for a given spin can be obtained by insisting on the maximum possible M quantum number, i.e., $N_{\alpha}-N_{\beta}=2S$. The total number of determinants is then

$$N_{\mathrm{det}} = N^{\alpha}_{\mathrm{det}} N^{\beta}_{\mathrm{det}} = \binom{m}{N_{\alpha}} \times \binom{m}{N_{\beta}}, \tag{2.33}$$

simply the product of the numbers of alpha and beta strings.

For the diagonal update in (2.22) (i.e., $\mu=\nu$), $(\hat{I} + \lambda \hat{E}^{\alpha}_{\mu\mu})$ will have the effect of multiplying determinants in which $\overset{+}{\phi}_{\mu}$ is occupied by the factor $(1 + \lambda)$ while

leaving other determinants unchanged (similarly for the beta-update). The number of alpha spin strings corresponding to this, $\binom{m-1}{N_\alpha-1}$, is very modest, and in the present strategy this information is stored in an indexing array. The schematic loop structure is shown in figure 2.1. The simplicity of this makes the code well suited for implementation on modern computer architectures (for example on a vector computer).

The case of a genuine orbital excitation in equation (2.22) (i.e., $\mu \neq \nu$) is only slightly more involved than the diagonal case. Again the outer loop is over spin strings where the νth (alpha) orbital is occupied. The most CPU-efficient strategy was found to involve introducing an intermediate set of $N_\alpha - 1$ alpha strings. This method serves to obtain both the phase factor and the indexing information very efficiently. It may be summarised as follows:

1. Rearrange the alpha string to get ϕ_ν first. This will give a phase-factor, P_1, related to the parity of the permutation.

2. 'Annihilate' ϕ_ν. This will give an intermediate alpha string with $N_\alpha - 1$ orbitals; index: $I^{\alpha-1}$.

3. 'Create' ϕ_μ. (Get index for final alpha string.)

4. Rearrange the alpha string to get ascending orbital order. (Get phase-factor P_2.)

Both the phase-factor and indexing information are conveniently kept in pre-calculated arrays. The dimensions of these will in all cases be reasonable, and insignificant compared with the size of the CI vector. For the case of N=14, S=0, for example, the number of alpha strings is $N^{\alpha}_{\mathrm{det}}$=3432, and the number of intermediate alpha strings $N^{\alpha-1}_{\mathrm{det}}$=3003, whereas 11778624 real words are required to store even a single CI vector (see table 2.2).

If ϕ_μ is already part of the $N_\alpha - 1$ string in 3. above, the corresponding index is zero, and so one may proceed directly to the next alpha string. The schematic loop structure is shown in figure 2.1. A significant saving is connected with the

$\boxed{\mu=\nu}$

Do I=1, $\binom{m-1}{N_\alpha-1}$

Lookup $I^\alpha(I)$

Do I^β=1, N^β_{det}

$\mathbf{c}(I^\alpha, I^\beta) := (1+\lambda) \times \mathbf{c}(I^\alpha, I^\beta)$

End Do (I^β)

End Do (I)

$\boxed{\mu\neq\nu}$

Do I=1, $\binom{m-1}{N_\alpha-1}$

Lookup $I^\alpha_{\text{from}}(I)$

Lookup $I^{\alpha-1}(I^\alpha_{\text{from}}, \nu)$

If ($I^{\alpha-1}$=0) proceed to next I

Lookup P_1=$P(I^\alpha_{\text{from}}, \nu)$

Lookup $I^\alpha_{\text{to}}(I^{\alpha-1}, \mu)$

Lookup P_2=$P(I^\alpha_{\text{to}}, \mu)$

Do I^β=1, N^β_{det}

$\mathbf{c}(I^\alpha_{\text{to}}, I^\beta) := \mathbf{c}(I^\alpha_{\text{to}}, I^\beta) + \lambda P_1 P_2 \times \mathbf{c}(I^\alpha_{\text{from}}, I^\beta)$

End Do (I^β)

End Do (I)

Figure 2.1: Schematic loop structures for the updates of an alpha orbital using determinants to span the CAS space.

fact that that the vector, $\mathbf{c}$, may be updated in place. (Standard programming would entail first copying the vector $\mathbf{c}\rightarrow\mathbf{c}'$ and subsequently carrying out the loop structure in figure 2.1 updating according to $\mathbf{c}':=\mathbf{c}'+\lambda P_1 P_2 \mathbf{c}$.) In the present case it is clear that no *recurrence* is possible since I^{α}_{from} and I^{α}_{to} belong to disjoint sets. For the case of Rumer functions (next section) special care must be taken to avoid recurrence.

An important simplification is possible for $S=0$, since in this case the CAS CI vector (and other vectors we shall consider later) will be symmetric with respect to the interchange of (all) alpha and beta spin orbitals. This means that there are only $N_{\text{det}}(N_{\text{det}}+1)/2$ unique expansion coefficients (these can be thought of as the upper, or lower, triangular matrix extracted from the square $\mathbf{c}$). It might not seem possible to limit the storage requirements to this extent, since after an alpha update, say, $\mathbf{c}$ will no longer posses the correct $\alpha\leftrightarrow\beta$ symmetry. However, if corresponding alpha and beta updates are done in pairs, one can still obtain the correct net effect of the total orbital transformation. The diagonal case of $\mu=\nu$ is uninteresting, since the correct result may be achieved by simply restricting the inside loop length in figure 2.1 to correspond to the upper triangle only. The loop structure for the $\nu\rightarrow\mu$ excitation is shown in figure 2.2. All the essentials of the problem are reflected by the transformation of the two-electron space $|\overset{+}{\phi}_{\mu}\overset{-}{\phi}_{\mu}|$, $|\overset{+}{\phi}_{\mu}\overset{-}{\phi}_{\nu}|$, $|\overset{+}{\phi}_{\nu}\overset{-}{\phi}_{\mu}|$, $|\overset{+}{\phi}_{\nu}\overset{-}{\phi}_{\nu}|$, and it is straightforward to verify for this case that the loop structure in figure 2.2 updating only the upper triangle is equivalent to that shown in figure 2.1. So the effect of this reformulation is a reduction of the storage requirements and computational effort by roughly a factor of two.

A good measure of the computational cost, as indicated for example by the number of floating point operations, is given by the number of inner iterations for the loop structures shown in figures 2.1 and 2.2. In table 2.2 this number is given for various numbers of active orbitals/electrons (including m diagonal updates and $m(m-1)$ off-diagonal). The rate of increase of this is only slightly bigger than that of the dimension of the CAS space. To exemplify this: for the *largest*

Alpha update

If ($I^{\alpha}_{\text{to}} > I^{\alpha}_{\text{from}}$) then

Do I^{β}=I^{α}_{to}, N^{β}_{det}

$$\mathbf{c}(I^{\alpha}_{\text{to}}, I^{\beta}) := \mathbf{c}(I^{\alpha}_{\text{to}}, I^{\beta}) + \lambda P_1 P_2 \times \mathbf{c}(I^{\alpha}_{\text{from}}, I^{\beta})$$

End Do (I^{β})

Else

Do I^{β}=I^{α}_{to}, I^{α}_{from}

$$\mathbf{c}(I^{\alpha}_{\text{to}}, I^{\beta}) := \mathbf{c}(I^{\alpha}_{\text{to}}, I^{\beta}) + \lambda P_1 P_2 \times \mathbf{c}(I^{\beta}, I^{\alpha}_{\text{from}})$$

End Do (I^{β})

Do I^{β}=I^{α}_{from}, N^{β}_{det}

$$\mathbf{c}(I^{\alpha}_{\text{to}}, I^{\beta}) := \mathbf{c}(I^{\alpha}_{\text{to}}, I^{\beta}) + \lambda P_1 P_2 \times \mathbf{c}(I^{\alpha}_{\text{from}}, I^{\beta})$$

End Do (I^{β})

End If

Beta update

If ($I^{\beta}_{\text{to}} > I^{\beta}_{\text{from}}$) then

Do I^{α}=1, I^{β}_{from}−1

$$\mathbf{c}(I^{\alpha}, I^{\beta}_{\text{to}}) := \mathbf{c}(I^{\alpha}, I^{\beta}_{\text{to}}) + \lambda P_1 P_2 \times \mathbf{c}(I^{\alpha}, I^{\beta}_{\text{from}})$$

End Do (I^{α})

Do I^{α}=I^{β}_{from}, I^{β}_{to}

$$\mathbf{c}(I^{\alpha}, I^{\beta}_{\text{to}}) := \mathbf{c}(I^{\alpha}, I^{\beta}_{\text{to}}) + \lambda P_1 P_2 \times \mathbf{c}(I^{\beta}_{\text{from}}, I^{\alpha})$$

End Do (I^{α})

Else

Do I^{α}=1, I^{β}_{to}

$$\mathbf{c}(I^{\alpha}, I^{\beta}_{\text{to}}) := \mathbf{c}(I^{\alpha}, I^{\beta}_{\text{to}}) + \lambda P_1 P_2 \times \mathbf{c}(I^{\alpha}, I^{\beta}_{\text{from}})$$

End Do (I^{α})

End If

Figure 2.2: Inner loop structures for the $\nu \rightarrow \mu$ updates in the case of S=0. (Outer loop structure as in figure 2.1.)

$N;S$	N_{det}	N_{L}	$N;S$	N_{det}	N_{L}
2 ; 0	4	8	8 ; 1	3136	1.19×10^5
2 ; 1	1	2	9 ; $\frac{1}{2}$	15876	7.78×10^5
3 ; $\frac{1}{2}$	9	63	10 ; 0	63504	1.91×10^6
4 ; 0	36	216	10 ; 1	44100	2.56×10^6
4 ; 1	16	160	11 ; $\frac{1}{2}$	213444	1.52×10^7
5 ; $\frac{1}{2}$	100	1.70×10^3	12 ; 0	853776	3.59×10^7
6 ; 0	400	4.80×10^3	12 ; 1	627264	5.14×10^7
6 ; 1	225	4.95×10^3	13 ; $\frac{1}{2}$	2944656	2.86×10^8
7 ; $\frac{1}{2}$	1225	3.80×10^4	14 ; 0	11778624	6.60×10^8
8 ; 0	4900	9.80×10^4	14 ; 1	9018009	9.92×10^8

Table 2.2: Number of determinants for N in N CAS and the number of loop iterations associated with a single structure transformation.

systems given in table 2.2, the computational effort associated with a structure transformation matrix becomes ca. 110 times the effort required for the simple evaluation of the scalar product between two CAS vectors. In other words, the effort is extremely reasonable, and much better than the $\approx N_{\text{CAS}}^2$ suggested by the brute force approach mentioned in section 2.4.

2.6.2 Structures based on Rumer functions

This approach was actually implemented before the determinant algorithms. A justification for this was that the number of structure transformation matrix elements was judged to be quite small (as shall be clear below) compared with say structures based on Kotani functions. However, the more practical consideration was that the non-orthogonal CI program available used these types of structures. This provided a very important test of our codes in the initial stages: If all (non-degenerate) solutions have been found using two different orbital sets we

have

$$\{\Psi_{\rm CAS}^{(\lambda)}\} = \{\Phi\}\mathbf{C} = \{\Phi'\}\mathbf{C}', \tag{2.34}$$

with the $N_{\rm CAS}$ eigenfunctions labelled by the eigenvalues, λ, giving

$$\{\Phi'\} = \{\Phi\}\mathbf{C}(\mathbf{C}')^{-1} = \{\Phi\}\mathbf{T}(\mathbf{O}), \tag{2.35}$$

with $\{\phi'\} = \{\phi\}\mathbf{O}$. Since each eigenvector has an arbitrary phase, care should be taken that the phases in the two cases are consistent. By making only small changes to the orbitals it was also possible to check the first and second order transformations in this way (see section 3.2).

As discussed briefly in section 2.2, with this type of strategy each structure in the many-electron space is defined by the spatial occupation of orbitals and the mode of spin-coupling, e.g.,

$$\Phi_{784} = \hat{\mathcal{A}}[\phi_3\phi_3\phi_7\phi_7\phi_1\phi_2\phi_5\phi_8{}^R\Theta_{0;2}^8]. \tag{2.36}$$

The ordering of the spatial configurations is determined on the basis of the number of doubly occupied orbitals, or ionicity, p. (In this work the configurations were arranged in order of increasing ionicity, and for each ionicity with a numerical ordering of the orbitals.) The main reason for this is, since doubly occupied orbitals must necessarily be singlet-coupled, e.g.,

$${}^R\Theta_{0;2}^8 = {}^R\Theta_0^2{}^R\Theta_0^2{}^R\Theta_{0;2}^4, \tag{2.37}$$

that a given configuration will contribute only f_S^{N-2p} functions to the structure list.

It is necessary for Rumer-based functions to combine the update of an alpha spin orbital and a beta spin orbital according to

$$\begin{aligned}(\hat{I} + \lambda\hat{E}_{\mu\nu}^{\alpha})(\hat{I} + \lambda\hat{E}_{\mu\nu}^{\beta}) &= \hat{I} + \lambda(\hat{E}_{\mu\nu}^{\alpha} + \hat{E}_{\mu\nu}^{\beta}) + \lambda^2\hat{E}_{\mu\nu}^{\alpha}\hat{E}_{\mu\nu}^{\beta} \\ &= \hat{I} + \lambda\hat{E}_{\mu\nu}^{(1)} + \lambda^2\hat{E}_{\mu\nu}^{(2)}.\end{aligned} \tag{2.38}$$

If $\mu{\neq}\nu$, then $\hat{E}_{\mu\nu}^{(1)}$ describes single excitations of the spatial orbital ϕ_ν and $\hat{E}_{\mu\nu}^{(2)}$ describes double excitations. $\hat{E}_{\mu\mu}^{(1)}$ is the 'number operator' for ϕ_μ with eigenvalues

of 0, 1 or 2 according to the occupancy of ϕ_μ, and $\hat{E}^{(2)}_{\mu\mu}$ similarly has eigenvalues 0 or 1 according to whether ϕ_μ is *doubly* occupied or not. In practice the effect of (2.38) is realised in the diagonal case by multiplying coefficients corresponding to structures where ϕ_μ is *only* singly occupied by $(1+\lambda)$, and coefficients for which ϕ_μ is doubly occupied by $(1+\lambda)^2$.

The matrix representation of $\hat{E}^{(2)}_{\mu\nu}$, $\mathbf{E}^{(2)}_{\mu\nu}$, is straightforward to evaluate. It combines structures containing $\phi^2_\nu\phi^0_\mu$ with structures of the same occupancy except for $\phi^0_\nu\phi^2_\mu$ with the simple factor of $+1$.

A *single* excitation, as represented by $\mathbf{E}^{(1)}_{\mu\nu}$, will often have the effect of generating a non-standard Rumer function. This will happen when the ordering of the orbitals in the newly created structure does not adhere to the ordering convention otherwise used for the structure set. After the permutation of the orbitals needed to get the right ordering, the resulting spin function will still be based on singlet-coupled pairs, but the corresponding Rumer diagram could have crossing lines. Thus the problem of expressing general, non-standard Rumer functions as linear combinations of the standard Rumer functions must be addressed. In the present work we used the relation

$$ {}^R\Theta^{\text{gen}} = \sum_{k,l} {}^R\Theta_k \mathbf{S}^{-1}_{kl} \langle {}^R\Theta_l | {}^R\Theta^{\text{gen}} \rangle. \tag{2.39} $$

$\mathbf{S}$ is here the overlap matrix between the standard Rumer functions, and ${}^R\Theta^{\text{gen}}$ is a general Rumer function. For obtaining overlaps between general Rumer functions, an algorithm based on superimposing the corresponding Rumer diagrams, as suggested in ref. [11], was employed. This approach, although conceptually very simple, is from a computational point of view quite demanding. It was therefore unavoidable to base the overall strategy on the precomputation, and storage of all the matrix elements of $\mathbf{E}^{(1)}_{\mu\nu}$.[1] Although these matrices are extremely sparse (see table 2.3 and below), the memory requirements associated herewith sets a sharp upper limit to the size of systems that may be studied (see below). Thus more

[1] Since the algorithm for obtaining these elements will therefore not influence the performance of the procedure, it will not be considered in any further detail here.

N ; S	N_{conf}	N_{CAS}	$n(\mathbf{E}^{(1)}_{\mu\nu})$ (avg.)	$n(\mathbf{E}^{(2)}_{\mu\nu})$	N_{L}
6 ; 0	141	175	105	20	4.50×10^3
6 ; 1	121	189	123	15	5.00×10^3
7 ; $\frac{1}{2}$	393	784	560	75	3.07×10^4
8 ; 0	1107	1764	1158	175	8.48×10^4
8 ; 1	1037	2352	1728	189	1.21×10^5
9 ; $\frac{1}{2}$	3139	8820	6845	784	6.08×10^5
10 ; 0	8953	19404	13991	1764	1.56×10^6
10 ; 1	8701	29700	24237	2352	2.61×10^6
11 ; $\frac{1}{2}$	25653	104544	88551	8820	1.16×10^7
12 ; 0	73789	226512	179578	19404	2.83×10^7
12 ; 1	72865	382239	344392	29700	5.28×10^7
13 ; $\frac{1}{2}$	212941	1288287	1192782	104544	2.15×10^8
14 ; 0	616227	2760615	2404525	226512	5.07×10^8
14 ; 1	612795	5010005	4965163	382239	1.03×10^9

Table 2.3: Scaling of key quantities for an N in N CAS based on Rumer functions, with increasing CAS space size.

viable schemes are likely to involve the calculation, in place, of the $\mathbf{E}^{(1)}_{\mu\nu}$ matrix elements. The simple structure of these matrices, with the few non-zero elements typically being simple powers of two, makes it seem likely that this should be possible. As far as we are aware, this problem has not previously been addressed in the literature. Alternatively, if the calculation of matrix elements for Kotani functions could be achieved more simply, this might be a deciding factor in the choice of expansion functions for the structure space.

As in the previous sections on determinants, we wish to avoid any unnecessary copying of the CAS vector. In order to achieve this we divide the single excitations into two types:

1. Single excitations—$\hat{E}^{(1)}_{\mu\nu}$—from structures where ϕ_ν is already singly occu-

pied.

2. Single excitations—$\hat{E}^{(1)}_{\mu\nu}$—from structures where ϕ_ν is doubly occupied.

It is clear that the second type of excitation will affect the coefficients needed in the first, whereas the reverse is not true. So by simply ensuring that type 1 excitations precede those of type 2, recurrence may be avoided, and the CAS vector updated in-place according to

$$\mathbf{c}(I_{\text{to}}) := \mathbf{c}(I_{\text{to}}) + \lambda \times \mathbf{E}^{(1)}_{\mu\nu}(I_{\text{to}}, I_{\text{from}}) \times \mathbf{c}(I_{\text{from}}). \tag{2.40}$$

Some important quantities indicative of the storage requirements and computational effort associated with a complete transformation are given in table 2.3. The sparseness of the single (and double) excitation matrix is worth noting. For the single excitation matrix the exact number of non-zero elements varies slightly according to the different values of μ and ν so we have here given the average number. In principle these matrices are of dimension $N_{\text{CAS}} \times N_{\text{CAS}}$ but in practice the number of non-zero elements is smaller than even N_{CAS}. That being said, since there are $m(m-1)$ of these matrices, the storage requirement for these is fairly large compared with that of the CI vector.[2] It will therefore very likely (depending on computer architecture) be this that sets the upper limit on the system size that may be treated using this approach.

The number of non-zero elements in the double-excitation matrix can be shown to be just the dimension of the CAS space with two electrons (and orbitals) less, i.e., $n(\mathbf{E}^{(2)}_{\mu\nu})=N_{\text{CAS}}(N-2,m-2,S)$. A general expression for N_{CAS} was given in equation (2.4). Similar expressions relating to the diagonal updates may be derived:

$$n(\mathbf{E}^{(1)}_{\mu\mu}=0) = N_{\text{CAS}}(N, m-1, S), \tag{2.41}$$

$$n(\mathbf{E}^{(1)}_{\mu\mu}=1) = N_{\text{CAS}}(N, m, S) - N_{\text{CAS}}(N, m-1, S) - N_{\text{CAS}}(N-2, m-1, S), \tag{2.42}$$

[2] In this work two integers (I_{from} and I_{to}) and one real was reserved for every matrix element of the $\mathbf{E}^{(1)}_{\mu\nu}$ and $\mathbf{E}^{(2)}_{\mu\nu}$ matrices. Although this might be reduced slightly due to the simple forms of the matrix elements, the basic conclusions are not altered.

and

$$n(\mathbf{E}^{(1)}_{\mu\mu}{=}2) = N_{\mathrm{CAS}}(N-2, m-1, S), \tag{2.43}$$

where only diagonal elements of $\mathbf{E}^{(1)}_{\mu\mu}$ have been counted.

By combining these numbers, the total number of loop iterations associated with one transformation may be estimated (table 2.3). Comparing these numbers with the corresponding ones for determinants given in table 2.2, the similarity is surprising in view of the large difference in the dimensions of the function spaces. The two methods can of course strictly not be compared using only this measure, since the loop structure for Rumer functions is rather more involved. Of note is the extra multiplication in the loop structure for the off-diagonal updates and the essentially random access to the $\mathbf{c}$ vector. So on most architectures it seems likely that for the structure transformation the determinant approach will be the more efficient.

2.6.3 Structures based on Kotani functions

This approach has not been coded in the present work. We shall nevertheless briefly consider what would be involved in such a procedure. Particularly researchers in the fields of MCSCF or CI are likely to have much of the technology already available, in which case the implementation would of course be much simplified.

A large part of the discussion in the last section concerning structures based on Rumer functions will apply equally well to structures based on Kotani functions. For example, the same ordering of the spatial configurations can be adopted, and for each configuration there are, as before, f_S^{N-2p} structures.

The main advantages of using Kotani functions would be the orthogonality of the spin functions, and possibly a less involved evaluation of matrix elements of $\mathbf{E}_{\mu\nu}$. A significant amount of effort has been devoted to the efficient, in-place, evaluation of these matrix elements. As in the case of Rumer functions, the important step is the evaluation of matrix elements between spin functions—the

so-called coupling coefficients. A detailed account of a possible approach may be found in ref. [12].

The single-excitation matrix is not likely to be quite as sparse for Kotani functions as was the case for Rumer functions. So, particularly for smaller systems, in terms of computational effort this approach might not attain maximum efficiency.

Bibliography

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

[2] J. Paldus, in *Theoretical Chemistry: Advances and Perspectives* (Eds. H. Eyring and D. G. Henderson), Vol. 2, Academic Press, New York (1976).

[3] K. Ruedenberg, M. W. Schmidt, M. M. Gilbert and S. T. Elbert, Chem. Phys., **71**, 41, 51, 61 (1982).
M. W. Schmidt, B. Lam, S. T. Elbert and K. Ruedenberg, Theoret. Chim. Acta, **68**, 69 (1985).

[4] J. J. W. McDouall and M. A. Robb, Chem. Phys. Lett., **132**, 319 (1986).

[5] B. Levy, Chem. Phys. Lett., **4**, 17 (1969).

[6] D. J. Thouless, *The Quantum Mechanics of Many-Body Systems*, Academic Press, New York (1961).

[7] J. J. W. McDouall and M. A. Robb, Chem. Phys. Lett., **142**, 131 (1987).

[8] R. B. Murphy and R. P. Messmer, J. Chem. Phys., **98**, 7958 (1993).

[9] P. Å. Malmqvist, Int. J. Quantum Chem. **30**, 479 (1986).

[10] P. Å. Malmqvist and B. O. Roos, Chem. Phys. Lett. **155**, 189 (1989).

[11] I. L. Cooper and R. McWeeny, J. Chem. Phys. **45**, 226 (1966).
I. L. Cooper and R. McWeeny, J. Chem. Phys. **45**, 3484 (1966).

[12] P. J. Knowles and H.-J. Werner, Chem. Phys. Lett., **145**, 514 (1988).

Chapter 3

Valence bond interpretations of CAS wavefunctions

3.1 Optimisation criteria

The algorithms for obtaining the structure transformation matrix described in chapter 2 are completely general, and we believe they may be of interest to other researchers wishing to change the orbital representation of full-CI structure spaces.[1] This chapter will be concerned with the interpretation of CAS wavefunctions (particularly of the N in N type), based on the form of the spin-coupled wavefunction. We shall rely heavily on the ability to transform the structure space efficiently in doing so. Similarly to section 2.3, once the transformation matrix, $\mathbf{T}(\mathbf{O})$, appropriate to a particular orbital transformation, $\mathbf{O}$, is known, the change to the overlap and Hamiltonian matrices, and to any given vector in the CAS space, may be determined according to:

$$\mathbf{S}'(\mathbf{O}) = \mathbf{T}^{\dagger}(\mathbf{O})\mathbf{S}\mathbf{T}(\mathbf{O})$$

$$\mathbf{H}'(\mathbf{O}) = \mathbf{T}^{\dagger}(\mathbf{O})\mathbf{H}\mathbf{T}(\mathbf{O}) \tag{3.1}$$

[1] For orbital optimisation schemes the discussion in section 3.2 of first and second order transformation matrices might also be of interest.

and

$$\mathbf{c}'(\mathbf{O}) = \mathbf{T}^{-1}(\mathbf{O})\mathbf{c}. \tag{3.2}$$

The untransformed quantities, **S**, **H** and **c**, are defined in terms of the orthogonal CAS orbitals, and may thus be obtained very efficiently. In this work **c** will always describe the lowest-energy state, but the theory (particularly the 'overlap-based' criteria below) could be applied equally well to, for example, excited states.

Our interpretation of the CAS will be based on the form of the spin-coupled wavefunction. One formulation of our basic aim could therefore be to find the best possible approximation of a function of 'spin-coupled form' to the CAS wavefunction. This leads directly to what we shall call an 'overlap-based' criterion of the form

$$\text{maximise} \quad S_{\text{cov}} = \frac{\langle \Psi_{\text{CAS}} | \Psi_{\text{cov}} \rangle}{(\langle \Psi_{\text{cov}} | \Psi_{\text{cov}} \rangle)^{1/2}}. \tag{3.3}$$

We have included the normalisation factor for the spin-coupled-like[2] wavefunction for generality. The overlap between this and the CAS wavefunction is then maximised with respect to the parameters defining Ψ_{cov}.

An alternative to this is possible when the CAS function describes the ground state (or the lowest-energy state for a given symmetry). In that case we can use the variation principle to define a good approximation to the CAS wavefunction according to

$$\text{minimise} \quad E_{\text{cov}} = \frac{\langle \Psi_{\text{cov}} | \hat{H} | \Psi_{\text{cov}} \rangle}{\langle \Psi_{\text{cov}} | \Psi_{\text{cov}} \rangle}. \tag{3.4}$$

We will refer to this as an 'energy-based' criterion. Generalisations of this to excited states may also be envisaged, but we shall not consider that particular problem any further here.

The quantities defined in (3.3) and (3.4) may now be optimised with respect to the parameters defining the spin-coupled-like wavefunction. Since we are after an

[2]Since strictly the term 'spin-coupled wavefunction' should only be used for the variationally optimised wavefunction as defined in section 1.2, we have introduced the term 'spin-coupled-like'. We take this to mean a wavefunction of the spin-coupled functional form—irrespective of the similarity with the normal SC wavefunction.

interpretation of the CAS active orbital set, we expand the spin-coupled orbitals in terms of these only:

$$\{\phi^{sc}\} = \{\phi^{cas}\}\mathbf{O}, \tag{3.5}$$

and this orbital transformation matrix can then be used in (3.1) and (3.2).

The spin-coupling coefficients may, as for the spin-coupled wavefunction, be treated as an additional set of parameters in the optimisation. It is useful both for the analysis and the optimisation to separate the two parameter sets. As such, the two criteria above immediately suggest four optimisation schemes for the spin-coupled wavefunction. Equations (3.3) and (3.4) may be optimised simultaneously with respect to the orbital and spin parameters as in what will be referred to as CASVB1 and 3 below (two 'VB versions' of the CAS wavefunction), but also 'mixed' schemes, where one set of parameters are optimised on the basis of energy, the other on the basis of overlap, would be possible.[3] We have in this initial investigation considered the more obvious optimisation criteria, but it should be clear from the calculations, chapter 4, that employing the 'mixed' criteria could be useful in some cases.

Besides treating the spin-coupling coefficients as merely optimisation parameters another very appealing approach is to extract them from the CAS wavefunction defined in terms of the $\{\phi^{sc}\}$ orbital set (i.e., from $\mathbf{c}'$ in (3.2)). In this case we use the fact that the CAS function is invariant under any linear transformation of the orbital set, so this can be viewed as a trivial re-expression of the wavefunction. This idea will be pursued in CASVB2 and 4.

3.1.1 CASVB1

This criterion is the maximisation of (3.3) with respect to both orbital and spin-coupling parameters. As such the approach is very similar to the spin-coupled procedure in nature, but the optimisation is based on 'overlap' rather than energy.

[3] We are indebted to Prof. Raimondi of Milan for pointing this independence of the parameter sets out to us.

In order to use the technology of structure transformations we express the spin-coupled-like wavefunction in terms of CAS structures, according to

$$\Psi_{\rm cov} = \sum_k c_{Sk}\Psi_k = \{\Phi\}\mathbf{T}(\mathbf{O})\mathbf{P}_{\rm cov}\mathbf{c}'_S. \tag{3.6}$$

Here $\mathbf{c}'_S$ is a vector of covalent coefficients corresponding to the spin-coupling coefficients c_{Sk}. Although not strictly necessary in this expression, we have introduced the projection matrix $\mathbf{P}_{\rm cov}$ which is diagonal with 1's in the positions relevant to the covalent structures and 0's elsewhere (including it here will be of use in section 3.2). We can now write the complete expression (3.3) as

$$S_{\rm cov} = \frac{\mathbf{c}'^{\dagger}\mathbf{S}'\mathbf{P}_{\rm cov}\mathbf{c}'_S}{\left(\mathbf{c}'^{\dagger}_S\mathbf{P}_{\rm cov}\mathbf{S}'\mathbf{P}_{\rm cov}\mathbf{c}'_S\right)^{1/2}} = \frac{\mathbf{c}^{\dagger}\mathbf{S}\mathbf{T}\mathbf{P}_{\rm cov}\mathbf{c}'_S}{\left(\mathbf{c}'^{\dagger}_S\mathbf{P}_{\rm cov}\mathbf{T}^{\dagger}\mathbf{S}\mathbf{T}\mathbf{P}_{\rm cov}\mathbf{c}'_S\right)^{1/2}}. \tag{3.7}$$

This expression may now be optimised with respect to the parameters defining $\mathbf{O}$ (and hence $\mathbf{T}$), as well as the f_S^N spin-coupling coefficients. We will consider the evaluation of first and second derivatives of this expression in section 3.2.

3.1.2 CASVB2

To justify properly the idea of taking the spin-coupling coefficients from the CAS wavefunction in the case of the overlap-based criterion, we start by changing the perspective slightly. Given that a spin-coupled-like form of the CAS wavefunction is sought, it is natural simply to maximise the weight of its covalent part. We thus write the CAS wavefunction as

$$\Psi_{\rm CAS} = c_{\rm cov}\Psi_{\rm cov} + c_{\rm ion}\Psi_{\rm ion} \tag{3.8}$$

where all the Ψ's are normalised. This separation implicitly depends on the orbital representation used. Since the covalent and ionic part of the wavefunction will not in general be orthogonal, we must address the problem of assigning weights to non-orthogonal functions. Various schemes for this were discussed in appendix 1.b, and any of these would form a plausible basis for an orbital optimisation scheme. How appropriate any of these are for the orbital optimisation

can only be determined by trial and error. Provided that we are prepared to disregard the exact form of $\Psi_{\rm ion}$, a plausible scheme is to orthogonalise the ionic part onto the covalent part of the wavefunction:

$$\Psi_{\rm CAS} = c^{\perp}_{\rm cov}\Psi_{\rm cov} + c_{\rm ion}\Psi^{\perp}_{\rm ion} \tag{3.9}$$

where

$$\Psi^{\perp}_{\rm ion} = \Psi_{\rm ion} - \langle\Psi_{\rm cov}|\Psi_{\rm ion}\rangle\Psi_{\rm cov}. \tag{3.10}$$

The covalent weight is then just $c^{\perp 2}_{\rm cov}$, with $c^{\perp}_{\rm cov}$ given by

$$c^{\perp}_{\rm cov} = c_{\rm cov} + c_{\rm ion}\langle\Psi_{\rm cov}|\Psi_{\rm ion}\rangle. \tag{3.11}$$

This is of course equivalent to equation (3.3) in that $c^{\perp}_{\rm cov}$=$S_{\rm cov}$ if the spin-coupling coefficients of the spin-coupled-like wavefunction equal the covalent CAS coefficients. As it is slightly simpler we choose to optimise $S_{\rm cov}$ directly rather than $c^{\perp 2}_{\rm cov}$. In terms of the vectors and matrices defined previously equation (3.11) becomes

$$S_{\rm cov} = \frac{\mathbf{c}'^{\dagger}\mathbf{S}'\mathbf{P}_{\rm cov}\mathbf{c}'}{\left(\mathbf{c}'^{\dagger}\mathbf{P}_{\rm cov}\mathbf{S}'\mathbf{P}_{\rm cov}\mathbf{c}'\right)^{1/2}} = \frac{\mathbf{c}^{\dagger}\mathbf{S}\mathbf{T}\mathbf{P}_{\rm cov}\mathbf{T}^{-1}\mathbf{c}}{\left(\mathbf{c}^{\dagger}\mathbf{T}^{-1\dagger}\mathbf{P}_{\rm cov}\mathbf{T}^{\dagger}\mathbf{S}\mathbf{T}\mathbf{P}_{\rm cov}\mathbf{T}^{-1}\mathbf{c}\right)^{1/2}}. \tag{3.12}$$

3.1.3 CASVB3

In section 2.1 we listed evidence for the similarity of the spin-coupled and CAS wavefunctions. It is likely that such a similarity will be reflected in a corresponding similarity of the spaces spanned by the active orbital sets (this will be corroborated further in the calculations, chapter 4). With this in mind, an obvious way of obtaining spin-coupled-like orbitals is to perform a standard spin-coupled calculation, but expanding the active orbitals only in the space defined by the CAS orbitals. If the spaces were in fact identical, this would yield the set of spin-coupled orbitals exactly.

This optimisation may be performed using the standard spin-coupled code, or as before by employing the technology of CAS structure transformations. Using

the quantities defined above we write the energy expression as

$$E_{\rm cov} = \frac{\mathbf{c}_S'^{\dagger}\mathbf{P}_{\rm cov}\mathbf{H}'\mathbf{P}_{\rm cov}\mathbf{c}_S'}{\mathbf{c}_S'^{\dagger}\mathbf{P}_{\rm cov}\mathbf{S}'\mathbf{P}_{\rm cov}\mathbf{c}_S'} \ . \tag{3.13}$$

This expression must then be minimised with respect to the orbital parameters, as well as the $f_S^N - 1$ free parameters associated with the spin-coupling coefficients. Verifying that the two methods gave identical converged solutions was a convincing test of the newly developed code.[4]

Since the CAS and spin-coupled active spaces will in general differ slightly, the spin-coupled wavefunction thus obtained will not be optimal with respect to the core-active, core-virtual and active-virtual orbital rotations. If the 'proper' variational spin-coupled wavefunction is required, the approach suggested in ref. [8] may be adopted to remedy this fact. The spin-coupling coefficients can be transformed to the basis of CAS structures using

$$\mathbf{c}_S = \mathbf{T}\mathbf{c}_S' \ , \tag{3.14}$$

after which the orbitals may be optimised with respect to the core-active, core-virtual and active-virtual rotations, using standard MCSCF procedures but keeping the CI expansion coefficients fixed. Then CASVB3 may be applied to this new orbital space, and this two-step process continued until self-consistency is reached.

While this obviously may not be the most efficient approach imaginable,[5] it can be implemented with very little effort. Also, as we have suggested earlier, the CAS and spin-coupled orbitals span virtually the same space, which means that (if one starts from a CAS calculation) the two-step procedure is likely to converge rapidly.

[4]Regarding the relative efficiency of the two approaches, see section 3.3.

[5]For example, a one-step procedure (which also might be feasible using this general type of approach) is normally preferable to a two-step one.

3.1.4 CASVB4

This is just the energy-equivalent of CASVB2. The energy expectation value for the covalent part of the CAS wavefunction is

$$E = \frac{\mathbf{c}'^{\dagger}\mathbf{P}_{\mathrm{cov}}\mathbf{H}'\mathbf{P}_{\mathrm{cov}}\mathbf{c}'}{\mathbf{c}'^{\dagger}\mathbf{P}_{\mathrm{cov}}\mathbf{S}'\mathbf{P}_{\mathrm{cov}}\mathbf{c}'} \tag{3.15}$$

which we may optimise with respect to the orbital coefficients. Compared to CASVB3, this approach has no degrees of freedom associated with the covalent spin-space. In this way it may be compared to the one spin function approximation (often perfect pairing) occasionally adopted in spin-coupled calculations. The spin coupling is constrained to be that of the covalent part of the CAS wavefunction expressed by the given orbitals. How appropriate this is, is likely to be connected with the amount of relaxation of the covalent coefficients associated with augmenting the spin-coupled wavefunction with ionic structures. This will be explored further in the calculations.

3.2 Derivative expressions

Orbital optimisation is generally a sufficiently complex problem as to warrant the evaluation of first and second derivatives, and our experience with the four optimisation criteria discussed above has if anything verified the necessity of utilising also second derivatives in order to obtain reliable convergence. The optimiser actually used here is very similar to the GQT method described in section 1.3.2 used in the spin-coupled orbital optimisation. Due to the (relatively) large expense associated with the Hessian evaluation, some experiments with optimisation of the length of the update vector were also carried out. (In this way the objective function may be evaluated more than once for a given gradient/Hessian pair. Using this idea one can for example optimise the α-parameter used in the GQT method—equation (1.37).)

All the orbital derivatives may be obtained by considering the simple updates of the form (2.22). We do this, for the first order update, by writing the

total transformation matrix in either the form $\mathbf{O}\mathbf{O}_{\mu\nu}(\lambda)$ or $\mathbf{O}_{\mu\nu}(\lambda)\mathbf{O}$. The corresponding orbital gradients are trivially related via a similarity transformation. The more common way of defining orbital derivatives by writing the total transformation as $\mathbf{O}+\lambda\boldsymbol{\delta}_{\mu\nu}$ would not be so well suited to the strategy of structure transformations.

The structure transformation matrix corresponding to the *spatial* update (2.22) (i.e., identical updates to the alpha and beta spin orbitals) was in section 2.6.2 shown to be

$$\mathbf{T}(\mathbf{O}_{\mu\nu}(\lambda)) = \mathbf{I} + \lambda\mathbf{E}^{(1)}_{\mu\nu} + \lambda^2\mathbf{E}^{(2)}_{\mu\nu}, \tag{3.16}$$

which means that the first order transformation matrix can be identified as the single-excitation matrix $\mathbf{E}^{(1)}_{\mu\nu}$, i.e.,

$$\left[\frac{\partial\mathbf{T}(\mathbf{O}_{\mu\nu}(\lambda))}{\partial\lambda}\right]_{\lambda=0} = \mathbf{E}^{(1)}_{\mu\nu}. \tag{3.17}$$

Expressions for the second order transformation matrices may be derived by considering an orbital transformation matrix of the form

$$\mathbf{O}_{\mu\nu,\sigma\tau}(\lambda_1,\lambda_2) = \mathbf{I} + \lambda_1\boldsymbol{\delta}_{\mu\nu} + \lambda_2\boldsymbol{\delta}_{\sigma\tau}. \tag{3.18}$$

This may be recast as a sequence of simple updates (2.22) by using

$$(\mathbf{I} + \lambda_1\boldsymbol{\delta}_{\mu\nu})(\mathbf{I} + \lambda_2\boldsymbol{\delta}_{\sigma\tau}) = \mathbf{I} + \lambda_1\boldsymbol{\delta}_{\mu\nu} + \lambda_2\boldsymbol{\delta}_{\sigma\tau} + \lambda_1\lambda_2\delta_{\nu\sigma}\boldsymbol{\delta}_{\mu\tau}. \tag{3.19}$$

In this way, it is straightforward to verify that

$$\mathbf{O}_{\mu\nu,\sigma\tau}(\lambda_1,\lambda_2) = \mathbf{O}_{\mu\nu}(\lambda_1)\mathbf{O}_{\sigma\tau}(\lambda_2) \qquad \text{for } \nu \neq \sigma, \tag{3.20}$$

$$\mathbf{O}_{\mu\nu,\nu\sigma}(\lambda_1,\lambda_2) = \mathbf{O}_{\nu\sigma}(\lambda_2)\mathbf{O}_{\mu\nu}(\lambda_1) \qquad \text{for } \sigma \neq \mu, \tag{3.21}$$

and

$$\mathbf{O}_{\mu\nu,\nu\mu}(\lambda_1,\lambda_2) = \mathbf{O}_{\mu\nu}(\lambda_1)\mathbf{O}_{\mu\mu}(-\lambda_1\lambda_2)\mathbf{O}_{\nu\mu}(\lambda_2). \tag{3.22}$$

This gives, by inserting into the expression (3.16) and differentiating:

$$\left[\frac{\partial^2\mathbf{T}(\mathbf{O}_{\mu\nu,\sigma\tau}(\lambda_1,\lambda_2))}{\partial\lambda_1\partial\lambda_2}\right]_{\lambda_1=0,\lambda_2=0} = \mathbf{E}^{(1)}_{\mu\nu}\mathbf{E}^{(1)}_{\sigma\tau} \qquad \text{for } \nu \neq \sigma, \tag{3.23}$$

$$\left[\frac{\partial^2 \mathbf{T}(\mathbf{O}_{\mu\nu,\nu\sigma}(\lambda_1,\lambda_2))}{\partial\lambda_1\partial\lambda_2}\right]_{\lambda_1=0,\lambda_2=0} = \mathbf{E}^{(1)}_{\nu\sigma}\mathbf{E}^{(1)}_{\mu\nu} \qquad \text{for } \sigma \neq \mu, \tag{3.24}$$

and

$$\left[\frac{\partial^2 \mathbf{T}(\mathbf{O}_{\mu\nu,\nu\mu}(\lambda_1,\lambda_2))}{\partial\lambda_1\partial\lambda_2}\right]_{\lambda_1=0,\lambda_2=0} = \mathbf{E}^{(1)}_{\mu\nu}\mathbf{E}^{(1)}_{\nu\mu} - \mathbf{E}^{(1)}_{\mu\mu}, \tag{3.25}$$

for the second order structure transformation matrices. The evaluation of the orbital gradient and Hessian is in this way expressed solely in terms of the action of single-excitation matrices, $\mathbf{E}^{(1)}_{\mu\nu}$, the technology for which was developed in sections 2.4–2.6.

For the derivatives with respect to spin-coupling coefficients (used in the optimisation in CASVB1 and 3), with structures based on Rumer functions, the simplest approach is to let also the spin functions defining the spin-coupled wavefunction be Rumer functions. The derivative is then just

$$\frac{\partial \mathbf{c}'_S}{\partial c_{Sk}} = \boldsymbol{\delta}^{(k)}, \tag{3.26}$$

where $\boldsymbol{\delta}^{(k)}$ is a vector with 1 in the kth position and 0's elsewhere (since the f_S^N first structures are covalent). In the case of determinants, we write the f_S^N linear independent spin-eigenfunctions (usually Rumer or Kotani) that define the spin-coupled wavefunction, as in (1.14), as linear combinations of the N_{det} spin strings defining the determinants, according to

$$\Theta^N_{S;k} = \sum_l c_l^{(k)}\Theta_l. \tag{3.27}$$

Then, similarly to above, the derivative may be written

$$\frac{\partial \mathbf{c}'_S}{\partial c_{Sk}} = \mathbf{c}^{(k)}, \tag{3.28}$$

where the $\mathbf{c}^{(k)}$ vector is just the vector of expansion coefficients as defined in (3.27). We note that both $\boldsymbol{\delta}^{(k)}$ and $\mathbf{c}^{(k)}$ are independent of the form of the orbitals or spin coupling, i.e., there are no second derivatives of $\mathbf{c}'_S$.

Having in general terms defined the derivatives with respect to the orbital and spin parameters, we are now in a position to derive explicit expressions for the gradients and Hessians corresponding to the four criteria as defined in equations (3.7), (3.12), (3.13) and (3.15).

As mentioned above there is a freedom in the definition of the orbital gradient (and Hessian), using either $\mathbf{O}\mathbf{O}_{\mu\nu}(\lambda)$ or $\mathbf{O}_{\mu\nu}(\lambda)\mathbf{O}$ as the total orbital transformation. The first order structure transformation will then be either $\mathbf{T}(\mathbf{O})\mathbf{E}^{(1)}_{\mu\nu}$ or $\mathbf{E}^{(1)}_{\mu\nu}\mathbf{T}(\mathbf{O})$. In the expressions for the four criteria considered here, the structure transformation matrix always occurs in the combination $\mathbf{T}\mathbf{P}_{\rm cov}$, and this is a crucial factor making the former definition much simpler computationally. With this, the single-excitation matrices will occur as $\mathbf{E}^{(1)}_{\mu\nu}\mathbf{P}_{\rm cov}$, and the computational effort associated with evaluating this is trivial compared with the general application of a single-excitation matrix. For the case of Rumer functions, only the elements with

$$I_{\rm to} \leq f_S^N$$

survive, and for determinants only the case

$$I^\beta = N^\beta_{\rm det} - I^\alpha_{\rm to} + 1$$

(for the update of an α orbital), thus making the inner β-loops in figures 2.1 and 2.2 completely redundant. Taking the efficiency of this step into account is important in view of the large number of Hessian matrix elements.

For the occurrence of the inverse of the transformation matrix, in CASVB2 and 4, we expand according to

$$\Big(\mathbf{I} + \lambda_1\mathbf{T}_1 + \lambda_2\mathbf{T}_2 + \lambda_1\lambda_2\mathbf{T}_{12}\Big)^{-1} =$$

$$\mathbf{I} - \lambda_1\mathbf{T}_1 - \lambda_2\mathbf{T}_2 + \lambda_1\lambda_2\Big(\mathbf{T}_1\mathbf{T}_2 + \mathbf{T}_2\mathbf{T}_1 - \mathbf{T}_{12}\Big) + \cdots \qquad (3.29)$$

where terms containing λ_1^2 or λ_2^2 have been left out. So the first and second order transformations associated with $\mathbf{T}^{-1}$ are trivially related to those of $\mathbf{T}$. Furthermore, the inverse always occurs in the combination $\mathbf{T}\mathbf{P}_{\rm cov}\mathbf{T}^{-1}$, and so it is convenient to write

$$\left[\frac{\partial}{\partial\lambda}\Big(\mathbf{T}(\mathbf{O}_{\mu\nu}(\lambda))\mathbf{P}_{\rm cov}\mathbf{T}^{-1}(\mathbf{O}_{\mu\nu}(\lambda))\Big)\right]_{\lambda=0} = \Big[\mathbf{E}^{(1)}_{\mu\nu}, \mathbf{P}_{\rm cov}\Big]. \qquad (3.30)$$

Similarly, the various cases of second order changes, equations (3.23), (3.24) and (3.25), can be written as double commutators, for example

$$\left[\frac{\partial^2}{\partial\lambda_1\partial\lambda_2}\Big(\mathbf{T}(\mathbf{O}_{\mu\nu,\sigma\tau}(\lambda_1,\lambda_2))\mathbf{P}_{\rm cov}\mathbf{T}(\mathbf{O}_{\mu\nu,\sigma\tau}(\lambda_1,\lambda_2))^{-1}\Big)\right]_{\lambda_1=0,\lambda_2=0} =$$

$$[\mathbf{E}^{(1)}_{\mu\nu},[\mathbf{E}^{(1)}_{\sigma\tau},\mathbf{P}_{\rm cov}]] \qquad \text{for}\, \nu\neq\sigma. \tag{3.31}$$

For the cases in (3.23)–(3.25) where the single-excitation matrices do not commute, care should be taken to preserve the ordering when writing down the double commutator.

The advantage of considering the changes to $\mathbf{T}$ and $\mathbf{T}^{-1}$ collectively in this fashion, is that several elements cancel out. For example, for the first order change the cases where $I_{\rm from}$ and $I_{\rm to}$ are either both 'covalent' or both 'ionic' cancel. In particular contributions from the diagonal updates, $\mu{=}\nu$, can be seen to vanish completely.

In the following we shall dispense with writing the orbital transformations explicitly. When the derivative with respect to λ is taken we have the update (2.22) in mind, whereas when the second derivative is taken with respect to λ_1 and λ_2, (3.18) is the underlying orbital transformation. Taking CASVB2 as the first case, we have, for the norm of the spin-coupled-like wavefunction:

$$\left[\frac{\partial}{\partial\lambda}\mathbf{c}'^{\dagger}\mathbf{P}_{\rm cov}\mathbf{S}'\mathbf{P}_{\rm cov}\mathbf{c}'\right]_{\lambda=0} = 2\mathbf{c}'^{\dagger}\mathbf{P}^{\dagger}_{\rm cov}\mathbf{S}'\left[\mathbf{E}^{(1)}_{\mu\nu},\mathbf{P}_{\rm cov}\right]\mathbf{c}', \tag{3.32}$$

and for the overlap with the CAS wavefunction:

$$\left[\frac{\partial}{\partial\lambda}\mathbf{c}'^{\dagger}\mathbf{S}'\mathbf{P}_{\rm cov}\mathbf{c}'\right]_{\lambda=0} = \mathbf{c}'^{\dagger}\mathbf{S}'\left[\mathbf{E}^{(1)}_{\mu\nu},\mathbf{P}_{\rm cov}\right]\mathbf{c}'. \tag{3.33}$$

By combining (3.32) and (3.33) it is straightforward to construct the complete gradient corresponding to equation (3.12). For the second derivative we get

$$\left[\frac{\partial^2}{\partial\lambda_1\partial\lambda_2}\mathbf{c}'^{\dagger}\mathbf{P}_{\rm cov}\mathbf{S}'\mathbf{P}_{\rm cov}\mathbf{c}'\right]_{\lambda_1=0,\lambda_2=0} =$$

$$2\mathbf{c}'^{\dagger}\mathbf{P}^{\dagger}_{\rm cov}\mathbf{S}'[\mathbf{E}^{(1)}_{\mu\nu},[\mathbf{E}^{(1)}_{\sigma\tau},\mathbf{P}_{\rm cov}]]\mathbf{c}' + 2\mathbf{c}'^{\dagger}[\mathbf{E}^{(1)}_{\mu\nu},\mathbf{P}_{\rm cov}]^{\dagger}\mathbf{S}'[\mathbf{E}^{(1)}_{\sigma\tau},\mathbf{P}_{\rm cov}]\mathbf{c}' \tag{3.34}$$

and

$$\left[\frac{\partial^2}{\partial\lambda_1\partial\lambda_2}\mathbf{c}'^{\dagger}\mathbf{S}'\mathbf{P}_{\rm cov}\mathbf{c}'\right]_{\lambda_1=0,\lambda_2=0} = \mathbf{c}'^{\dagger}\mathbf{S}'[\mathbf{E}^{(1)}_{\mu\nu},[\mathbf{E}^{(1)}_{\sigma\tau},\mathbf{P}_{\rm cov}]]\mathbf{c}'. \tag{3.35}$$

Here, and in the following, we assume the right ordering of the excitation matrices in the double commutators according to the three cases (3.23)–(3.25). These two expressions may then be combined with (3.32) and (3.33) to give the complete Hessian.

For CASVB1 the orbital gradient and Hessian may be constructed using very similar expressions as for CASVB2. The double commutators should be substituted by simple products of single excitation matrices and $\mathbf{P}_{\text{cov}}$, and occurrences of $\mathbf{P}_{\text{cov}}\mathbf{c}'$ by $\mathbf{P}_{\text{cov}}\mathbf{c}'_S$. The only exception from this is in the occurrence of $-\mathbf{E}^{(1)}_{\mu\mu}$ in equation (3.25). However, by recognising that the contribution of $\mathbf{E}^{(1)}_{\mu\mu}$ to the *total* Hessian is just the orbital gradient corresponding to $\mathbf{O}_{\mu\mu}$ (which is zero), this term may be omitted without introducing any errors in the final expressions.

For the derivatives with respect to spin-coupling coefficients we have

$$\frac{\partial}{\partial c_{Sk}}\mathbf{c}'^{\dagger}_S\mathbf{P}_{\text{cov}}\mathbf{S}'\mathbf{P}_{\text{cov}}\mathbf{c}'_S = 2\mathbf{c}'^{\dagger}_S\mathbf{P}_{\text{cov}}\mathbf{S}'\mathbf{P}_{\text{cov}}\mathbf{c}^{(k)} \tag{3.36}$$

(with $\mathbf{c}^{(k)} = \boldsymbol{\delta}^{(k)}$ for Rumer functions), and

$$\frac{\partial}{\partial c_{Sk}}\mathbf{c}'^{\dagger}\mathbf{S}'\mathbf{P}_{\text{cov}}\mathbf{c}'_S = \mathbf{c}'^{\dagger}\mathbf{S}'\mathbf{P}_{\text{cov}}\mathbf{c}^{(k)}, \tag{3.37}$$

for the first derivatives, and

$$\frac{\partial^2}{\partial c_{Sk}\partial c_{Sl}}\mathbf{c}'^{\dagger}_S\mathbf{P}_{\text{cov}}\mathbf{S}'\mathbf{P}_{\text{cov}}\mathbf{c}'_S = 2\mathbf{c}^{(k)\dagger}\mathbf{P}_{\text{cov}}\mathbf{S}'\mathbf{P}_{\text{cov}}\mathbf{c}^{(l)} \tag{3.38}$$

contributing to the second derivative. Finally we have the following contributions to the spin-orbital part of the Hessian:

$$\left[\frac{\partial^2}{\partial\lambda\partial c_{Sk}}\mathbf{c}'^{\dagger}_S\mathbf{P}_{\text{cov}}\mathbf{S}'\mathbf{P}_{\text{cov}}\mathbf{c}'_S\right]_{\lambda=0} =$$

$$2\mathbf{c}'^{\dagger}_S\mathbf{P}_{\text{cov}}\mathbf{S}'\mathbf{E}^{(1)}_{\mu\nu}\mathbf{P}_{\text{cov}}\mathbf{c}^{(k)} + 2\mathbf{c}'^{\dagger}_S\mathbf{P}_{\text{cov}}\mathbf{E}^{(1)\dagger}_{\mu\nu}\mathbf{S}'\mathbf{P}_{\text{cov}}\mathbf{c}^{(k)}, \tag{3.39}$$

and

$$\left[\frac{\partial^2}{\partial\lambda\partial c_{Sk}}\mathbf{c}'^{\dagger}\mathbf{S}'\mathbf{P}_{\text{cov}}\mathbf{c}'_S\right]_{\lambda=0} = \mathbf{c}'^{\dagger}\mathbf{S}'\mathbf{E}^{(1)}_{\mu\nu}\mathbf{P}_{\text{cov}}\mathbf{c}^{(k)}. \tag{3.40}$$

The derivative expressions for the energy-based criteria are rather simpler than those described above. The derivatives of the overlap integrals occurring

in both denominators are the same as that for CASVB1 and 2 respectively. For the un-normalised energy in the two cases, we can simply substitute the overlap matrix in the overlap integral by the Hamiltonian matrix. The resulting first and second derivative expressions may then be combined to give the complete gradient and Hessian in the same manner as was discussed in section 1.3.

3.3 Computational considerations

In this section we discuss some of the basic considerations concerned with the implementation of the criteria and their first and second derivatives. Particularly for CASVB3 and 4 the basic strategy is governed by the necessity for transforming one index in the Hamiltonian at a time. So, for example, for the numerator in CASVB4, the vector

$$\mathbf{T(O)P}_{\text{cov}}\mathbf{T}^{-1}\mathbf{(O)c}$$

is constructed in a right-to-left fashion, before $\mathbf{H}$ (in the untransformed CAS structure basis) is multiplied on from the left. In this way the application of the Hamiltonian is reduced to an N_{CAS}^2 (or in the case of determinants N_{det}^2) process. It is still, however, the time-limiting step for larger calculations, since, as was demonstrated in sections 2.6.1 and 2.6.2, the scaling of the structure transformation is roughly *linear* with the number of structures. The similar multiplication by the overlap matrix (for the case of Rumer-based expansion functions) can in comparison with this be treated as scaling linearly with N_{CAS}.

Once $\mathbf{H'P}_{\text{cov}}\mathbf{c'}$ in this way has been constructed going right to left we may multiply from the left by

1. $\mathbf{c'}^{\dagger}\mathbf{P}_{\text{cov}}$ to get the un-normalised energy.

2. $\mathbf{c'}^{\dagger}[\mathbf{E}^{(1)}_{\mu\nu}, \mathbf{P}_{\text{cov}}]^{\dagger}$ to get the first derivative of the un-normalised energy.

3. $\mathbf{c'}^{\dagger}[\mathbf{E}^{(1)}_{\mu\nu}, [\mathbf{E}^{(1)}_{\sigma\tau}, \mathbf{P}_{\text{cov}}]]^{\dagger}$ to get contributions to the second derivative of the un-normalised energy.

The same can be done for the denominator by constructing first $\mathbf{S}'\mathbf{P}_{\rm cov}\mathbf{c}'$.

For the CASVB3 we may similarly construct $\mathbf{H}'\mathbf{P}_{\rm cov}\mathbf{c}'_S$ and $\mathbf{S}'\mathbf{P}_{\rm cov}\mathbf{c}'_S$ and multiply from the left by

1. $\mathbf{c}'^{\dagger}\mathbf{P}_{\rm cov}$ to get the un-normalised energy and overlap integral.

2. $\mathbf{c}'^{\dagger}\mathbf{P}_{\rm cov}\mathbf{E}^{(1)\dagger}_{\mu\nu}$ to get the first derivatives of these.

3. $\mathbf{c}'^{\dagger}\mathbf{P}_{\rm cov}(\mathbf{E}^{(1)}_{\mu\nu}\mathbf{E}^{(1)}_{\sigma\tau})^{\dagger}$ to get contributions to the second derivatives.

For CASVB1, $\mathbf{S}'\mathbf{P}_{\rm cov}\mathbf{c}'_S$ and $\mathbf{S}'\mathbf{P}_{\rm ion}\mathbf{c}'_S$ may be constructed and subsequently multiplied from the left by the same quantities as in the case of CASVB3.

The computational effort associated with evaluating these quantities is thus very reasonable, involving at most a single multiplication by the Hamiltonian. It is the remaining parts of the Hessian which represent the major time consuming expressions. In these cases a single-excitation matrix (or its commutator) occurs on both sides of the Hamiltonian (or overlap matrix), so it is not possible to have only one right-hand vector. For CASVB2 there will be $N(N-1)+f_S^N-1$ such right-hand vectors[6] of the forms

$$\mathbf{T}(\mathbf{O})[\mathbf{E}^{(1)}_{\mu\nu},\mathbf{P}_{\rm cov}]\mathbf{c}'_S$$

and

$$\mathbf{T}(\mathbf{O})\mathbf{P}_{\rm cov}\mathbf{c}^{(k)}$$

which will be multiplied by the Hamiltonian. For CASVB1 and 3 there will only be the $N(N-1)$ vectors associated with the orbital variations. So the basic scaling properties of the energy-based criteria will, in the absence of any molecular symmetry, be:

CASVB3: $N^2_{\rm CAS}\times(N(N-1)+f_S^N)$

CASVB4: $N^2_{\rm CAS}\times(N(N-1)+1)$

[6]The diagonal orbital updates have been excluded since these are known to correspond to redundant parameters.

To put this in context, CASVB3 was for the test cases of N=6, S=0 and N=8, S=0 found to be competitive with our standard spin-coupled codes, but slightly slower. Since the computer-intensive part of the code is essentially a simple matrix multiplication there does not seem to be much scope for optimisation of the algorithm. For some particular optimisation problems, however, the computational effort may be somewhat reduced. In the case of molecular systems possessing some degree of symmetry the Hamiltonian will be block-diagonal, and so the scaling properties of the methods will be significantly reduced compared with N_{CAS}^2. Also when the number of free parameters are reduced for reasons other than symmetry (when redundancies occur or constraints are applied), the computational effort may be reduced accordingly. Since only right-hand vectors associated with the *free* parameter set need to be constructed, the computational effort for the structure transformation based strategy is in direct proportion to the number of these. The approach suggested in section 3.1.3 could as such be the optimal way of doing also full spin-coupled calculations in either of these cases. This would depend on the CASSCF program's efficiency in dealing with the core-active, core-virtual and active-virtual orbital rotations.

For the present calculations the limiting factor in the case of CASVB3 and 4 in the calculations was the storage requirements for the Hamiltonian, not the actual computational effort. Efficient algorithms for applying the Hamiltonian on a given vector without having to store it explicitly have been known for some time in the context of direct CI calculations (first proposed by Roos [1]). In the present work, however, the technology for doing this was not readily available, and the programming from scratch of such a scheme was judged too time-consuming a task. Implementation of this would be necessary for extending the utility of CASVB3 and 4 to larger systems.

For CASVB1 and 2 the fact that no Hamiltonian enters the expressions makes the algorithms applicable to rather larger systems. We shall describe the strategy currently in use for determinants (the scheme for Rumer functions would be very

similar). Beginning with the slightly simpler CASVB2, the basic strategy is as follows:

1. Construct $\mathbf{c}' = \mathbf{T}(\mathbf{O}^{-1})\mathbf{c}$.

2. Project to get $\mathbf{P}_{\text{cov}}\mathbf{c}'$.

Now we use the orthogonality of the spin strings to write

$$\mathbf{S}' = \mathbf{T}^{\dagger}(\mathbf{O})\mathbf{T}(\mathbf{O}) = \mathbf{T}(\mathbf{O}^{\dagger})\mathbf{T}(\mathbf{O}) = \mathbf{T}(\mathbf{s}), \tag{3.41}$$

where s is the orbital overlap matrix. In this way we can write

3. Apply $\mathbf{T}(\mathbf{s})$ on $\mathbf{P}_{\text{cov}}\mathbf{c}'$ to get $\mathbf{S}'\mathbf{P}_{\text{cov}}\mathbf{c}'$.

4. Multiply $\mathbf{S}'\mathbf{P}_{\text{cov}}\mathbf{c}'$ from the left by $\mathbf{c}'^{\dagger}\mathbf{P}_{\text{cov}}$ and $\mathbf{c}'^{\dagger}$ to get S_{cov}.

So in this way two structure transformations are sufficient to obtain a single function evaluation.

5. Apply $\mathbf{T}(\mathbf{s})$ on $\mathbf{c}'$ (or $\mathbf{T}(\mathbf{O})$ on $\mathbf{c}$) to get $\mathbf{S}'\mathbf{c}'$.

6. Multiply from the left by $\mathbf{c}'^{\dagger}[\mathbf{E}^{(1)}_{\mu\nu}, \mathbf{P}_{\text{cov}}]^{\dagger}$ and $\mathbf{c}'^{\dagger}[\mathbf{E}^{(1)}_{\mu\nu}, [\mathbf{E}^{(1)}_{\sigma\tau}, \mathbf{P}_{\text{cov}}]^{\dagger}$ to obtain the gradient and contributions to the Hessian.

7. Generate successively the right-hand vectors $[\mathbf{E}^{(1)}_{\sigma\tau}, \mathbf{P}_{\text{cov}}]\mathbf{c}'$, apply $\mathbf{T}(\mathbf{s})$ and multiply from the left by $\mathbf{c}'^{\dagger}[\mathbf{E}^{(1)}_{\mu\nu}, \mathbf{P}_{\text{cov}}]^{\dagger}$ to get remaining part of the Hessian.

If we assume that a single structure transformation scales as N_{L}, values for which are given in table 2.2, and that the structure transformations are actually the time limiting step, we get an overall scaling property as

CASVB2: $(N(N-1)+2) \times N_{\text{L}}$.

In the case of CASVB1, S_{cov} can be evaluated using only a single structure transformation, according to

1. Construct $\mathbf{T}(\mathbf{O})\mathbf{P}_{\text{cov}}\mathbf{c}'_S$.

2. Evaluate the norm, $(\mathbf{T}(\mathbf{O})\mathbf{P}_{\rm cov}\mathbf{c}'_S)^\dagger\mathbf{T}(\mathbf{O})\mathbf{P}_{\rm cov}\mathbf{c}'_S$, and overlap with the CAS wavefunction, $\mathbf{c}'\mathbf{S}'\mathbf{P}_{\rm cov}\mathbf{c}'_S = \mathbf{c}\mathbf{T}(\mathbf{O})\mathbf{P}_{\rm cov}\mathbf{c}'_S$, to get $S_{\rm cov}$.

Alternatively, as $\mathbf{T}(\mathbf{O})\mathbf{P}_{\rm cov}\mathbf{c}'_S$ is not used in the evaluation of the first and second derivatives, $S_{\rm cov}$ may be calculated from the vectors generated in 3. below. This would save a structure transformation in cases where it is known in advance that the gradient and Hessian will be needed (i.e., if a rejection is not possible). We use the term 'rejection' to signify updates that lead to an increase in energy or decrease in the covalent overlap. Such updates are discarded, and a new updated instead attempted with the previous gradient/Hessian pair. Particularly for schemes that optimise the length of each individual update, it is useful that the cost of a single function evaluation is minimised as here.

3. Apply $\mathbf{T}(\mathbf{O})$ on $\mathbf{c}$ to get $\mathbf{S}'\mathbf{c}'$ and $\mathbf{T}(\mathbf{s})$ to $\mathbf{P}_{\rm cov}\mathbf{c}'_S$.

4. Multiply the two vectors above from the left by $\mathbf{c}'^\dagger_S\mathbf{E}^{(1)\dagger}_{\mu\nu}$, $\mathbf{c}'^\dagger_S(\mathbf{E}^{(1)}_{\mu\nu}\mathbf{E}^{(1)}_{\sigma\tau})^\dagger$ and $\mathbf{c}^{(k)}$ to get the gradient and contributions to the Hessian.

4. Generate successively the right-hand vectors $\mathbf{E}^{(1)}_{\sigma\tau}\mathbf{c}'_S$, apply $\mathbf{T}(\mathbf{s})$ and multiply from the left by $\mathbf{c}'^\dagger_S\mathbf{E}^{(1)\dagger}_{\mu\nu}$ and $\mathbf{c}^{(k)\dagger}$ to get contributions to the orbital-orbital and spin-orbital parts of the Hessian.

5. Generate successively the right-hand vectors $\mathbf{c}^{(l)}$, apply $\mathbf{T}(\mathbf{s})$ and multiply from the left by $\mathbf{c}^{(k)\dagger}$ to get contributions to the spin-spin part of the Hessian.

With the same assumptions as for CASVB2 we get an overall scaling property as

CASVB1: $\left(N(N-1) + f_S^N + 2\right) \times N_{\rm L}$.

In conclusion we can see that, for all four criteria, the cost of a function evaluation is small compared with the evaluation of the complete Hessian. If we take the example of N=10, S=0, the ratios of the computational effort between the two would be 1:134, 1:46, 1:132 and 1:91 for the four criteria respectively. Several schemes that seek to reduce the number of complete Hessian evaluations

are possible. As mentioned in section 3.2 one may attempt several updates for a single gradient-Hessian pair. Another possibility is to perform the calculation with an approximate Hessian, leaving out the expensive part. In our attempts at this, the algorithm showed a good ability to go from a 'wild guess' to a 'feasible region', but neither in the intermediate region nor close to convergence did this approach perform well. There is still scope for experimentation in this area, for example evaluating the complete Hessian only every, say, three iterations.

Bibliography

[1] B. O. Roos, Chem. Phys. Lett. **15**, 153 (1972).

Chapter 4

Applications of CAS to VB transformations

For all molecular systems considered in this chapter, the integral evaluation, SCF and CAS [1, 2] calculations were carried out using the MOLPRO suite of programs [3]. All other calculations described here were performed using our own codes.

4.1 Singlet methylene

The geometry chosen for methylene was: r(C–H)=1.117Å and $\angle$(H–C–H)=102.4°, as used by Bauschlicher and Taylor [4]. Correlation consistent pVTZ basis sets by Dunning were used for C/H consisting of (10s5p2d/5s2p) Cartesian gaussians contracted to [4s3p2d/3s2p] [5].

In the '6 in 6' CAS calculation described here, the core consisted of one optimised doubly occupied orbital ($1a_1$, being roughly a 1s orbital on C), 6 orbitals were active giving 36 virtual orbitals. In the natural orbital representation the active MOs thus defined are $2a_1$, $1b_2$, $3a_1$, $1b_1$, $2b_2$ and $4a_1$, in order of decreasing occupation number. The active orbitals and their occupation numbers are

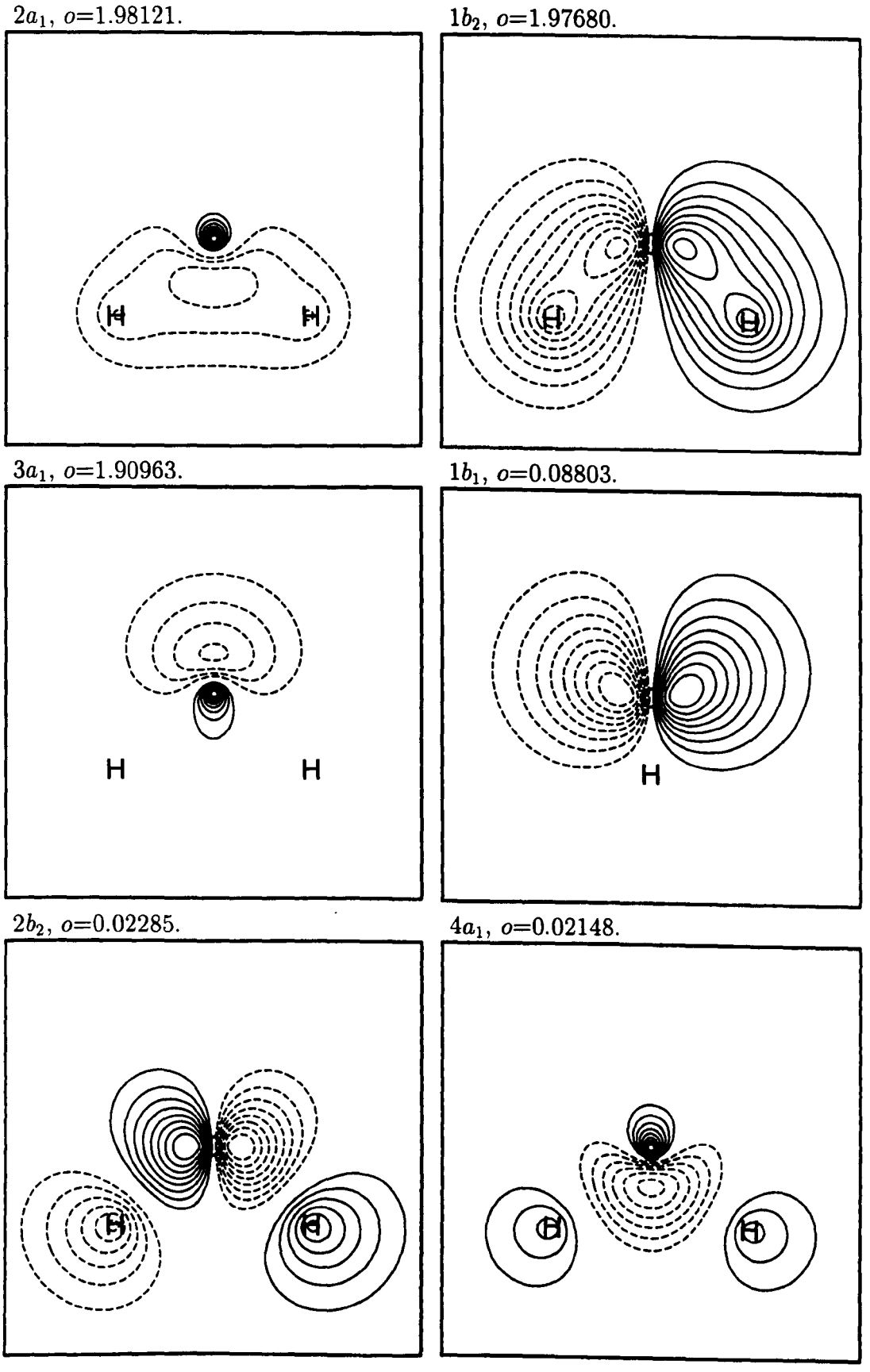

Figure 4.1: Natural orbital representation of active CAS MOs in CH_2 and their occupation numbers.

	E/hartree	$(E-E_{\mathrm{CAS}})$/millihartree
SCF	−38.891714	61.59
$\mathrm{SC}^{\mathrm{CAS}}$	−38.948156	5.15
SC	−38.948765	4.54
SC+CI	−38.952771	0.54
CAS	−38.953306	0.00

Table 4.1: Energies of some of the wavefunctions considered for methylene.

shown in figure 4.1.[1] $1b_1$ is plotted in the plane orthogonal to the molecular plane, bisecting the H–C–H angle, with the other orbitals in the molecular plane. The three first MOs have occupation numbers close to 2 and the corresponding configuration with those orbitals doubly occupied dominated strongly the CAS wavefunction with a weight of 93.49%. Since the corresponding occupied Hartree-Fock orbitals are qualitatively similar, this would suggest that the SCF picture is basically correct, but it is well established [4] that a realistic treatment of this state requires at least a two-configuration description of the form $(\cdots 3a_1^2)-\lambda(\cdots 1b_1^2)$.

We can see that particularly for $2a_1$ and $3a_1$ the role in bond and lone-pair formation is not clear, and the manner in which $1b_1$, $2b_2$ and $4a_1$ serve to improve this picture is not at all obvious. In short: the MO picture alone is rather unsatisfactory for interpreting the chemical bonding. A total correlation energy of 61.59 millihartree was recovered in the CAS calculation (see table 4.1).

In the subsequent spin-coupled calculation, the spin-coupled orbitals were expanded in the active and virtual CAS MOs. In order to facilitate the comparison

[1] All of the contour plots in this thesis depict representations of ϕ with positions of the nuclei projected onto the page and indicated by means of their chemical symbols. It is convenient to define $F_1=\min(|\phi_{\mathrm{min}}|, \frac{1}{2}|\phi_{\mathrm{max}}|/n_{ctr})$ and $F_2=|\phi_{\mathrm{max}}|$—we have used $n_{ctr}=6$ or $n_{ctr}=8$. The plots were constructed by requesting n_{ctr} equally spaced contour heights (full lines) between F_1 and F_2, and a further n_{ctr} equally spaced contour heights (dashed lines) between $-F_2$ and $-F_1$. Adjacent contour heights differ by $(F_2-F_1)/(n_{ctr}+1)$.

SC	LP1	LP2	SP1	H1	SP2	H2
LP1	1	0.67262	0.19725	0.13170	0.19725	0.13170
LP2		1	0.19725	0.13170	0.19725	0.13170
SP1			1	0.80601	0.31433	0.12042
H1				1	0.12042	-0.05517
SP2					1	0.80601
H2						1

Table 4.2: Overlap matrix for the spin-coupled orbitals of singlet methylene.

with CASVB1–4, the core for this calculation was taken from the CAS wavefunction. Contour plots of the spin-coupled orbitals are shown in figure 4.2, and the corresponding overlap matrix is given in table 4.2.[2] The lone-pair orbitals are plotted in the plane bisecting the H–C–H angle, the bond-forming orbitals in the molecular plane.

The two C–H bonds are described by sp^x-hybrids on carbon pointing towards slightly deformed 1s orbitals on the hydrogens. The lone-pair orbitals can also be classified as sp^x-hybrids, but they are much more localised on carbon than the bonding orbitals (i.e., they have more s-character). The lobes point out of the molecular plane so as to form a rough tetrahedron with the C–H bonds, thus enlarging the average inter-electronic distance. It is particularly the description of the lone-pair which, being markedly different from that afforded by conventional MO theory, is interesting when comparing different VB methods.

The overlap of 0.806 between SP1 and H1 is a typical value between orbitals forming a σ-bond. Similarly, the overlap between the lone-pair orbitals (for which the corresponding electrons are also singlet-coupled) of 0.673 is significantly larger than for the other sp^x hybrid pairs. This is a consequence of their greater s-

[2]Since the phases of the orbitals are chosen such that the largest positive contours take its maximum value, this may not necessarily coincide with the relative phases of the orbitals in the overlap matrix.

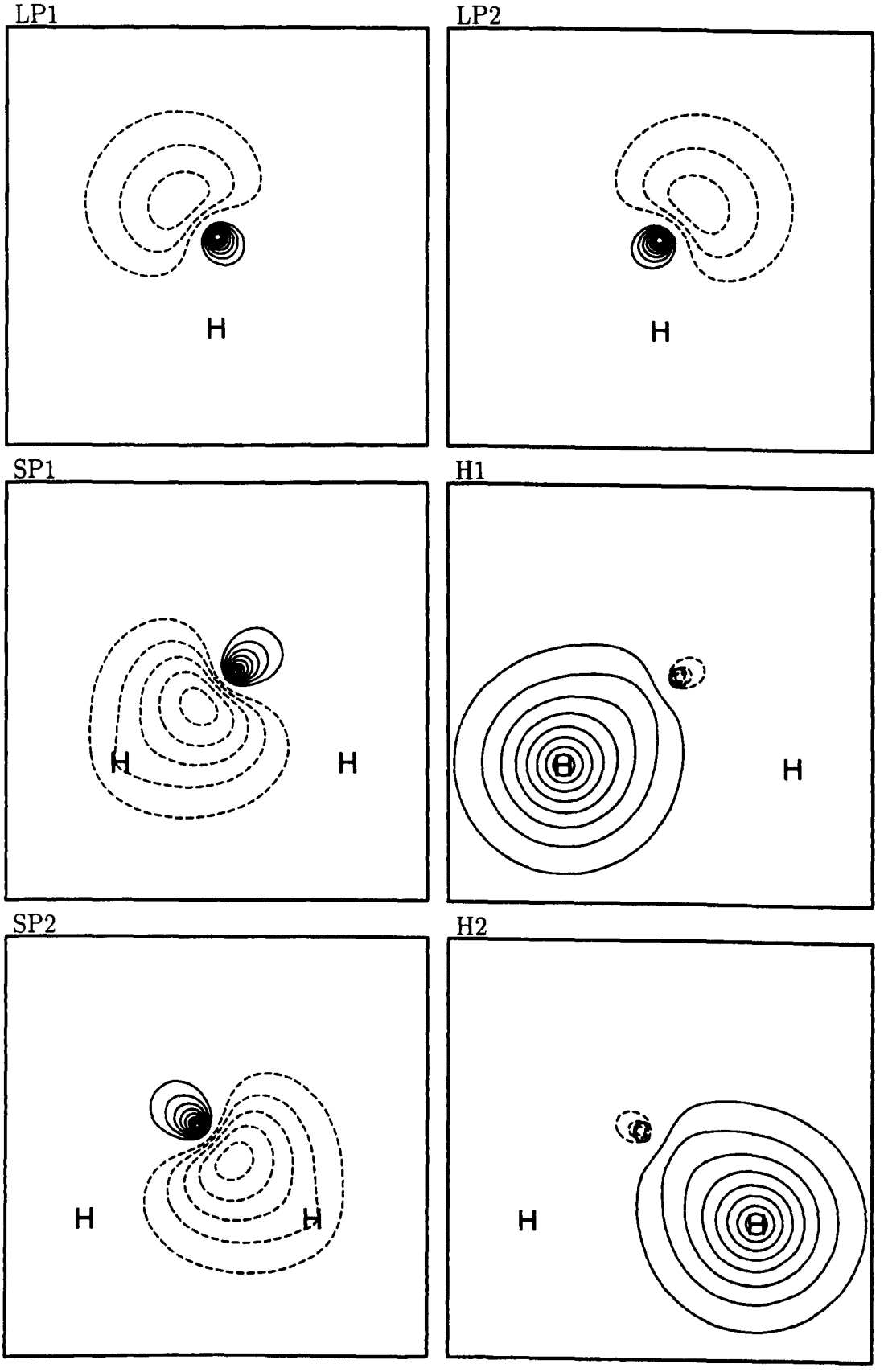

Figure 4.2: Spin-coupled orbitals of CH_2 (singlet).

character.

The electrons associated with the lone-pair are *exactly* singlet-coupled. This is necessary to give a symmetry-pure wavefunction since only the lone-pair orbitals contain components of the B_1 irreducible representation (cf. section 1.6.1). As a consequence there are only two non-vanishing coefficients in the Kotani basis (based on the orbital order shown in figure 4.2). The perfect-pairing spin function ($^K\Theta_5$) is overwhelmingly dominant, contributing 98.81% to the total spin function (see also table 4.5).

The spin-coupled wavefunction recovered 92.63% of the correlation energy attained by the CAS (see table 4.1). The discrepancy of about 6 millihartree can be reduced by an order of magnitude by adding to the covalent structures of the spin-coupled orbitals all the ionic (see table 4.1—this energy is given under the heading: SC+CI). The final difference of about 0.5 millihartree may be attributed to the difference in the orbital spaces spanned by the spin-coupled and CAS orbitals respectively. Another way of indirectly obtaining a measure of this difference would be to perform a spin-coupled calculation in the space of the CAS active MOs (see table 4.1 under SC^{CAS}). The energy difference between this and the full spin-coupled calculation is similarly around 0.6 millihartree. It is a general finding in the following that the difference in the orbital spaces is of relatively minor importance compared with the inclusion of ionic structures

To quantify this difference we examined the distribution of each of the spin-coupled orbitals between the active and virtual orbital spaces. We found that 0.0063%, 0.0901% and 0.1024%, of the orbital was expanded in the virtual CAS MOs, for the lone-pair, sp^x-bonding hybrids and hydrogen orbitals respectively. Clearly, since all our criteria are based on the transformation of the active CAS MOs, this is the minimum deviation from the spin-coupled orbitals that we can hope to achieve (in the sense that the overlap between equivalent orbitals will be maximum), so it is encouraging that the numbers are so small. Projecting the spin-coupled orbitals onto the active CAS space (by simply truncating the

coefficient vectors) would give the 'ideally' most spin-coupled-like orbital set, so we will in the following compare this set with the orbitals obtained from the various criteria. This orbital set will be referred to as SC^{trunc}.

All four criteria have proven extremely successful in reproducing the spin-coupled picture for this system. The orbital plots shown in figures 4.3–4.6 are difficult to distinguish from the spin-coupled orbitals by eye. The clearest differences may be found in the diffuse low contours of the sp^x type hybrids. The CASVB1 hybrids, for example, are clearly not as diffuse as those of the other orbital sets. Among the others CASVB4 seems to give the most diffuse hybrids whereas CASVB2 and spin-coupled are very similar indeed. Similar but less obvious differences in the amount of deformation may be ascertained for the hydrogen and lone-pair orbitals. The spin-coupled solutions seems to give slightly more localised hydrogen orbitals, but slightly less localised lone-pair orbitals whereas the difference amongst the various criteria are minute.

Examining the overlap matrices (tables 4.3 and 4.4) gives a clearer picture of the pattern in the variations between the orbital sets. As indicated by the plots, all the methods agree very closely on the description of the lone-pair orbitals, and the variation in the LP1–LP2 overlap is less than 0.5% in all the cases. If we in this way assume the variations of the lone-pair orbitals to be negligible, the variations of the LP–SP and LP–H overlaps must be attributed to changes in the sp^x type hybrids and hydrogen orbitals. For the sp^x hybrids this thus gives an indication of the localisation of the orbital (but there will also be an angular dependence of the overlap). For the bonding orbitals, CASVB2 and 4 reproduce the spin-coupled picture almost exactly, whereas CASVB3 and particularly CASVB1 differ somewhat. Of particular note is the decrease of the overlap between the bond-forming orbitals.

We note that it is the inclusion of the $1b_1$ orbital in the CAS wavefunction that gives the spin-coupled-like description of the lone-pair. The need to go beyond simple HF (in which there is no B_1 component) to describe this feature

should therefore be obvious. Furthermore, the great s-character of the lone-pair orbitals may be rationalised from the small occupation number of $1b_1$ (0.08803).

Turning from the analysis of the orbitals to the spin-space, we note that, as discussed in section 3.1, the spin coupling may be obtained on the basis of three different criteria—the overlap based and energy based optimisations, or taken from the CAS wavefunction expressed in terms of the given orbital sets. We have tabulated weights for the three possibilities in table 4.5 (with superscripts s, e and c respectively). We note that Θ^c weights for the spin-coupled orbitals are based on the wavefunction obtained by performing a full CI using the SC orbitals, *not* the CAS wavefunction (with the present version of our codes is was not possible to overlap-optimise the spin coupling in this case).

As a general trend, the perfect-pairing function clearly has the largest weight when the spin coupling is determined by the CAS wavefunction. Generally the weight is smallest when the spin coupling is energy-optimised, but the differences in this case are not as large. With the greater importance of the perfect-pairing function it is clear that the *orbitals* for CASVB2 and 4 adapt accordingly—by increasing the overlap between the bond-forming orbital pairs. That the spin coupling is generally different for CASVB1 and 3 is a consequence of the fact that the orbitals in these two cases differ the most from the other orbital sets. Again it is interesting that for the spin coupling used in the two optimisations the weights of the perfect-pairing functions are quite low.

In table 4.6 we have tabulated values of $S_{\rm cov}$ and $E_{\rm cov}$ corresponding to the various orbital sets and spin-coupling coefficients. As such, CASVB1–4 optimise $S^s_{\rm cov}$, $S^c_{\rm cov}$, $E^e_{\rm cov}$ and $E^c_{\rm cov}$ respectively—$S^e_{\rm cov}$ and $E^s_{\rm cov}$ would lead to the 'mixed' criteria mentioned in section 3.1.

All orbital sets are of an excellent all-round quality when compared with the optimal values $S^s_{\rm cov}$(CASVB1)=0.99862 and $E^e_{\rm cov}$(CASVB3)=−38.948156 hartree. The CASVB1 orbitals are perhaps a little behind the others, particularly in the value for $E^c_{\rm cov}$. It is worth noting, that the change in energy and the spin-coupling

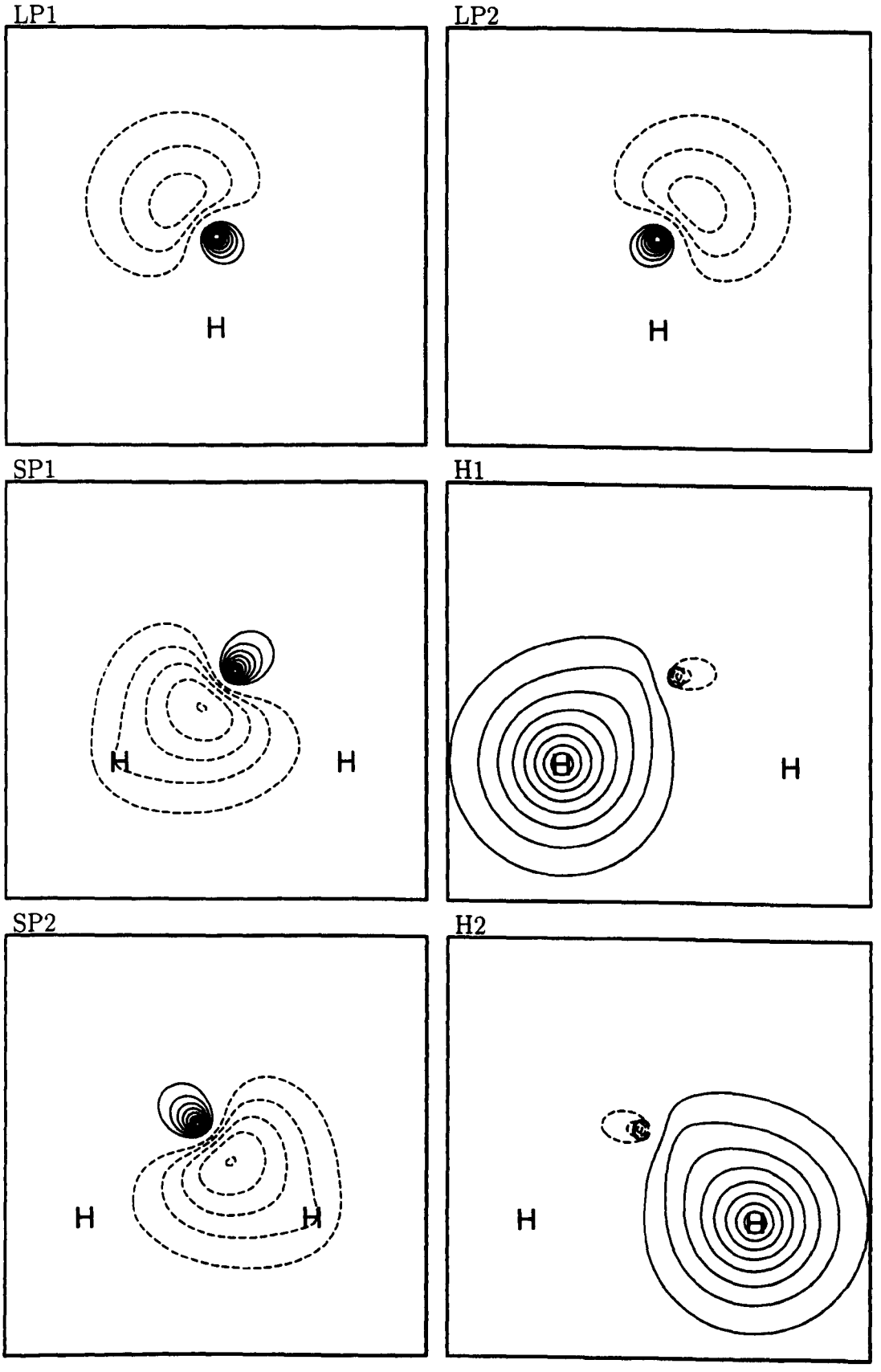

Figure 4.3: CASVB1 orbitals of CH_2 (singlet).

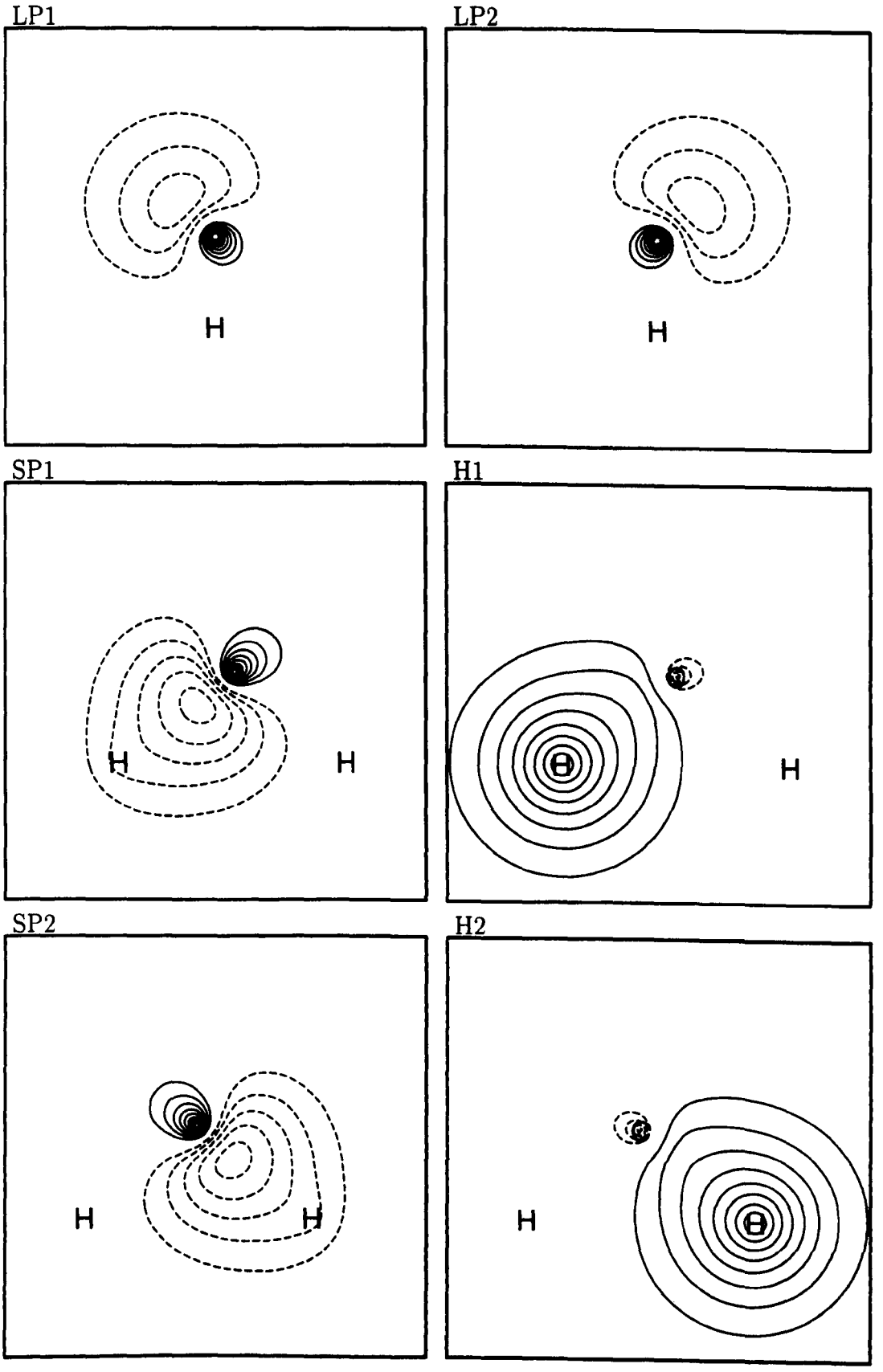

Figure 4.4: CASVB2 orbitals of CH_2 (singlet).

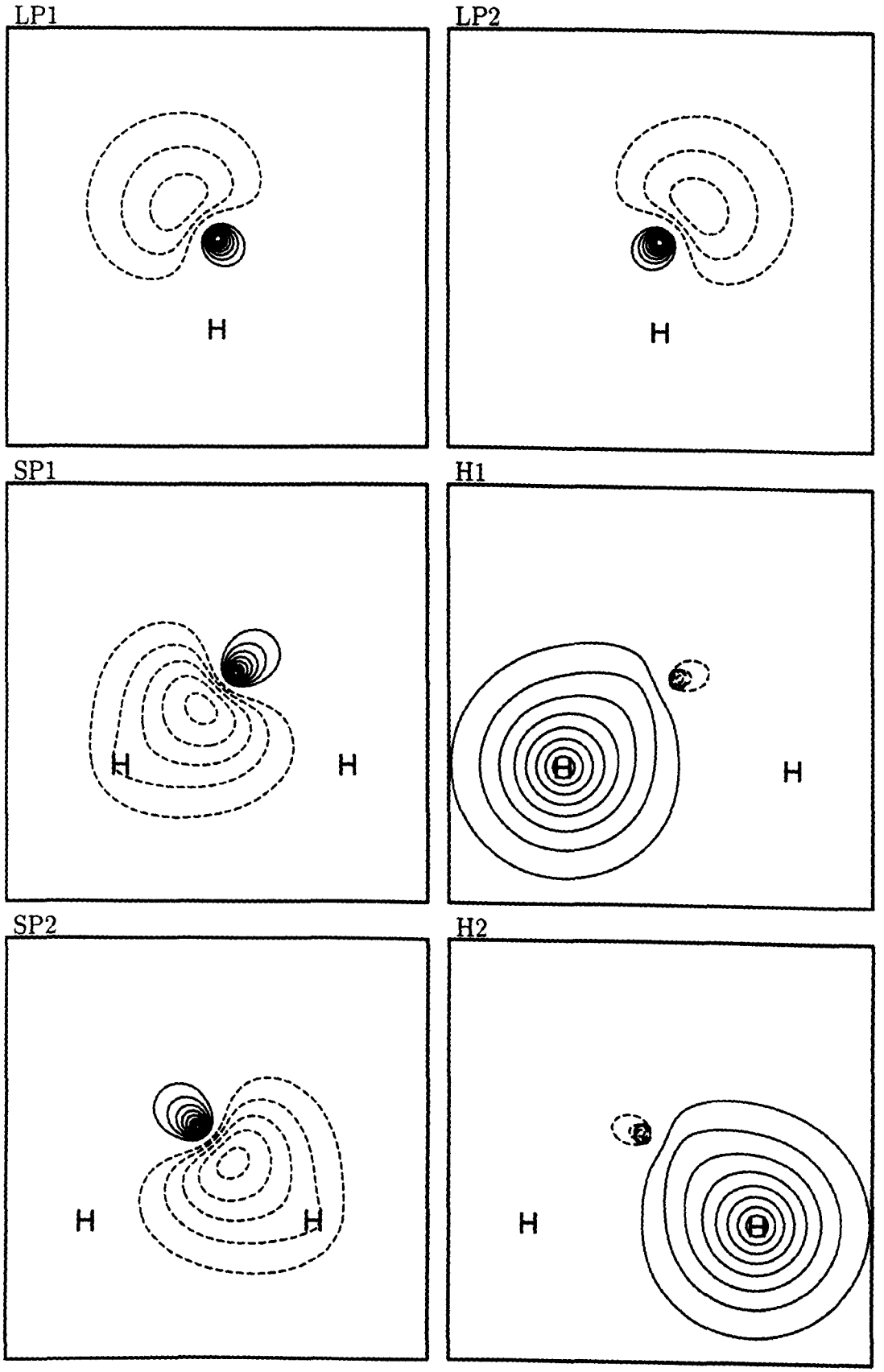

Figure 4.5: CASVB3 orbitals of CH_2 (singlet).

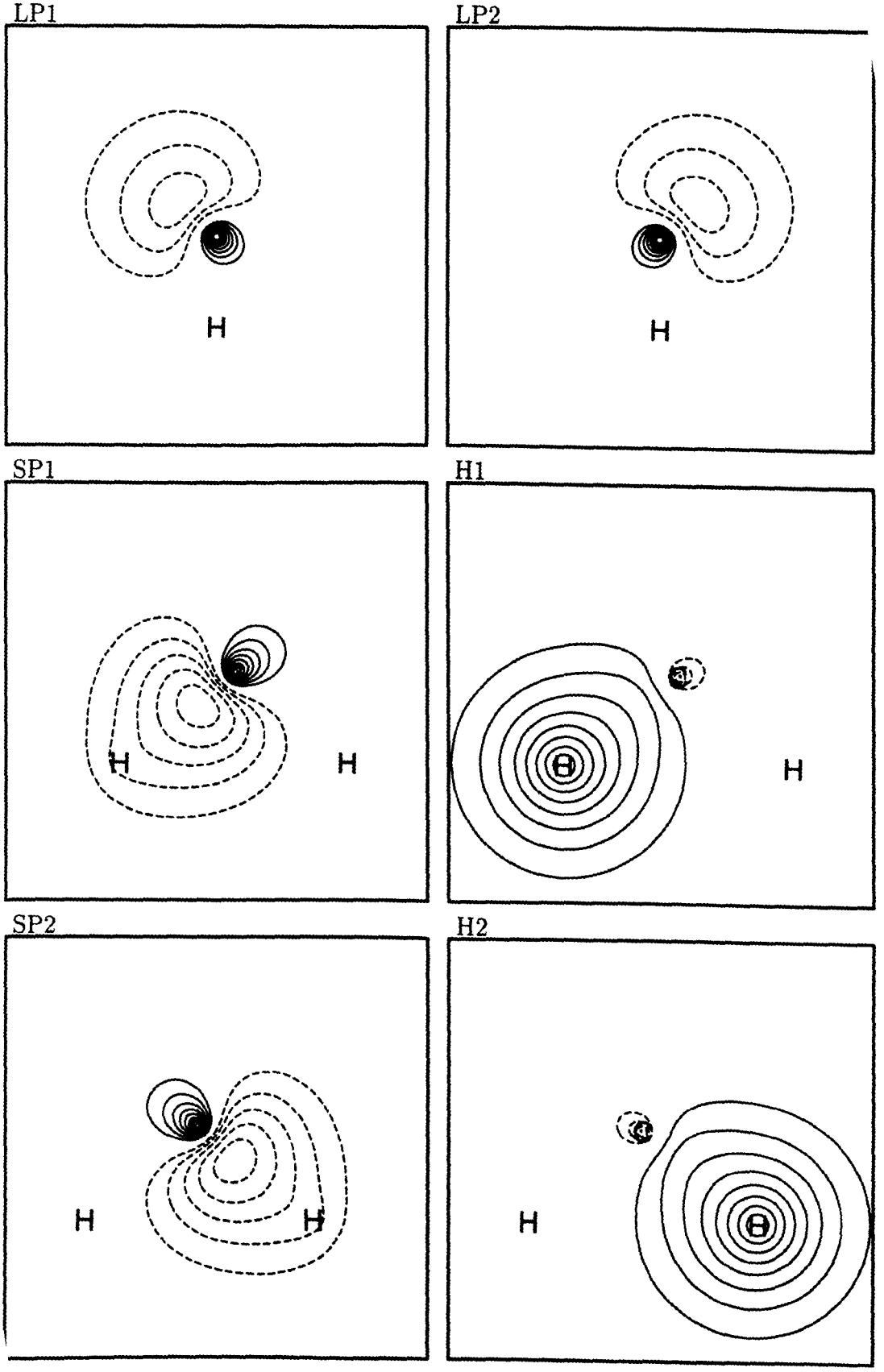

Figure 4.6: CASVB4 orbitals of CH_2 (singlet).

SC^{trunc}	LP1	LP2	SP1	H1	SP2	H2
LP1	1	0.67272	0.19735	0.13178	0.19735	0.13178
LP2		1	0.19735	0.13178	0.19735	0.13178
SP1			1	0.80777	0.31426	0.12087
H1				1	0.12087	-0.05554
SP2					1	0.80777
H2						1
CASVB1	LP1	LP2	SP1	H1	SP2	H2
LP1	1	0.67317	0.24466	0.13304	0.24466	0.13304
LP2		1	0.24466	0.13304	0.24466	0.13304
SP1			1	0.74579	0.51738	0.11675
H1				1	0.11675	-0.29320
SP2					1	0.74579
H2						1
CASVB2	LP1	LP2	SP1	H1	SP2	H2
LP1	1	0.67344	0.22814	0.15780	0.22814	0.15781
LP2		1	0.22814	0.15780	0.22814	0.15781
SP1			1	0.80943	0.31012	0.11663
H1				1	0.11663	-0.07398
SP2					1	0.80943
H2						1

Table 4.3: Overlap matrices for the truncated spin-coupled, CASVB1 and CASVB2 orbitals respectively.

CASVB3	LP1	LP2	SP1	H1	SP2	H2
LP1	1	0.67092	0.22123	0.12949	0.22123	0.12949
LP2		1	0.22123	0.12949	0.22123	0.12949
SP1			1	0.78157	0.40317	0.11027
H1				1	0.11027	-0.17451
SP2					1	0.78157
H2						1
CASVB4	LP1	LP2	SP1	H1	SP2	H2
LP1	1	0.67102	0.21018	0.14280	0.21019	0.14280
LP2		1	0.21018	0.14280	0.21019	0.14280
SP1			1	0.81085	0.26078	0.10507
H1				1	0.10507	-0.03747
SP2					1	0.81085
H2						1

Table 4.4: Overlap matrices for the CASVB3 and CASVB4 orbitals respectively.

	${}^{K}\Theta_3^s$	${}^{K}\Theta_5^s$	${}^{K}\Theta_3^e$	${}^{K}\Theta_5^e$	${}^{K}\Theta_3^c$	${}^{K}\Theta_5^c$
SC	–	–	1.19%	98.81%	0.78%	99.22%
SC^{trunc}	1.20%	98.80%	1.34%	98.66%	0.83%	99.17%
CASVB1	2.10%	97.90%	2.30%	97.70%	0.24%	99.76%
CASVB2	1.08%	98.92%	1.28%	98.72%	0.82%	99.18%
CASVB3	1.47%	98.53%	1.62%	98.38%	0.60%	99.40%
CASVB4	1.01%	98.99%	1.18%	98.82%	0.90%	99.10%

Table 4.5: Weights of the Kotani functions for methylene.

	S^s_{cov}	S^e_{cov}	S^c_{cov}
$\mathrm{SC}^{\mathrm{trunc}}$	0.99855	0.99855	0.99854
CASVB1	0.99862	0.99814	0.99814
CASVB2	0.99860	0.99859	0.99859
CASVB3	0.99858	0.99858	0.99851
CASVB4	0.99856	0.99856	0.99856
	E^s_{cov}/hartree	E^e_{cov}/hartree	E^c_{cov}/hartree
$\mathrm{SC}^{\mathrm{trunc}}$	−38.948108	−38.948111	−38.948070
CASVB1	−38.948061	−38.948066	−38.946893
CASVB2	−38.948080	−38.948085	−38.948051
CASVB3	−38.948153	−38.948156	−38.947945
CASVB4	−38.948125	−38.948129	−38.948117

Table 4.6: Covalent overlaps and energies for the different orbital sets for methylene.

coefficients upon adding all ionic structures, or when going form the (full) SC orbital set to the active CAS orbital space, is large compared to most of the differences between the various orbital sets.

In conclusion, methylene can be viewed as a straightforward case, where the correspondence between the spin-coupled wavefunction and CAS is fairly clear-cut. All four criteria reproduce the spin-coupled picture, in particular the description of the two non-bonding orbitals. We can therefore claim with some justification, that the CAS method vindicates the spin-coupled interpretation of the bonding in CH_2.

	E/hartree	$(E-E_{\mathrm{CAS}})$/millihartree
SCF	−230.764056	72.77
SC^{CAS}	−230.829233	7.59
SC	−230.829331	7.49
SC+CI	−230.836729	0.09
CAS	−230.836822	0.00

Table 4.7: Benzene energies for the various wavefunctions considered.

4.2 Benzene

The geometry used for benzene was r(C–C)=1.3964 Å and r(C–H)=1.0831 Å, as determined in ref. [7] by vibration-rotation spectroscopy. The basis set was the same pVTZ basis as used for methylene [5], except for the polarisation functions which were in this case a single Cartesian d Gaussian on C, with ζ=0.8, and a single p Gaussian on H, ζ=1.0.[3]

The MOs in the CAS calculation were partitioned into 18 optimised core orbitals of σ symmetry, 6 active orbitals of π symmetry giving 96 and 30 virtuals of σ and π symmetry respectively. The converged active natural orbitals were in order of decreasing occupation number: $1a_{1u}$, the two components of $1e_{1g}$, the two components of $1e_{2u}$, and $1b_{1g}$ (see figure 4.7). As in the case of CH_2, the occupation numbers for the first three MOs were close to 2 (and the weight of the determinant with all three orbitals doubly occupied was 88.43%) suggesting no fundamental shortcomings of the Hartree-Fock picture (for which the basic MO picture is the same). The correlation energy retrieved by the CAS wavefunction amounted to 72.8 millihartree (see table 4.7).

For the purpose of the spin-coupled calculation, the 6 active electrons were expanded in the 36 MOs of π symmetry. As for CH_2 the core was chosen to be identical to the CAS core, i.e., it was not reoptimised for the spin-coupled

[3]Approximately the geometric mean of the corresponding polarisation function exponents used for CH_2. This was done to limit the total number of basis functions.

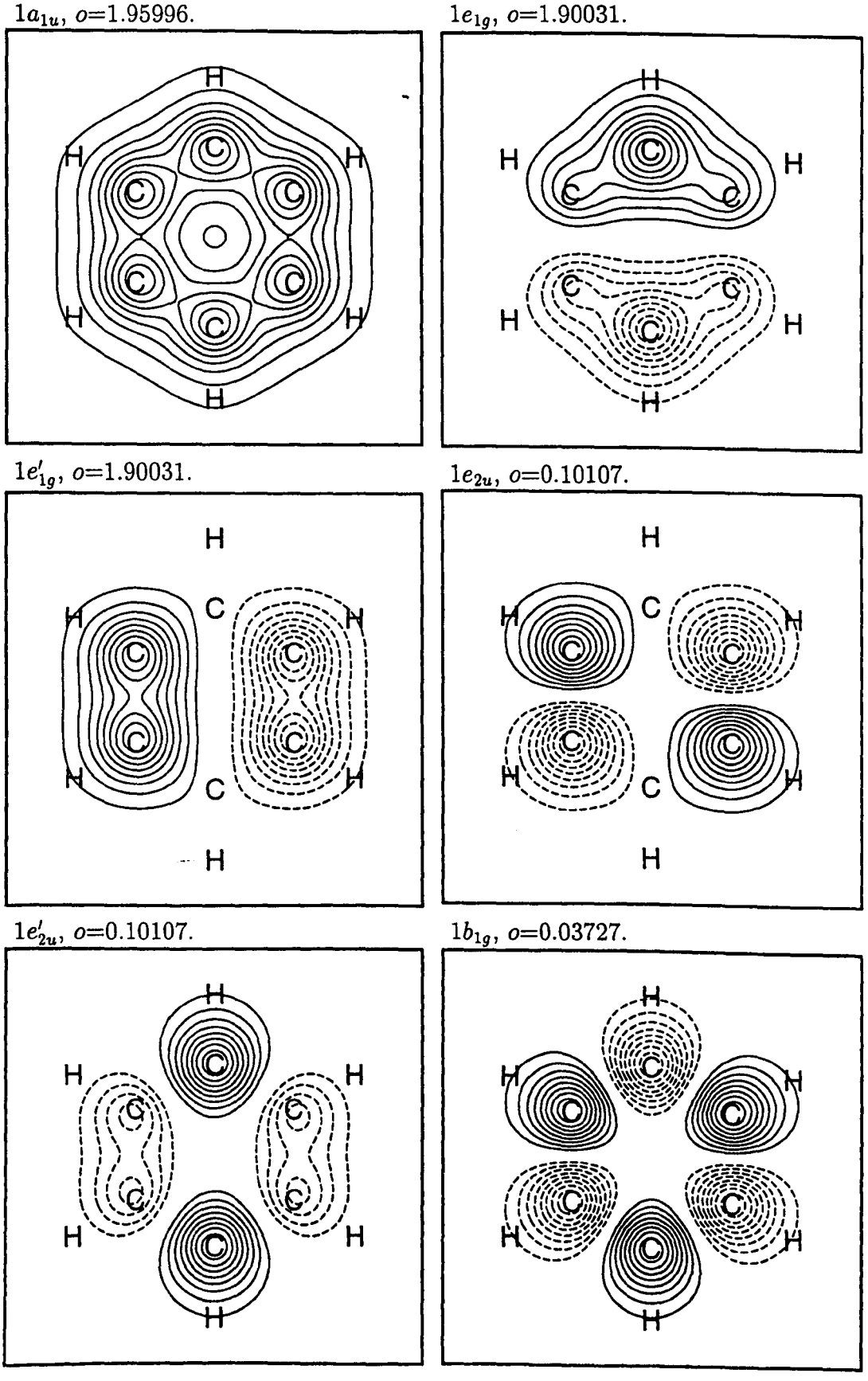

Figure 4.7: Natural orbital representation of active CAS MOs in benzene and their occupation numbers.

active-electron density. The spin-coupled solution consists of six highly localised, equivalent orbitals. A contour plot of $|\phi_1|^2$, 1 bohr above the molecular plane, is shown in figure 4.8. The spin-coupled description of benzene and a range of other aromatic molecules has been reported previously [6]. As can be seen, this orbital is essentially a 2p function centred on C, with slight delocalisation towards the neighbouring carbon atoms. The five other orbitals can be generated by successive C_6 rotations. The coupling of the associated electron spins is conveniently expressed in the Rumer basis, where the two 'Kekulé' structures are strongly dominant (having Chirgwin-Coulson weights of 40.6% each), with only slight contributions from the three so-called 'Dewar' structures (of 6.3% each) (see table 4.10).

About 0.0022% of each spin-coupled orbital lay outside the CAS active space. This relatively low value is likely to be associated with the high symmetry of benzene, and the consequently reduced number of free parameters (we found this value to be heavily dependent on the basis set size, for example). So the major part of the energy difference can be attributed to adding the ionic structures (see table 4.7). The spin-coupled wavefunction recovered 89.71% of the non-dynamical correlation energy, but including the ionic structures took this number to 99.87%. Even though the energy gain is comparable to that for methylene, is seems that adding ionic structures has a more profound effect in the case of benzene. A clear indication of this fact is the change in the relative weights of spin-coupling as shown in table 4.10. It is clear that inclusion of the ionic structures significantly stabilises the Dewar structures relative to the Kekulé structures.

In the optimisation of the orbitals according to the four criteria, a tendency to break symmetry was encountered. This was particularly pronounced for CASVB2 and 4, and the solutions presented here were reproduced with several different basis sets.

Even in standard variational methods, solutions with broken symmetry or other features known not to be in agreement with the true physical state are

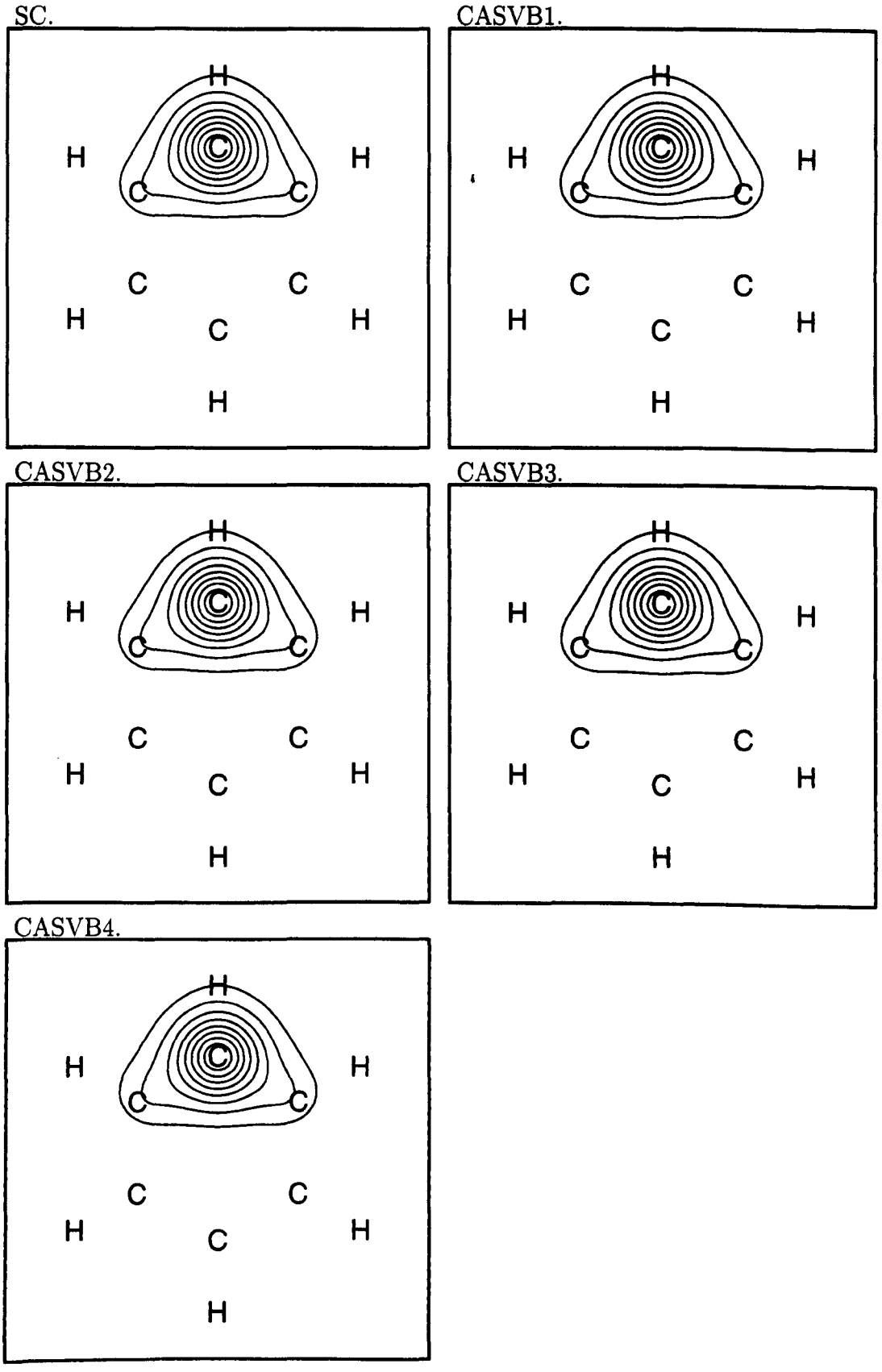

Figure 4.8: Orbitals of all symmetric orbital sets.

often encountered. Constraints are routinely imposed to ensure that the overall wavefunction is of the correct symmetry. The fact that the optimisation is overlap-based or that particular constraints on the spin coupling are imposed should not affect these considerations. We have nevertheless included also the symmetry-broken solutions here, in order to, if possible, gain further insights into the nature of the optimisation schemes.

The obvious problem when imposing constraints is how to do so without jeopardising the flexibility of the spin-coupled-like wavefunctions. When the symmetry-broken solutions, as in the present case, can be viewed as merely perturbations of symmetry-pure solutions (and we expect this will often be the case), this is in practice not a problem. By running the unconstrained optimisation first, the basic nature of the solution will become obvious.

When optimised in this way, CASVB1, 2 and 4 all had near-singular Hessians at convergence. In all cases there were two very small eigenvalues relating to variations in the orbital parameters only. For CASVB2 the eigenvalues were positive, for CASVB4 negative, thus rendering the symmetry-pure solutions unstable. In the case of CASVB1 the Hessian was (doubly) singular within the numerical accuracy of our codes (it is possible that increased accuracy would reveal a preference for CASVB2- or 4-like solutions). For CASVB3, two low eigenvalues could similarly be identified, but they were in this case positive giving a stable symmetric solution.

As can be seen from the orbital plots, figures 4.8, 4.9 and 4.10, all criteria reproduced the spin-coupled picture remarkably accurately. In the case of the symmetry-broken solutions for CASVB2 and 4 some variation in the deformation of the orbitals is present, but the symmetric versions of these criteria (which approximately represents the 'average' amount of deformation of the un-symmetric orbitals) resemble the spin-coupled orbitals very closely. We note that the non-symmetric solutions both have mirror planes. For CASVB2 this goes through C1 and C4 so that ϕ_1 is symmetric and ϕ_2, say, can be reflected into ϕ_6. For

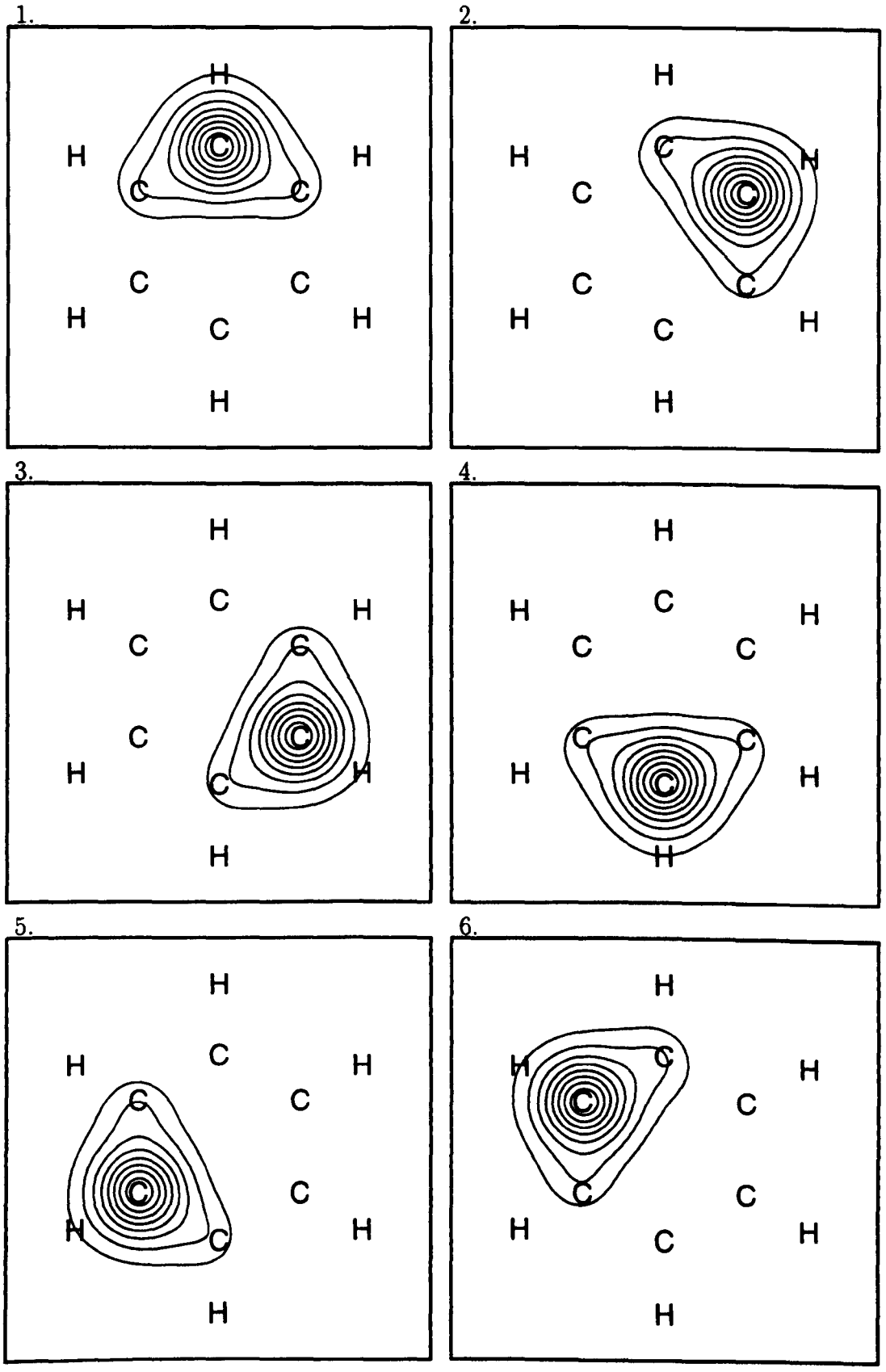

Figure 4.9: CASVB2 orbitals.

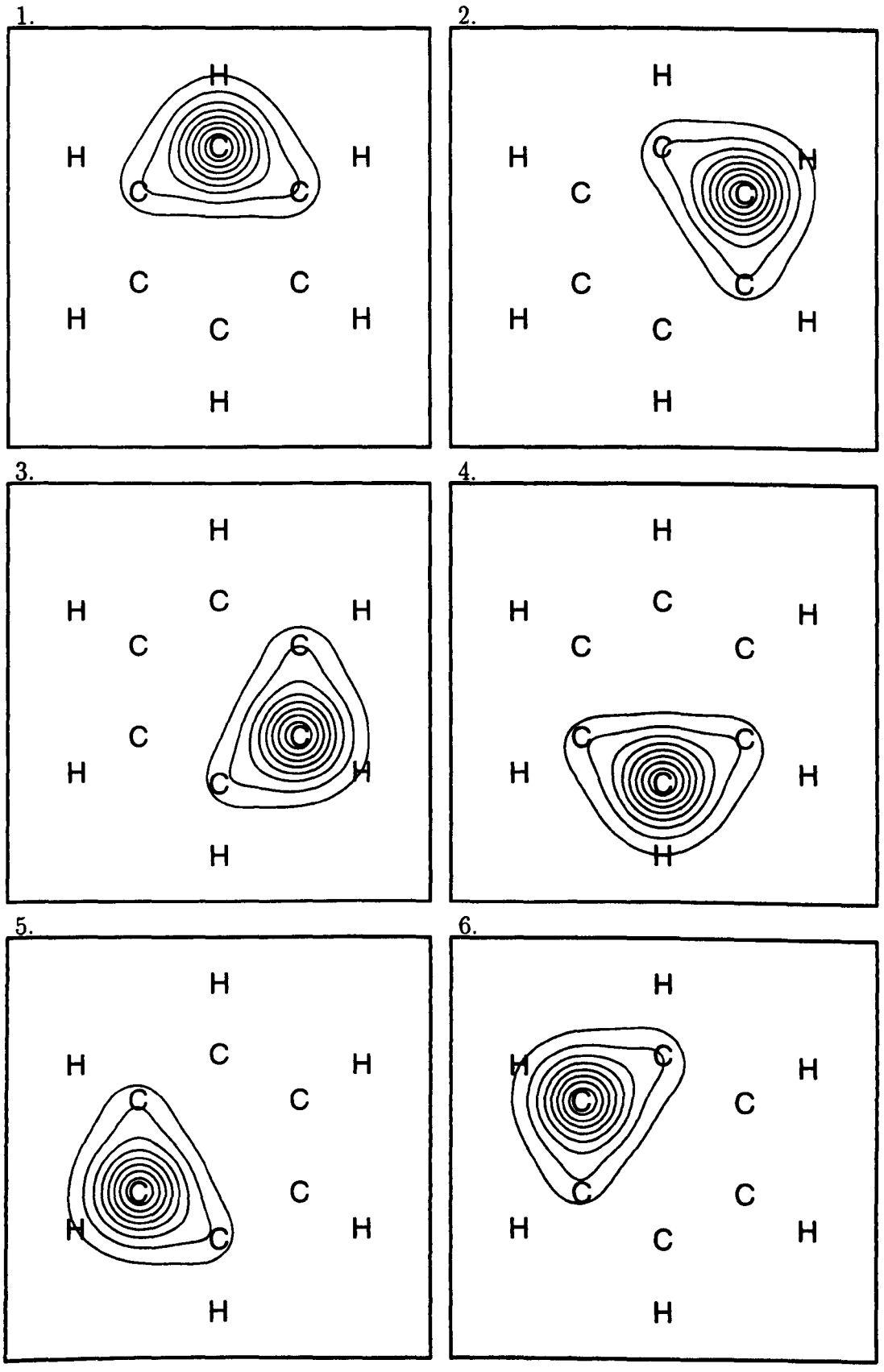

Figure 4.10: CASVB4 orbitals.

	$\langle\phi_1\|\phi_2\rangle$	$\langle\phi_1\|\phi_3\rangle$	$\langle\phi_1\|\phi_4\rangle$
SC	0.52383	0.02939	-0.15700
CASVB1	0.51564	-0.01664	-0.21672
CASVB2	0.49919	-0.07186	-0.27225
CASVB3	0.52355	0.02809	-0.15804
CASVB4	0.50056	-0.04674	-0.23127

Table 4.8: Unique overlap integrals for the symmetry-pure orbital sets of benzene.

CASVB4 the mirror plane bisects the C1–C2 and C4–C5 bonds. This can be verified by examining the overlap matrices (table 4.9), but is also reflected in a partial symmetry in the spin-coupling coefficients (table 4.10).

Examining the values for the overlap, the apparent success of CASVB3 in reproducing spin-coupled is remarkable. This, of course, is closely related to the relatively small difference between the CAS and SC orbital spaces. CASVB1 also gives very spin-coupled-like values, whereas the symmetric CASVB2 and 4 seems to agree more with each other than with the other orbital sets.

These findings are reflected in the variations in the spin-coupling coefficients between the orbital sets (table 4.10). In this table we have used '(s)' for the orbital sets that were constrained to be symmetry-pure. The weights of the Kekulé spin functions are largest for the energy-optimised spin couplings, and smallest when the spin coupling is determined by the CAS wavefunction, and this rationalises the decreased values for $\langle\phi_1|\phi_2\rangle$ for CASVB2 and 4, and the slightly higher value for CASVB3 seen in table 4.8.

Values for S_{cov} and E_{cov} are given in table 4.11. Again CASVB3 seems to be in best agreement with the (projected) spin-coupled values, but the variations between all the orbital sets are very slight. For the energy, for example, the maximum deviation amounts to no more than 0.5 millihartree.

CASVB2	1	2	3	4	5	6
1	1	0.58673	0.05608	-0.24512	0.05608	0.58673
2		1	0.48906	-0.18195	-0.24708	0.16706
3			1	0.42168	-0.24581	-0.24708
4				1	0.42168	-0.18195
5					1	0.48906
6						1
CASVB4	1	2	3	4	5	6
1	1	0.56136	0.09910	-0.21989	-0.04748	0.54559
2		1	0.54559	-0.04749	-0.21989	0.09911
3			1	0.44734	-0.16611	-0.21464
4				1	0.45757	-0.16611
5					1	0.44734
6						1

Table 4.9: Overlap matrices for non-symmetric orbital sets.

	Θ^s_{K1}	Θ^s_{K2}	Θ^s_{D1}	Θ^s_{D2}	Θ^s_{D3}
SC^{trunc}	36.92%	36.92%	8.72%	8.72%	8.72%
CASVB1	36.53%	36.53%	8.98%	8.98%	8.98%
CASVB2	33.72%	33.72%	10.64%	10.96%	10.96%
CASVB2 (s)	34.33%	34.33%	10.45%	10.45%	10.45%
CASVB3	36.87%	36.87%	8.75%	8.75%	8.75%
CASVB4	33.66%	34.40%	10.48%	10.98%	10.48%
CASVB4 (s)	34.51%	34.51%	10.33%	10.33%	10.33%
	Θ^e_{K1}	Θ^e_{K2}	Θ^e_{D1}	Θ^e_{D2}	Θ^e_{D3}
SC	40.60%	40.60%	6.27%	6.27%	6.27%
SC^{trunc}	40.70%	40.70%	6.20%	6.20%	6.20%
CASVB1	40.96%	40.96%	6.03%	6.03%	6.03%
CASVB2	36.83%	36.83%	8.71%	8.82%	8.82%
CASVB2 (s)	37.86%	37.86%	8.09%	8.09%	8.09%
CASVB3	40.62%	40.62%	6.26%	6.26%	6.26%
CASVB4	35.75%	37.40%	8.81%	9.22%	8.81%
CASVB4 (s)	37.30%	37.30%	8.47%	8.47%	8.47%
	Θ^c_{K1}	Θ^c_{K2}	Θ^c_{D1}	Θ^c_{D2}	Θ^c_{D3}
SC	31.56%	31.56%	12.29%	12.29%	12.29%
SC^{trunc}	31.50%	31.50%	12.33%	12.33%	12.33%
CASVB1	31.37%	31.37%	12.42%	12.42%	12.42%
CASVB2	32.92%	32.92%	11.48%	11.33%	11.33%
CASVB2 (s)	32.37%	32.37%	11.75%	11.75%	11.75%
CASVB3	31.52%	31.52%	12.32%	12.32%	12.32%
CASVB4	32.09%	33.57%	11.48%	11.38%	11.48%
CASVB4 (s)	32.56%	32.56%	11.63%	11.63%	11.63%

Table 4.10: Weights of the Rumer functions for benzene.

	S^s_{cov}	S^e_{cov}	S^c_{cov}
$\mathrm{SC}^{\mathrm{trunc}}$	0.99522	0.99517	0.99513
CASVB1	0.99546	0.99540	0.99539
CASVB2	0.99545	0.99542	0.99545
CASVB2 (s)	0.99544	0.99540	0.99543
CASVB3	0.99521	0.99517	0.99513
CASVB4	0.99518	0.99516	0.99517
CASVB4 (s)	0.99516	0.99514	0.99515
	E^s_{cov}/hartree	E^e_{cov}/hartree	E^c_{cov}/hartree
$\mathrm{SC}^{\mathrm{trunc}}$	−230.829185	−230.829233	−230.828956
CASVB1	−230.828917	−230.828980	−230.828690
CASVB2	−230.828812	−230.828849	−230.828792
CASVB2 (s)	−230.828835	−230.828877	−230.828776
CASVB3	−230.829186	−230.829233	−230.828962
CASVB4	−230.829087	−230.829113	−230.829060
CASVB4 (s)	−230.829102	−230.829129	−230.829051

Table 4.11: Covalent overlaps and energies for the different orbital sets for benzene.

	E/hartree	$(E-E_{\mathrm{CAS}})$/millihartree
SCF	−40.212541	83.00
$\mathrm{SC}^{\mathrm{CAS}}$	−40.275607	19.93
SC	−40.278158	17.38
SC+CI	−40.291257	4.28
CAS	−40.295536	0.00

Table 4.12: Methane energies for some of the wavefunctions considered.

4.3 Methane

In the calculations on methane the carbon-hydrogen bond length was: r(C–H)= 2.065 bohr. We used the same correlation-consistent pVTZ basis set as for CH_2 [5].

The MOs in the CAS calculation were partitioned as follows:

- One optimised core orbital, to describe the two core electrons ($\approx$C($1s^2$));
- Eight active orbitals, to accommodate the eight valence electrons
- 52 virtual orbitals.

The '8 in 8' CAS energy defined in this way (see table 4.12) represents an improvement over the SCF description of 83.0 millihartree. The eight active MOs and their occupation numbers are given in figures 4.11 and 4.12. The four first occupation numbers are as before close to 2 (and the weight of the configuration with these orbitals doubly occupied is 96.06%) indicating the quality of the SCF description.

A spin-coupled calculation was then carried out for the eight valence orbitals of CH_4, with the two core electrons accommodated in the core orbital from the CAS calculation (without further relaxation). In spite of it being only a single configuration wavefunction, the spin-coupled energy accounts for 79.06% of the correlation energy recovered by the CAS wavefunction. This is a somewhat lower

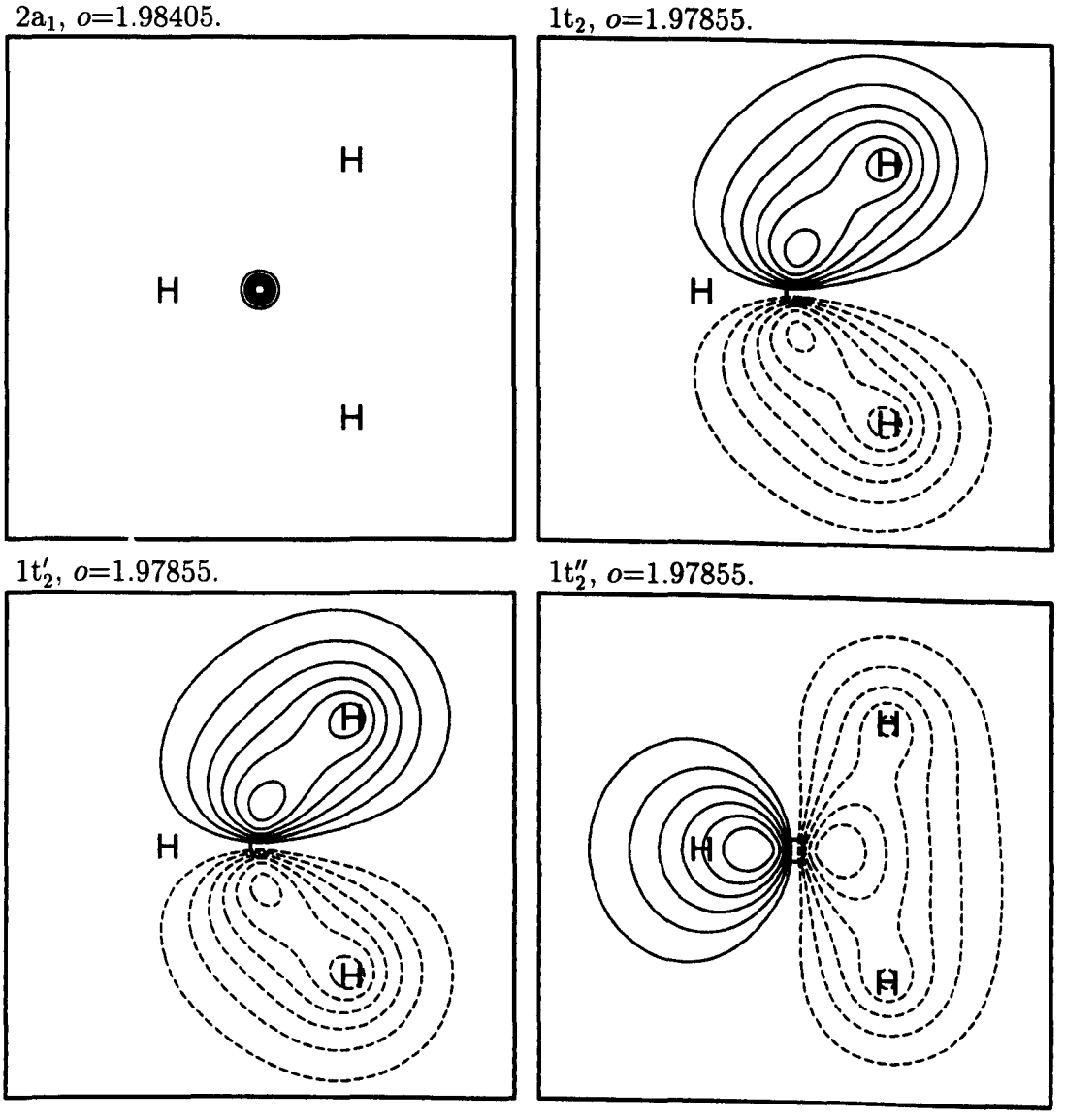

Figure 4.11: MOs of CH_4.

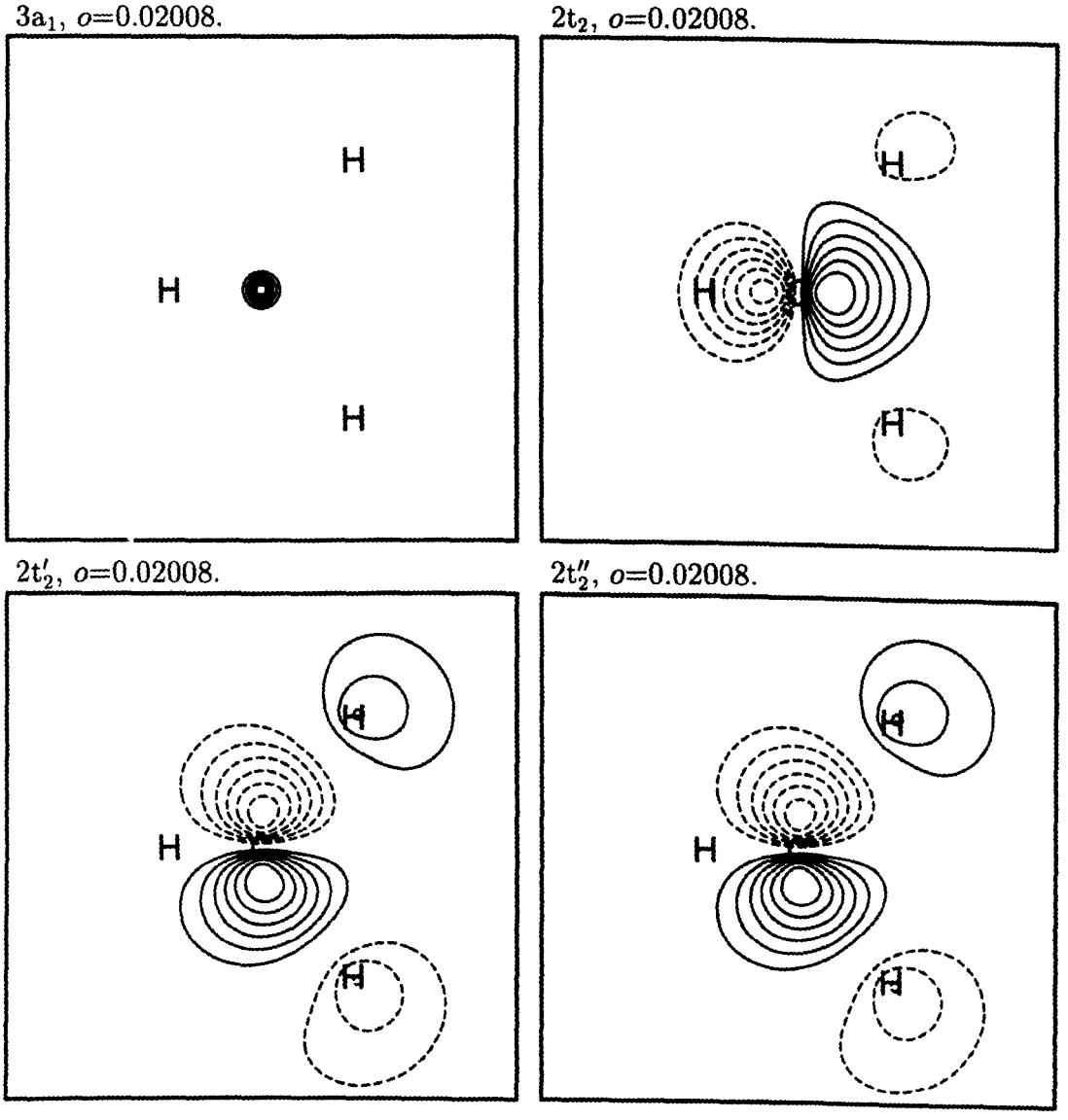

Figure 4.12: MOs of CH_4.

proportion than found for methylene and benzene, and this is a likely consequence of the increased number of active electrons (and thus the increased difference in the number of variational parameters between the two methods as indicated in table 2.1).

The spin-coupled description of methane, calculated with a variety of basis sets, has been presented on many occasions [9]. Although no such constraints were imposed, the converged spin-coupled wavefunction consists of four symmetry-related pairs of orbitals, each associated with a different C–H bond. The symmetry-unique orbitals are shown in figure 4.13 (the plots are generated in the same way as for CH_2 and benzene). Orbital 1 takes the form of a deformed sp^x-like hybrid on carbon and orbital 2 resembles a distorted H(1s) function. The overlap between these two is 0.69 (see table 4.13), with predominant singlet-coupling of the associated electron spins. Orbitals 3, 5 and 7 are the carbon hybrids in the other C–H bonds, orbitals 4, 6 and 8 respectively, are the hydrogen orbitals to which they point.

For a high-symmetry system such as methane it can be particularly beneficial to express the total spin function in the Serber basis. Each of the f_S^N=14 Serber functions for N=8 and S=0 may be represented [10]

$$(((s_{12}s_{34})S_4; s_{56})S_6; s_{78}), \tag{4.1}$$

in which s_{12}, s_{34}, s_{56} and s_{78} may be 0 or 1 (for singlet and triplet electron pairs) and the S_{2p} denote the total spin of the first p electron pairs. We find that there are only four unique non-zero spin-coupling coefficients for CH_4, such that

$$\begin{aligned}\Theta_0^8 = &+b_1(((11)2;1)1;1) + b_2(((11)0;1)1;1)\\ &-b_3(((10)1;0)1;1) - b_3(((01)1;0)1;1) - b_3(((00)0;1)1;1)\\ &-b_3(((10)1;1)0;0) - b_3(((01)1;1)0;0) - b_3(((11)0;0)0;0)\\ &+b_4(((00)0;0)0;0)\end{aligned} \tag{4.2}$$

with all the b_i positive. The Serber functions are orthogonal and so the corresponding weights of the different modes of spin coupling are simply b_i^2, as given in

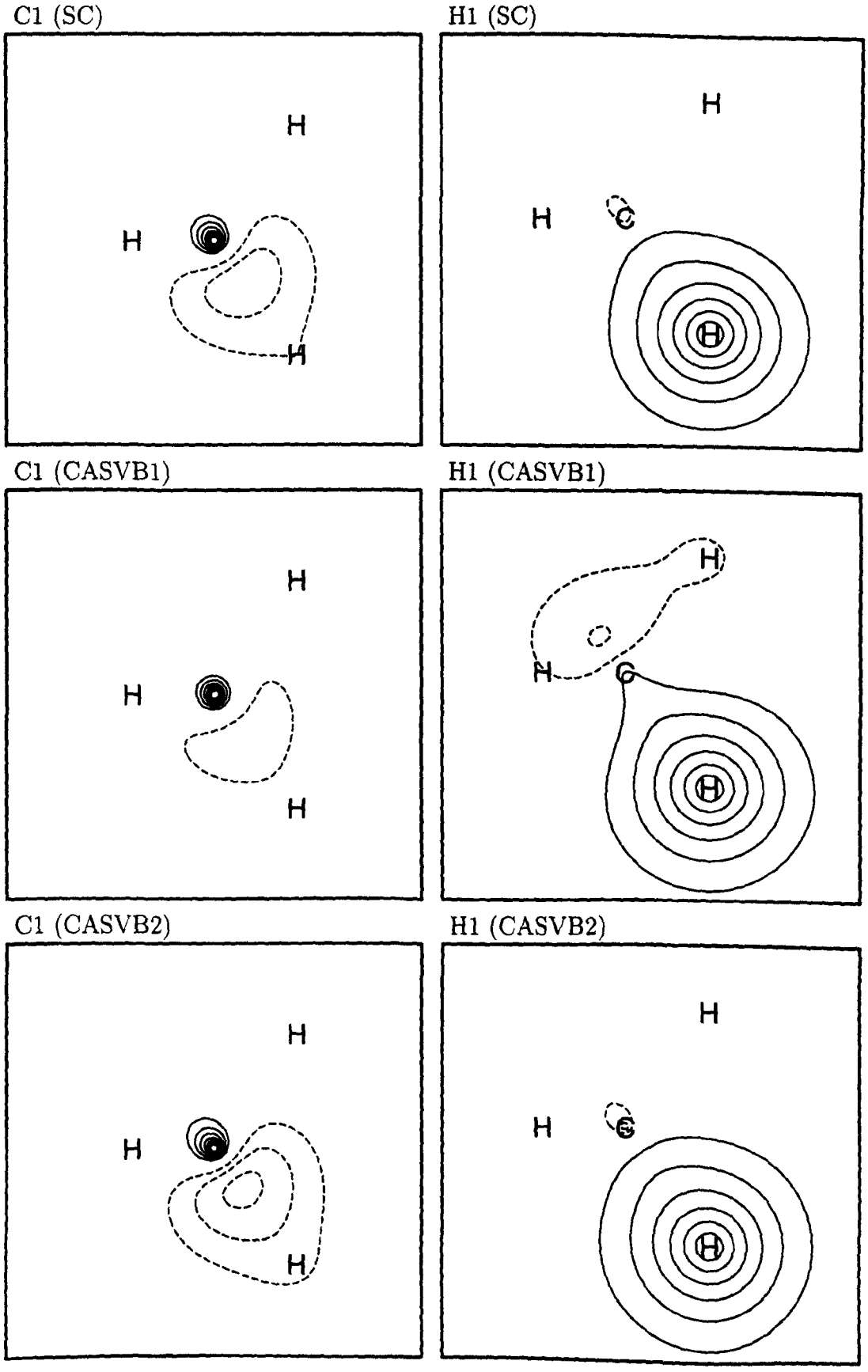

Figure 4.13: Orbitals of CH_4.

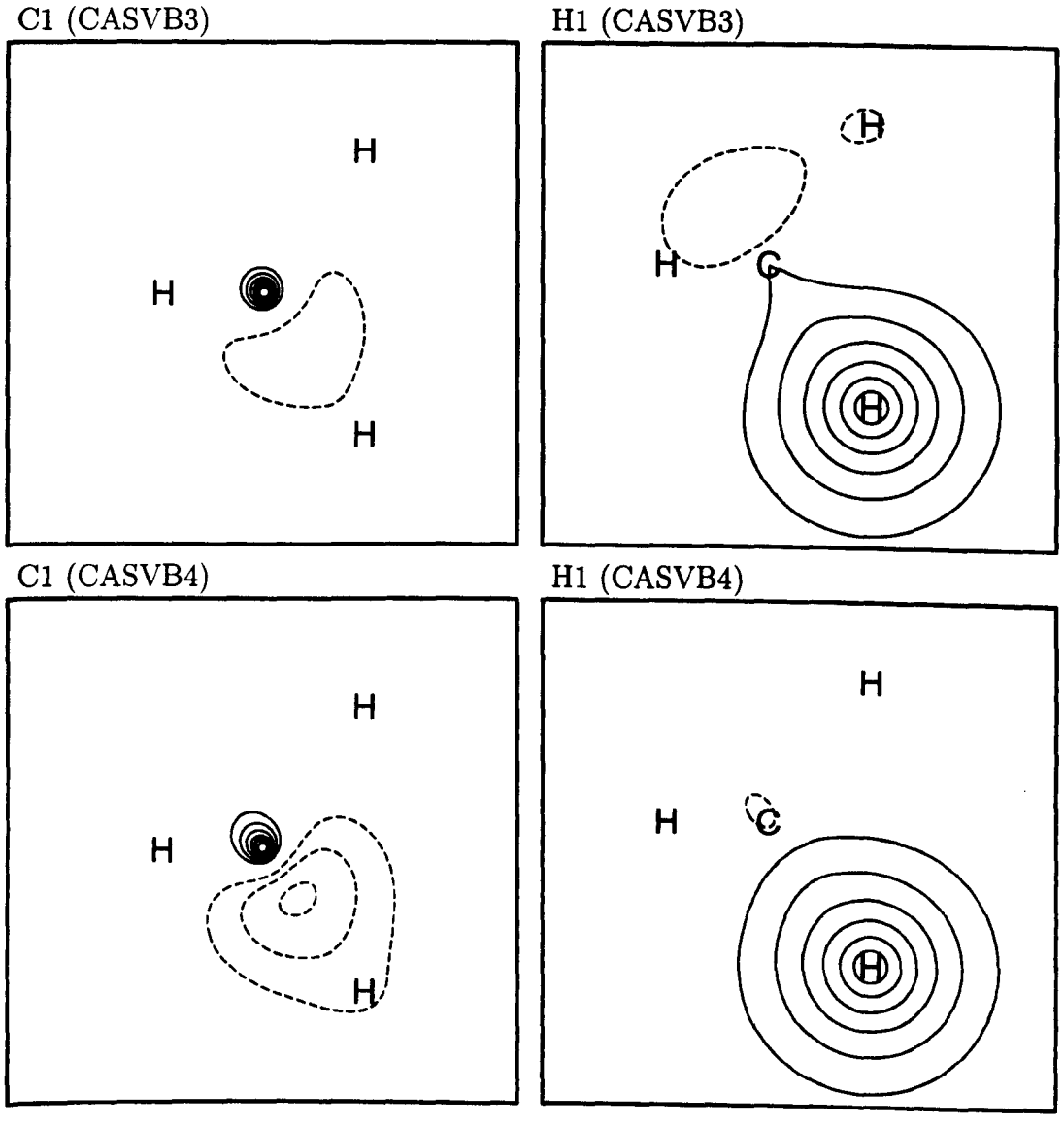

Figure 4.14: Orbitals of CH_4.

table 4.14. The spin coupling for SC is dominated by the perfect-pairing function (90.0%), but the weights of the other modes (particularly those in which only two pairs are triplet-coupled) are not insignificant.

For methane we found that 0.35% and 0.56%, for the carbon and hydrogen orbitals respectively, was outside the CAS active space. Compared with the corresponding values for methylene and benzene these numbers are quite large. The energy difference between SC+CI and the CAS energy was correspondingly larger, 4.3 millihartree, and the difference between the 'full' spin-coupled calculation and the spin-coupled calculation performed in the CAS active orbital space was of a similar magnitude, 2.6 millihartree.

The orbitals obtained from the four different criteria were somewhat less successful for methane than was the case for CH_2 and benzene. CASVB2 and 4 agree qualitatively with the spin-coupled picture, but CASVB1 and 3 can only be described as giving anomalous, 'rogue' solutions. (All the plots are given in figures 4.13 and 4.14.) In both cases the sp^x hybrids on C are extremely localised, and the overlaps between the (supposedly) bond-forming orbitals are a mere 0.20 and 0.27 respectively (see table 4.13). We have previously encountered such anomalous solutions, exhibiting marked differences from the 'full' spin-coupled solution, only in cases in which the total number of orbital free parameters was severely restricted.

The main difference between the solutions based on CASVB2 and 4 and spin-coupled is the larger delocalisation of the bond-forming orbitals for the CAS transformations. This gives rise to an increase in the C1–H1 overlap (table 4.13), and this may again be rationalised from the fact that the weight of the perfect-pairing function is increased when the spin coupling is taken from the CAS wavefunction (see table 4.14).

Compared with the difference from the CAS spin coupling, the difference between the overlap- and energy-optimised spin couplings are minute. The form of the orbitals optimised with CASVB2 and 4 seems to favour the perfect-pairing

mode of spin coupling in all cases. Compared with this the spin couplings for CASVB1 and 3 are very odd, being quite far removed from the perfect-pairing situation. To a good approximation the spin coupling corresponds to coupling all the electrons associated with the C sp^x-type hybrids constructively (to a quintet) and likewise for the electrons associated with the H orbitals. This mode of spin coupling has weights of 92.39% and 88.14% for CASVB1 and 3 respectively. For comparison the corresponding weight for spin-coupled would be 63%, and for CASVB2 and 4 around 48%. It is noticeable that even for the anomalous CASVB1 and 3 orbital sets the perfect-pairing function becomes strongly dominant when the spin coupling is taken from the CAS wavefunction.

Looking at the covalent overlaps and energies given in table 4.15 it is clear that CASVB1 and 3 do quite a bit better than the other orbital sets, generally around 0.001 in S_{cov} and 2–3 millihartree in E_{cov}. That this is actually slightly better than might be expected was already indicated in table 4.12. For CH_2 and benzene we had the approximate identity

$$E_{\text{SC+CI}} - E_{\text{CAS}} \approx E_{\text{SC}^{\text{CAS}}} - E_{\text{SC}},$$

fulfilled within 0.1 millihartree. In this case the discrepancy amounts to 1.73 millihartree.

Basing the spin coupling on the CAS wavefunction (i.e., imposing a near-perfect-pairing function) is totally inappropriate for CASVB1 and 3, which is reflected in the nonsensical values for S^c_{cov} and E^c_{cov}.

For CASVB2 and 4 it is interesting that both the overlaps and energies are actually inferior to the best SC^{trunc} values. It therefore seems that decreasing the weight of the perfect-pairing function slightly, has a positive effect on these quantities, but it is likely that these two criteria don't have the flexibility to do so for reasonable orbital sets.

In conclusion it is difficult to disregard the success of CASVB1 and 3 in obtaining very good values for the covalent overlaps and energies. The orbitals do not, however, conform to the standard valence bond picture of the bonding,

	$\langle\phi_{C1}\|\phi_{H1}\rangle$	$\langle\phi_{C1}\|\phi_{C2}\rangle$	$\langle\phi_{C1}\|\phi_{H2}\rangle$	$\langle\phi_{H1}\|\phi_{H2}\rangle$
SC	0.69189	0.52904	0.09591	−0.15228
SC^{trunc}	0.69949	0.53129	0.09569	−0.15211
CASVB1	0.20131	0.91557	−0.06164	−0.32211
CASVB2	0.78266	0.32257	0.05914	−0.10900
CASVB3	0.27192	0.88261	−0.03813	−0.31588
CASVB4	0.75235	0.37043	0.05805	−0.13052

Table 4.13: Unique overlap integrals for the various orbitals of methane.

and we therefore regard them as 'anomalous' solutions probably caused, at least in part, by the small number of free orbital parameters. The 'full' spin-coupled energy surface was investigated in the regions around both converged solutions, but no evidence of local minima, or even stationary points, were found. This type of solution, it seems, simply does not exist for the standard spin-coupled calculations. It is clear that basing the spin coupling on the CAS wavefunction, as in CASVB2 and 4, goes a long way towards alleviating this problem by restricting the spin-coupled-like wavefunctions to physically more realistic parts of the energy surface.

	$(b_1^s)^2$	$(b_2^s)^2$	$(b_3^s)^2$	$(b_4^s)^2$
SC^{trunc}	0.19%	0.23%	1.55%	90.29%
CASVB1	0.83%	1.04%	6.72%	57.82%
CASVB2	0.00%	0.00%	0.84%	94.95%
CASVB3	0.64%	0.80%	5.66%	64.58%
CASVB4	0.00%	0.00%	0.82%	95.08%
	$(b_1^e)^2$	$(b_2^e)^2$	$(b_3^e)^2$	$(b_4^e)^2$
SC	0.19%	0.24%	1.59%	90.00%
SC^{trunc}	0.13%	0.16%	1.56%	90.36%
CASVB1	0.79%	0.98%	6.64%	58.37%
CASVB2	0.00%	0.01%	0.89%	94.61%
CASVB3	0.60%	0.75%	5.63%	64.85%
CASVB4	0.00%	0.01%	0.87%	94.77%
	$(b_1^c)^2$	$(b_2^c)^2$	$(b_3^c)^2$	$(b_4^c)^2$
SC	0.01%	0.01%	0.48%	97.08%
SC^{trunc}	0.00%	0.00%	0.31%	98.13%
CASVB1	0.13%	0.16%	1.09%	93.17%
CASVB2	0.01%	0.01%	0.53%	96.79%
CASVB3	0.07%	0.09%	0.58%	96.38%
CASVB4	0.01%	0.01%	0.50%	96.99%

Table 4.14: Weights of the Serber functions for methane.

	S^s_{cov}	S^e_{cov}	S^c_{cov}
$\mathrm{SC}^{\mathrm{trunc}}$	0.99529	0.99529	0.99457
CASVB1	0.99613	0.99612	0.98279
CASVB2	0.99522	0.99521	0.99517
CASVB3	0.99605	0.99605	0.98478
CASVB4	0.99515	0.99515	0.99509
	E^s_{cov}/hartree	E^e_{cov}/hartree	E^c_{cov}/hartree
$\mathrm{SC}^{\mathrm{trunc}}$	−40.273644	−40.273646	−40.271765
CASVB1	−40.275304	−40.275324	−40.236634
CASVB2	−40.273082	−40.273085	−40.272922
CASVB3	−40.275602	−40.275607	−40.243954
CASVB4	−40.273424	−40.273427	−40.273219

Table 4.15: Covalent overlaps and energies for the different orbital sets for methane.

4.4 Ozone

In the calculations on ozone the bond length was chosen as r(O–O)= 1.2717 Å, and the angle as $\angle$(O–O–O)=121.6085°. For the three oxygens we added to a Huzinaga basis set [11] of (10s6p) Cartesian gaussians contracted to [5s3p] a polarisation function in the form of a d-function with exponent ζ=0.85.

We considered here treating only the four π-electrons of ozone as active. There were thus ten σ orbitals as core (6 of A_1 symmetry and 4 of B_2), leaving 35 and 11 virtuals of σ and π symmetry respectively. It was immediately clear that such a treatment leads to two CAS solutions very similar in energy, corresponding to active MOs of symmetries $1b_1$, $2b_1$, $3b_1$, $1a_2$ (solution A) or $1b_1$, $2b_1$, $1a_2$, $2a_2$ (solution B). The MOs and occupation numbers for the CAS calculations with optimised core are shown in figures 4.15 and 4.16. Both solutions indicate the inadequacy of the SCF description in the marked differences in the occupation numbers from 2 for the lowest-lying MOs and from 0 for the highest-lying, and this is also reflected in the relatively large amount of correlation energy obtained with only four active electrons. With the core orbitals taken from the SCF solution the two CAS wavefunctions are almost degenerate—only 22 microhartree being in favour of solution A (see table 4.16). Optimising the core accentuates this difference slightly, giving an overall difference of 2.2 or 2.5 millihartree between the two solutions with the two cores respectively, again in favour of type A. The fact that there are near-degenerate solutions in this way, suggests that the '4 in 4' CAS may not lead to a satisfactory description of the bonding for this system. This will be discussed further in the following.

The traditional spin-coupled description using four active electrons, with the core taken from the SCF, is shown in figure 4.17. The central oxygen here forms two π-bonds in the same plane, and in this sense behaves 'hypervalent'. The perfect-paring mode of spin coupling dominates with a weight of 95.49%.

When we examined the lowest-lying CAS solution (A), a very different picture emerged. We have shown the orbitals obtained by applying the four different

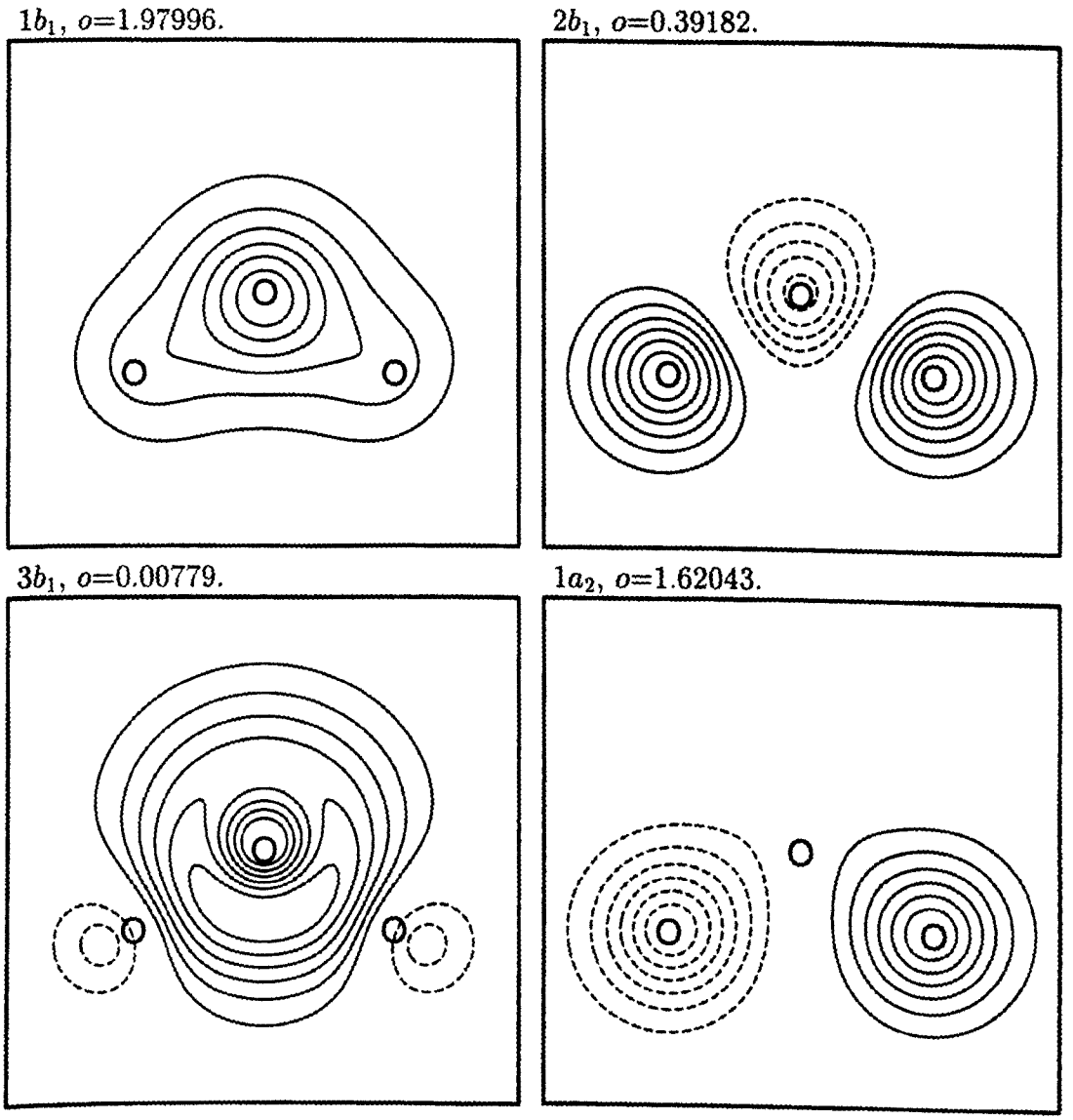

Figure 4.15: MOs of ozone obtained from CAS A.

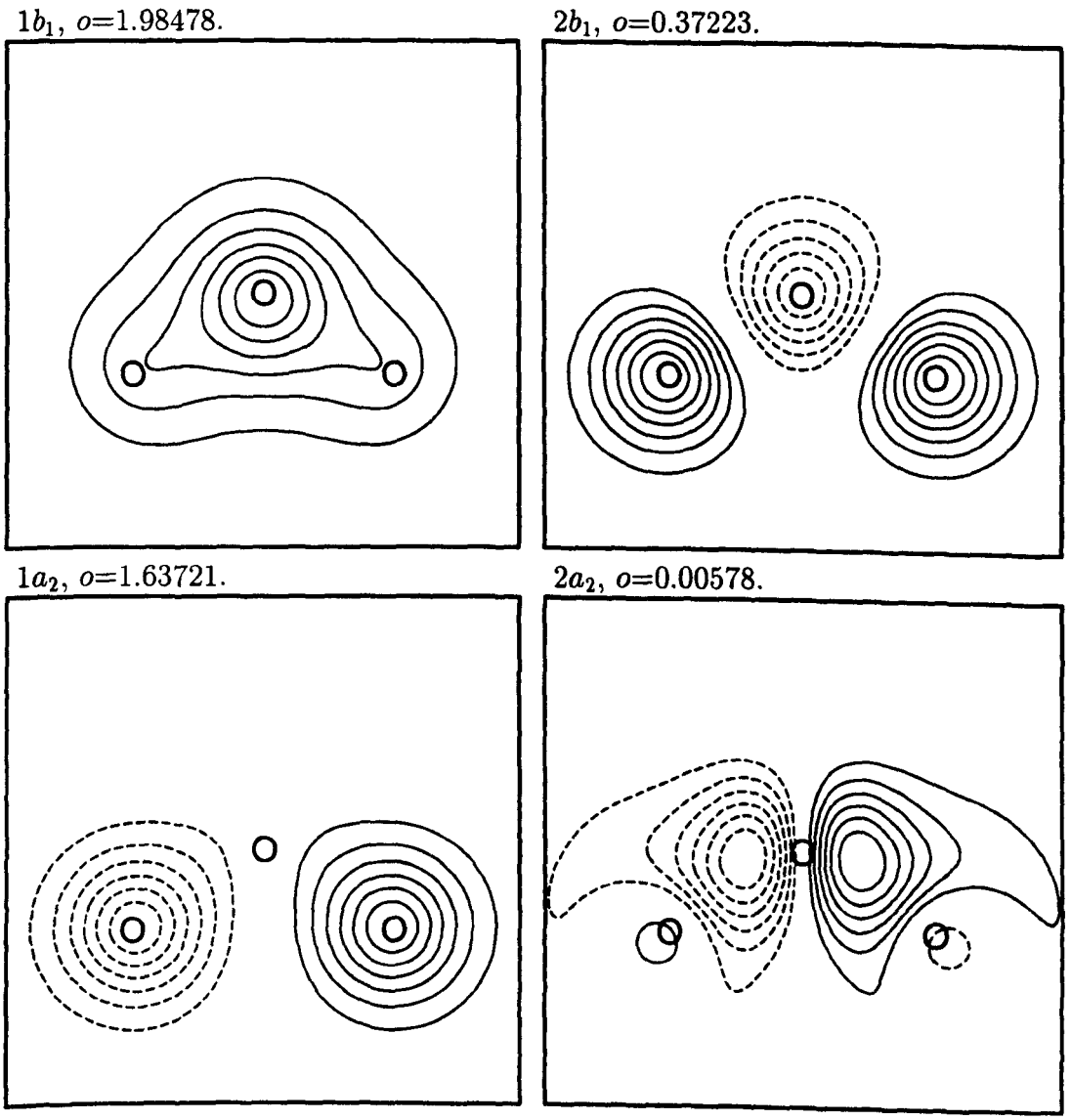

Figure 4.16: MOs of ozone obtained from CAS B.

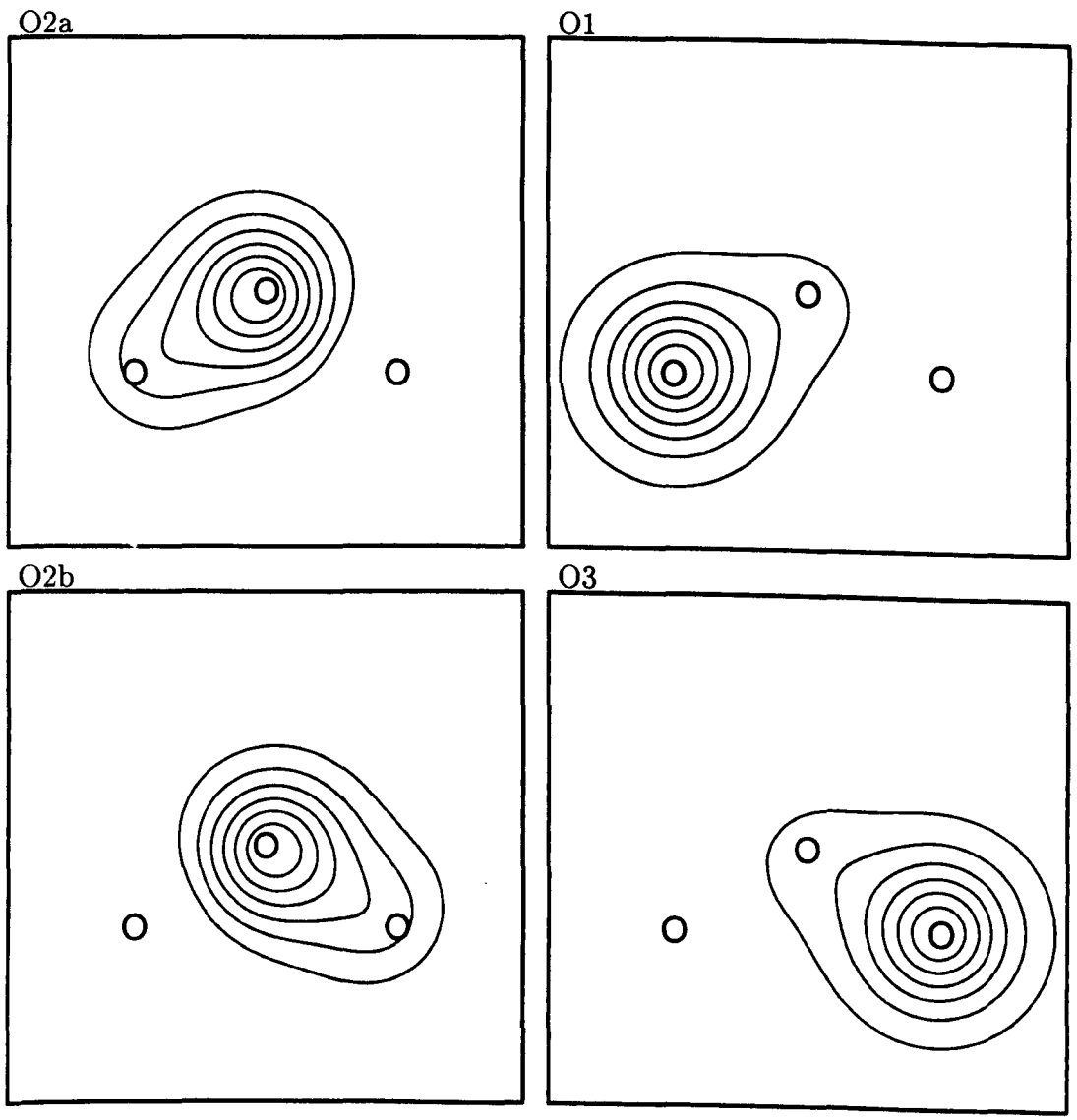

Figure 4.17: Spin-coupled orbitals of ozone (SCF core).

	E/hartree	$(E_{SCF}-E)$/millihartree
SCF	−224.319107	0.00
SC A; SCF core	−224.417478	98.37
SC B; SCF core	−224.417579	98.47
CAS B; SCF core	−224.417910	98.80
CAS A; SCF core	−224.417932	98.83
SC B; A core	−224.424705	105.60
SC B; B core	−224.424833	105.73
CAS B; opt. core	−224.425112	106.01
SC A; B core	−224.427020	107.91
SC A; A core	−224.427173	108.07
CAS A; opt. core	−224.427541	108.43

Table 4.16: Ozone energies for some of the wavefunctions considered.

criteria in figures 4.18 and 4.19. Since only the O1 and O3 orbitals contain components of the A_2 irreducible representation, the corresponding electrons must be perfectly singlet-coupled. Since there are thus no free parameters associated with the spin-space there is therefore no difference between CASVB1 and 2, nor between CASVB3 and 4. The description of the bond-formation between the central and terminal oxygens is here rather different from the standard valence bond description in terms of orbital pairs with associated singlet-coupled electrons. Most of the energy-lowering is likely to be associated with the large delocalisation of the O2b orbital, and in this way the picture is very MO-like. Since the interaction between the terminal oxygens is likely to be quite modest, this solution is often referred to as being of 'diradical' type.

Examining the 'full' spin-coupled energy surface again, the solution shown in figure 4.20 was found. This solution is in excellent agreement with the CAS interpretation as signified by the orbital plots and the overlap matrices given in table 4.17 below. This spin-coupled wavefunction produced 99.67% of the

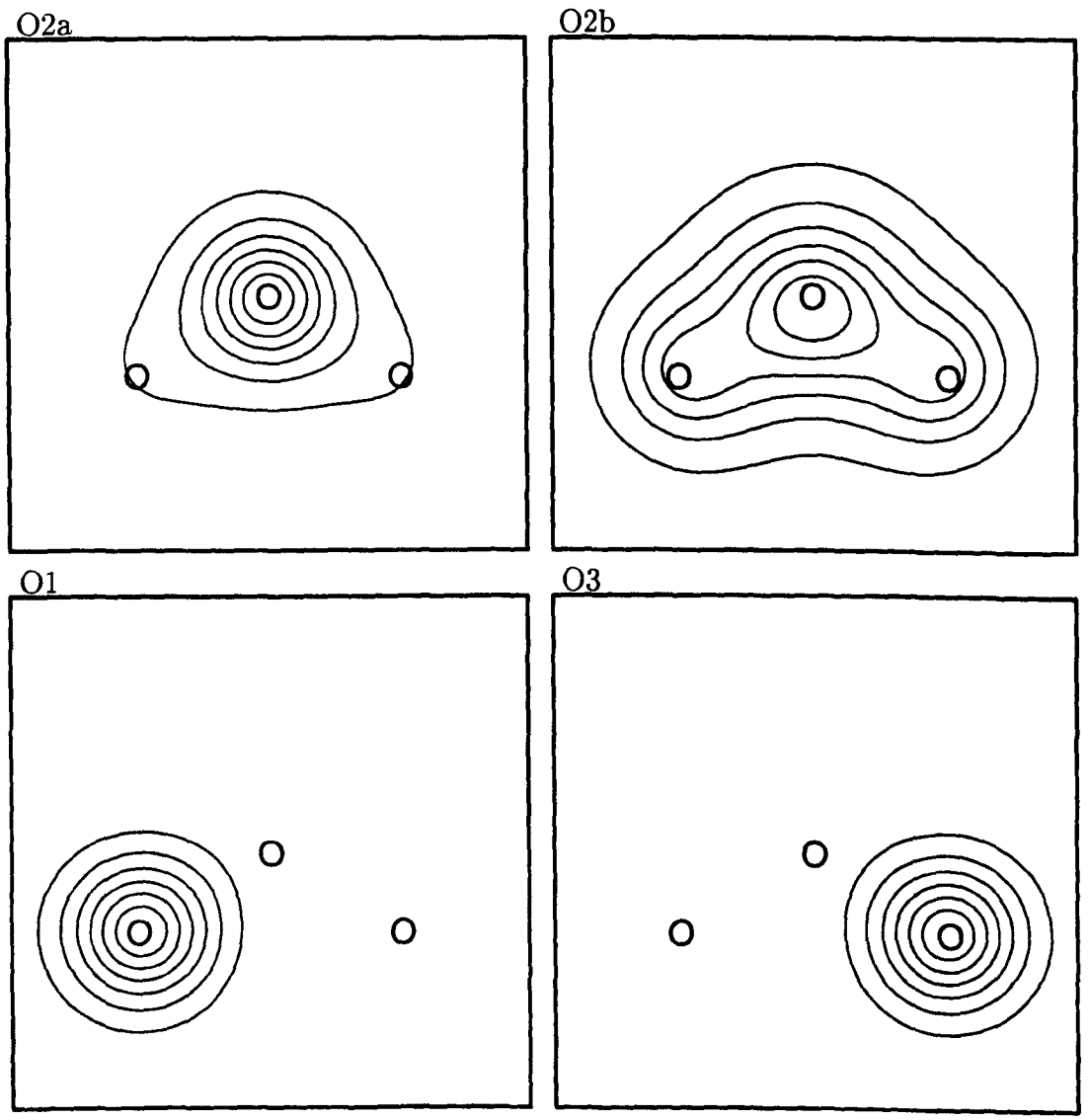

Figure 4.18: CASVB1 (or 2) interpretation of CAS A (ozone).

	⟨O2a|O2b⟩	⟨O1|O2a⟩	⟨O1|O2b⟩	⟨O1|O3⟩
SC	0.80625	0.19919	0.47042	-0.16651
SC^{trunc}	0.80628	0.19919	0.47043	-0.16651
CASVB1 or 2	0.80519	0.19666	0.47416	-0.16639
CASVB3 or 4	0.80767	0.19894	0.46958	-0.16714

Table 4.17: Unique overlap integrals for ozone based on CAS A.

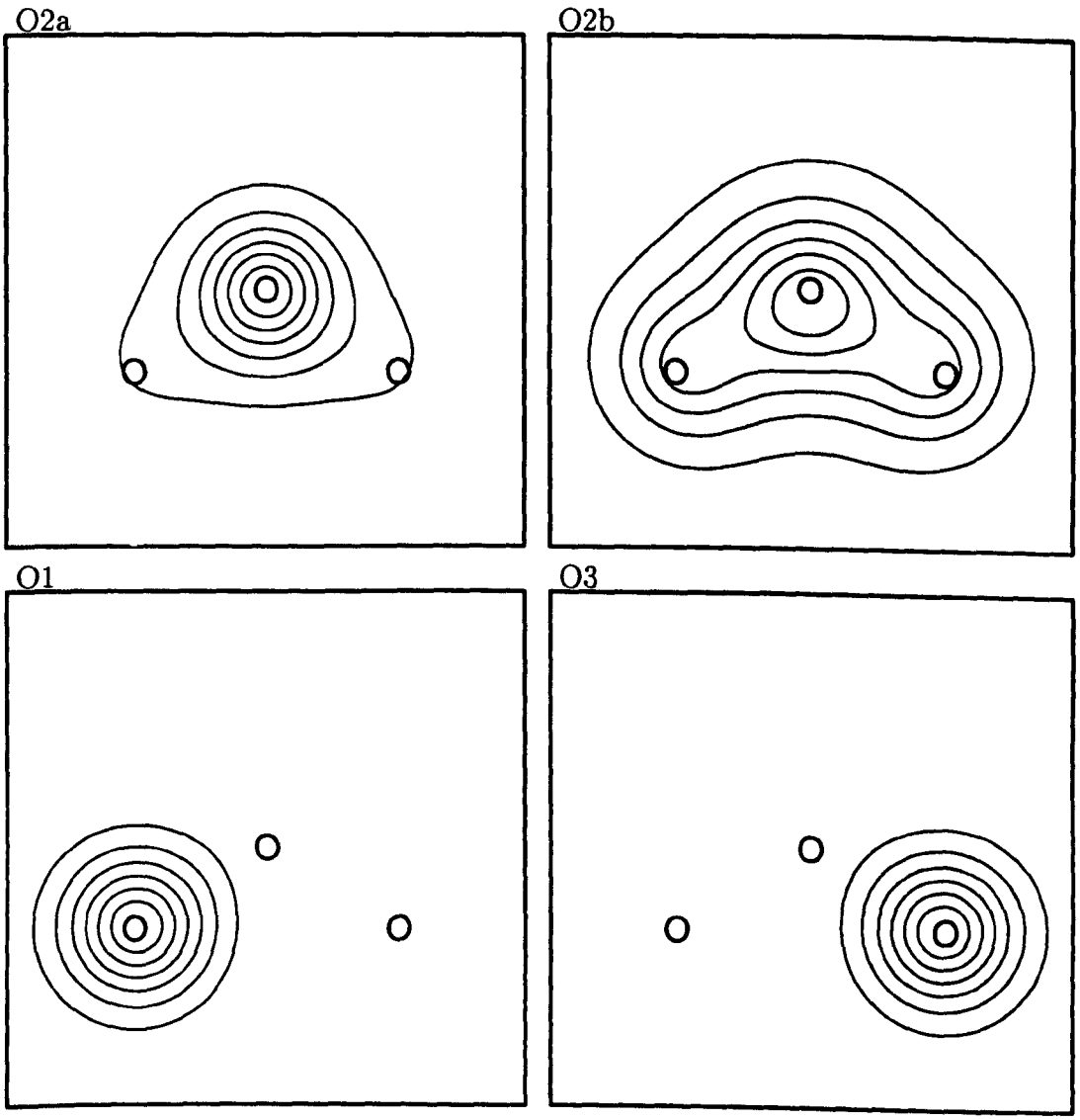

Figure 4.19: CASVB3 (or 4) interpretation of CAS A (ozone).

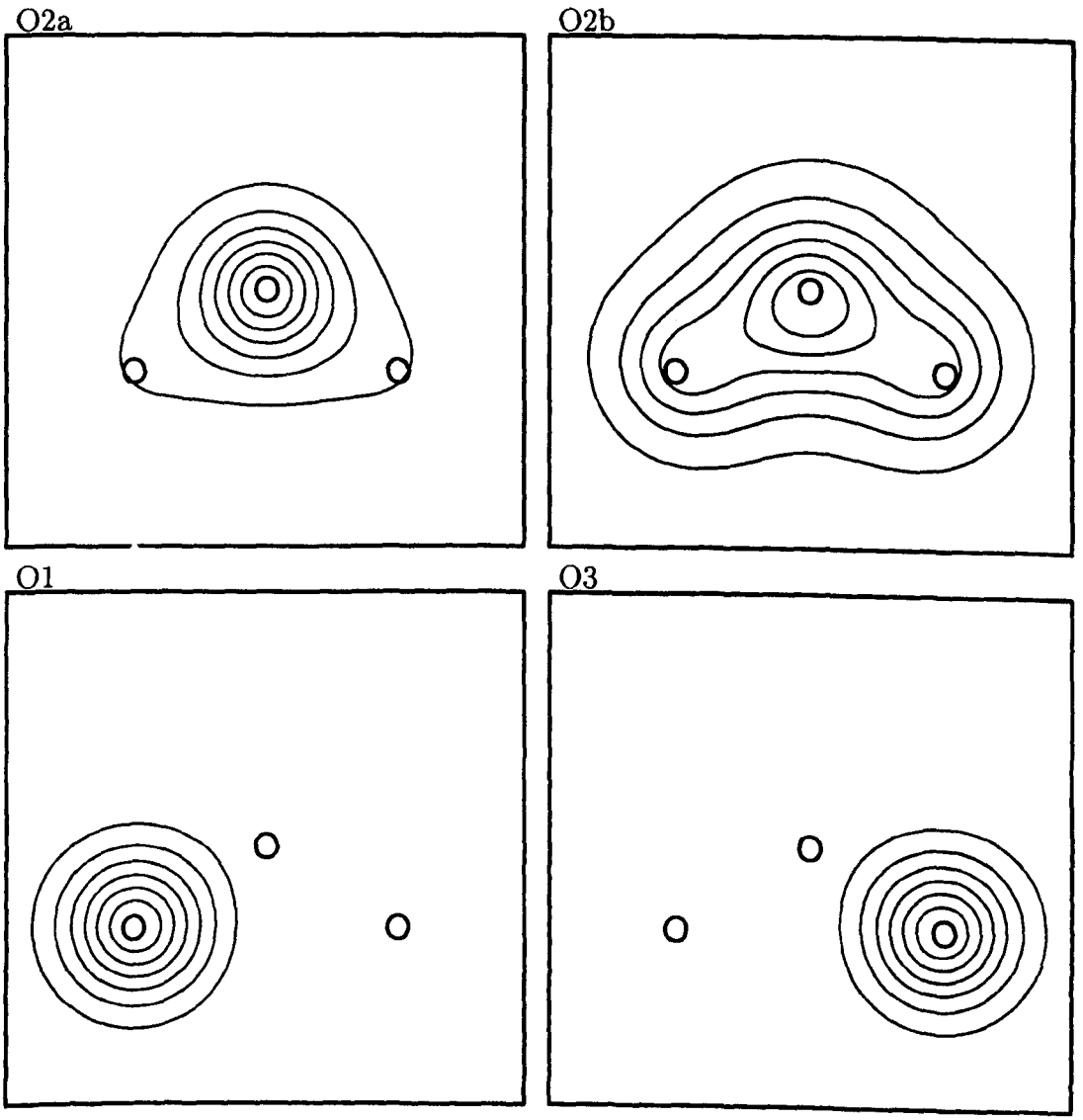

Figure 4.20: Spin-coupled orbitals of ozone (solution A).

	S_{cov}	E_{cov}/hartree
SC^{trunc}	0.9999051	−224.427160
CASVB1 or 2	0.9999145	−224.427139
CASVB3 or 4	0.9999051	−224.427162

Table 4.18: Covalent overlaps and energies for the interpretation of CAS A.

correlation energy obtained by the CAS wavefunction. Only 0.0017%, for both O2a and O2b, and 0.000022% for the terminal orbitals, was outside the CAS space.

Values for S_{cov} and E_{cov} for the various orbital sets are given in table 4.18 (again due to the lack of free spin-parameters there is only one set of values in each case). The values for S_{cov} are probably of a quality that is to be expected when there are only four active electrons (considering the small difference in the number of free parameters, cf. table 2.1). The values for E_{cov} compare well with the full spin-coupled energy of E_{SC}=−224.427173 hartree.

When the second CAS was examined, the spin-coupled interpretations were found to correspond very well to the already well known 'hypervalent' solution (see figures 4.21 and 4.22). So there are in this way not only two very close CAS solutions, but to each of these a corresponding spin-coupled solutions exists. This correspondence is unique since the projection of the spin-coupled solution A on CAS B (or vice versa) would give linearly dependent orbitals. In other words: the distribution of CAS MOs among the irreducible representations *alone* determines the types of spin-coupled-like solutions that are possible. The different energies of the two solutions with various choices of core are shown in table 4.16. The spin-coupled solution B recovered 99.74% of the correlation energy produced by CAS B. For this set of spin-coupled orbitals only 0.00022% and 0.00018%, for the central and terminal orbitals respectively, was outside the CAS space. The overlap matrices, spin-coupling coefficients and covalent overlaps and energies based on CAS B are given in tables 4.19–4.21. The weight of the perfect-pairing func-

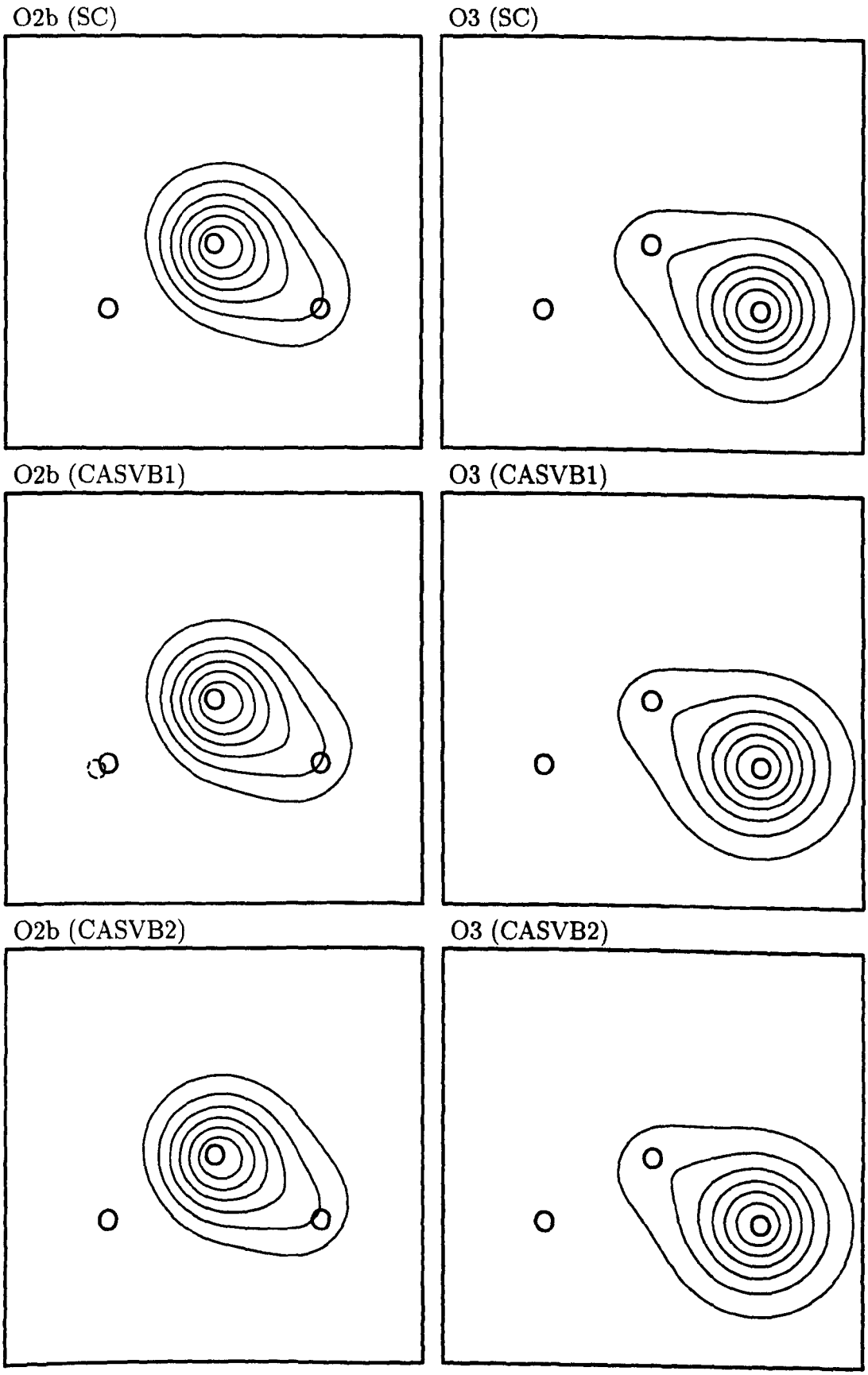

Figure 4.21: Spin-coupled solution B with the CASVB1 and 2 interpretations of CAS B.

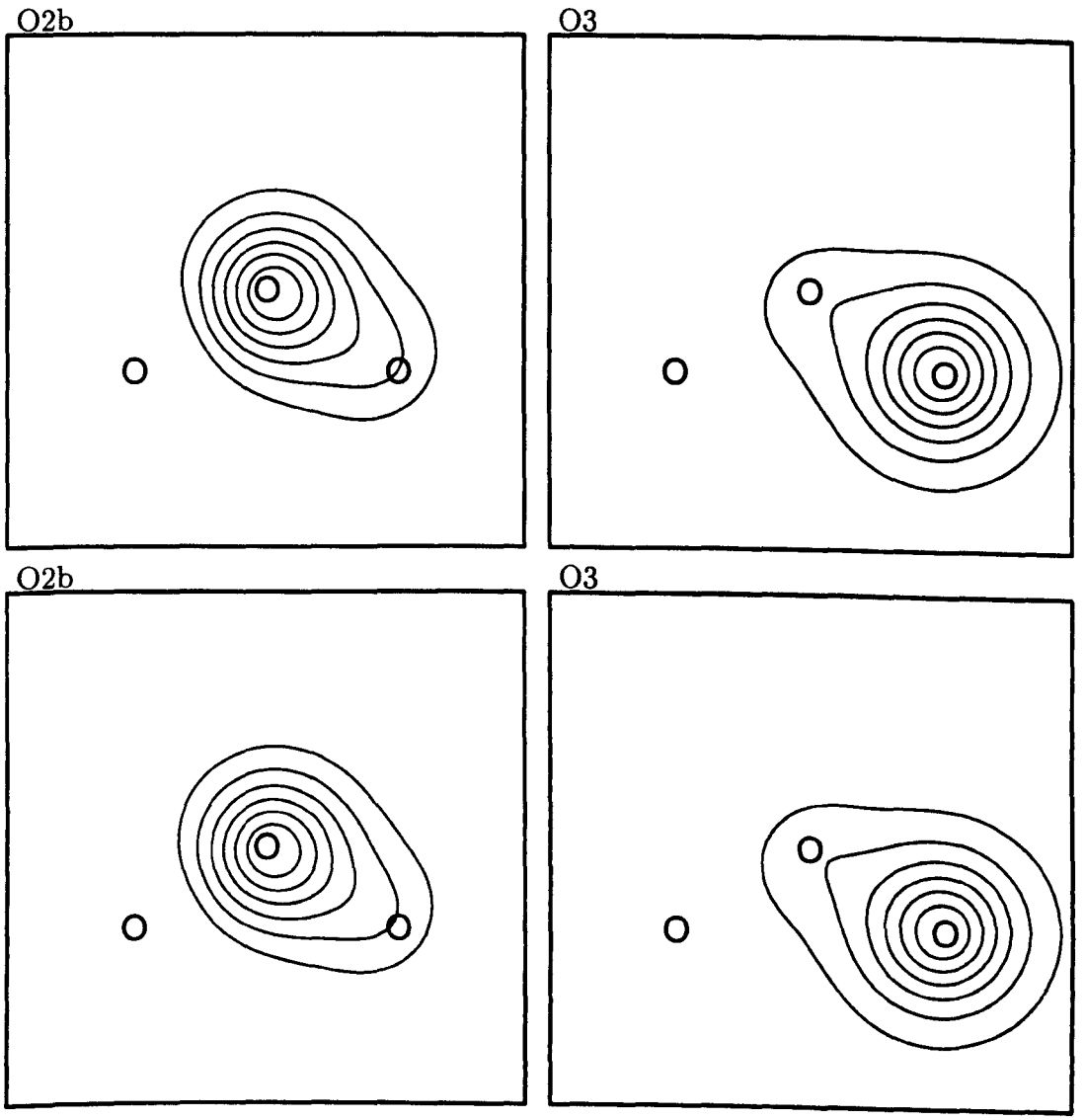

Figure 4.22: CASVB3 and 4 interpretations of CAS B.

tion is slightly bigger when the spin coupling is taken from the CAS wavefunction, but between the overlap- and energy-optimised spin couplings the differences are minute. The values of $S_{\rm cov}$ and $E_{\rm cov}$ are all reasonable, and similar to those obtained from CAS A, except in those cases where the CAS spin coupling has been *imposed* (i.e., $S^{\rm c}_{\rm cov}$ and $E^{\rm c}_{\rm cov}$ for the SC$^{\rm trunc}$, CASVB1 and CASVB3 orbital sets). For CASVB2 and 4 it seems the orbitals can easily adapt to the high weight of the perfect-pairing mode associated with the CAS spin coupling, but for other orbital sets this is not so appropriate.

As a final argument regarding the close correspondence between the two CAS and spin-coupled solutions we have in figures 4.23 and 4.24 shown the natural orbitals of the two competing spin-coupled solutions. It is clear that the form of the orbitals, as well as the values for their occupation numbers, are very similar between the corresponding CAS and SC cases. It is worth noting that the *form* of the spin-coupled solution is sufficient to determine the distribution of the natural orbitals between the irreducible representations, and this may be utilised to investigate various alternative bonding-pictures. In the present case, for example, it seems that the convergence characteristics of the standard spin-coupled methods are such that it is unlikely that the diradical solution will be found. But if such a solution was suspected it would then be trivial to construct the corresponding CAS, and hence the corresponding spin-coupled-like, and possibly full spin-coupled, solutions.

Near-degeneracy, as encountered in this case, is an indication that two differing descriptions may play a role in a realistic description of the molecular system. The natural thing to do, is thus to base an improved calculation on the corresponding two configurations—in this case SC A and SC B. This was done in a simple CI step which gave an energy of $E_{\rm 2cnf}$=−224.431396 hartree (the lowest-energy A-core was used in this case). This represents a lowering of 4.2 millihartree from SC A which is quite significant compared with the spacing between the two spin-coupled solutions (2.5 millihartree). As expected the two configurations

	⟨O2a|O2b⟩	⟨O1|O2a⟩	⟨O1|O2b⟩	⟨O1|O3⟩
SC	0.72540	0.57259	0.10621	0.13718
SC^{trunc}	0.71729	0.57165	0.09675	0.13358
CASVB1	0.70673	0.56751	0.07920	0.12001
CASVB2	0.80097	0.55848	0.16013	0.12406
CASVB3	0.71729	0.57165	0.09675	0.13357
CASVB4	0.79560	0.57062	0.17368	0.15033

Table 4.19: Unique overlap integrals for ozone based on CAS B.

	${}^K\Theta_1^s$	${}^K\Theta_2^s$	${}^K\Theta_1^e$	${}^K\Theta_2^e$	${}^K\Theta_1^c$	${}^K\Theta_2^c$
SC^{trunc}	5.52%	94.48%	5.62%	94.38%	2.61%	97.39%
CASVB1	6.28%	93.72%	6.38%	93.62%	5.09%	94.91%
CASVB2	3.23%	96.77%	3.33%	96.67%	3.22%	96.78%
CASVB3	5.80%	94.20%	5.90%	94.10%	2.52%	97.48%
CASVB4	3.19%	96.81%	3.29%	96.71%	3.23%	96.77%

Table 4.20: Weights of the Kotani functions for CAS B interpretations of ozone.

contribute comparable amounts to the two-configuration solution, the diradical dominating only slightly with 56.3% (Chirgwin-Coulson).

One may of course also treat the two-configuration wavefunction in the context of orbital optimisation, and work along these lines is currently ongoing. Judging from the form of the natural orbitals, the '4 in 6' CAS suggests itself as a sensible calculation.

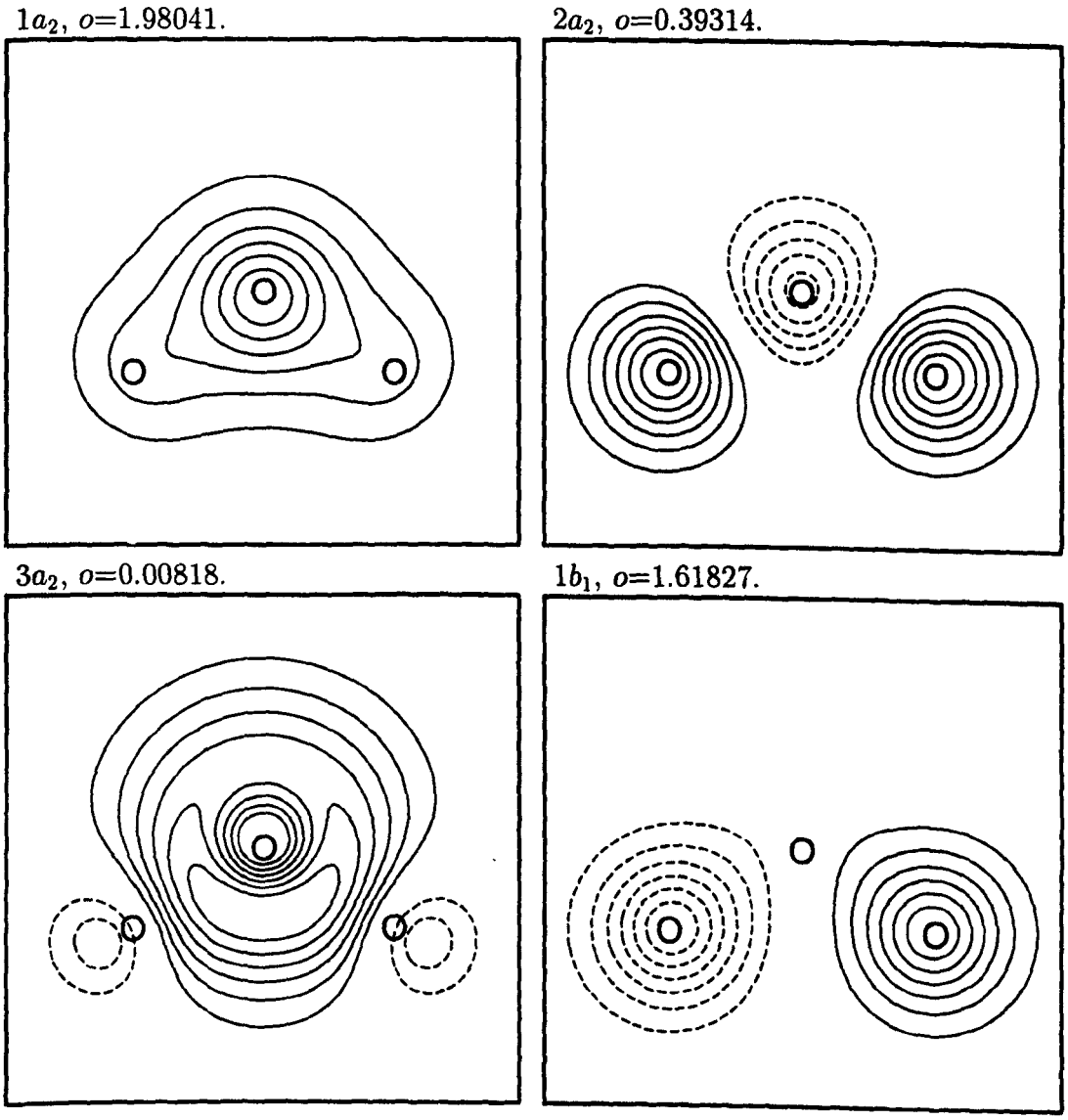

Figure 4.23: Natural orbitals for the 'SC-A' solution.

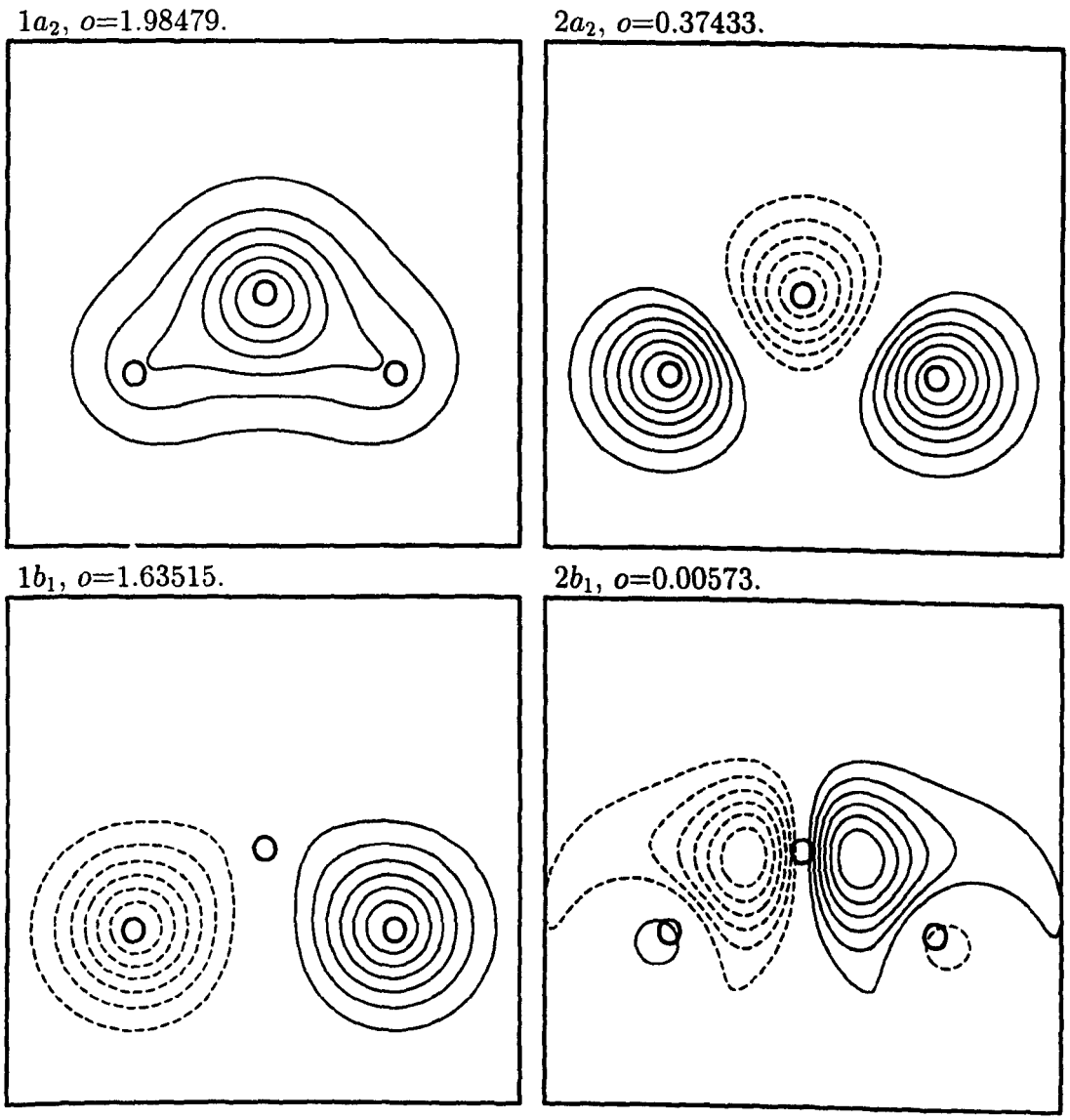

Figure 4.24: Natural orbitals for the 'SC-B' solution.

	S^s_{cov}	S^e_{cov}	S^c_{cov}
$\mathrm{SC}^{\mathrm{trunc}}$	0.9999447	0.9999438	0.9989869
CASVB1	0.9999458	0.9999450	0.9982762
CASVB2	0.9999398	0.9999385	0.9999397
CASVB3	0.9999452	0.9999443	0.9987433
CASVB4	0.9999386	0.9999373	0.9999384
	E^s_{cov}/hartree	E^e_{cov}/hartree	E^c_{cov}/hartree
$\mathrm{SC}^{\mathrm{trunc}}$	−224.424826	−224.424826	−224.424005
CASVB1	−224.424825	−224.424826	−224.423396
CASVB2	−224.424804	−224.424805	−224.424804
CASVB3	−224.424826	−224.424827	−224.423795
CASVB4	−224.424806	−224.424807	−224.424807

Table 4.21: Covalent overlaps and energies for the different orbital sets for CAS B of ozone.

MO	Occupation	MO	Occupation
$2a_g$	1.98694	$2b_{1u}$	1.98276
$3a_g$	1.98018	$3b_{1u}$	0.02294
$4a_g$	0.01849	$4b_{1u}$	0.01322
$1b_{2g}$	0.02238	$1b_{2u}$	1.98035
$1b_{3g}$	1.97852	$2b_{2u}$	0.01978
$2b_{3g}$	0.01648	$1b_{3u}$	1.97796

Table 4.22: Active MOs and occupation numbers for B_2H_6.

4.5 Diborane

In diborane the terminal hydrogens and the boron atoms are all coplanar, whereas a perpendicular plane is defined by the boron atoms and the bridging hydrogens. The precise geometry was chosen as r(B–B)=1.7665 Å, r(B–H^t)= 1.2000 Å, r(B–H^b)= 1.3260 Å and $\angle$(B–B–H^t)=121° [12]. For B/H we employed correlation consistent pVDZ basis sets by Dunning [5], consisting of (9s4p1d/4s1p) Cartesian gaussians contracted to [3s2p1d/2s1p], as stored internally in MOLPRO.

Keeping the $1s^2$ on both boron atoms as core, there were 12 active electrons and orbitals, and 46 virtuals. The list of MOs and occupation numbers for the lowest-lying CAS wavefunction ('CAS A') is given in table 4.22. Of particular note is the distribution of the active MOs between the irreducible representations, as this will determine the nature of the spin-coupled-like representations that are possible. Judging from the CAS, the SCF description seems quite reasonable; the occupation numbers for MOs of symmetries corresponding to the occupied SCF MOs are close to two, the corresponding configuration dominates the CAS wavefunction by 94.47%, and the correlation energy obtained was relatively modest when taking into account that twelve electrons were correlated—90.32 millihartree (see table 4.23).

We have interpreted the CAS wavefunction described above, using CASVB1,

	E/hartree	$(E_{SCF}-E)$/millihartree
SCF	−52.815220	0.00
SC	−52.893038	77.82
CAS B	−52.898012	82.79
CAS A	−52.905543	90.32

Table 4.23: Energies of some of the wavefunctions considered for diborane.

2 and 3. The CASVB3 solution was obtained with our standard spin-coupled codes, but employing CASVB4 was not possible, since the present version of our codes requires storage of the complete Hamiltonian. All three criteria converged onto symmetry-broken solutions as was also seen for benzene. The orbitals are shown in figures 4.25–4.27. As for benzene partial symmetry is retained, here in the form of a vertical mirror plane perpendicular to the bridging plane. The question for diborane is whether this symmetry-breaking can be considered 'accidental' (as we argued was the case for benzene) or a consequence of the inadequacy of a single-configuration description. By 'accidental' we mean symmetry-breaking brought on by the constraints imposed on the spin and/or orbital parameters. It is clear that a symmetry-pure two-configuration description can easily be constructed from solutions of the forms seen in figures 4.25–4.27, so that the symmetry-breaking could be viewed as a consequence of not having a two-configuration description (cf. the discussion in section 1.2). Various evidence suggests that the latter interpretation is more reasonable: the deformation of the orbitals is clearly larger than was the case for benzene, the differences in the covalent overlaps and energies between symmetric and symmetry-broken solutions correspondingly larger (see below), and there is also the fact that even CASVB3 leads to symmetry-breaking for B_2H_6. We have nevertheless included also the symmetry-pure orbital sets for CASVB1 and 2 in figure 4.28 (imposing the necessary symmetry constraints was not possible with the present version of the spin-coupled code, so the symmetric CASVB3 solution has not been included).

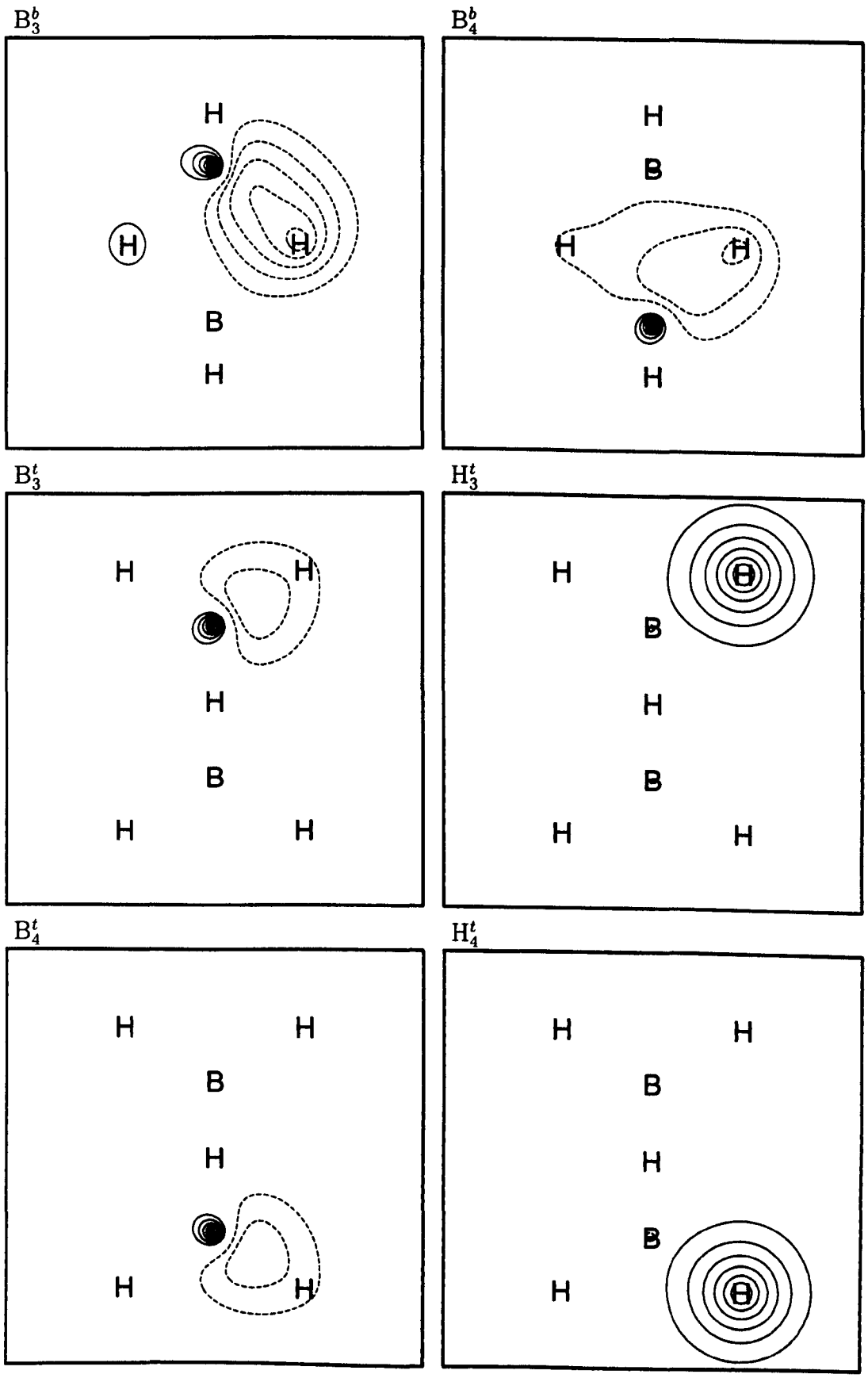

Figure 4.25: CASVB1 orbitals of B_2H_6.

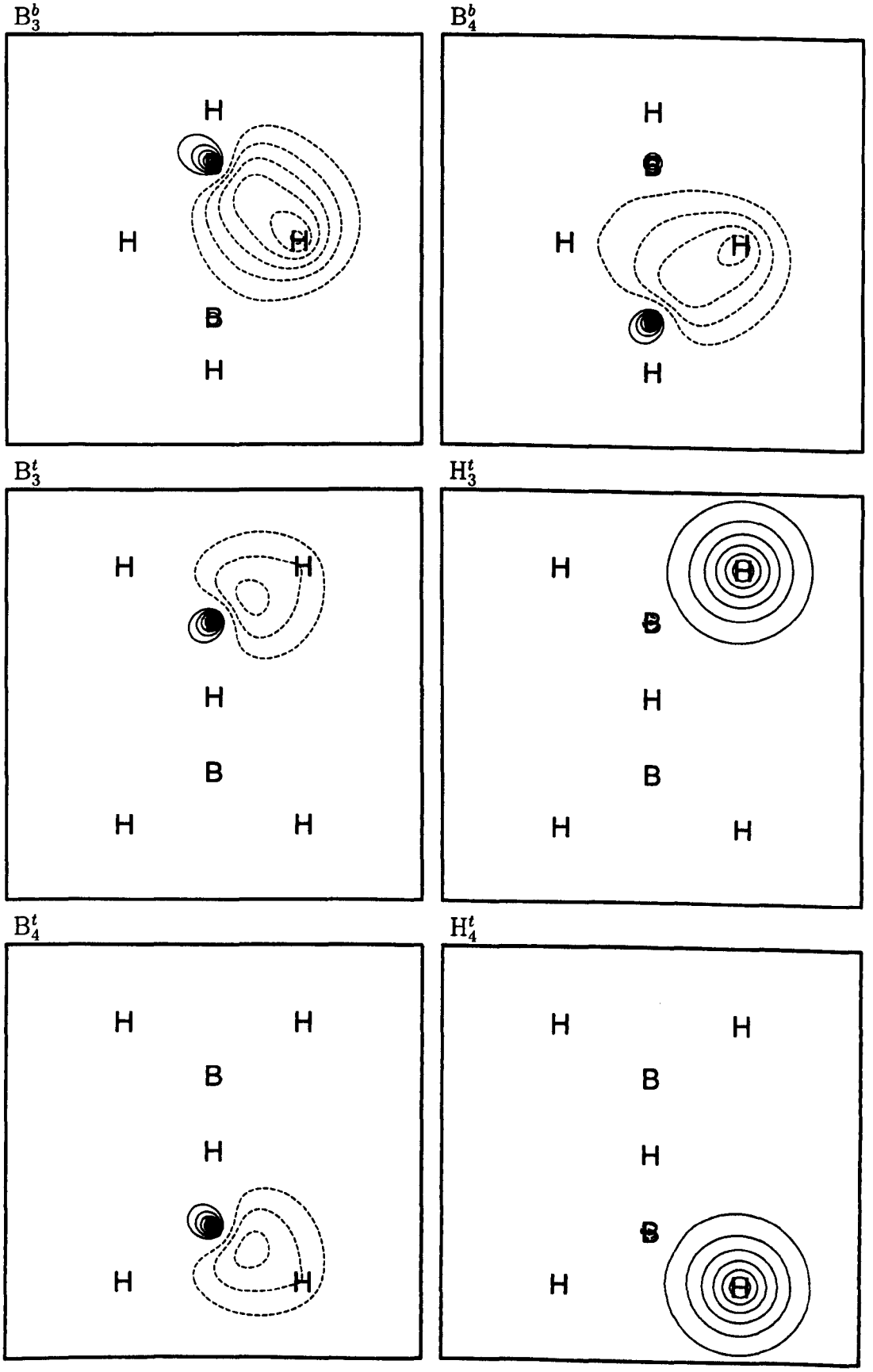

Figure 4.26: CASVB2 orbitals of B_2H_6.

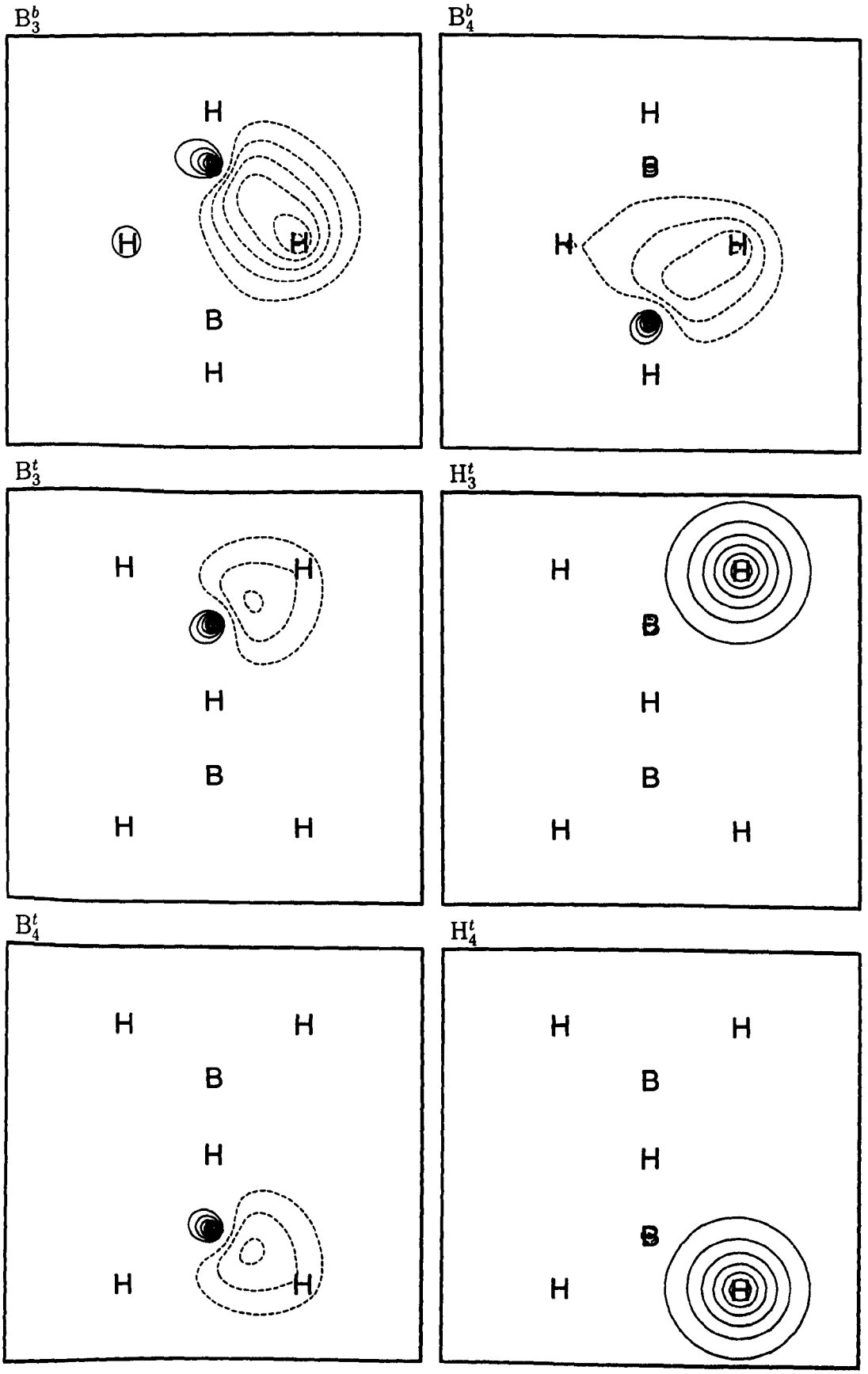

Figure 4.27: CASVB3 orbitals of B_2H_6.

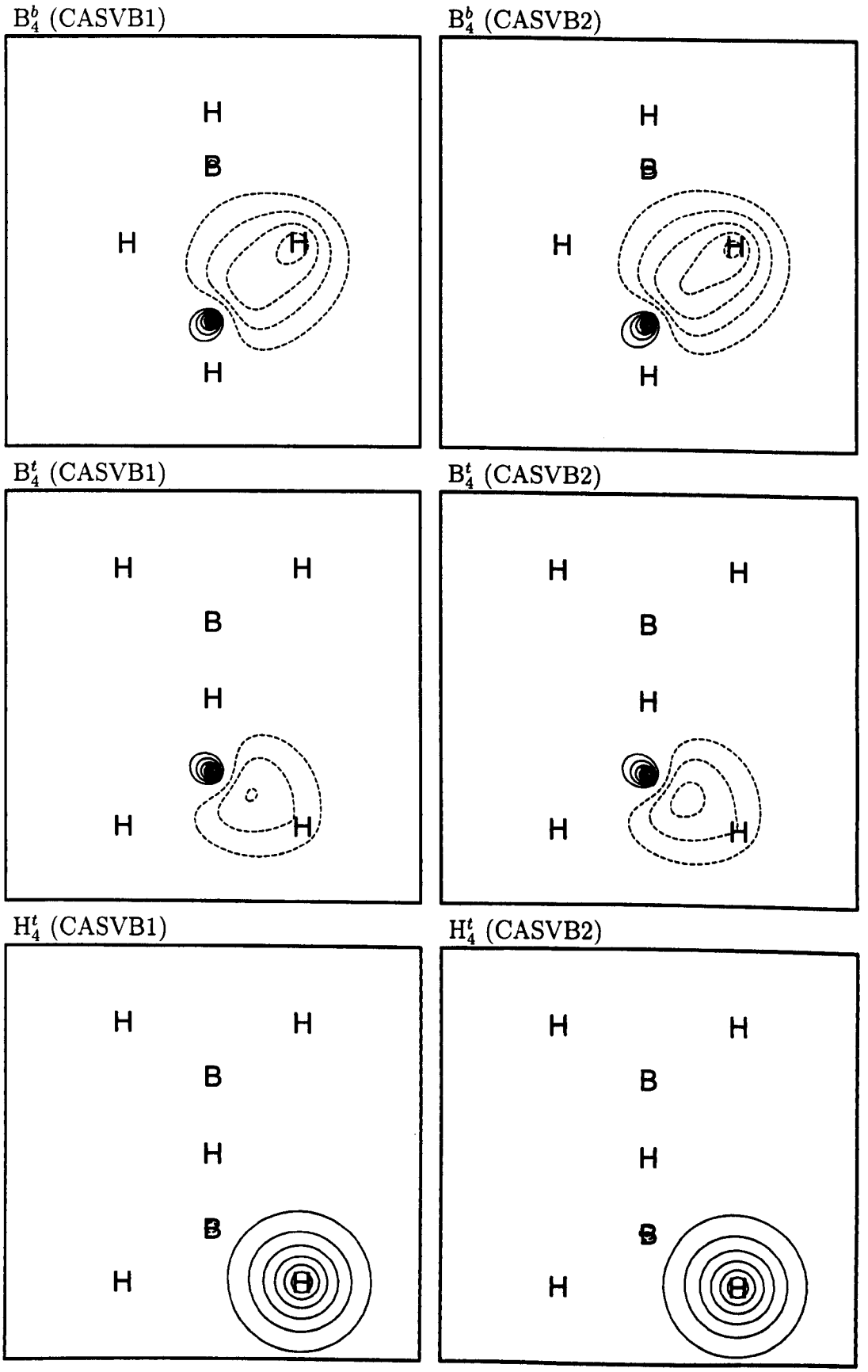

Figure 4.28: Unique symmetry-pure orbitals of B_2H_6 (CASVB1 and 2).

In all the cases the terminal B–H bonds are described very much as would be expected from a standard spin-coupled calculation, by sp^x-type hybrids on B pointing towards what are essentially 1s orbitals on H. The corresponding electron spins are to a good approximation singlet-coupled (see below). The description of the bridging three-centre two-electron bonds, however, is at odds with the more commonly accepted valence bond and LMO descriptions. These include SOPP-GVB calculations [13], four-electron spin-coupled treating the bridging region only [14], and twelve-electron spin-coupled calculations in a two-spin-function approximation [15]. The bonding in the bridging region is, in the interpretations of the CAS wavefunction, described by four orbitals, all being admixtures of sp^x-type hybrids on boron and s-type functions on hydrogen. For the symmetry-pure solutions all orbitals are of course equivalent, whereas for the other orbital sets they are only pair-wise equivalent. The orbitals on the upper B atom in the figures here show very great H character, the orbitals on the lower B atom show quite little H character. The electrons in orbitals pointing towards the same hydrogen are near-singlet-coupled (see below).

The orbitals for CASVB2 show greater deformation in the bond-forming directions, and this is reflected in greater overlaps for the bond-forming orbital pairs (see tables 4.24 and 4.25—the orbitals are ordered: upper left=1, lower left=2, upper right=3, lower right=4). This is likely to be associated with the higher weight of the perfect pairing function for CASVB2 (see table 4.26). CASVB3 seems to be intermediate between the two overlap-based criteria with respect to both the amount of deformation of bond-forming orbitals, as well as the weight of the perfect-pairing mode of spin coupling.

The values for S_{cov} and E_{cov} are given in table 4.27. As for the other systems studied so far, all values are quite reasonable except for the cases where the CAS spin coupling is *imposed*. The difference between the overlap and energy based optimisations seems slightly greater than has been the case for the other systems we have examined. CASVB3 is lower than CASVB1 by about 0.4 millihartree,

	$\langle B_1^b \vert B_2^b \rangle$	$\langle B_1^b \vert B_3^b \rangle$	$\langle B_1^b \vert B_4^b \rangle$	$\langle B_2^b \vert B_4^b \rangle$
CASVB1	0.49282	-0.36283	0.14605	0.83338
CASVB2	0.77989	-0.10135	0.26491	0.57243
CASVB3	0.65508	-0.34432	0.15566	0.68008
CASVB1 (s)	0.77907	0.25364	0.18885	0.25364
CASVB2 (s)	0.83506	0.21135	0.20927	0.21135

Table 4.24: Overlaps for the bridging region of B_2H_6.

	$\langle B_1^t \vert H_1^t \rangle$	$\langle B_2^b \vert H_2^t \rangle$
CASVB1	0.70113	0.70739
CASVB2	0.80969	0.81182
CASVB3	0.78539	0.80069
CASVB1 (s)	0.76968	0.76968
CASVB2 (s)	0.80959	0.80959

Table 4.25: Overlaps between bond-forming terminal orbitals for B_2H_6.

whereas CASVB1 is superior in the covalent overlap by about 2×10^{-4}. These differences may not be large, but the consequence is that diborane does not show the excellent agreement between overlap- and energy-based optimisations seen for the other systems. Also the effect of imposing symmetry constraints is larger for this system. S_{cov} decreased by 2×10^{-4} for CASVB1 and by 6×10^{-5} for CASVB2, and E_{cov} increased by 0.21 millihartree and 0.05 millihartree respectively when these constraints were applied. For benzene the corresponding values were 2×10^{-5} and 0.02 millihartree for CASVB2 and 2×10^{-5} and 0.01 millihartree for CASVB4.

We now examine the more commonly accepted description of the bonding region. Considering the bridging region alone, our new picture of the three-centre two-electron bonds transforms as $A_g+B_{2g}+B_{1u}+B_{3u}$,[4] whereas the 'standard' description (discussed below) will transform according to $2A_g+2B_{3u}$. It is thus

[4] Of course strictly only the symmetry-pure solutions.

	Θ^s_{pp}	Θ^e_{pp}	Θ^c_{pp}
CASVB1	85.13%	84.93%	95.28%
CASVB2	94.90%	94.80%	96.05%
CASVB3	91.18%	91.39%	96.43%
CASVB1 (s)	92.62%	92.86%	96.15%
CASVB2 (s)	94.80%	94.72%	95.52%

Table 4.26: Weights of the perfect-pairing mode of spin coupling (Kotani) for the different orbital sets for diborane.

	S^s_{cov}	S^e_{cov}	S^c_{cov}
CASVB1	0.99571	0.99570	0.98227
CASVB2	0.99554	0.99553	0.99552
CASVB3	0.99557	0.99557	0.99235
CASVB1 (s)	0.99549	0.99548	0.99539
CASVB2 (s)	0.99548	0.99547	0.99546
	E^s_{cov}/hartree	E^e_{cov}/hartree	E^c_{cov}/hartree
CASVB1	−52.889322	−52.889349	−52.868926
CASVB2	−52.889250	−52.889263	−52.889200
CASVB3	−52.889762	−52.889782	−52.884695
CASVB1 (s)	−52.889112	−52.889127	−52.888863
CASVB2 (s)	−52.889174	−52.889181	−52.889148

Table 4.27: Covalent overlaps and energies for the different orbital sets for diborane.

	$\langle B_1^b \vert H_1^b \rangle$	$\langle B_1^b \vert B_2^b \rangle$	$\langle B_1^b \vert H_2^b \rangle$	$\langle H_1^b \vert H_2^b \rangle$	$\langle B_1^t \vert H_1^t \rangle$
CASVB1	0.84963	0.45686	0.26690	0.20090	0.80613
CASVB2	0.85314	0.40036	0.24616	0.21447	0.82199
CASVB3	0.85020	0.43526	0.25613	0.20175	0.82052

Table 4.28: Overlaps between bridging and bond-forming terminal orbitals of B_2H_6.

	Θ_{pp}^s	Θ_{pp}^e	Θ_{pp}^c
CASVB1	94.38%	94.56%	95.96%
CASVB2	95.27%	95.34%	95.76%
CASVB3	95.00%	95.14%	95.78%

Table 4.29: Weights of the perfect-pairing mode of spin coupling for the different interpretations of CAS B.

simple to identify the CAS wavefunction with the correct transformational properties, with active orbitals transforming as $4A_g+2B_{3g}+2B_{1u}+2B_{2u}+2B_{3u}$. This solution, 'CAS B', is 7.5 millihartree above the lowest CAS (see table 4.23). The interpretations of this CAS according to CASVB1–3, however, reproduce the conventional picture of the bonding (see figures 4.29 and 4.30).

It is worth mentioning that although the energy gap was smaller for the corresponding 4 in 4 CAS solutions (-52.840084 hartree and -52.838508 hartree for A and B respectively), solution A was also in this case lowest. Valence bond interpretations of these two solutions reproduced the same two descriptions of the bridging region that the twelve electron calculations in this work have produced.

The overlap integrals and spin coupling show generally the same type of variations between CASVB1–3 as seen for the other systems studied in this work (see tables 4.28 and 4.29). The values are significantly more uniform in this case, however, to the extent for example that imposing the CAS spin coupling is not associated with significantly inferior covalent overlaps or energies. It is remarkable

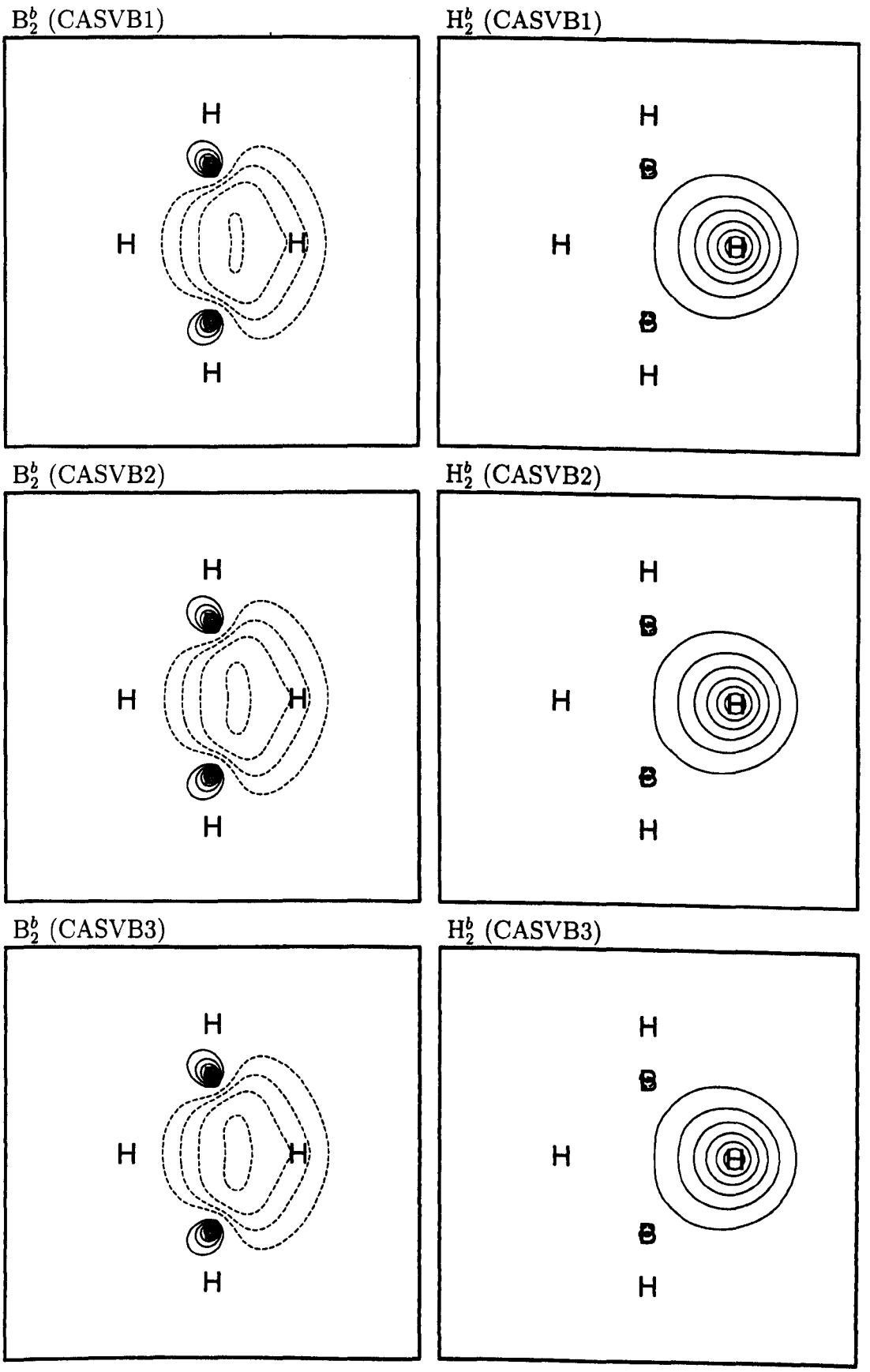

Figure 4.29: CASVB1–3 interpretations of solution B; bridging orbitals.

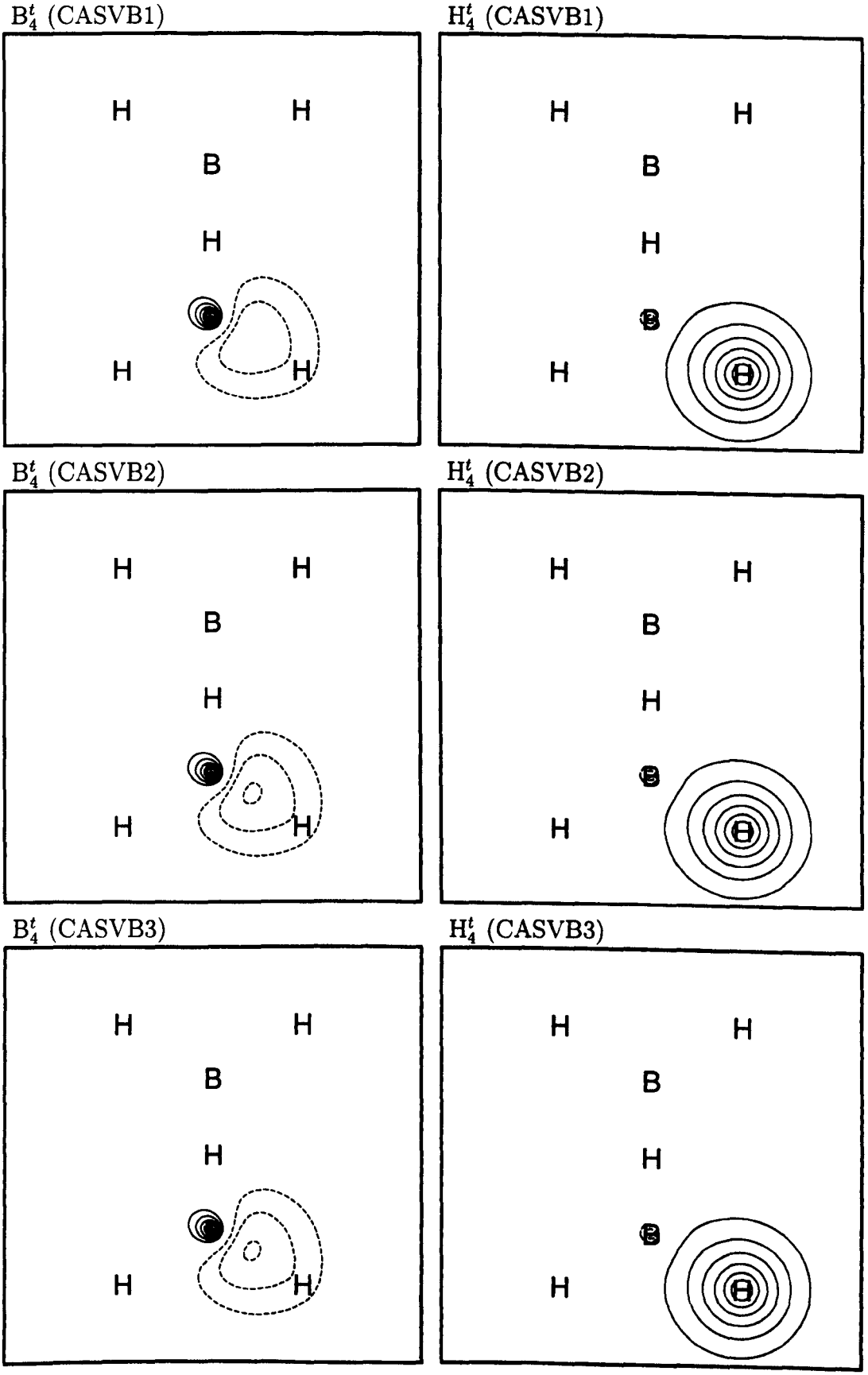

Figure 4.30: CASVB1–3 interpretations of solution B; terminal orbitals.

	S^s_{cov}	S^e_{cov}	S^c_{cov}
CASVB1	0.99862	0.99862	0.99858
CASVB2	0.99861	0.99861	0.99861
CASVB3	0.99861	0.99861	0.99860
	E^s_{cov}/hartree	E^e_{cov}/hartree	E^c_{cov}/hartree
CASVB1	−52.892407	−52.892409	−52.892307
CASVB2	−52.892424	−52.892425	−52.892413
CASVB3	−52.892449	−52.892450	−52.892422

Table 4.30: Covalent overlaps and energies for the different orbital sets for diborane.

that the values for S_{cov} and even E_{cov} are superior to the solutions based on CAS A (see table 4.30). The conclusion must be that this form of wavefunction 'lends' itself far better to a spin-coupled-like interpretation. Since the energy difference between CAS A and B is fairly significant, one must say that for this system the CAS and SC models do *not* agree in all particulars.

Finally we have investigated the standard spin-coupled energy surface in order, if possible, to establish the effect of orbital constraints on the converged orbital sets. It is worth mentioning that spin-coupled calculations with 12 active electrons at this point in time is considered computationally quite intensive. Here we were able to reduce the total computer time significantly by using excellent starting guesses in the form of the converged CASVB1 or 2 solutions. In most cases convergence was reached in a handful of iterations. The important question to be asked is whether—in analogy with what was found for ozone—there corresponds a local, or global, minimum in the energy surface to the new description of the bonding region. Since there are significant differences between the symmetry-pure and symmetry-broken solutions both were investigated. This was done by simply starting from the appropriate CASVB solution, but in neither case was any evidence of a local minimum be found. A curious finding, however, was that

	$\langle B_1^b \vert H_1^b \rangle$	$\langle B_1^b \vert B_2^b \rangle$	$\langle B_1^b \vert H_2^b \rangle$	$\langle H_1^b \vert H_2^b \rangle$	$\langle B_1^t \vert H_1^t \rangle$
SC	0.85143	0.42939	0.24966	0.18963	0.82711

Table 4.31: Spin-coupled overlaps between bridging and bond-forming terminal orbitals of B_2H_6.

the symmetry-broken solutions converged extremely rapidly towards the standard solution. So the reason for the symmetry-breaking itself is perhaps to be found in the discrepancies between the CAS and SC predictions. It may be that in real terms the two descriptions are not so different as the form of the orbitals suggests.

The (standard) spin-coupled orbitals with core orbitals taken from CAS A are shown in figure 4.31. It is clear, from the plots and overlaps (table 4.31), that this solution is in very good agreement with the interpretations of CAS B given above. Of course the energy difference between the SC wavefunction (see table 4.23) and the CAS interpretations, about 6 millihartree, should be kept in mind. The weight of the perfect-pairing spin function was in this case 95.68%.

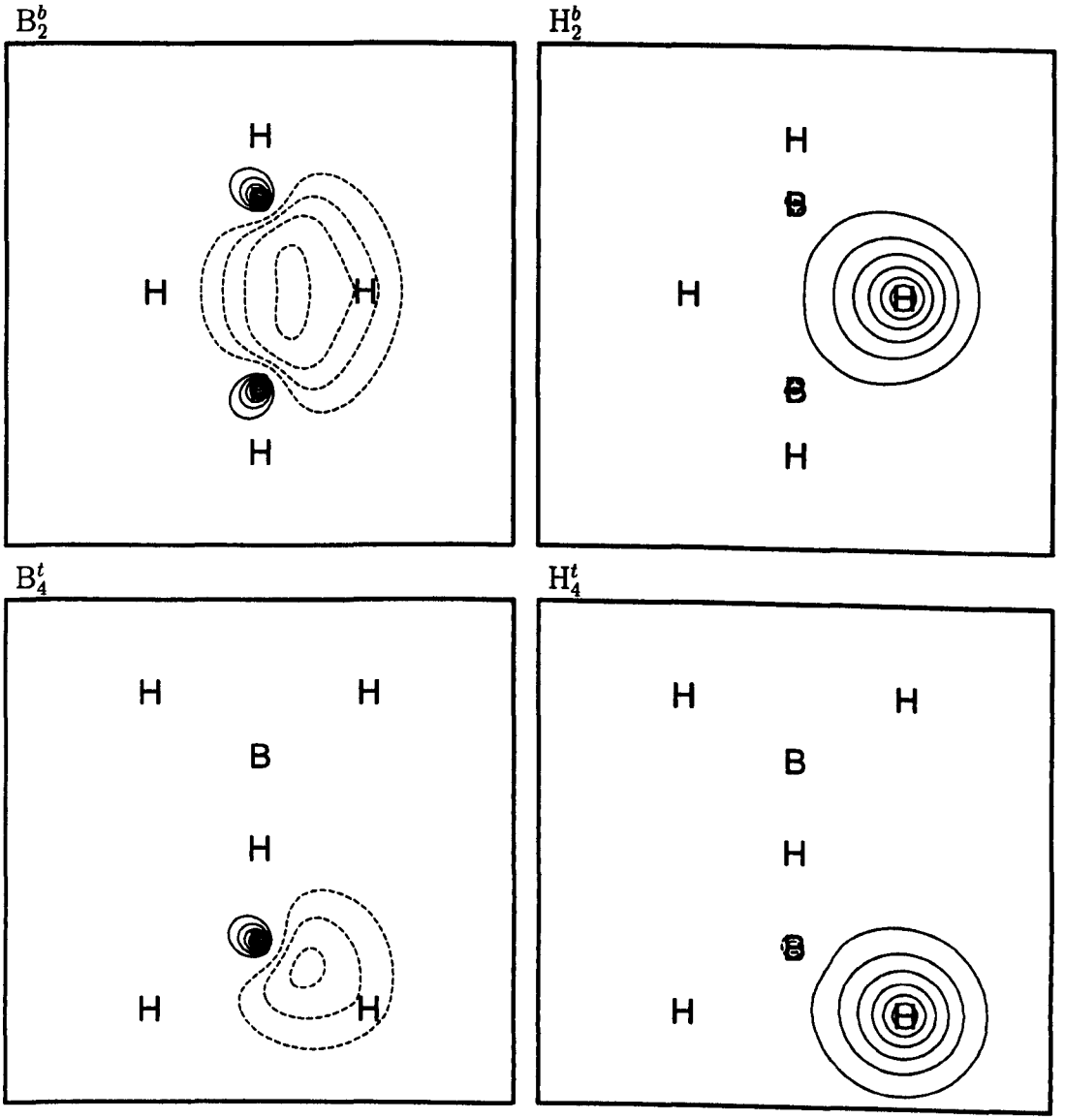

Figure 4.31: Spin-coupled orbitals of B_2H_6.

4.6 Discussion

The simple idea of arranging *sequential* updates so as to obtain the effect of simultaneous ones we believe may be a useful alternative to the approaches often adopted in, say, MCSCF theory. This would be true even for structure expansions not necessarily of full-CI form. The theory is of course applicable also to unitary orbital transformations, but transformations to non-orthogonal orbital sets are likely to be particularly useful in view of the computational cost associated with setting up and solving the eigenvalue problem in this case. Much of the technology related to the application of the single excitation matrices, for example evaluation of coupling-coefficients, is already widely available. So implementation of the suggested strategy would be very straightforward in conjunction with many of the already existing MCSCF codes.

The criteria we have presented here are all based on the form of the spin-coupled wavefunction. It is of course perfectly straightforward to define similar criteria based on other functional forms, provided that these lie inside a general 'N in m' CAS space. This might involve substituting $\mathbf{P}_{\text{cov}}$ with a more general matrix, but as mentioned in section 3.2, it is important for the efficiency particularly of the Hessian evaluation that such a matrix can be identified and included in the expressions. In view of the simplicity of the procedures, *general* non-orthogonal MCSCF procedures based on such a procedure seem perfectly viable. These might, as in this work, attempt to interpret the CAS wavefunction, or, alternatively, with the two-step procedure adopted in ref. [8], be adapted to locate the fully variational solutions.

These methods are quite simple to implement, but nevertheless have proven very competitive in relation to the standard spin-coupled codes. This is particularly remarkable in view of the concentrated effort that has been put into the optimisation of this procedure. There is even the possibility that a CAS driven two-step procedure, as discussed above, could be significantly more efficient for systems containing some degree of symmetry. One drawback of the existing SC

codes is that no utilisation of symmetry is possible. It is quite likely that the present work could pave the way for this area to be developed further. There is the argument that rigorous constraints on the orbitals will predispose the calculation towards certain types of solution, but a not so restrictive alternative would be to specify the distribution of natural orbitals among the irreducible representations. This would be equivalent to what is often done in CAS calculations utilising symmetry.

This work has confirmed the previously existing notions regarding the similarity between equivalent CAS and SC calculations. If anything the link between the two methods has been considerably strengthened. We have presented various pieces of evidence for the similarity between the two wavefunctions, where of particular note is the similarity between the active orbital spaces, and the impressive values for S_{cov} that may be obtained with a function of spin-coupled form.

Comparing the four criteria, the main, surprising, finding is the small difference between the overlap- and energy-based criteria. In the systems studied here, for example, constraining the spin coupling as in CASVB2 and 4 seemed to have a far greater effect. Since the overlap-based criteria are considerably cheaper computationally, this suggests that future work on the interpretation of CAS wavefunctions will be mostly concerned with these. Alternatively, if an energy-based solution is required, the corresponding overlap-optimised orbital set is likely to provide an excellent guess. We can take this one step further. We may regard the spin-coupled wavefunction as the best energy-based approximation to the CAS wavefunction, but with a *general* expansion of the orbitals, as in equation 1.15. It is easy to envisage the equivalent overlap-based optimisation, and it is likely that the resulting solution would be a very affordable, excellent approximation (or guess) to the spin-coupled wavefunction. Such an approach might also serve to alleviate some of the problems apparently associated with the small number of free orbital parameters in some cases.

Two problems were encountered in the orbital optimisations: the 'rogue' so-

lutions for methane and the symmetry-breaking found to occur for benzene and diborane. At this stage we have to admit that satisfactory explanations have eluded us for either. The shortness of the expansion of the spin-coupled-like orbitals is one likely cause of the very different solutions found for methane. This is, of course, in spite of the small difference between the CAS and SC active orbital spaces. We have not encountered other such rogue solutions, so it is difficult at this stage adequately to gauge the extent of the problem. We believe the practical approach to this problem is always to perform calculations employing CASVB1 and CASVB2 in tandem. If these two solutions agree on a qualitative level it is likely that a physically sensible solution is located, whereas if there are significant discrepancies a certain amount of care should be exercised in interpreting the results.

The small number of orbital parameters is probably also the main factor in the symmetry-breaking found for benzene and diborane, but the constraints involved in CASVB2 and 4, and maybe also the nature of overlap-optimisation, could be contributory factors. In all the cases the breaking of symmetry was associated with a singular, or near-singular, Hessian, so although the improvement in S_{cov} or E_{cov} was relatively modest, the distortion of the orbitals was quite significant. Nevertheless, it was in both cases clear what the corresponding symmetry-pure solutions would be, so in practice it should be simple to impose the applicable constraints once convergence onto a symmetry-broken solution is diagnosed. Occurrence of these solutions is perhaps not aesthetically very pleasing, but we believe that they do not in practice present any serious problems. The main problem is, as in the case of B_2H_6, to determine whether the symmetry-breaking may be a genuine feature of the solution. This is to a certain extent a matter of interpretation, but we have in section 4.5 considered some of the evidence that may influence the decision one way or the other.

By and large the agreement between the CAS interpretations and equivalent spin-coupled results has been excellent. Only for diborane can one say that

real disagreement between the two methods has occurred. It has always been possible to reproduce the qualitative SC solutions, but for ozone and diborane the interpretation of the CAS has suggested alternative (and for ozone probably more 'correct') descriptions of the bonding. It is slightly paradoxical that the constrained, and computationally cheaper, CASVB2 and 4, generally are closest to the spin-coupled solutions. This finding may be related to the fact that the weight of the perfect-pairing mode of spin coupling was generally higher for these two criteria. Of course, being close to SC is not necessarily an end in itself. It would be interesting to investigate which orbital set performs better in subsequent configuration interaction steps, for example.

Another question that raises itself in relation to this, is the adaptation of other developments in spin-coupled theory to the spin-coupled-like wavefunctions. It should for example be straightforward to define virtual orbitals for the spin-coupled-like wavefunctions by constructing Fock operators in a similar fashion to the scheme employed for the spin-coupled wavefunction.

In conclusion, we restate that it is the ability to transform exactly and efficiently the CAS CI space that has made this whole research practical. We believe that the utility of this method extends to many other areas. The link between the CAS and spin-coupled methods has been much strengthened—to the extent that one can almost call SC the 'little brother' of CAS—and it is our hope that this 'duality' will be considered by researchers employing *both* methods in the future.

Bibliography

[1] H.-J. Werner and P. J. Knowles, J. Chem. Phys. **82**, 5053 (1985).

[2] P. J. Knowles and H.-J. Werner, Chem. Phys. Lett. **115**, 259 (1985).

[3] MOLPRO is a package of *ab initio* programs written by H.-J. Werner and P. J. Knowles, with contributions of J. Almlöf, R. Amos, S. Elbert, K. Hampel, W. Meyer, K. Peterson, R. Pitzer, and A. Stone.

[4] C. W. Bauschlicher and P. R. Taylor, J. Chem. Phys. **85**, 6510 (1986).

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

[6] D. L. Cooper, J. Gerratt and M. Raimondi Nature, **323**, 699 (1986).

J. Gerratt, Chem. in Britain **23**, 327 (1987).

For a review: D. L. Cooper, J. Gerratt and M. Raimondi, in "*Advances in the Theory of Benzenoid Hydrocarbons*", eds. I. Gutman and S. J. Cyvin, Topics in Current Chemistry **153**, 41 (1990).

[7] A. Cabana, J. Bachaud and J. Giguère, Can. J. Phys., **52**, 1949 (1974).

[8] D. L. Cooper, J. Gerratt and M. Raimondi, Nature **323**, 699 (1986).

[9] F. Penotti, D. L. Cooper, J. Gerratt and M. Raimondi, J. Mol. Struc. (THEOCHEM) **169**, 437 (1988).

P. C. Hiberty and D. L. Cooper, J. Mol. Struc. (THEOCHEM) **169**, 437 (1988).

M. Sironi, D. L. Cooper, J. Gerratt and M. Raimondi, J. Am. Chem. Soc. **112**, 5054 (1990).

D. L. Cooper, J. Gerratt and M. Raimondi, Chem. Rev. **91**, 929 (1991).

[10] R. Pauncz, "*Spin Eigenfunctions*", Plenum Press, New York (1979).

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

[12] "*Tables of interatomic distances and configurations in molecules and ions*", The Chemical Society, London (1958).

[13] S. Wilson and J. Gerratt, J. Mol. Phys. **30**, 765 (1975).

[14] D. L. Cooper, J. Gerratt and M. Raimondi, J. Mol. Struc. (THEOCHEM) **229**, 155 (1991).

[15] M. Sironi, M. Raimondi, D. L. Cooper and J. Gerratt, J. Phys. Chem. **95**, 10617 (1991).

Chapter 5

The biorthogonal method

5.1 Introduction

As we have seen evidence for in the past chapters, non-orthogonal methods, in particular those employing orbital optimisation, have been given a rebirth. Especially from a chemists viewpoint these methods are immensely attractive. Unfortunately the many advantages to be obtained from valence bond (VB) approaches are limited to relatively small systems, mainly due to the poor scaling properties of the computational effort with the number of active electrons—the so-called $N!$ problem. The biorthogonal approximation is an attempt at circumventing this, making the computational effort similar to that of orthogonal methods, at the expense of the variational property of the solutions.

The biorthogonal method in electronic structure calculations has received attention on and off for the last 25 years. The use of a dual basis in quantum chemistry was first considered as an efficient way of calculating the matrix elements between two Slater determinants by direct use of the Löwdin formula [1]. The foundation for the present work was laid in 1971 when, in a second quantised formalism, Moshinsky and Seligman considered using a dual basis for optimising many-electron functions [2]. Since then, the convergence properties of the method with respect to increasing number of configurations have been investigated [3],

and special attention has been given to the features of the variation principle [4]. Recently optimisation of non-orthogonal orbitals using the biorthogonal method has been carried out [5, 6].

The concept of biorthogonal sets, or reciprocal spaces, pervades many areas of mathematics, physics and chemistry. But the utility of non-symmetric expectation values (as put forward in section 5.3) is not limited to addressing the non-orthogonality problem. A noteworthy application in quantum chemistry is the optimisation of explicitly correlated wavefunctions by the so-called transcorrelated method [7]. The optimisation of

$$\Psi = \prod_{i>j} f(\mathbf{r}_i, \mathbf{r}_j)\Phi, \tag{5.1}$$

where $f(\mathbf{r}_i, \mathbf{r}_j)$ is a general correlation factor between electrons i and j and Φ is a wavefunction lacking explicit correlation (e.g., a linear combination of Slater determinants), is exceedingly difficult when the standard variation principle is employed. The main technical difficulty lies in calculating the very complicated integrals over the one- and two-electron operators and correlation factors. This problem is much simplified if a left-hand function of the form

$$\Psi' = \prod_{i>j} f(\mathbf{r}_i, \mathbf{r}_j)^{-1}\Phi \tag{5.2}$$

is used [7]. The non-symmetric expectation value

$$\langle E' \rangle = \frac{\langle \Psi' | \hat{H} | \Psi \rangle}{\langle \Psi' | \Psi \rangle} \tag{5.3}$$

may then be optimised with respect to all non-linear parameters in the same way as discussed for the spin-coupled wavefunction in section 5.7.

Since orbital optimisation is imperative to the success of valence bond methods, we will in this chapter take a look at the biorthogonal method in this context, and for some simple systems attempt to gauge the validity of the approximations involved. The main way this differs from previous work is the possibility we have of quantifying the errors associated with the biorthogonal method, by comparing it with the equivalent calculations where non-orthogonality is incorporated fully, without approximations.

5.2 The biorthogonal orbital set

The bottleneck in a large number of post-SCF methods is the calculation of matrix elements of the Hamiltonian between many-electron wavefunctions. For our purposes it is sufficient to consider matrix elements between Slater determinants, and we consider here the expensive part of the Hamiltonian: the two-electron operator. The efficiency of our method will be governed largely by the amount of work associated with calculating terms of the form

$$\Big\langle |\phi_{\mu_1}\phi_{\mu_2}\cdots\phi_{\mu_N}| \Big| \sum_{i<j}^{N} \hat{g}(i,j) \Big| |\phi_{\nu_1}\phi_{\nu_2}\cdots\phi_{\nu_N}| \Big\rangle, \tag{5.4}$$

where the bra and ket are Slater determinants.

Inspecting the Löwdin formula (see appendix 1.a) we can see that *second order* cofactors of the overlap matrix between the two orbital sets are needed. This gives on the order of N^4 terms for each pair of determinants. For most non-trivial sets of Slater determinants the calculation of the cofactors is not in itself significant—the expensive part is the combination of cofactors with two-electron integrals. Thus the computational effort scales roughly as $N^4 N_{\text{det}}^2$ for methods employing non-orthogonal orbitals directly.[1]

By comparison we have in the orthogonal case only N^2 elements for two identical determinants, and N terms for determinants differing in one spin-orbital. The computational effort will thus scale as $N^2 N_{\text{det}}$ or as an upper estimate $N N_{\text{det}}^2$. There will, of course, be many vanishing matrix elements for determinants differing in more than two spin orbitals. It should be clear from this why from a computational point of view orthogonality is so attractive.

In the biorthogonal method we consider two orbital sets with the property

$$\langle \tilde{\phi}_\mu | \phi_\nu \rangle = \delta_{\mu\nu} \tag{5.5}$$

where it is understood that neither of the orbital sets, $\{\tilde{\phi}\}$ or $\{\phi\}$, will in general

[1] A reduction in the amount of work necessary can be achieved by taking into account the orthogonality of the alpha and beta spin functions and the point-group symmetry of the molecule. This however will not affect the *scaling* properties of the methods.

be orthogonal. It is easy to show that, if the two sets are taken to span identical spaces, the biorthogonal, or 'dual', set is related to the direct set by the linear transformation

$$\{\tilde{\phi}\} = \{\phi\}\mathbf{s}^{-1} \tag{5.6}$$

where $\mathbf{s}_{\mu\nu}=\langle\phi_\mu|\phi_\nu\rangle$.

A 'left' and a 'right' set of determinants can now be chosen, such that any 'left' determinant is constructed from $\tilde{\phi}$ orbitals, and any 'right' determinant is constructed from ϕ orbitals. This means that the two sets of Slater determinants will also satisfy the biorthogonality condition

$$\langle\tilde{\Phi}_i|\Phi_j\rangle = \delta_{ij}, \tag{5.7}$$

but as they in general span different parts of the full CI-space, the term 'biorthogonal' has slightly different connotations here than in the case of the two orbital sets.

Examining the Löwdin formula in appendix 1.a, it is clear that since the overlap matrices between the left- and right-hand orbitals are identical in the biorthogonal and orthogonal cases, the cofactors, density matrices etc. entering the energy expressions will be exactly the same for the two. In this work all matrix elements were evaluated by considering pairs of Slater determinants whereby the simple Slater-Condon rules may be employed. More sophisticated schemes involving the direct evaluation of matrix elements between configuration state functions (CSFs) would also be possible. For those the basic procedure in the biorthogonal method would also be essentially unchanged. The main difference, from a computational point of view, between the standard orthogonal and biorthogonal approaches lies in the one- and two-electron integrals. In the biorthogonal method the non-symmetric integrals

$$\langle\tilde{\phi}_\mu|\hat{h}|\phi_\nu\rangle,$$

$$\langle\tilde{\phi}_\mu\tilde{\phi}_\nu|\hat{g}|\phi_\sigma\phi_\tau\rangle \tag{5.8}$$

are required. So the one- and two-electron integrals loses the normal left↔right symmetry associated with real orbitals. This in turn makes the Hamiltonian matrix non-symmetric.[2]

5.3 The energy expression

We have concluded that Hamiltonian matrix elements between biorthogonal orbital sets can be obtained very cheaply—essentially at the cost of the equivalent orthogonal calculation—but so far we have not presented any arguments relating to the utility of such a scheme. To do this we consider first the non-symmetric form of the energy expectation value taking the form

$$\tilde{E} = \frac{\langle \tilde{\Psi} | \hat{H} | \Psi \rangle}{\langle \tilde{\Psi} | \Psi \rangle}. \tag{5.9}$$

In general $\tilde{\Psi}$ and Ψ can be essentially unrestricted (i.e., we could also consider right- and left-hand functions not constructed from biorthogonal many-electron functions), as long as their overlap remains non-vanishing. For such a quantity to be sensibly interpreted as an energy of the system one of two following conditions must be satisfied; either

$$\tilde{\Psi} \approx \Psi \tag{5.10}$$

or

$$\hat{H}\Psi \approx E\Psi. \tag{5.11}$$

In the biorthogonal method it is the second condition which is of most interest. In that case we have

$$\frac{\langle \tilde{\Psi} | \hat{H} | \Psi \rangle}{\langle \tilde{\Psi} | \Psi \rangle} \approx E \frac{\langle \tilde{\Psi} | \Psi \rangle}{\langle \tilde{\Psi} | \Psi \rangle} = E \tag{5.12}$$

so that no matter how poor a wavefunction is $\tilde{\Psi}$, we may still interpret $\tilde{E}$ as a meaningful approximation to the energy associated with the wavefunction Ψ.

[2]If the left-hand and right-hand determinant sets span different spaces, the eigenvalues of the non-symmetric Hamiltonian will not be upper bounds to the energy. One of the main drawbacks of the biorthogonal method is this breakdown of the variation principle.

It should be borne in mind that this kind of non-symmetric expectation value only makes sense if Ψ (or $\tilde{\Psi}$) is close to being an eigenfunction of the operator in question. I.e., the expression

$$\frac{\langle\tilde{\Psi}|\hat{O}|\Psi\rangle}{\langle\tilde{\Psi}|\Psi\rangle} \tag{5.13}$$

cannot be sensibly interpreted as an expectation value for $\hat{O}$ unless Ψ happens to be a near-eigenfunction for this operator. Otherwise the standard symmetric expression will have to be used. The biorthogonal method as formulated here is thus restricted to optimisation of the wavefunction and evaluation of, specifically, the energy. In the context in which we shall consider the biorthogonal method it is the optimisation step that requires the majority of the computational effort with standard methods. For the size of systems we intend to study here analysing the wavefunction in terms of, say, one- and two-electron properties will in general be quite straightforward using the standard expressions.

5.4 The non-symmetric eigenvalue problem

With no exceptions of which we are aware, previous work utilising the biorthogonal approximation has been based on solving the associated *linear* variational problem. It therefore seems appropriate to go into some detail here in describing the non-symmetric generalisation of the standard eigenvalue problem.

As discussed in section 5.3, our aim is for the relation

$$\hat{H}\Psi \approx E\Psi \tag{5.14}$$

to be satisfied as accurately as possible. We can write that as

$$\hat{H}\Psi = E\Psi + \Phi. \tag{5.15}$$

A necessary condition for Φ to be zero is then

$$\langle\tilde{\Psi}_i|\Phi\rangle = 0. \tag{5.16}$$

Constructing the matrix representation, $\tilde{\mathbf{H}}$, of the Hamiltonian operator in the basis of the left- and right-hand function sets, $\{\tilde{\Psi}\}$ and $\{\Psi\}$ respectively, this condition may be expressed as

$$\tilde{\mathbf{H}}_{i\neq 1,1} = \langle \tilde{\Psi}_{i\neq 1} | \hat{H} | \Psi_1 \rangle = 0 \tag{5.17}$$

with Ψ_1 being the optimal wavefunction. This result is analogous to the corresponding condition normally put on the symmetric Hamiltonian. In the traditional case the component of Φ in the direct space is forced to zero, whereby Ψ becomes the projection of the exact solution onto the direct space. In the biorthogonal method it is the component of Φ in the *biorthogonal* space which is zero.

We wish to find a general non-singular transformation of the N_{dim} right-hand structures, according to

$$\{\Psi'\} = \{\Psi\}\mathbf{T}, \tag{5.18}$$

such that (5.17) is satisfied. The transformation of the left-hand structures must then, in order to preserve biorthogonality, be

$$\{\tilde{\Psi}'\} = \{\tilde{\Psi}\}\mathbf{T}^{-1\dagger}. \tag{5.19}$$

In this way the overlap matrix, which we can assume to be the identity for the untransformed direct and biorthogonal sets (if CSFs based on orthogonal spin functions are used), is unchanged:

$$\mathbf{S}' = \mathbf{T}^{-1}\mathbf{I}\mathbf{T} = \mathbf{I}. \tag{5.20}$$

As an illustrative example it is useful to consider the solution to a simple two-state problem with Hamiltonian

$$\tilde{\mathbf{H}} = \begin{pmatrix} \tilde{H}_{11} & \tilde{H}_{12} \\ \tilde{H}_{21} & \tilde{H}_{22} \end{pmatrix}, \tag{5.21}$$

where in general $\tilde{H}_{12} \neq \tilde{H}_{21}$. We can solve this for the first eigenvector by transforming according to

$$\tilde{\mathbf{H}}' = \mathbf{T}^{-1}\tilde{\mathbf{H}}\mathbf{T} = \begin{pmatrix} 1 & 0 \\ -\lambda_1 & 1 \end{pmatrix} \begin{pmatrix} \tilde{H}_{11} & \tilde{H}_{12} \\ \tilde{H}_{21} & \tilde{H}_{22} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ \lambda_1 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} \tilde{H}_{11} - \lambda_1 \tilde{H}_{12} & \tilde{H}_{12} \\ \tilde{H}_{21} + \lambda_1(\tilde{H}_{11} - \tilde{H}_{22}) + \lambda_1^2 \tilde{H}_{12} & \tilde{H}_{22} + \lambda_1 \tilde{H}_{12} \end{pmatrix}. \tag{5.22}$$

In this way solving the equation

$$\tilde{H}'_{21} = \tilde{H}_{21} + \lambda_1(\tilde{H}_{11} - \tilde{H}_{22}) + \lambda_1^2 \tilde{H}_{12} = 0 \tag{5.23}$$

for λ_1, will lead to the required solution for the first (right-hand) eigenvector. We have two comments to this statement. Firstly, we may consider solving the equation to first order in λ_1. Ignoring terms containing λ_1^2 above, we get

$$\lambda_1 = -\frac{\tilde{H}_{21}}{\tilde{H}_{11} - \tilde{H}_{22}}, \tag{5.24}$$

and thus an associated energy to first order as

$$\tilde{E}^{(0)} + \tilde{E}^{(1)} = \tilde{H}_{11} + \frac{\tilde{H}_{12}\tilde{H}_{21}}{\tilde{H}_{11} - \tilde{H}_{22}}. \tag{5.25}$$

This is analogous to what may be obtained by a first order perturbation treatment of the symmetric case. Secondly, we note that complex eigenvalues and roots are possible, in the present case if

$$(\tilde{H}_{11} - \tilde{H}_{22})^2 - 4 \times \tilde{H}_{12}\tilde{H}_{21} < 0. \tag{5.26}$$

The complex eigenvalues will always occur as complex conjugate pairs with complex conjugate eigenvectors.[3] A given eigenvalue will be real if the matrix $\tilde{\mathbf{H}}$ is not very 'far' from being symmetric and if the closest eigenvalue evaluated from the symmetrised $\tilde{\mathbf{H}}$ (i.e., $\frac{1}{2}(\tilde{\mathbf{H}}+\tilde{\mathbf{H}}^\dagger)$) is not near-degenerate [8].

In the same way as was done for the first eigenvector, we may obtain the second eigenvector by transforming $\tilde{\mathbf{H}}'$. This will give a total (direct) transformation matrix as

$$\mathbf{T} = \begin{pmatrix} 1 & \lambda_2 \\ \lambda_1 & 1 + \lambda_1\lambda_2 \end{pmatrix}. \tag{5.27}$$

[3]We can view the roots as being solutions to the secular polynomial (or: characteristic equation) of N_{dim}th degree: $\det(\tilde{\mathbf{H}} - \lambda\mathbf{I}) = 0$. The complex roots of a polynomial with only real coefficients will always occur as conjugate pairs.

with

$$\lambda_2 = -\frac{\tilde{H}'_{12}}{\tilde{H}'_{11} - \tilde{H}'_{22}} = -\frac{\tilde{H}_{12}}{\tilde{H}_{11} - \tilde{H}_{22} - 2\lambda_1 \tilde{H}_{12}}. \tag{5.28}$$

The diagonal elements of $\tilde{\mathbf{H}}'$ will be unaffected by this second transformation since $\tilde{H}'_{21}$=0. In the general case it is thus sufficient to bring the Hamiltonian to triangular form in order to obtain the correct eigenvalues. If all the eigenvectors are also required, however, the Hamiltonian must be completely diagonalised.

Once the Hamiltonian has been diagonalised, the right-hand eigenvectors are the best possible approximations in $\{\Psi\}$ to the corresponding exact eigenstates, in the sense that all components of $\hat{H}\Psi_i$ in the biorthogonal space $\{\tilde{\Psi}\}$ are zero except for $\tilde{\Psi}_i$. Similarly the left hand eigenvectors will be the best possible approximations in $\{\tilde{\Psi}\}$ to the corresponding eigenstates, since when the Hamiltonian is completely diagonal the conjugate of (5.17) is also fulfilled.

As mentioned before, the eigenvalues of the non-symmetric eigenvalue problem do not adhere to the variation principle, in the sense that $\tilde{E}_i > E_i^{ex}$ will not necessarily hold for any given i. In spite of this it is generally assumed that the *ordering* of the eigenvalues is basically correct. Thus the ith eigenvector is interpreted as the best approximation to the ground state if

$$\tilde{E}_i < \tilde{E}_j, \qquad \text{for all } \; j \neq i. \tag{5.29}$$

This will pursued further in sections 5.6 and 5.7.

In practice most computational subroutine libraries have procedures available for carrying out such diagonalisations of non-symmetrical matrices. Special considerations may be necessary for larger dimensions of the matrix involved, as often encountered in quantum chemical problems. Normally only the lowest-lying roots are required, and this is utilised in a generalisation of Nesbett's algorithm [8], and in a modification of Davidson's scheme [9].

5.5 Error analysis

The energies obtained using the biorthogonal method are in error for two reasons: the error in the biorthogonal energy evaluation, and the error in the biorthogonal optimisation criterion. It has been shown, [10], that the error in the biorthogonal energy is proportional to the errors in the dual and direct wavefunctions respectively, as

$$\delta E \propto |\tilde{\Delta}| \times |\Delta|, \tag{5.30}$$

where the exact eigenfunction in the full CI space is written as

$$\Psi_{\mathrm{ex}} = \tilde{\Psi} + \tilde{\Delta} = \Psi + \Delta, \tag{5.31}$$

Ψ_{ex}, $\tilde{\Psi}$ and Ψ all being normalised. With the assumption that $|\tilde{\Delta}|$ is big, and $|\Delta|$ is small (since Ψ is close to being an eigenfunction), we can see that the error in the energy is *first* order in the error of the wavefunction—not second order as is the case for traditional methods.

The next step is to ask: how big is the actual error in the biorthogonally optimised wavefunction defined by the condition (5.17), and what is the difference between the biorthogonally optimised wavefunction and the traditionally optimised one? We will attempt to answer these questions in the calculations.

5.6 A super-CI approach for orbital optimisation

The super-CI approach is a widely used method for optimising orbitals, using a sequence of CI calculations [11, 12, 13], and the scheme described below has some points in common with the strategy for optimising non-orthogonal orbitals, employed by van Lenthe and coworkers [14, 15, 16]. The super-CI approach is often preferred in circumstances where using the explicit conditions on the orbitals would in some way be impractical. For the biorthogonal method there are

two good reasons to prefer such a scheme: the non-symmetric eigenvalue problem, as described in section 5.4, is very well understood, and a straightforward minimisation of the energy functional is not feasible due to the break-down of the variation principle.

We will here concentrate on optimising a wavefunction of spin-coupled form

$$\Psi_{\mathrm{SC}} = \sum_{k}^{f_S^N} c_{Sk} \hat{\mathcal{A}} \{ \phi_1 \phi_2 \cdots \phi_N \Theta_{SM;k}^N \}. \tag{5.32}$$

We aim to optimise this function with respect to the spin-coupling coefficients, c_{Sk}, and the orbital parameters, in a similar manner as was described in section 1.3.

We first consider the relationship between the direct and dual orbital sets for this case. It is useful, as done in section 2.2, to partition the direct orbital set according to

1. Frozen core
2. Optimised core
3. Active orbitals
4. Virtual orbitals

We may assume, without loss of generality, that the core and virtual orbitals are orthogonal amongst themselves as well as to the active orbital set. The only necessary non-orthogonality is between different active orbitals. This means that the dual active orbitals are given as

$$\{\tilde{\phi}\} = \{\phi\}\mathbf{s}^{-1}, \tag{5.33}$$

with $\mathbf{s}_{\mu\nu}=\langle\phi_\mu|\phi_\nu\rangle$. The dual core and virtual orbitals coincide with the direct ones. In the following we will for simplicity only allow *frozen* core orbitals, which we will assume have been effectively removed from the problem as discussed in section 1.4, but it would be fairly straightforward to extend the procedures described bellow to allow also core optimisation. In the remainder of this chapter

we will use μ, ν, $\cdots$ for occupied orbitals, α, β for virtual orbitals, and η, θ for orbitals that may be either occupied or virtuals.

The super-CI scheme adopted in this work may be classified as a two-step procedure, with alternating updates of spin-coupling coefficients and orbital parameters. This is useful to limit the size of the non-symmetric Hamiltonian that has to be set up and diagonalised. Other schemes have taken this one step further, updating one orbital at a time, for example (see ref. [17]).

The spin coupling coefficients, c_{Sk}, are obtained by solving the $f_S^N \times f_S^N$ secular problem, ensuring that

$$\langle \tilde{\Psi}_{i\neq 1} | \hat{H} | \Psi_1 \rangle = 0. \tag{5.34}$$

It should be noted that the biorthogonal Hamiltonian matrix of spin-coupled structures *is* symmetric. This is due to the fact that only the spin-part of the wavefunction varies between the different functions. However, since there is also in this case a difference in the spaces spanned by the left- and right-hand structures, the energies obtained will still not be variational.

For this set of spin coupling coefficients we now consider, as in section 2.5, orbital updates of the form

$$\phi_\mu \rightarrow \phi_\mu + \lambda c_\mu^\eta \phi_\eta. \tag{5.35}$$

As discussed in section 3.2 this give rise to a change in the wavefunction according to

$$\Psi_{\rm SC} \rightarrow \Psi_{\rm SC} + c_\mu^\alpha \hat{E}_{\eta\mu}^{(1)} \Psi_{\rm SC} + (c_\mu^\alpha)^2 \hat{E}_{\eta\mu}^{(2)} \Psi_{\rm SC}. \tag{5.36}$$

Indeed any arbitrary combination of orbital updates gives rise to a change in the wavefunction of the form

$$\Psi_{\rm SC} \rightarrow \Psi_{\rm SC} + \left(\sum_{\mu,\eta} c_\mu^\eta \hat{E}_{\eta\mu}^{(1)} \Psi_{\rm SC} \right) + \mathcal{O}(c^2), \tag{5.37}$$

where $\mathcal{O}(c^2)$ contains the terms of second and higher orders in the orbital coefficients. The next step is to set up and solve the secular problem in the space of $\Psi_{\rm SC}$ and all the singly ionic and singly excited structures. Assuming first order

behaviour in equation 5.37, we may then obtain the orbital parameters from the corresponding coefficient of the given eigenvector.

Since the higher order terms of the orbital parameters will be negligible only close to convergence, it is necessary to iterate until self-consistency. All these considerations are equivalent to the standard super-CI case, the only difference being that a *non-symmetric* secular problem is solved in order to obtain the right-hand CI-vector.

A number of problems were encountered in the practical optimisation, which are not normally associated with variational super-CI schemes. Various 'convergence accelerating' schemes, along the lines of those described in ref. [16], were attempted, but none gave any considerable improvement. An important complication was the finding that the converged solution did not always correspond to the lowest root of the secular problem (this will be discussed further in the following section). So by consistently choosing the lowest-lying CI-vector for the next update, convergence onto the correct solution is by no means ensured. An alternative strategy attempted, was to choose the CI-vector that is changed the *least* from the previous iteration, but this instead risks convergence onto higher-lying solutions.

These complications may perhaps be overcome, but the main consideration making a more sophisticated treatment of the optimisation problem warranted is the exceptionally poor convergence characteristics of the super-CI scheme. This made convergence literally impossible when the numbers of active electrons and spin functions took reasonable values.

5.7 A non-symmetric Newton-Raphson scheme

The convergence characteristics of the super-CI approach has proven to be unsatisfactory—especially since systems with a reasonable number of orbitals and spin functions often exhibit oscillatory convergence characteristics. We now be-

lieve the poor convergence characteristics to be a consequence of the inherent assumption regarding the transformation of the dual structure set as laid out in equation (5.19). After the orbitals are updated in a real iteration, both the direct and dual structures will span entirely different spaces, and this is what leads to the break-down of that assumption.

In order to lay the foundation for a more rigorous treatment of the optimisation problem we start by generalising the case of linear variation as outlined in section 5.4 to the case of general non-linear variables. In order to do this we consider again the linear two state problem as in section 5.4. We transform the direct wavefunction as

$$\Psi_1' = \Psi_1 + \lambda\Psi_2 \tag{5.38}$$

and the dual wavefunction similarly according to

$$\tilde{\Psi}_1' = \frac{1}{1-\lambda\tilde{\lambda}}(\tilde{\Psi}_1 + \tilde{\lambda}\tilde{\Psi}_2). \tag{5.39}$$

The factor in front is included to keep the overlap between Ψ_1' and $\tilde{\Psi}_1'$ normalised. This will give a biorthogonal energy as

$$\begin{aligned}\tilde{E} = \tilde{H}_{11}' &= \frac{1}{1-\lambda\tilde{\lambda}}(\tilde{H}_{11} + \lambda\tilde{H}_{12} + \tilde{\lambda}\tilde{H}_{21} + \lambda\tilde{\lambda}\tilde{H}_{22}) \\ &= \tilde{H}_{11} + \lambda\tilde{H}_{12} + \tilde{\lambda}\tilde{H}_{21} + \lambda\tilde{\lambda}(\tilde{H}_{22} - \tilde{H}_{11}) + \mathcal{O}(\lambda^3),\end{aligned} \tag{5.40}$$

where we have expanded the normalisation factor. With this formulation the variables λ and $\tilde{\lambda}$ will of course in general be interdependent. The necessary condition for Ψ_1 to be an eigenfunction, that $\tilde{H}_{21}$=0, can with this formulation be expressed as

$$\frac{\partial\tilde{E}}{\partial\tilde{\lambda}} = 0. \tag{5.41}$$

Generalising this to an arbitrary number of variables, we can write this condition as

$$\tilde{\mathbf{g}} = \mathbf{0}, \tag{5.42}$$

with

$$\tilde{g}_i = \frac{\partial\tilde{E}}{\partial\tilde{\lambda}_i}.$$

Note that this gradient is defined as the *partial* derivative of the biorthogonal energy. The ordinary gradient, on the other hand, would take the relationship between the direct and dual parameter sets into account (see below).

For a variational method the Hessian must be positive (semi-) definite since otherwise an infinitesimal change of the parameter set could lead to a decrease in the energy. Examining (5.40) it is clear that this is equivalent to the condition $\tilde{H}_{22}>\tilde{H}_{11}$. So generally, a variational super-CI scheme must always converge onto a situation where the solution is lower in energy than all other roots of the secular problem. As was seen in the last section, this is not always the case for the biorthogonal super-CI scheme, and this feature must be accepted if a non-variational optimisation is undertaken. This also means that one can put no conditions on the biorthogonal Hessian—it may have, and indeed often has, a number of negative eigenvalues at convergence. We note that it is of course *possible* to converge onto stationary points with a positive definite Hessian, but the energy for such a situation may very well be quite a bit higher than for the situation with one or more negative eigenvalues.

Our basic aim is to solve (5.42) using a simple Newton-Raphson procedure. To achieve this we define a Hessian matrix according to

$$\tilde{G}_{ij} = \frac{\partial \mathrm{d}\tilde{E}}{\partial \bar{\lambda}_i \mathrm{d}\lambda_j}. \tag{5.43}$$

Note that this is the *ordinary* derivative with respect to the direct parameter. In this way an update may be found, simply by solving

$$\tilde{\mathbf{G}}\boldsymbol{\delta} = -\tilde{\mathbf{g}}. \tag{5.44}$$

This is the basic underlying equation for the optimisation procedures considered in the present work. We will first outline how the gradient and Hessian entering this expression may be evaluated for the case of the spin-coupled optimisation problem, after which further details of the optimisation schemes actually employed will be provided.

For the (ordinary) derivative with respect to a right-hand parameter it is necessary to use the chain rule:

$$\frac{\mathrm{d}}{\mathrm{d}\lambda_i} = \frac{\partial}{\partial\lambda_i} + \sum_j \frac{\partial\tilde{\lambda}_j}{\partial\lambda_i}\frac{\partial}{\partial\tilde{\lambda}_j}. \tag{5.45}$$

In the optimisation we may consider the optimisation of spin coupling coefficients and orbital parameters respectively. The present implementation of this approach is, like the super-CI procedure described in the last section, a two-step procedure with alternating updates of the spin-coupling and orbital parameter sets. We here derive sufficient expressions to make it clear how also a complete second-order procedure could be employed.

Optimisation of the *linear* spin-coupling coefficients is of course little different from the considerations given in section 5.4. As such the optimal coefficients may be determined simply by solving the secular problem. In the context of an iterative orbital optimisation procedure, however, step-size control is likely to be imperative (see section 1.3).

As mentioned in the last section, the biorthogonal Hamiltonian of the spin-coupled structures is symmetric. The secular problem (cf. section 5.4) may thus be solved by a unitary transformation, or: identical transformations of the direct and dual structure basis. So there is no difference between the two sets of spin-coupling coefficients, and (5.7) reduces to the trivial

$$\frac{\mathrm{d}}{\mathrm{d}c_{Sk}} = \frac{\partial}{\partial c_{Sk}} + \frac{\partial}{\partial\tilde{c}_{Sk}} = 2\frac{\partial}{\partial c_{Sk}}. \tag{5.46}$$

The derivative of the biorthogonal energy with respect to a spin coupling coefficient is (similarly to the discussion in section 3.2) given as

$$\frac{\partial\tilde{E}}{\partial c_{Sk}} = \langle\tilde{\Psi}_{\mathrm{SC}}|\hat{H}|\Psi_k\rangle \tag{5.47}$$

with

$$\Psi_k = \hat{\mathcal{A}}\{\phi_1\phi_2\cdots\phi_N\Theta^N_{SM;k}\}. \tag{5.48}$$

In order to obtain expressions for the derivatives with respect to the orbital parameters, we consider again simple updates of the form

$$\phi_\mu \rightarrow \phi_\mu + \lambda\phi_\eta. \tag{5.49}$$

It is useful to distinguish between what might be termed 'internal' updates of the form

$$\phi_\mu \rightarrow \phi_\mu + c_\mu^\nu \phi_\nu, \tag{5.50}$$

and orbital updates utilising the virtual orbital space

$$\phi_\mu \rightarrow \phi_\mu + c_\mu^\alpha \phi_\alpha. \tag{5.51}$$

The former type of update is the simpler, since the change in the dual orbital set is just

$$\tilde{\phi}_\nu \rightarrow \tilde{\phi}_\nu - c_\mu^\nu \tilde{\phi}_\mu. \tag{5.52}$$

This may be verified by confirming that the dual set defined in this way does in fact satisfy the biorthogonality condition (5.5). As shown in ref. [4], if the direct orbital set on one hand and the space spanned by the dual set on the other are both defined, the biorthogonal set is unique. I.e., only one set of dual functions exists that satisfies (5.5). This leads to the expression

$$\frac{\partial \tilde{c}_\nu^\mu}{\partial c_\mu^\nu} = -1, \tag{5.53}$$

with the remaining such derivatives all being zero.

For the second type of update the change in the dual set, by use of (5.6), can be shown to be

$$\tilde{\phi}_\nu \rightarrow \tilde{\phi}_\nu + c_\mu^\alpha (\mathbf{s}^{-1})_{\nu\mu} \phi_\alpha \tag{5.54}$$

to first order in c_μ^α. There will also be higher-order terms due to the change in the direct overlap matrix of the form

$$\mathbf{s}_{\mu\mu} \rightarrow \mathbf{s}_{\mu\mu} + (c_\mu^\alpha)^2, \tag{5.55}$$

but these are not required for our purposes. This then gives

$$\left[\frac{\partial \tilde{c}_\nu^\alpha}{\partial c_\mu^\alpha}\right]_{c_\mu^\alpha = 0} = (\mathbf{s}^{-1})_{\nu\mu}. \tag{5.56}$$

Using this first to obtain expressions for the ordinary gradient, $\mathbf{g}$, gives

$$g_{c_\mu^\nu} = \left[\frac{\partial \tilde{E}}{\partial c_\mu^\nu}\right]_{c_\mu^\nu = 0} - \left[\frac{\partial \tilde{E}}{\partial \tilde{c}_\nu^\mu}\right]_{\tilde{c}_\nu^\mu = 0} \tag{5.57}$$

and

$$g_{c_\mu^\alpha} = \left[\frac{\partial \tilde{E}}{\partial c_\mu^\alpha}\right]_{c_\mu^\alpha=0} + \sum_\nu (\mathbf{s}^{-1})_{\nu\mu} \left[\frac{\partial \tilde{E}}{\partial \tilde{c}_\nu^\alpha}\right]_{\tilde{c}_\nu^\alpha=0}. \tag{5.58}$$

Similarly, for the Hessian matrix, $\tilde{\mathbf{G}}$, defined as in (5.43) one gets

$$\tilde{G}_{\tilde{c}_\mu^\eta c_\nu^\sigma} = \left[\frac{\partial^2 \tilde{E}}{\partial \tilde{c}_\mu^\eta \partial c_\nu^\sigma}\right]_{\tilde{c}_\mu^\eta=0, c_\nu^\sigma=0} - \left[\frac{\partial^2 \tilde{E}}{\partial \tilde{c}_\mu^\eta \partial \tilde{c}_\sigma^\nu}\right]_{\tilde{c}_\mu^\eta=0, c_\sigma^\nu=0} \tag{5.59}$$

and

$$\tilde{G}_{\tilde{c}_\mu^\eta c_\nu^\alpha} = \left[\frac{\partial^2 \tilde{E}}{\partial \tilde{c}_\mu^\eta \partial c_\nu^\alpha}\right]_{\tilde{c}_\mu^\eta=0, c_\nu^\alpha=0} + \sum_\sigma (\mathbf{s}^{-1})_{\sigma\nu} \left[\frac{\partial^2 \tilde{E}}{\partial \tilde{c}_\mu^\eta \partial \tilde{c}_\sigma^\alpha}\right]_{\tilde{c}_\mu^\eta=0, \tilde{c}_\sigma^\alpha=0}. \tag{5.60}$$

The first (partial) derivatives of the biorthogonal energy with respect to the orbital parameters may be expressed as

$$\left[\frac{\partial \tilde{E}}{\partial c_\mu^\eta}\right]_{c_\mu^\eta=0} = \langle \tilde{\Psi}_{\mathrm{SC}} | \hat{H} | \hat{E}_{\eta\mu}^{(1)} \Psi_{\mathrm{SC}} \rangle \tag{5.61}$$

and

$$\left[\frac{\partial \tilde{E}}{\partial \tilde{c}_\mu^\eta}\right]_{\tilde{c}_\mu^\eta=0} = \langle \hat{E}_{\tilde{\eta}\tilde{\mu}}^{(1)} \tilde{\Psi}_{\mathrm{SC}} | \hat{H} | \Psi_{\mathrm{SC}} \rangle. \tag{5.62}$$

For the second derivatives required we have

$$\left[\frac{\partial^2 \tilde{E}}{\partial \tilde{c}_\mu^\eta \partial c_\nu^\theta}\right]_{\tilde{c}_\mu^\eta=0, c_\nu^\theta=0} = \langle \hat{E}_{\tilde{\eta}\tilde{\mu}}^{(1)} \tilde{\Psi}_{\mathrm{SC}} | \hat{H} | \hat{E}_{\theta\nu}^{(1)} \Psi_{\mathrm{SC}} \rangle \quad \text{for} \quad (\mu\eta) \neq (\nu\theta), \tag{5.63}$$

$$\left[\frac{\partial^2 \tilde{E}}{\partial \tilde{c}_\mu^\eta \partial c_\mu^\eta}\right]_{\tilde{c}_\mu^\eta=0, c_\mu^\eta=0} = \langle \hat{E}_{\tilde{\eta}\tilde{\mu}}^{(1)} \tilde{\Psi}_{\mathrm{SC}} | \hat{H} | \hat{E}_{\eta\mu}^{(1)} \Psi_{\mathrm{SC}} \rangle - \langle \tilde{\Psi}_{\mathrm{SC}} | \hat{H} | \Psi_{\mathrm{SC}} \rangle \tag{5.64}$$

and

$$\left[\frac{\partial^2 \tilde{E}}{\partial \tilde{c}_\mu^\eta \partial \tilde{c}_\nu^\theta}\right]_{\tilde{c}_\mu^\eta=0, \tilde{c}_\nu^\theta=0} = \langle \hat{E}_{\tilde{\eta}\tilde{\mu},\tilde{\theta}\tilde{\nu}}^{(2)} \tilde{\Psi}_{\mathrm{SC}} | \hat{H} | \Psi_{\mathrm{SC}} \rangle, \tag{5.65}$$

with $\hat{E}_{\tilde{\eta}\tilde{\mu},\tilde{\theta}\tilde{\nu}}^{(2)}$ describing the corresponding double excitation. All these derivatives were evaluated by expanding the direct and dual wavefunctions in terms of Slater determinants. In this case it is simple to identify the respective values by using the Slater-Condon rules for matrix element evaluation.

For the 'pure' Newton-Raphson scheme, employing 5.44 directly, the convergence characteristics were very encouraging indeed. Trust region control is important for the stability of the optimisation procedure (as discussed in section 1.3), and in order to implement this, some measure of the 'success' of an

update is in general needed. For the 'pure' Newton-Raphson, the quantity

$$k = \sum_i |\tilde{g}_i| \tag{5.66}$$

was defined (the *norm* of this gradient would be an alternative choice). In this fashion, updates that increase k may be rejected, and the improvement of k may be contrasted to that predicted by the second order model. In this way the trust region size may be adjusted in much the same fashion as for variational procedures.

The pure Newton-Raphson scheme will not in general converge onto the lowest-energy solution. To amend this fact, a modification in the spirit of the GQT method described in section 1.3 has proven very successful. The GQT method may be viewed as a linear combination of a steepest-descent approach and the Newton-Raphson scheme. In the variational variant of this procedure this linear combination is made such as to favour updates that obtain a positive definite Hessian, but this is not applicable to the biorthogonal case. Instead we have had good experience with the simple combination

$$\boldsymbol{\delta}' = \alpha(\boldsymbol{\delta} - \beta \mathbf{g}). \tag{5.67}$$

Here $\boldsymbol{\delta}$ is obtained from (5.44), β is adjusted during the optimisation procedure, and α is determined so as to obtain an update within the trust region size. An inherent problem with this, as for any non-variational procedure, is how to determine trust sphere sizes and values for the parameter β. The pure steepest-descent method, when applied to biorthogonal orbital optimisation, obtains essentially arbitrarily small energies, this generally being associated with linear dependence in the direct orbital set. Thus if β above is chosen too big this kind of divergence may occur. On the other hand if the value for β is too small, convergence onto an excited state is likely. We have not found this to be a major problem, there seems to be scope for quite a bit of variation in the choice of β. A feasible starting value in many cases, is given by

$$\beta^2 = \frac{|\boldsymbol{\delta}|^2}{|\mathbf{g}|^2}. \tag{5.68}$$

Close to a stationary point the modification of the pure Newton-Raphson scheme will of course impede convergence. When k defined in (5.66) becomes sufficiently small, the value of β is thus adjusted down and eventually set to zero.

A consideration we have not touched upon so far is the cost of integral transformations during the optimisation procedure. Implementing the algorithm outlined above literally, would require a full transformation of all two-electron integrals every iteration which would soon become computationally prohibitive. To circumvent this we consider updates of the form

$$\phi_\mu \rightarrow \phi_\mu + \lambda\chi_p, \tag{5.69}$$

where χ_p is a generic basis function. To obtain the the orbital gradient, the two following partial derivatives are then required:

$$\left[\frac{\partial\tilde{E}}{\partial c_\mu^p}\right]_{c_\mu^p=0}$$

and

$$\left[\frac{\partial\tilde{E}}{\partial\tilde{c}_\mu^p}\right]_{\tilde{c}_\mu^p=0}.$$

Similarly, for the evaluation of the Hessian partial derivatives of the forms

$$\left[\frac{\partial^2\tilde{E}}{\partial\tilde{c}_\mu^p\partial c_\nu^q}\right]_{\tilde{c}_\mu^p=0,c_\nu^q=0}$$

and

$$\left[\frac{\partial^2\tilde{E}}{\partial\tilde{c}_\mu^p\partial\tilde{c}_\nu^q}\right]_{\tilde{c}_\nu^q=0}$$

are needed. Once these quantities are evaluated it is trivial to transform them to obtain expressions for the partial and ordinary gradient defined in terms of occupied/virtual orbitals. In the evaluation the following partially transformed integrals will be required

- One-electron integrals: $\langle\chi_p|\hat{H}_1|\phi_\mu\rangle$ and $\langle\tilde{\phi}_\mu|\hat{H}_1|\chi_p\rangle$.
- Two-electron integrals: $\langle\chi_p\tilde{\phi}_\mu|\hat{H}_2|\phi_\nu\phi_\sigma\rangle$, $\langle\tilde{\phi}_\mu\tilde{\phi}_\nu|\hat{H}_2|\chi_p\phi_\sigma\rangle$, $\langle\chi_p\chi_q|\hat{H}_2|\phi_\mu\phi_\nu\rangle$ and $\langle\chi_p\tilde{\phi}_\mu|\hat{H}_2|\chi_q\phi_\nu\rangle$.

This may be compared with the similar quantities needed in the (variational) optimisation of the spin-coupled wavefunction (see section 1.3). Also for the present case there is little extra cost involved in obtaining the partially transformed integrals since the fully transformed integrals are normally constructed by transforming one index at a time. Thus the cost of the integral transformations is comparable to that for standard spin-coupled calculations.

5.8 Triplet methylene

A series of test calculations were performed on the triplet state of the CH_2 molecule. CH_2 has C_{2v} symmetry; the geometry was taken to be $\angle$(H–C–H)=130° and r(C–H)=1.082Å. For the basis set a standard double zeta (DZ) set [18] was chosen, consisting of (9s5p/4s) gaussians contracted to [4s2p/2s] for C/H. The lowest-lying SCF MO (corresponding approximately to $1s^2$ on C) was kept doubly occupied (frozen core), while the 6 valence electrons were active, the corresponding orbitals expanded in the basis of the remaining 13 MOs.

The spin-coupled orbitals for this system are shown in figure 5.1 (six positive contours were plotted). The bonding orbitals are plotted in the molecular plane, the non-bonding in the perpendicular plane. The corresponding overlaps are given in table 5.1. The description of both C–H bonds resembles what was found for the singlet state of this molecule (section 4.1), i.e., sp^x-type hybrids on carbon pointing towards deformed 1s orbitals on hydrogen. (Since SP2 and H2 are related to SP1 and H1 by symmetry, they have not been included in figure 5.1.) The corresponding electron pairs are very nearly singlet-coupled. The non-bonding orbitals, LP1 and LP2, are in this case slightly different. LP1 is an in-plane sp^x-type hybrid, whereas LP2 is a π-type orbital (of B_1 symmetry). The electrons associated with the non-bonding orbitals are in this case very nearly triplet-coupled.

The spin coupling for CH_2 is shown in table 5.2. As for methane, the Serber

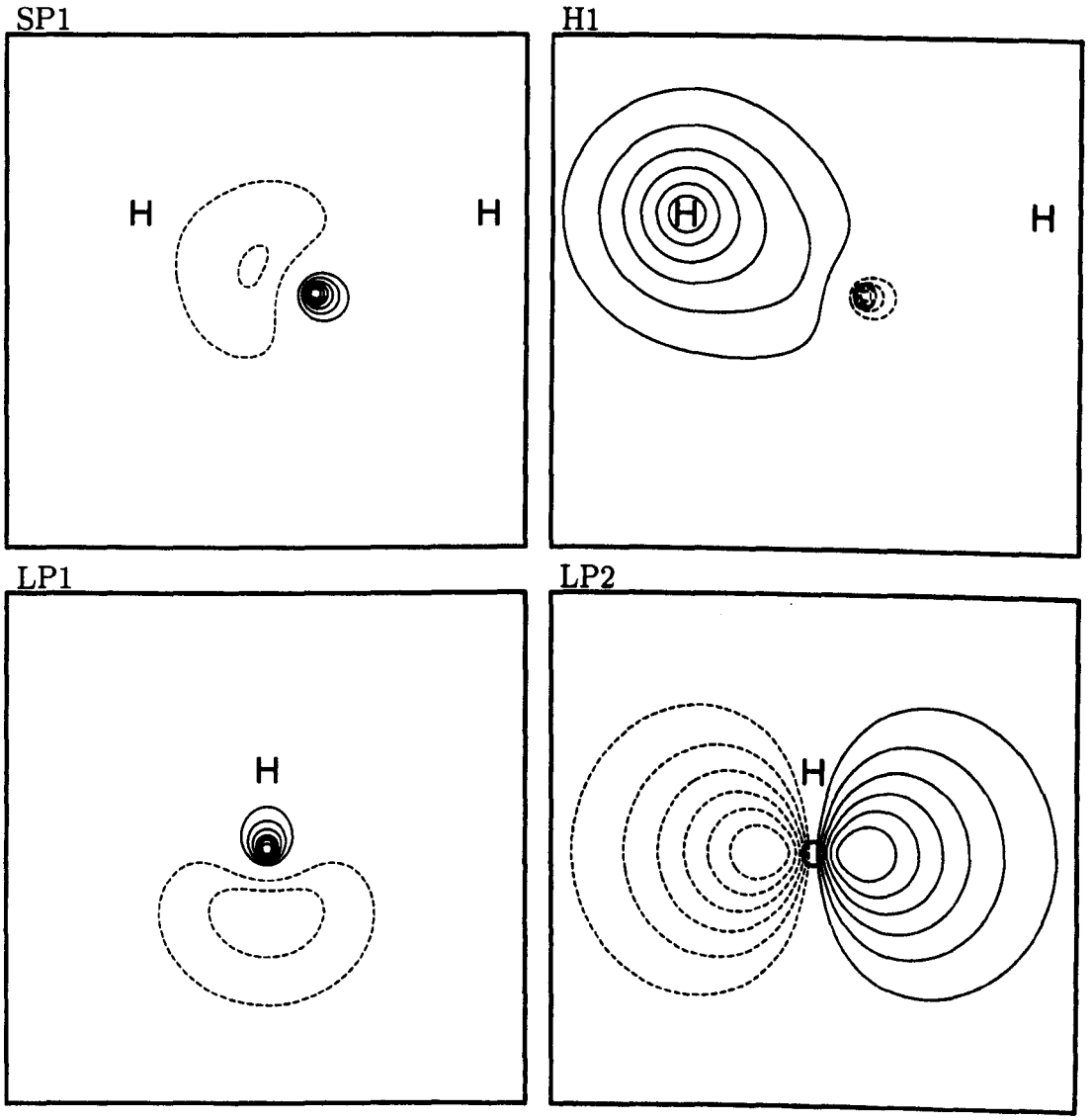

Figure 5.1: Spin-coupled orbitals of CH_2 (triplet).

	⟨SP1|H1⟩	⟨SP1|SP2⟩	⟨SP1|H2⟩	⟨SP1|LP1⟩	⟨H1|H2⟩	⟨H1|LP1⟩
SC	0.79257	0.50835	0.19120	0.44087	-0.05886	0.21367
BO	0.82807	0.32601	0.18761	0.30293	0.11655	0.15005

Table 5.1: Non-zero, symmetry-unique overlap integrals for CH_2 triplet.

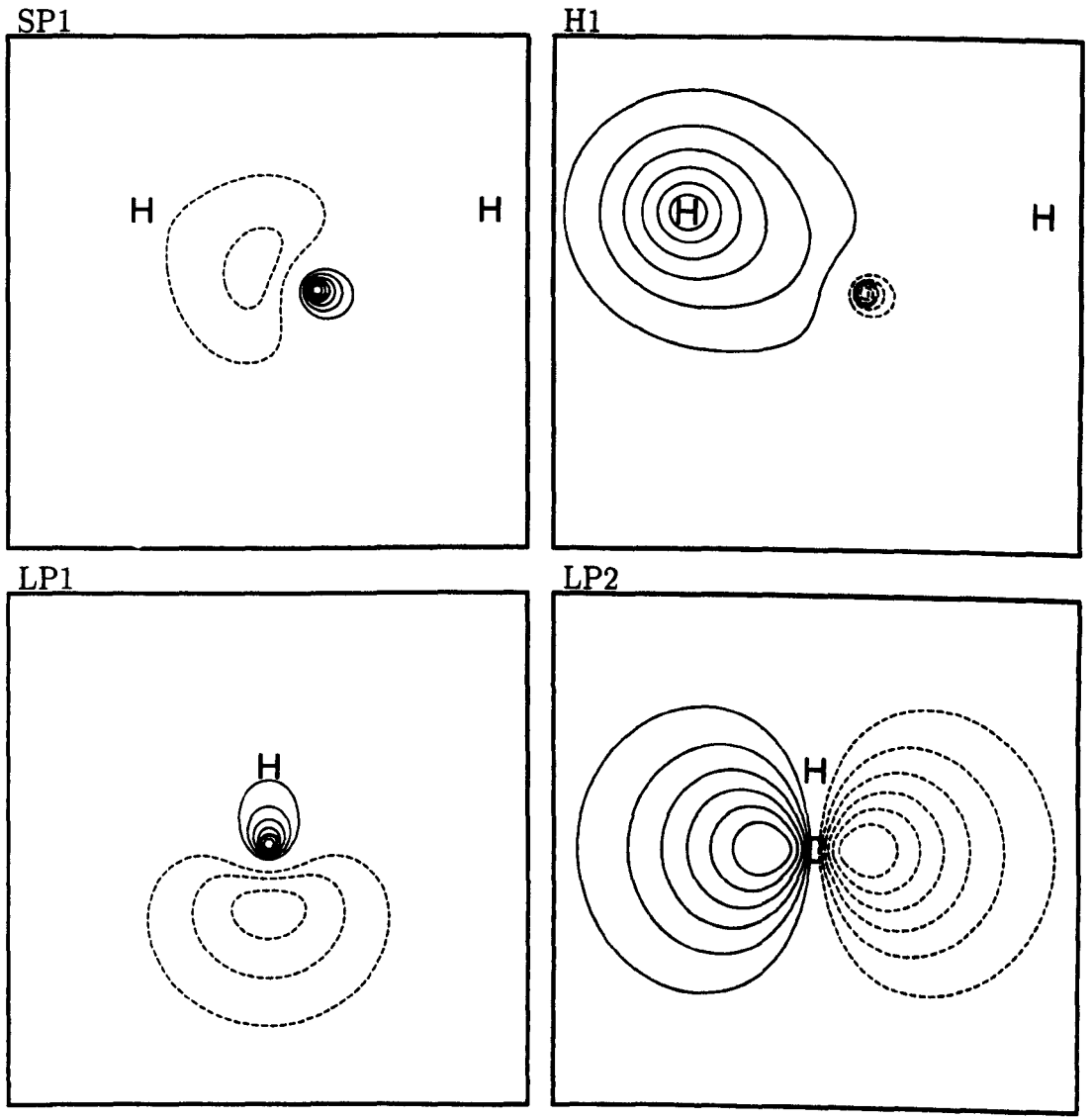

Figure 5.2: Biorthogonal orbitals of CH_2 (triplet).

	SC	BO
((11)2;1)	0.07%	0.03%
((10)1;0)	(−)0.06%	(−)0.05%
((01)1;0)	(−)0.06%	(−)0.05%
((11)1;0)	0.00%	0.00%
((10)1;1)	(−)2.59%	(−)3.04%
((01)1;1)	(−)2.59%	(−)3.04%
((11)1;1)	0.00%	0.00%
((11)0;1)	(−)0.74%	(−)0.40%
((00)0;1)	93.90%	93.38%

Table 5.2: Weights of Serber spin functions for the SC and BO wavefunctions (phase in brackets).

basis is convenient for highlighting the relationship between symmetry-equivalent functions. The perfect-pairing mode of spin coupling is dominating, and adding the modes where only one of the SP–H pairs is triplet-coupled accounts for 99.1% of the total spin function. The tendency for the non-bonding orbitals to couple to a singlet is very small indeed.

The biorthogonal orbitals for CH_2 are shown in figure 5.2. As can be seen there is quite good qualitative agreement between the spin-coupled and biorthogonal orbitals. The bonding orbitals exhibit a slightly greater deformation towards each other, which is reflected in their greater overlap, but this is in spite of a smaller weight of the perfect-pairing function for the biorthogonal wavefunction. Also the non-bonding sp^x-type hybrid is slightly more diffuse for BO, whereas the two π orbitals are very similar.

The biorthogonal wavefunction had at convergence 9 negative eigenvalues of the orbital-orbital Hessian defined according to

$$G_{\tilde{c}_\mu^\eta c_\nu^\theta} = \left[\frac{\partial^2 \tilde{E}}{\partial \tilde{c}_\mu^\eta \partial c_\nu^\theta}\right]_{\tilde{c}_\mu^\eta=0, c_\nu^\theta=0} . \tag{5.70}$$

(Since the spin-spin Hessian is symmetric the lowest root must of course be the first.) This would correspond to convergence onto root number 10 in the super-CI strategy. This illustrates one of the problems in the super-CI strategy that a number of roots may have to be investigated before the correct solution is obtained.

To illustrate the relationship between the SC and BO wavefunctions further, we show here Hamiltonian and overlap matrices evaluated between the structures Ψ_{SC}, Ψ_{BO}, $\tilde{\Psi}_{SC}$ and $\tilde{\Psi}_{BO}$ respectively:

$$\mathbf{H} = \begin{pmatrix} -38.95026 & -38.88167 & -1.15049 & -1.18677 \\ -38.88167 & -38.94768 & -1.18476 & -1.22989 \\ -1.15049 & -1.18476 & -35.19663 & -34.73075 \\ -1.18677 & -1.22989 & -34.73075 & -35.15383 \end{pmatrix}$$

$$\mathbf{S} = \begin{pmatrix} 1 & 0.99824 & 0.02950 & 0.03043 \\ 0.99824 & 1 & 0.03043 & 0.03159 \\ 0.02950 & 0.03043 & 1 & 0.98731 \\ 0.03043 & 0.03159 & 0.98731 & 1 \end{pmatrix}$$

As can be seen the quality of the biorthogonal wavefunction is quite good, it is only separated from the spin-coupled wavefunction by some 3 millihartree. One must say that the dual wavefunctions are exceptionally poor energy-wise. Their purpose is to facilitate the matrix element evaluation, but they can be attributed no real physical meaning. The similarity of the spin-coupled and biorthogonal wavefunctions is also indicated by their large overlap—0.998. Comparing equivalent orbitals for the two cases much the same conclusion emerged, the smallest overlap being 0.989 between the two equivalent LP1 orbitals.

From the matrices above we may also obtain the biorthogonal energies for the spin-coupled and biorthogonal wavefunctions, as

$$\tilde{E}_{SC} = \mathbf{H}(3,1)/\mathbf{S}(3,1) = -1.15049/0.02950 \text{ hartree} = -38.99403 \text{ hartree}$$

and

$$\tilde{E}_{BO} = \mathbf{H}(4,2)/\mathbf{S}(4,2) = -1.22989/0.03159 \text{ hartree} = -38.93630 \text{ hartree}.$$

That these values agreed with those obtained using the biorthogonal energy evaluation was a reassuring check of the new program. It is clear that the biorthogonal energy for the biorthogonal wavefunction is rather more realistic than the corresponding energy for the spin-coupled wavefunction. In an attempt to understand the underlying reasons for this we follow the error analysis by Boys [10].

We note that the 'error', measured as the length of the difference vector, is related to the overlap integral according to

$$|\Delta| = \sqrt{2 - 2|S|} \tag{5.71}$$

for normalised wavefunctions. With the spin-coupled orbitals (full CI in the space of the 6 orbitals) an energy of $E_{\text{SC+CI}}$=−38.95375 hartree is obtained. The error in the spin-coupled wavefunction is

$$\langle \Psi_{ex} | \Psi_{SC} \rangle = 0.99897 \rightarrow |\Delta| = 0.04541$$

and in its dual

$$\langle \tilde{\Psi}_{ex} | \tilde{\Psi}_{SC} \rangle = 0.03016 \rightarrow |\tilde{\Delta}| = 1.39167.$$

Using these values then gives estimates for the proportionality constant, ϵ as may be defined from equation (5.30), as ϵ=1.72, ϵ=1.94 and ϵ=−0.64, for the spin-coupled wavefunction, its dual, and the biorthogonal energy respectively. The biorthogonal energy for the spin-coupled wavefunction is in this case *lower* than the full-CI energy, so ϵ must thus be negative.

We may also perform this analysis for the biorthogonal wavefunction. The energy for the full-CI wavefunction in this case was $E_{\text{BO+CI}}$=−38.95285 hartree. The error in the biorthogonal wavefunction is

$$\langle \Psi_{ex} | \Psi_{SC} \rangle = 0.99777 \rightarrow |\Delta| = 0.06680$$

and in its dual

$$\langle \tilde{\Psi}_{ex} | \tilde{\Psi}_{SC} \rangle = 0.03094 \rightarrow |\tilde{\Delta}| = 1.39216.$$

This then gives ϵ=1.16, ϵ=1.96 and ϵ=0.18 for the three cases respectively.

It seems that these variations in ϵ do not conform to any logical patterns, and one must conclude that the errors in the optimised wavefunctions are too big to make such analysis very useful.

5.9 Naphthalene

As an example of a system with slightly more active electrons we have chosen the π-electron system of naphthalene, $C_{10}H_8$, (singlet). The considerations regarding the validity of σ-π separation are of course analogous to the case of benzene. Naphthalene is planar with D_{2h} symmetry, the precise geometry used in this work was as determined by rotational spectroscopy in ref. [19]. The same basis set as for CH_2 (triplet) was employed [18]. There were 29 core orbitals and the 10 active orbitals were expanded in 20 functions of π symmetry.

The spin-coupled description of naphthalene has previously been reported in ref. [20]. The symmetry-unique spin-coupled and biorthogonal orbitals are shown in figure 5.3 (8 positive contours were plotted 1 bohr above the molecular plane). The description is reminiscent of that found in benzene with highly localised orbitals associated with each carbon atom. Each shows slight deformation towards its nearest neighbours.

We remark that for naphthalene the Hessian defined as in (5.70) had 10 negative eigenvalues. Thus if a super-CI scheme was employed, the solution would at convergence be root number 11.

The biorthogonal orbitals seem also in this case to be a very good approximation to the spin-coupled orbitals. Particularly ϕ_{10} seems to be slightly more localised for the BO orbital. This may be verified by examining the overlap matrix, table 5.3 (the orbitals are here numbered consecutively round the ring). We have in this case chosen to present the spin coupling as the expectation values of $(\hat{s}_\mu+\hat{s}_\nu)^2$, between the symmetry-unique orbital pairs—table 5.4, but we have also

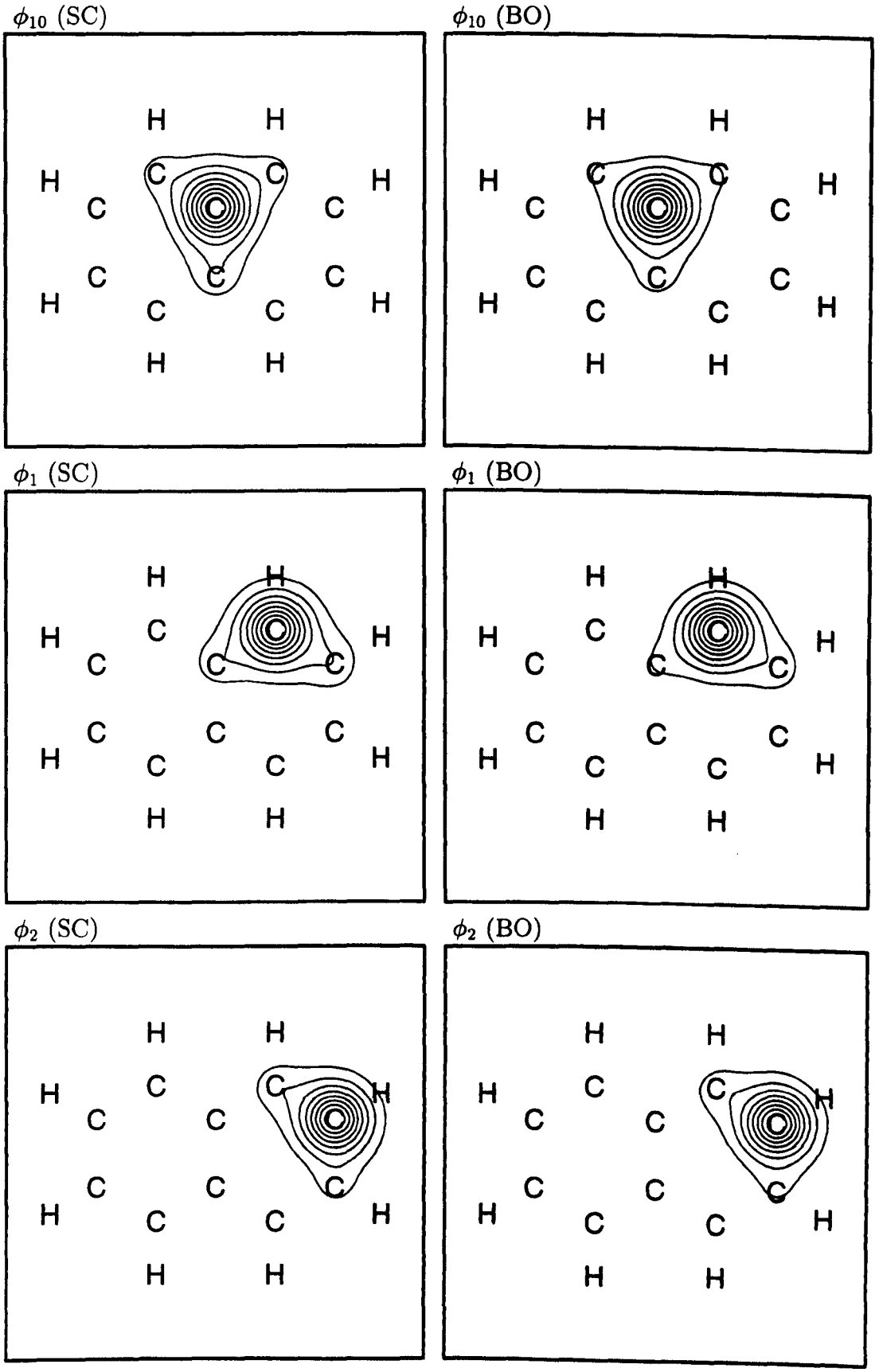

Figure 5.3: SC and BO orbitals for naphthalene.

	$\langle 1\|2\rangle$	$\langle 1\|3\rangle$	$\langle 1\|4\rangle$	$\langle 1\|5\rangle$	$\langle 1\|6\rangle$
SC	0.60682	0.06326	-0.02668	0.06780	-0.07032
BO	0.60602	0.11714	0.04941	0.10290	-0.04997
	$\langle 1\|7\rangle$	$\langle 1\|8\rangle$	$\langle 1\|9\rangle$	$\langle 1\|10\rangle$	$\langle 2\|3\rangle$
SC	-0.05099	0.00330	0.09933	0.42746	0.39764
BO	-0.01871	0.02654	0.07784	0.36890	0.38036
	$\langle 2\|4\rangle$	$\langle 2\|5\rangle$	$\langle 2\|6\rangle$	$\langle 2\|7\rangle$	$\langle 2\|8\rangle$
SC	0.06326	-0.12157	-0.05099	0.01706	0.02213
BO	0.11714	0.01615	-0.01871	-0.00628	0.00879
	$\langle 2\|9\rangle$	$\langle 2\|10\rangle$	$\langle 3\|10\rangle$	$\langle 4\|10\rangle$	$\langle 5\|10\rangle$
SC	0.00330	0.04319	-0.12157	0.06780	0.59086
BO	0.02654	0.10208	0.01615	0.10290	0.58812

Table 5.3: Symmetry-unique overlap integrals for naphthalene.

	$\langle 1\|2\rangle$	$\langle 1\|3\rangle$	$\langle 1\|4\rangle$	$\langle 1\|5\rangle$	$\langle 1\|6\rangle$
SC	0.30660	1.85126	1.10894	1.79468	1.40106
BO	0.23589	1.86780	1.06717	1.83922	1.32013
	$\langle 1\|7\rangle$	$\langle 1\|8\rangle$	$\langle 1\|9\rangle$	$\langle 1\|10\rangle$	$\langle 2\|3\rangle$
SC	1.61266	1.35912	1.65881	0.90687	0.86627
BO	1.66733	1.31664	1.68929	0.99652	0.95098
	$\langle 2\|4\rangle$	$\langle 2\|5\rangle$	$\langle 2\|6\rangle$	$\langle 2\|7\rangle$	$\langle 2\|8\rangle$
SC	1.85126	1.20556	1.61266	1.36278	1.61375
BO	1.86780	1.12839	1.66733	1.32918	1.66047
	$\langle 2\|9\rangle$	$\langle 2\|10\rangle$	$\langle 3\|10\rangle$	$\langle 4\|10\rangle$	$\langle 5\|10\rangle$
SC	1.35912	1.82199	1.20556	1.79468	0.54180
BO	1.31664	1.84332	1.12839	1.83922	0.38508

Table 5.4: Symmetry-unique expectation values of $(\hat{s}_\mu+\hat{s}_\nu)^2$ based on the total spin function.

	SC	BO
${}^{R}\Theta_1$: (1–2,3–4,5–6,7–8,9–10)	13.96%	3.15%
${}^{R}\Theta_{20}$: (1–2,3–4,5–10,6–7,8–9)	30.95%	32.36%
${}^{R}\Theta_{23}$: (1–10,2–3,4–5,6–7,8–9)	13.96%	3.15%

Table 5.5: Chirgwin-Coulson weights of the 'Kekulé' spin functions for naphthalene.

given the Chirgwin-Coulson weights of the three Kekulé-type Rumer functions—table 5.5. The smaller overlaps between the bond-forming orbitals in the case of BO is contradicted by the larger tendency towards singlet coupling between these orbitals, i.e., smaller values for $\langle(\hat{s}_\mu+\hat{s}_\nu)^2\rangle$. This is the reverse of what was found for CH_2 where larger bond-forming overlaps were associated with a smaller weight of the perfect-pairing mode of spin coupling.

We have again constructed Hamiltonian and overlap matrices in the basis of Ψ_{SC}, Ψ_{BO}, $\tilde{\Psi}_{SC}$ and $\tilde{\Psi}_{BO}$ respectively:

$$\mathbf{H} = \begin{pmatrix} -383.34272 & -373.43529 & -1.41123 & -2.44918 \\ -373.43529 & -383.32426 & -2.44677 & -4.05072 \\ -1.41123 & -2.44677 & -379.81846 & -369.13665 \\ -2.44918 & -4.05072 & -369.13665 & -379.96811 \end{pmatrix}$$

$$\mathbf{S} = \begin{pmatrix} 1 & 0.97415 & 0.00368 & 0.00639 \\ 0.97415 & 1 & 0.00638 & 0.01057 \\ 0.00368 & 0.00638 & 1 & 0.97170 \\ 0.00639 & 0.01057 & 0.97170 & 1 \end{pmatrix}$$

The separation between the two wavefunctions is 18 millihartree which is somewhat disappointing, but a likely consequence of the biorthogonal approximation for this number of active electrons. The overlap between the two wavefunctions, 0.97415, might also have been higher. The individual overlaps between equivalent orbitals were slightly better, the smallest being 0.99328 between the central

orbitals. The biorthogonal energies for the two wavefunctions are

$$\tilde{E}_{SC} = \mathbf{H}(3,1)/\mathbf{S}(3,1) = -383.57261 \text{ hartree}$$

and

$$\tilde{E}_{BO} = \mathbf{H}(4,2)/\mathbf{S}(4,2) = -383.22066 \text{ hartree.}$$

The uncertainty in these values is of course unacceptable, so it seems imperative that the biorthogonal energy expression should be used *only* during the optimisation.

5.10 Discussion

A fairly substantial amount of work has been invested in evolving an acceptable optimisation procedure, and we believe that this has been achieved to a large extent. Certain inescapable problems are connected with the fact that the optimisation is not a simple minimisation problem, but these may be circumvented fairly straightforwardly. One should, however, exercise some critical sense in verifying that the state obtained is in fact reasonable, but this probably holds true for any optimisation procedure.

In this work we have assumed that the optimal biorthogonal root would have the lowest biorthogonal energy. This has been true for the examples given here, but in view of the large discrepancies between the energies evaluated using the variational and biorthogonal expressions, this may not always be the case. In practice the nature of the orbitals and spin coupling are likely to suggest the existence of lower solutions. A possible way to take this into account could be to search for the stationary point, as defined by (5.41), that has lowest *standard* energy. These roots might for example be classified by the number of negative eigenvalues of the biorthogonal Hessian, and this would suggest schemes more in the spirit of traditional GQT schemes in order to locate different solutions.

The biorthogonal optimisation criterion seems to have been fairly successful for the cases treated here. This is in spite of the quite large variations in the

biorthogonal energy between similar wavefunctions. The biorthogonal approximation may not in all cases give quantitatively satisfactory results particularly for larger numbers of active electrons, but in view of the small computational cost the results are still very encouraging. It is likely to be an excellent way of getting qualitative descriptions for systems outside the scope of the variational spin-coupled or CAS methods. The biorthogonal method may be the only viable way of obtaining spin-coupled-like pictures for 16 or 18 active electrons and the full spin space.

5.a Appendix: Biorthogonality as a notational tool

This discussion does not fit in with the main test of this chapter, but it nevertheless seems useful to point out how biorthogonality may be utilised, not as a method of computational convenience, but as a convenient notation when non-orthogonality is encountered.

As pointed out in [21], a Hamiltonian matrix, say, may be defined in two quite distinctly different manners. The most commonly used definition is according to

$$(\mathbf{H}^{\mathbf{S}})_{ij} = \langle \Phi_i | \hat{H} | \Phi_j \rangle. \tag{5.72}$$

In many cases however, the Hamiltonian matrix defined as the transformation matrix of the set of functions $\{\Phi\}$, i.e., according to

$$\{\hat{H}\Phi\} = \{\Phi\}\mathbf{H}, \tag{5.73}$$

will be required. For the case of orthogonal functions, the two definitions coincide, but this seems to be the source of no little confusion for the corresponding case of non-orthogonal functions. The general relationship between the two definitions is

$$\mathbf{H} = \mathbf{S}^{-1}\mathbf{H}^{\mathbf{S}}. \tag{5.74}$$

So by defining a dual set of functions in the usual manner,

$$\{\tilde{\Phi}\} = \{\Phi\}\mathbf{S}^{-1}, \tag{5.75}$$

it is possible to define **H** also in the 'standard' form via its matrix elements according to

$$(\mathbf{H})_{ij} = \langle\tilde{\Phi}_i|\hat{H}|\Phi_j\rangle. \tag{5.76}$$

To give just one example where such a strategy may be very useful, we take the second quantised formalism which is widespread in the standard MCSCF literature. In the standard fashion we define the set of creation and annihilation operators corresponding to the spin orbital set $\{\varphi\}$ according to

$$a_r^\dagger|\ \rangle = |\varphi_r\rangle \tag{5.77}$$

and

$$a_r|\varphi_r\rangle = |\ \rangle. \tag{5.78}$$

The anticommutation relations

$$\{a_r^\dagger, a_s^\dagger\} = 0 \tag{5.79}$$

and

$$\{a_r, a_s\} = 0 \tag{5.80}$$

are a simple consequence of the antisymmetry of the wavefunction. For the relationship between a creation and an annihilation operator we have

$$\langle\ |a_r a_s^\dagger|\ \rangle = \langle\varphi_r|\varphi_s\rangle = \mathbf{s}_{rs}, \tag{5.81}$$

which leads to the anticommutation relation

$$\{a_r, a_s^\dagger\} = \mathbf{s}_{rs}. \tag{5.82}$$

With this definition, many of the second quantised relations normally used for orthogonal orbitals will hold also for non-orthogonal orbitals. The energy expectation value, for example, may be evaluated using a Hamiltonian on second

quantised form according to

$$\hat{H}^{\text{2nd}} = \sum_{rs} \langle \varphi_r | \hat{H}_1 | \varphi_s \rangle a_r^\dagger a_s + \frac{1}{2} \sum_{rstu} \langle \varphi_r \varphi_s | \hat{H}_2 | \varphi_t \varphi_u \rangle a_r^\dagger a_s^\dagger a_t a_u. \tag{5.83}$$

This formulation is of course entirely equivalent to what may be derived from considering standard Hamiltonian matrix elements between antisymmetric wavefunctions. As such the Löwdin formula may be derived from the anticommutation relation (5.82) directly.

The above definitions are useful when we are evaluating just the $\mathbf{H}^{\mathbf{s}}$ representation of the Hamiltonian (or any other operator), but one inadequacy becomes clear if we consider, say, the excitation operator $\hat{E}_{rs}$. With orthogonal orbitals this is defined according to

$$\hat{E}_{rs} = a_r^\dagger a_s. \tag{5.84}$$

It is clear, that if the orbitals are orthogonal, we have

$$\hat{E}_{rs} | \cdots \varphi_s \cdots \rangle = | \cdots \varphi_r \cdots \rangle. \tag{5.85}$$

So $\hat{E}_{rs}$ may be interpreted as an excitation operator in the normal meaning of the word, in that it simply substitutes occurrences of φ_s by φ_r. This will not be the case for non-orthogonal orbitals, however, due to (5.82). In this case it makes more sense to define, as done in ref. [2], the annihilation operator from the *dual* orbital set $\{\tilde{\varphi}\}$ defined in the usual fashion

$$\{\tilde{\varphi}\} = \{\varphi\}\mathbf{s}^{-1}. \tag{5.86}$$

This means that the anticommutation relation analogous to 5.82 becomes identical to the orthogonal case

$$\{\tilde{a}_r, a_s^\dagger\} = \delta_{rs}. \tag{5.87}$$

So with the definition of the excitation operator according to

$$\hat{E}_{rs} = a_r^\dagger \tilde{a}_s, \tag{5.88}$$

the effect is, as in the orthogonal case, the simple substitution of φ_s by φ_r. This definition was used in chapter 2.

Bibliography

[1] F. Prosser and S. Hagstrom, Int. J. Quantum Chem. **2**, 89 (1968).

[2] M. Moshinsky and T. H. Seligman, Ann. Phys. **66**, 311 (1971).

[3] J. M. Norbeck and R. McWeeny, Chem. Phys. Lett. **34**, 206 (1975).

[4] P. W. Payne, J. Chem Phys. **77**, 5630 (1982).

[5] J. J. W. McDouall, Theor. Chim. Acta **83**, 339 (1992).

[6] J. J. W. McDouall, Theor. Chim. Acta **85**, 395 (1993).

[7] S. F. Boys and N. C. Handy, Proc. Roy. Soc. A, **310**, 43 (1969).

[8] C. F. Bender and I. Shavitt, J. Comput. Phys. **6**, 146 (1970).

[9] S. Rettrup, J. Comput. Phys. **45**, 100 (1982).

[10] S. F. Boys, Proc. Roy. Soc. A, **309**, 195 (1969).

[11] F. Grein and T. C. Chang, Chem. Phys. Lett. **12**, 44 (1971).

[12] A. Banerjee and F. Grein, Int. J. Quantum Chem. **10**, 123, (1976).

[13] T. C. Chang and W. H. E. Schwarz, Theor. Chim. Acta **44**, 45, (1977).

[14] J. H. van Lenthe and G. G. Balint-Kurti, Chem. Phys. Lett. **76**, 138 (1980).
J. H. van Lenthe and G. G. Balint-Kurti, J. Chem. Phys. **78**, 5699 (1983).

[15] J. H. van Lenthe, J. Verbeek and P. Pulay, Mol. Phys. **73**, 1159 (1991).

[16] J. Verbeek, PhD Thesis, University of Utrecht, Utrecht, (1990).

[17] G. D. Fletcher, G. Doggett and A. S. Howard, Phys. Rev. A **46**, 5459 (1992).

G. Doggett and G. D. Fletcher, J. Mol. Struc. **260**, 313 (1992).

G. Doggett, G. D. Fletcher and F. R. Manby, J. Mol. Struc. **300**, 191 (1993).

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

[19] K. K. Innes, J. E. Parkin, D. K. Ervin, J. M. Hollas and J. G. Ross, J. Mol. Spec. **16**, 406 (1965).

[20] G. Raos, J. Gerratt, D. L. Cooper and M. Raimondi, Chem. Phys. **186**, 251 (1994).

[21] R. McWeeny, "*Methods of Molecular Quantum Mechanics*", 2. ed., Academic Press, London (1989).

Or: R. McWeeny and B. T. Sutcliffe, "*Methods of Molecular Quantum Mechanics*", Academic Press, London (1969).

Chapter 6

Two-body potentials for modelling ionic solids

6.1 Introduction

Atomistic simulation have proven to be an important tool in the study of ionic solids, especially in areas that are not readily accessible for experiment such as: obtaining defect energies, describing surface effects and obtaining properties of crystals under extreme temperatures and pressures.

The present study concentrates on highly ionic systems of mono-atomic ions, mainly the oxides and halides of the first two main-group metals. Only when a satisfactory understanding of these fairly simple systems is obtained, is it appropriate to raise our ambitions and consider more challenging areas. There is a tendency at present to become over-ambitious very early. Thus attempts at describing systems such as diamond and graphite using ionic models have been made—before universal agreement has been reached on a 'simple' system such as MgO. The comparative success of calculations employing such strategies illustrates the power of present-day methods in describing covalent effects, but a lot of physics is lost from this inappropriate starting point.

In most simulations, energy estimates are required for relatively arbitrary

configurations of ions. A large amount of attention has been given to this, including, to mention just a few approaches, density functional theory (DFT) calculations on the super-molecule, MO super-molecule treatments, techniques based on perturbation theory, finite cluster calculations, periodic Hartree-Fock and Carr-Parrinello schemes.

In a large body of work [1, 2, 3, 4, 5, 6, 7, 8]. the interactions are divided according to the number of interacting species, into pair or two-body interactions, three-body, $\cdots$. By only including interactions up to a certain (normally quite low) order, an estimate of the total energy may be obtained. In this work we concentrate on obtaining suitable pair potentials.

There are inherent problems with this approach which should be pointed out. Although the above expansion is in principle exact if interactions up to N-body are included for the treatment of an N-particle system, there is no *unique* partitioning into the various many-body interactions. More serious, perhaps, is that no convergence in the energy estimate can be assured as more interactions are included in the expansion. For the pair interactions in particular it should be noted that there is no such thing as a 'correct' potential. During a practical calculation, a particular inter-ionic distance may be sampled under a wide range of circumstances (different surrounding geometries, varying degrees of shielding by species between the two ions in question etc.), and the best potential is therefore, more than anything else, the best *compromise* between all these types of interaction.

The pair potential is often written on the form [3]

$$V_{\mathrm{AB}}(R) = \frac{Z_{\mathrm{A}} Z_{\mathrm{B}}}{R} + U^{N}_{\mathrm{AB}}(R) + U^{P}_{\mathrm{AB}}(R; \alpha_{\mathrm{A}}, \alpha_{\mathrm{B}}), \tag{6.1}$$

where R is the inter-ionic distance, and Z_{A} and Z_{B} the (formal) charges of the ions. Since the Coulombic part of this expression in general will be strongly dominant, this term is often excluded for comparison purposes. It is generally most accurate to calculate the electronic polarisation energy directly from the polarisabilities of the ions, α_{A} and α_{B}, as indicated by $U^{P}_{\mathrm{AB}}(R; \alpha_{\mathrm{A}}, \alpha_{\mathrm{B}})$ in equation (6.1). A very

successful method for dealing with this term is represented by the shell model, first introduced by Dick and Overhauser [9], in which a mass-less charged shell is combined with a spring such that, with appropriate charge and spring constant, the correct polarisability is obtained. It is not in the scope of this work to describe further details of the shell model. A large amount of attention has been given in the to the problem of obtaining realistic polarisabilities of ions in crystals [10, 11, 12, 13]. The particular relevance of this work to the present is the necessary attention to the effects of the crystalline environment on the description of the ions. This will be pursued further in section 6.4. We shall not consider the actual calculation of polarisabilities further here, but one thing that it is important to keep in mind when combining a theoretically obtained potential with the shell model (or other models taking into account the electronic polarisation energy) is to ensure that the polarisation energy component is properly excluded from the calculation. In other words, only the quantity U_{AB}^{N} in equation (6.1) should be calculated.

Pair potentials may be obtained in a variety of ways, but may be loosely divided into empirically based methods, and non-empirical. In empirical procedures a functional form of the potential is normally assumed, and the corresponding parameters fitted such that a subsequent simulation would give realistic crystal properties (lattice parameters, phonon frequencies, cohesive energies, elasticity constants etc.). The basic drawback of this approach seems to be the poor transferability of pair potentials to different crystal systems, from bulk to surface properties, or to different temperatures and pressures. Also typical sampling distances should be taken into account. A potential based on bulk properties at normal temperatures and pressures may not be appropriate for the study of interstitial formation, for example. These facts mean that, although many structures may be *rationalised* using empirical potentials, only few actual *predictions* can be made with any degree of confidence.

A theoretical investigation is therefore important to clarify exactly which fac-

tors are important in determining the form of a potential. The basic problem of calculating the interaction energy for two ions has many features in common with intermolecular forces, as will become clear in section 6.3. We shall therefore briefly recapitulate a highly successful approach to intermolecular forces based on the spin-coupled model. Thereafter we consider the actual potential generation, before investigating methods of generating appropriate descriptions of the individual ions.

6.2 Intermolecular forces using a valence bond approach

The area of interest is here the interaction between two or more species that do not form actual chemical bonds. Often one or more of these will be closed-shell entities, but the actual charges of the interacting systems will not be of particular relevance to the development of the theory. Since the interactions considered here are very weak, especially in the intermediate and long range, it seems obvious that an approach based on unperturbed *fragments* as a starting point must be preferable. Ordering the different terms of the interaction energy, which can be done by perturbative methods, leads to electrostatic, induction and dispersion interactions. Furthermore, charge-transfer and correlation effects may play important roles for the interaction potential [14, 15]. Another consequence of the small magnitude of these interactions is that relativistic effects, which are customarily neglected when considering chemical interactions, may become important enough to affect the form of the inter-ionic potentials, particularly for the heavier ions (third row and beyond). These will not be considered further here, however.

In the basic MO approach, interaction energies are obtained from a calculation on the complete system of interacting species (the super-molecule). Since the interaction energy in the case of intermolecular forces is so small, a serious

problem in super-molecule calculations is the occurrence of basis set superposition error (BSSE). This occurs when the description of one fragment is improved due to the larger basis set available in the presence of another fragment. This artificial error, which would not occur with a complete basis set, is unfortunately highly geometry dependent—one may say that the quality of the basis set depends strongly on the geometrical arrangement. Various schemes for correcting this error exist, but they are not always very reliable. By allowing non-orthogonal orbitals, the wavefunctions can be constrained to be on separate fragments, and this problem may be completely avoided. In chemical bond formation BSSE can be neglected compared with the (large) chemical interaction energy, so the average quantum chemist is most likely to come across this term only when considering intermolecular forces problems.

The approach described here was first suggested by Wormer and coworkers [16] who showed that for the He-He system, very simple VB calculations reproduce the interaction energy obtained from fairly sophisticated MO-CI (super-molecule) treatments. This method has been further developed by Cooper and coworkers. A useful introduction may be found in the review, ref. [17], or most recently in work on the LiH··He system [18, 19]. For treating the interaction between two subsystems A and B, we define zeroth order wavefunctions Φ_A^0 and Φ_B^0, where Φ_A^0 is defined in terms of A orbitals, ϕ^A, and Φ_B^0 is defined in terms of B orbitals, ϕ^B. The zeroth order wavefunctions may vary in level of sophistication, but will normally be simple SCF or SC wavefunctions. In order to avoid BSSE, orbitals belonging to system A are constrained to be expanded only in terms of A basis functions, and similarly for the B orbitals. We will consider here briefly the example of the Li··He system, further details may be obtained from ref. [18, 19] or ref. [8]. Spin-coupled wavefunctions were obtained for the two fragments of the forms

$$\Phi_{\mathrm{LiH}}^0 = \mathcal{A}[\phi_1^{\mathrm{LiH}}\phi_2^{\mathrm{LiH}}\phi_2^{\mathrm{LiH}}\phi_3^{\mathrm{LiH}}\Theta^{\mathrm{LiH}}] \tag{6.2}$$

for LiH, and

$$\Phi^0_{\mathrm{He}} = \mathcal{A}[\phi_1^{\mathrm{He}}\phi_2^{\mathrm{He}}\Theta^{\mathrm{He}}]. \tag{6.3}$$

for the He subsystem. The ϕ^{LiH} orbitals are only expanded in basis functions centred on Li and H, and the ϕ^{He} orbitals similarly only in basis functions centred on He.

The electrostatic interaction between the two systems is found by simply constructing the antisymmetrised product

$$\mathcal{A}[\Phi^0_{\mathrm{LiH}}\Phi^0_{\mathrm{He}}] \tag{6.4}$$

and evaluating the energy. Besides the purely Coulombic interaction (between electron cloud A, electron cloud B, nuclei in A and nuclei in B) this includes the exchange interaction where an electron from system A is interchanged with one from B, stemming from the antisymmetriser in equation (6.4).

The induction energy can be attributed to the configurations $\mathcal{A}[\Phi^0_{\mathrm{LiH}}\Phi'_{\mathrm{He}}]$ and $\mathcal{A}[\Phi'_{\mathrm{LiH}}\Phi^0_{\mathrm{He}}]$ where the Φ's are singly excited configurations. The dispersion contribution to the interaction energy can analogously be attributed to configurations of the form $\mathcal{A}[\Phi'_{\mathrm{LiH}}\Phi'_{\mathrm{He}}]$. The form of the singly excited configurations in both of these expressions will of course strongly affect the rate of convergence of these different terms. It is therefore important to choose the virtual orbitals with some care. A recent publication that addresses this particular problem for the spin-coupled wavefunction may be found in ref. [19].

To determine the relative importance of these different terms, all configurations are included in a non-orthogonal CI calculation, and the relevant secular problem is solved. In this manner the interaction energy can in a consistent way be attributed to the electrostatic, induction and dispersion terms.

Further correlation can be added to the SCF or SC description by describing each *fragment* by a small CI (VB) expansion. The electrostatic energy is evaluated by considering the direct product of these two wavefunctions. The size of this expansion increases very rapidly as the expansions on each fragment are enlarged,

and it is therefore not practical to augment the study of correlation effects with that of the induction and dispersion terms.

The study of charge transfer effects requires a great deal of care, since these are closely connected with the BSSE mentioned earlier. A feasible approach is to optimise separately the ionised fragments A^+, B^-, A^- and B^+. These may then be combined, and the relative importance of the structures $\mathcal{A}[\Phi_A\Phi_B]$. $\mathcal{A}[\Phi_A^+\Phi_B^-]$ and $\mathcal{A}[\Phi_A^-\Phi_B^+]$ can be assessed by constructing the corresponding list of structures, and solving the secular problem $\mathbf{Hc} = E\mathbf{Sc}$. This is the physically most direct way of describing charge transfer, so it is likely that BSSE related errors will be minimised. Even so, the effect of increasing basis set size on the importance of charge transfer should be investigated.

It is important to realise that any error occurring from BSSE in the above scheme will tend to *overestimate* the importance of charge transfer. Thus if the energy improvement due to charge transfer and perhaps BSSE with a relatively small basis set is already negligible, this computation can be dispensed with in more sophisticated calculations.

6.3 Describing the super-system

It is useful to consider here the treatment of the super-system before we address the problems associated with describing the individual fragments. This will enable us in the next section to illustrate the effects of varying descriptions of the ions directly on the calculated potentials.

This work is mainly concerned with the calculation of the electrostatic interaction, or the simple Coulomb and exchange terms, as defined in equation (6.4). Optimisation of the super-system wavefunction presents special problems in the case of ionic systems—some of these will be considered in appendix 6.a. The types of fragment wavefunctions we consider will be of either SCF, SC or SCVB form. The optimisation of these wavefunctions will be performed in a simulated

crystal environment, as will be described in section 6.4, but once the set of variational parameters that define the given wavefunction has been determined, the subsequent theory will be indistinguishable from the case of intermolecular forces.

A technical point is the transformation of the spin functions used in the supermolecule calculation (e.g. Rumer, Kotani or Serber basis), to direct products of the spin functions from the calculations on the two subsystems. The transformation matrix was in these calculations found simply by considering the coefficients of the spin functions in the determinant basis. If the spin functions on subsystem A and B are defined in terms of determinants as

$$\Theta_k^{(A)} = \sum_I \Phi_I^{(A)} ({}^{(A)}\mathbf{b})_{Ik}$$

and

$$\Theta_k^{(B)} = \sum_I \Phi_I^{(B)} ({}^{(B)}\mathbf{b})_{Ik}, \tag{6.5}$$

then the direct product spin functions are

$$\Theta_{kl}^{(AB)} = \sum_{I,J} \hat{\mathcal{A}} (\Phi_I^{(A)} \Phi_J^{(B)}) ({}^{(A)}\mathbf{b})_{Ik} ({}^{(B)}\mathbf{b})_{Jl}, \tag{6.6}$$

which, expressed in terms of the new spin basis, becomes

$$\Theta_{kl}^{(AB)} = \sum_{I,J,m} ({}^{(AB)}\mathbf{b}^{-1})_{IJm} ({}^{(A)}\mathbf{b})_{Ik} ({}^{(B)}\mathbf{b})_{Jl} \Theta_m^{(AB)} \tag{6.7}$$

since

$$\Theta_m^{(AB)} = \sum_{I,J} \hat{\mathcal{A}} (\Phi_I^{(A)} \Phi_J^{(B)}) ({}^{(AB)}\mathbf{b})_{IJm}. \tag{6.8}$$

This rather 'brute force' way of obtaining the transformation matrix, has the advantage of being applicable irrespective of the spin basis used in the three calculations. It is generally possible, especially if one is using Rumer functions, to reduce the spin space used in the super-system calculation.[1] Finding the direct product basis is of course necessary when the fragments are not allowed to relax, but even in the case when the super-system is reoptimised, it is useful to

[1] In some cases the maximum reduction may be obtained if orbitals in some of the configurations are reordered.

distinguish between the direct product and other type of spin coupling. In many cases[2] restricting the spin space to be only direct products makes a great deal of sense.

6.4 Describing a single ion

A realistic description of an ion in a crystal is important for the determination of a range of properties, specifically the calculation of inter-ionic potentials. This has been addressed previously in for example ref. [20], but we also refer the reader to the general literature on calculation of inter-ionic potentials and ion polarisabilities. The present investigation is limited to mono-atomic ions, and we shall constrain these to be spherical. There is considerable evidence, both theoretical and experimental, that this is in fact an excellent first approximation. When considering the interaction between two non-spherical species, averaging over the angular variables would be necessary to obtain an interaction curve of high accuracy. The discussion is kept simpler and the computational effort more reasonable by insisting on spherical ions.

In this section MgO is taken as a generic example to illustrate the main issues involved in generating inter-ionic potentials. MgO, or periclase, has rock-salt structure with lattice constant a_{lat}=4.208 Å [21]. This being a highly ionic system means that an approach based on intermolecular forces is likely to be very well suited. For the same reason we can expect the effects of three-body (and higher) forces to be minimal. This being said there has for a number of years been disagreement on this system—particularly the description of the O^{2-} ion. Resolving these discrepancies could mean a big step forward for this field as a whole.

Unless otherwise stated the ions were described at the SCF level, both having electron configurations $1s^2 2s^2 2p^6$. The basis sets used in the two cases were of triple-zeta-valence (TZV) quality, a (10s6p)/[5s3p] basis for O^{2-} [22], and a

[2]E.g. if the iterative scheme proposed in the appendix (or something equivalent) was used.

(12s9p)/[6s5p] basis for Mg^{2+} [23].

To a first approximation we represent the surrounding ions by a simple lattice of point charges. For the highly ionic systems considered here it seems reasonable to neglect any effects related to charge-transfer, which means we can assign simple integer values to the point charges. For systems such as TiO_2 which are ionic, but for which charge transfer is likely to play an important role, one might reconsider this premise. So for MgO the Madelung potential is simulated by an array of alternating $+2$ and -2 charges. We have here followed a scheme first introduced by Evjen [24] of scaling point charges on faces, edges and corners by $\frac{1}{2}$, $\frac{1}{4}$ and $\frac{1}{8}$ respectively. The advantage of this is faster convergence with increasing lattice size of the Madelung potential in the vicinity of the central ion. Most molecular packages, including MOLPRO [25] which was used in the present work, have the ability to include point charges in the integral evaluation.

If the electron density is constrained to be spherical, any point charge lattice can be substituted by a spherical potential. Even though for technical reasons we have performed the calculations using point charges, it is illustrative to consider the radial part of the spherically symmetric potential. It is easy to show that a point charge q, at distance a from the origin contributes

$$E(R;q,a) = \begin{cases} -q/a, & R < a \\ -q/R, & R > a \end{cases} \tag{6.9}$$

to the energy of a single electron when averaged over the angular variables. In figure 6.1 the radial parts of spherically averaged Madelung potentials from 3×3×3, 5×5×5, 7×7×7 and finally 15×15×15 point charge lattices are shown. This was done to determine what lattice size would be sufficient to simulate accurately the infinite crystal. The Madelung potentials here are generated by putting point charge lattices around O^{2-} at the perfect lattice points. As can be seen the potentials differ in two ways. For smaller distances there is a relatively small (except for the 3×3×3 lattice), constant separation stemming from the different contributions from the edges to the first term in equation (6.9). Of course a constant

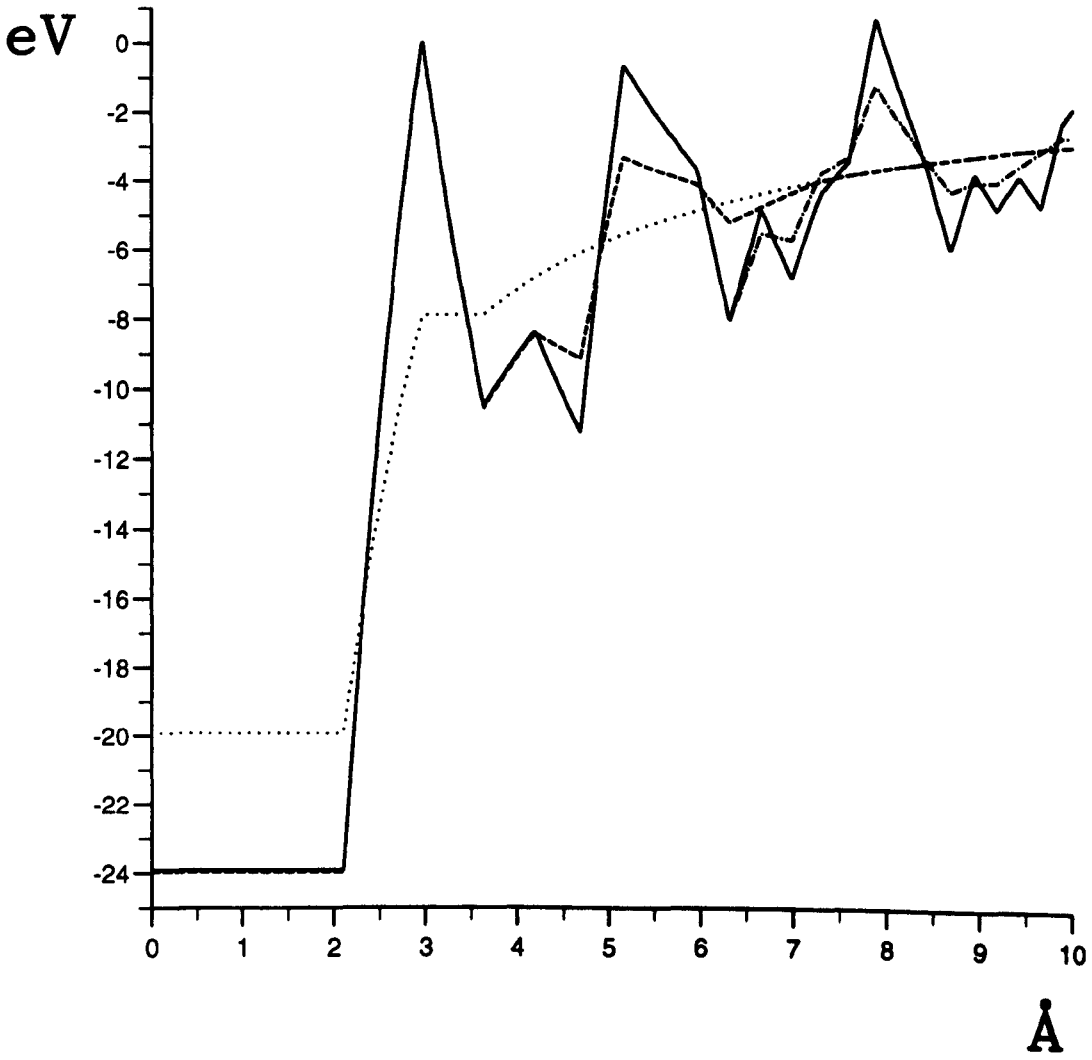

Figure 6.1: Convergence of Madelung potential with the size of the point charge lattice. 3×3×3 is dotted, 5×5×5 broken, 7×7×7 double-broken and 15×15×15 shown as full line.

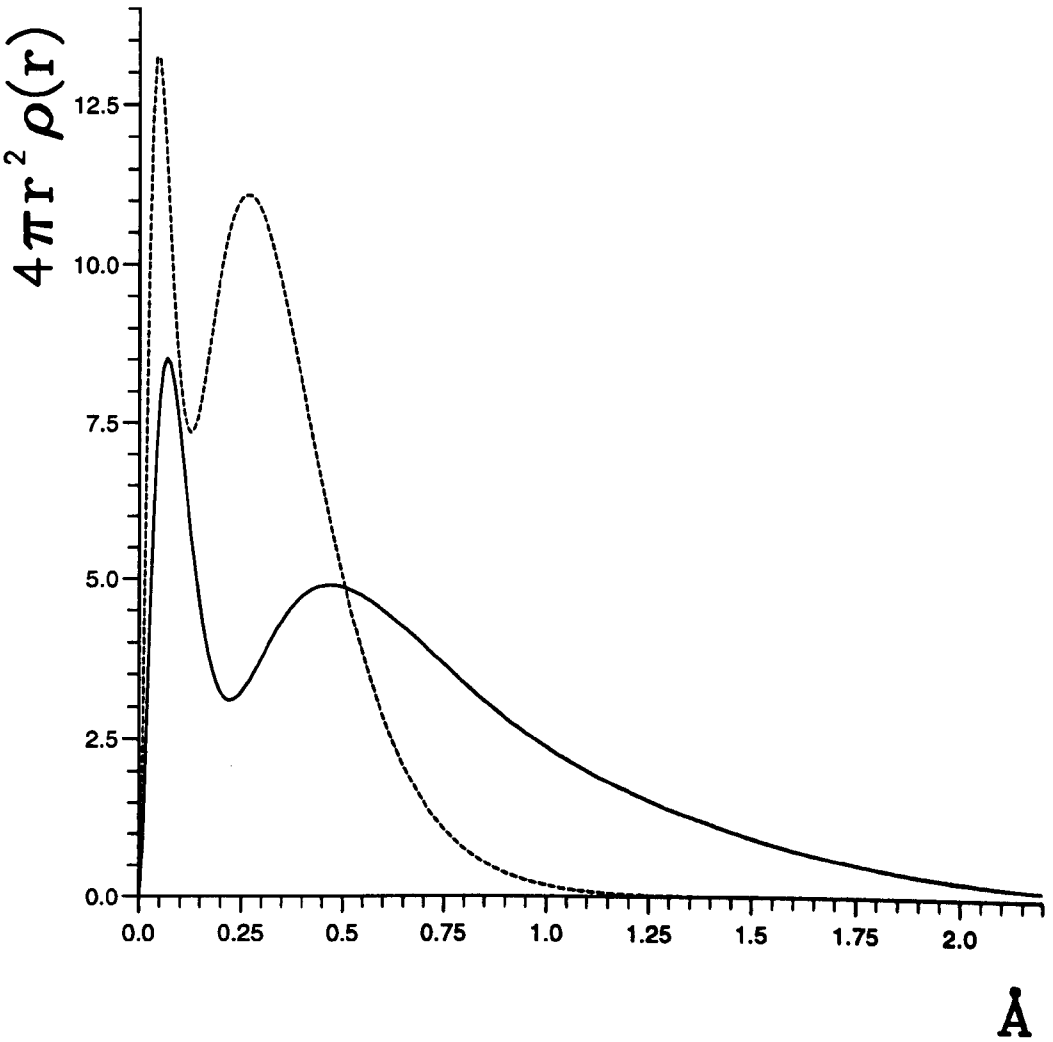

Figure 6.2: Radial distribution functions for O^{2-} (full line) and Mg^{2+} (broken line).

shift of the potential energy does not alter the electron distribution, or indeed any physical aspects of the problem. Only at distances where the magnitude of the point charges differ is there a significant deviation. The form of the 5×5×5 and 7×7×7 potentials is for example identical for distances below a_{lat}=4.208 Å. Thus only electron density beyond this distance will be affected by the difference between these two lattices. To put this in context we have in figure 6.2 shown the radial distribution functions (RDF) for Mg^{2+} and O^{2-} (SCF wavefunctions calculated in the field of a 15×15×15 point charge lattice). The electron density

more than a_{lat} from the nucleus corresponds to merely 10^{-5}% of the total for even the O^{2-} radial distribution.

To re-iterate: the description of the central ion converges very rapidly as the size of point charge lattice is increased. The Madelung potential is not converged in absolute terms (since there is a constant difference in the potential energy), but this will not affect the electron distribution. The central ion does not extend far enough to be affected by the fact that the point charge array is finite. An illustration of this is that the difference between the MOs calculated with 13×13×13 and 15×15×15 lattices is much smaller than the convergence threshold of the MOs themselves. Since no particular computational effort is associated with including the point charges, all the systems considered here have been generated using 15×15×15 point charge lattices.

There are two ways in which the potential generated by the point charge lattice may be significantly different from the above: the crystal system may not be simply cubic, or the lattice constant may be altered. In figure 6.3 we have shown the Madelung potentials experienced by an O^{2-} ion in systems with rock-salt, CsCl and anti-fluorite structures respectively. In order to allow a direct comparison, the CsCl and anti-fluorite Madelung potentials have been shifted so as to coincide with the periclase potential at the origin. This is admittedly rather arbitrary, since different crystal systems will generally have differing lattice constants. Out to a distance of about 4Å the forms of the potentials are surprisingly similar, and it is likely that the description of the central ion will not be significantly affected by the crystal geometry, provided that the lattice separations are chosen appropriately.

Varying compounds will of course have different lattice constants, and even for the same compound the lattice separation will depend to a certain degree on temperature and pressure. It is therefore useful to investigate the effect of the lattice constant on the radial distribution function and inter-ionic potentials. The RDFs for O^{2-} with a varied from 4 bohr to 8 bohr as well as the limit in which

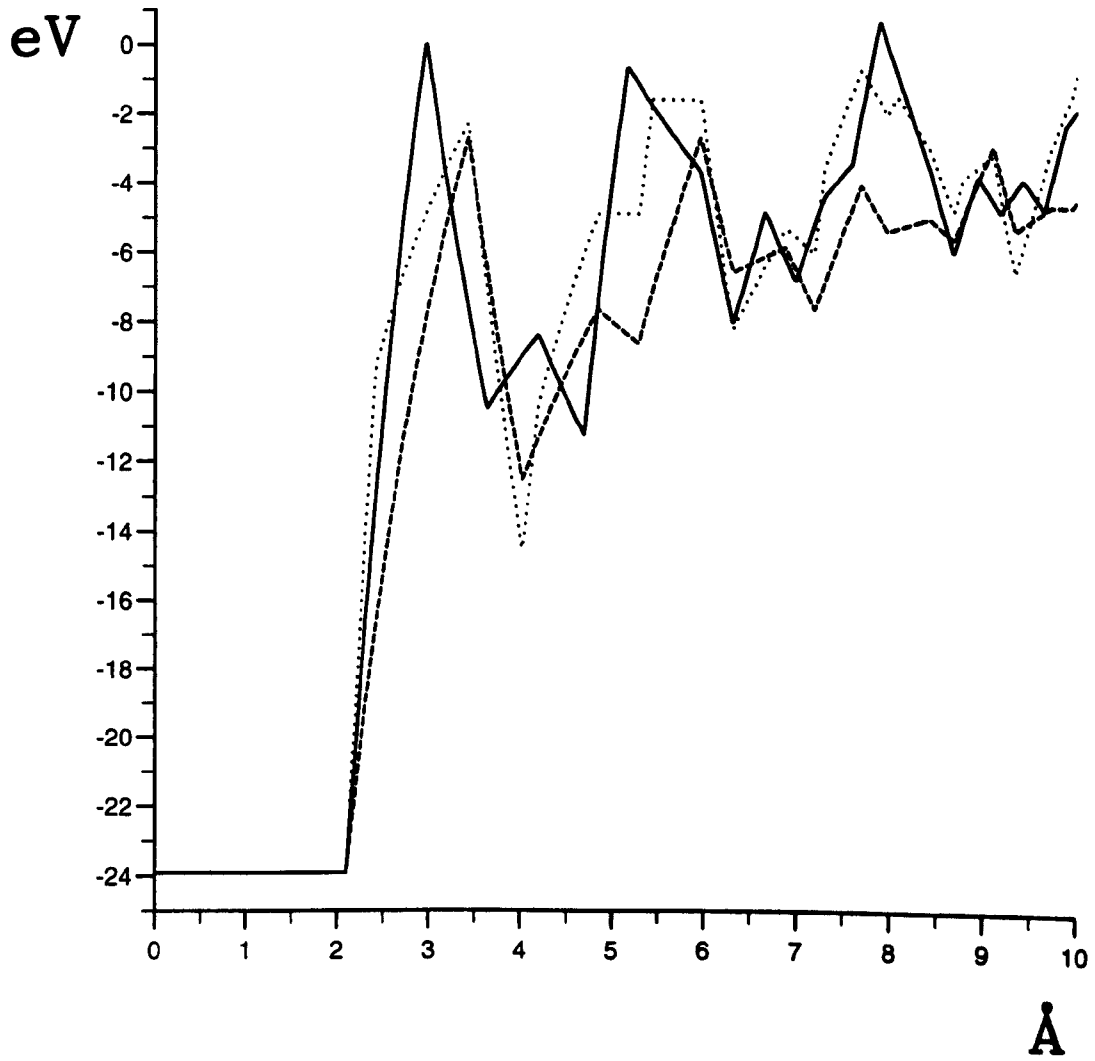

Figure 6.3: Madelung potentials for various crystal geometries. Full line=rock-salt structure, broken line=anti-fluorite structure and dotted line=CsCl structure.

the lattice is removed entirely, are shown in figure 6.4. There is no appreciable difference between the 8 bohr distribution and the distribution in the absence of a lattice (equivalent to infinite separation). Since 8 bohr is close to the actual lattice constant for periclase, one must conclude that the point charge lattice alone has only a very mildly stabilising effect for this case.

The corresponding interionic potentials between Mg^{2+} and O^{2-} are shown in figure 6.4. Of course the above variations in the lattice constant may be rather severe, but there seems nevertheless to be an appreciable effect on the potential. For special cases, such as interstitial formation for example, it may be advisable to adjust the interionic potentials to account for the stronger Madelung field.

Diminishing the lattice constant for the point charge array has to a good approximation the effect of shifting the potentials to the left. The interaction is predominantly repulsive, so this can be taken as a consequence of the increasingly localised charge distributions. We will return to the idea of 'shifted' potentials in section 6.5.

The corresponding O^{2-}–O^{2-}interaction curves are shown in figure 6.4. As can be seen there is a favourable interaction in this case—the potential even being attractive for some separations. The potentials are therefore arranged in the reverse order compared with the Mg^{2+}–O^{2-}curves—the most delocalised charge distributions generally leading to the *least* repulsive potentials.

More sophisticated ways of approximating the Madelung potential have been investigated in the literature. Instead of approximating the surrounding ions by point charges, one can depict them as spheres of sizes close to the ionic radii. Also a rough estimate of the exchange interaction with neighbouring ions can be incorporated [26]. In this work we improved the description of the crystalline environment by substituting the nearby point charges with actual ions. Computational cost normally limits the number of ions that can be included in this way. Thus for MgO the O^{2-} ion was described with the 6 nearest Mg^{2+} ions included, described at the SCF level. The effect of this, as is reflected by the RDFs shown

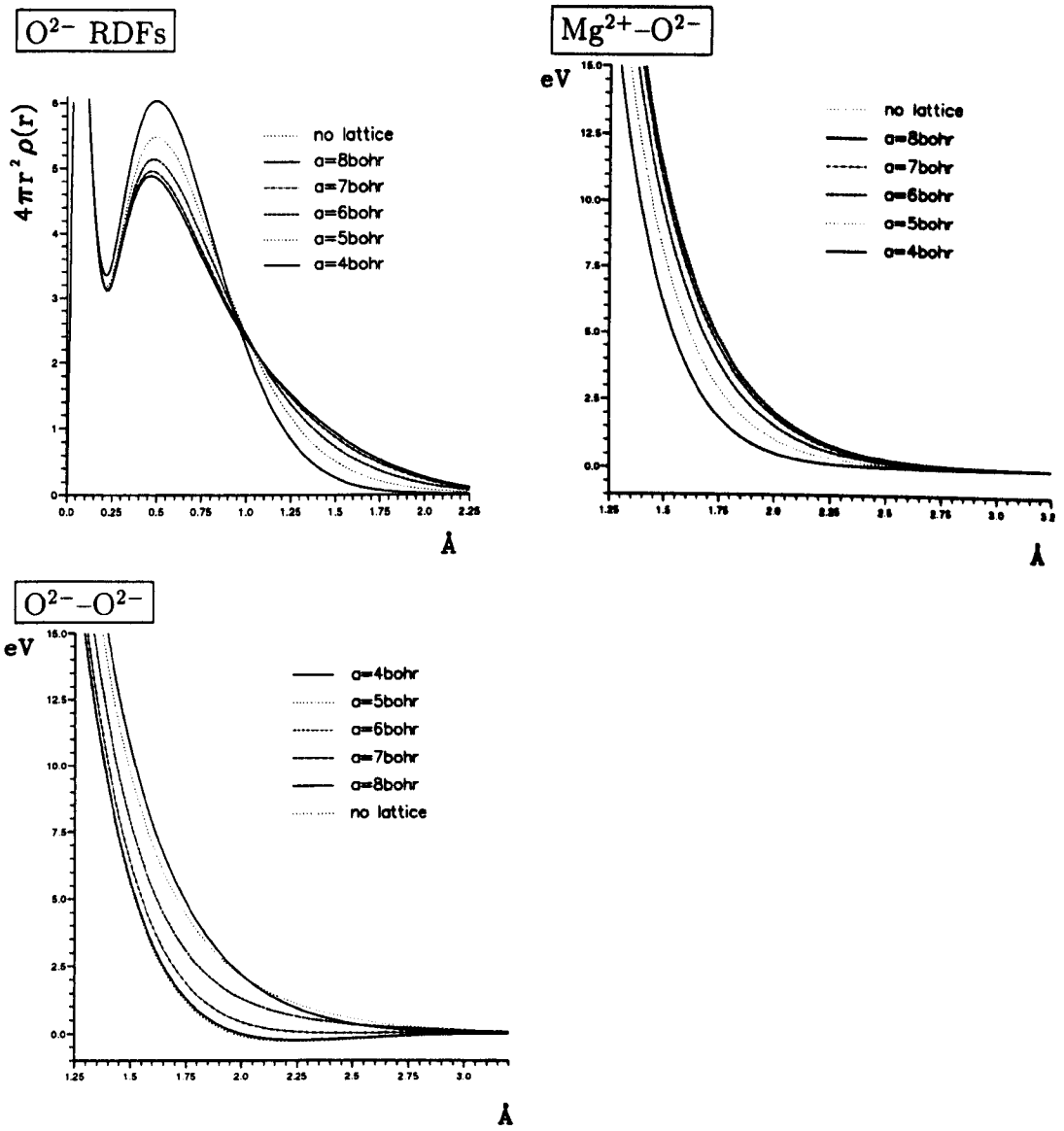

Figure 6.4: Radial distribution functions for O^{2-}, Mg^{2+}–O^{2-} potentials, and O^{2-}–O^{2-} potentials with varying lattice constants.

in figure 6.5, was to contract the O^{2-} slightly. The equivalent procedure was followed for the description of the Mg^{2+} ion, and as may be seen from figure 6.7 the effect was far smaller in this case. It is generally considered that, since anions are generally much more diffuse than cations (cf. figure 6.2), the description of the anion is the more important. Including nearest neighbours for also Mg^{2+} was an attempt at verifying this supposition.

To get a more precise idea of the effect including nearest neighbours, interaction potentials were calculated as shown in figure 6.6. For O^{2-} the effect of nearest neighbours is quite significant. The contraction of the electron density makes the potential clearly less repulsive. The effect of nearest neighbours for Mg^{2+} is, as suggested above, minute in comparison. Similarly to what occurs when the point charge separation is diminished, the change of the potential is in this case also roughly a leftward shift. It seems that the *form* of the stabilising Madelung field may be of minor importance compared with the approximate strength thereof. A very good approximation to the potential in figure 6.6 could be obtained by just decreasing the separation in the point charge lattice by approximately 15%.

The effect of nearest neighbours on the O^{2-}–O^{2-}potential can also be seen to be quite significant—in this case in the direction of a more repulsive potential (see figure 6.6). Also in this case the effect is roughly equivalent to a decrease in the point charge lattice separation, but in this case probably nearer 20%.

It is important also to investigate the role of correlation effects in determining the form of potentials. In this work the starting point has always been a restricted Hartree-Fock (RHF) calculation (most of the ions considered were closed-shell systems). Various spin-coupled wavefunctions have been investigated, with some or all electrons in the outermost shell active. For further electron correlation a non-orthogonal CI based on the spin-coupled orbitals might be performed. In this case, however, the super-molecule calculation tends to become prohibitively large very quickly.

In the most correlated description of O^{2-} considered here, an eight electron

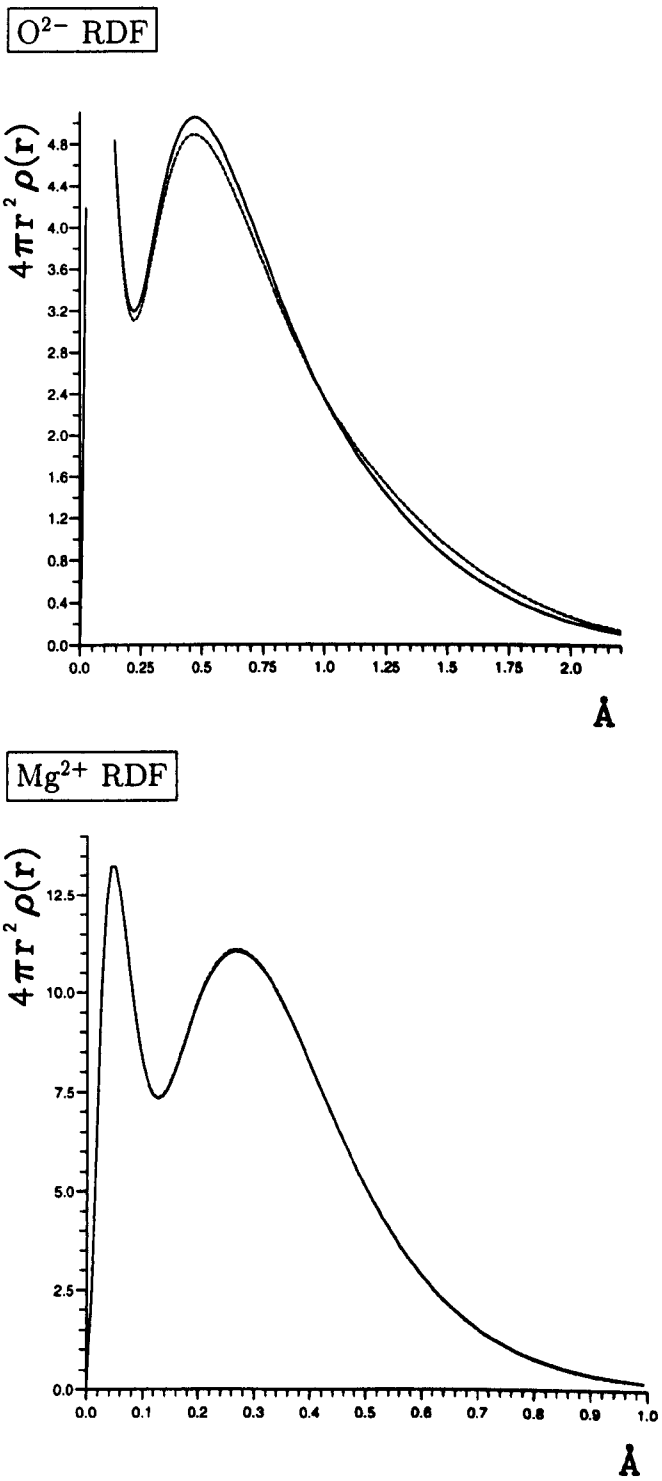

Figure 6.5: Radial distribution functions for O^{2-} and Mg^{2+}. Broken lines are point charge lattice only, full lines: point charge lattice and nearest neighbours.

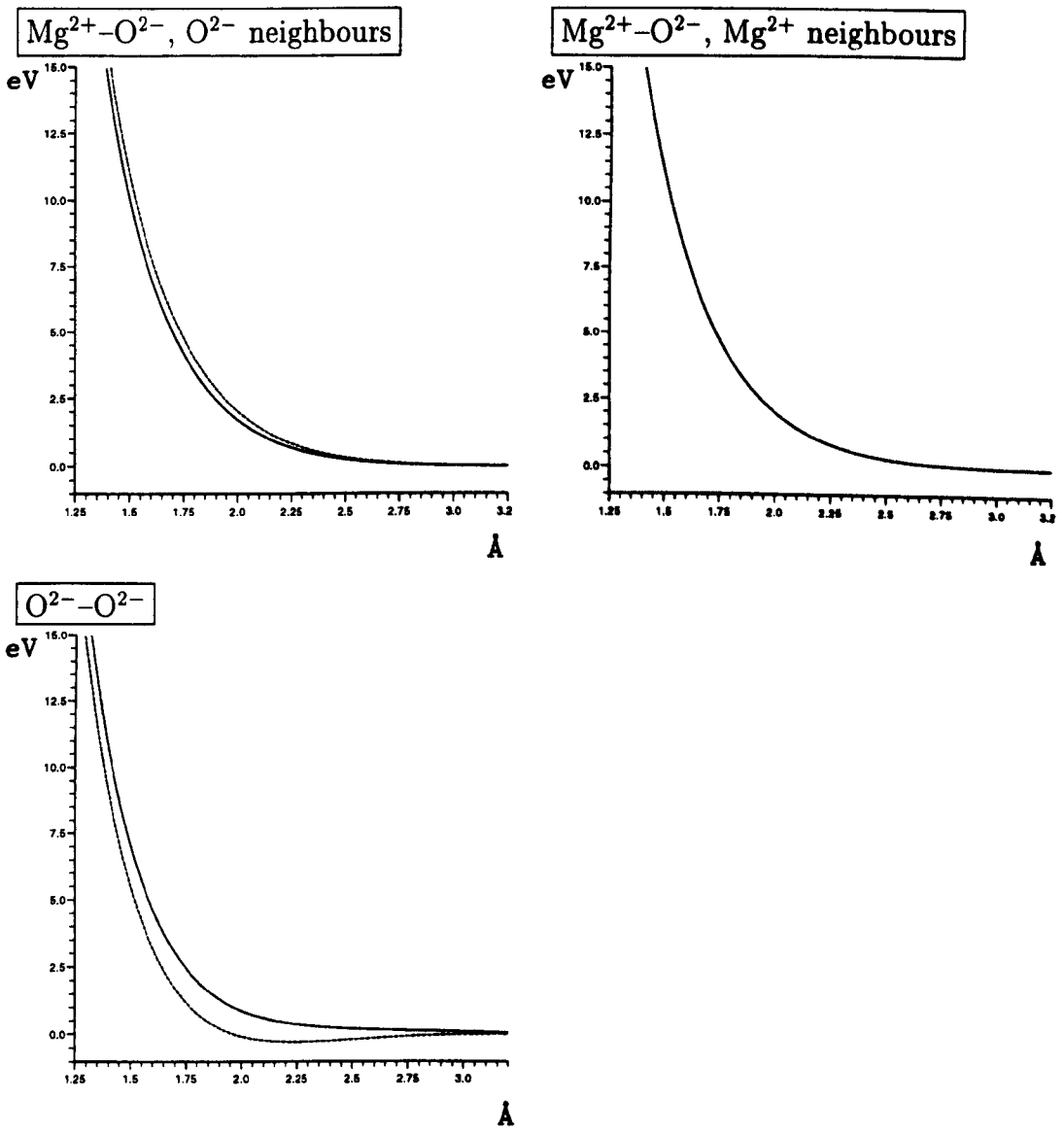

Figure 6.6: Mg^{2+}–O^{2-} potentials (nearest neighbours for O^{2-}), Mg^{2+}–O^{2-} potentials (nearest neighbours for Mg^{2+}), and O^{2-}–O^{2-} potentials. Full lines are potentials with nearest neighbours, broken lines without.

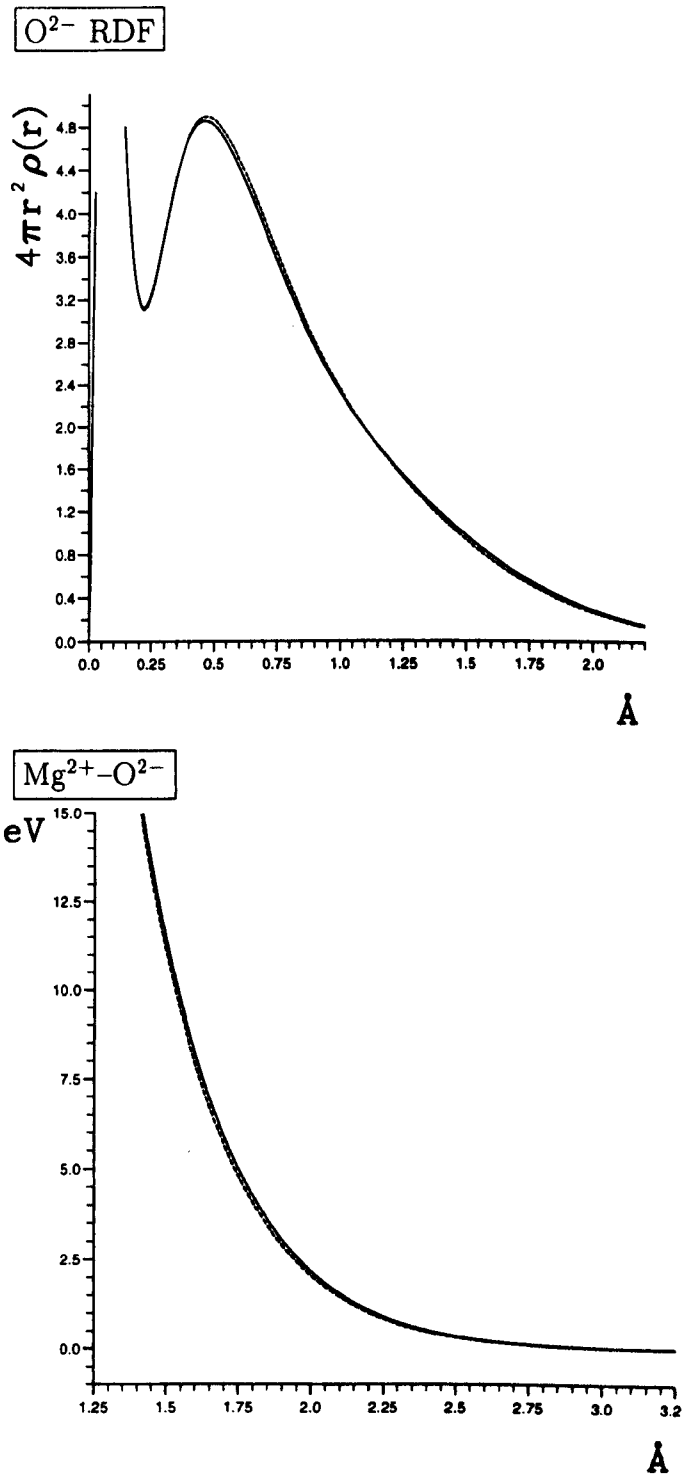

Figure 6.7: Radial distribution functions for O^{2-} and Mg^{2+}–O^{2-} potentials. Full lines are based on SCF descriptions of O^{2-}, broken lines are based on symmetrised 8 electron SC treatments.

spin-coupled calculation was performed. This gave a symmetry-broken solution, however, with the pair of s-orbitals mixing with a pair of p-orbitals in one of the three equivalent directions. Since it is *essential* for the potential calculation to have a spherically symmetric description of the fragment, the SC solution was 'symmetrised' by solving the secular problem spanned by the three possible symmetry-related SC solutions. As can be seen in figure 6.7 the effect of increased correlation is to contract the electron density slightly. This hence makes the Mg^{2+}–O^{2-}potential slightly more repulsive, in agreement with the general trends observed when varying the strength of the Madelung potential.

6.5 Test of potentials

The sections 6.2–6.4 we believe provide a plausible theoretical foundation for the calculation of potentials. It remains to be seen of course how well potentials calculated in this way perform compared with the existing empirical and non-empirical potentials.

Generally very good agreement between potentials calculated in the manner described above and the 'standard' potentials has been found for the cation–cation and cation–anion potentials. As an example of this we have shown in figure 6.8 a comparison with two already existing Mg^{2+}–O^{2-}potentials, an empirical and electron gas potential employed in ref. [27]. As can be seen there is very good qualitative agreement between the three curves, although the new potential is slightly more repulsive for shorter distances. We will come back to what might be done to improve the quality of our potentials.

The strongest test of any new method for potential generation must be the 'difficult' O^{2-}–O^{2-}potential. In figure 6.9 we have shown the comparison between the new potential and three previously employed potentials. The calculated potential in this figure, was obtained using an SCF wavefunction for $O^{2-}$with nearest neighbours of Mg^{2+}, and point charge lattice with a=4.208 Å. As can be seen

System	a/Å	System	a/Å
LiF	4.0270[29]	KCl	6.2788[33]
LiCl	5.1295[30]	MgO	4.2080[21]
NaF	4.6140[31]	CaO	4.7990[34]
NaCl	5.6420[32]		

Table 6.1: Lattice constants for highly ionic systems.

there are substantial differences between the existing potentials. The new method agrees with the electron gas approach on the very long range of the potential. They diverge already at about 3.3 Å, however, and this is most likely connected with the assumption of additivity of the electron densities, an approximation which will break down for shorter r. The new potential seems to agree tolerably well in the short-range region with the probably most reliable empirical potential due to Binks and Grimes [28]. This should probably be taken with a pinch of salt, however, since the errors in both potentials will be largest in this region. The fact remains that the new potential seems to be generally rather too repulsive.

An important consideration when comparing these potentials is the typical sampling distances used in simulations. It will be in this region that existing theoretical potentials have 'stood their ground', and also here that the empirical potentials are most accurately determined. In table 6.2 we have shown the typical sampling distances for a range of systems studied in ref. [27]. Experimental and calculated Hugoniot data were compared in that work, and this is normally considered a particularly severe test of inter-ionic potentials. As can be seen from table 6.2, however, even this application does not sample as small separations as might have been supposed.

The most important variable in the generation of these potentials is the 'strength' of the Madelung field, as related for example to the separation in the point charge array. Including next-nearest neighbours might for example have an appreciable effect on this particular aspect. A very attractive approach would be

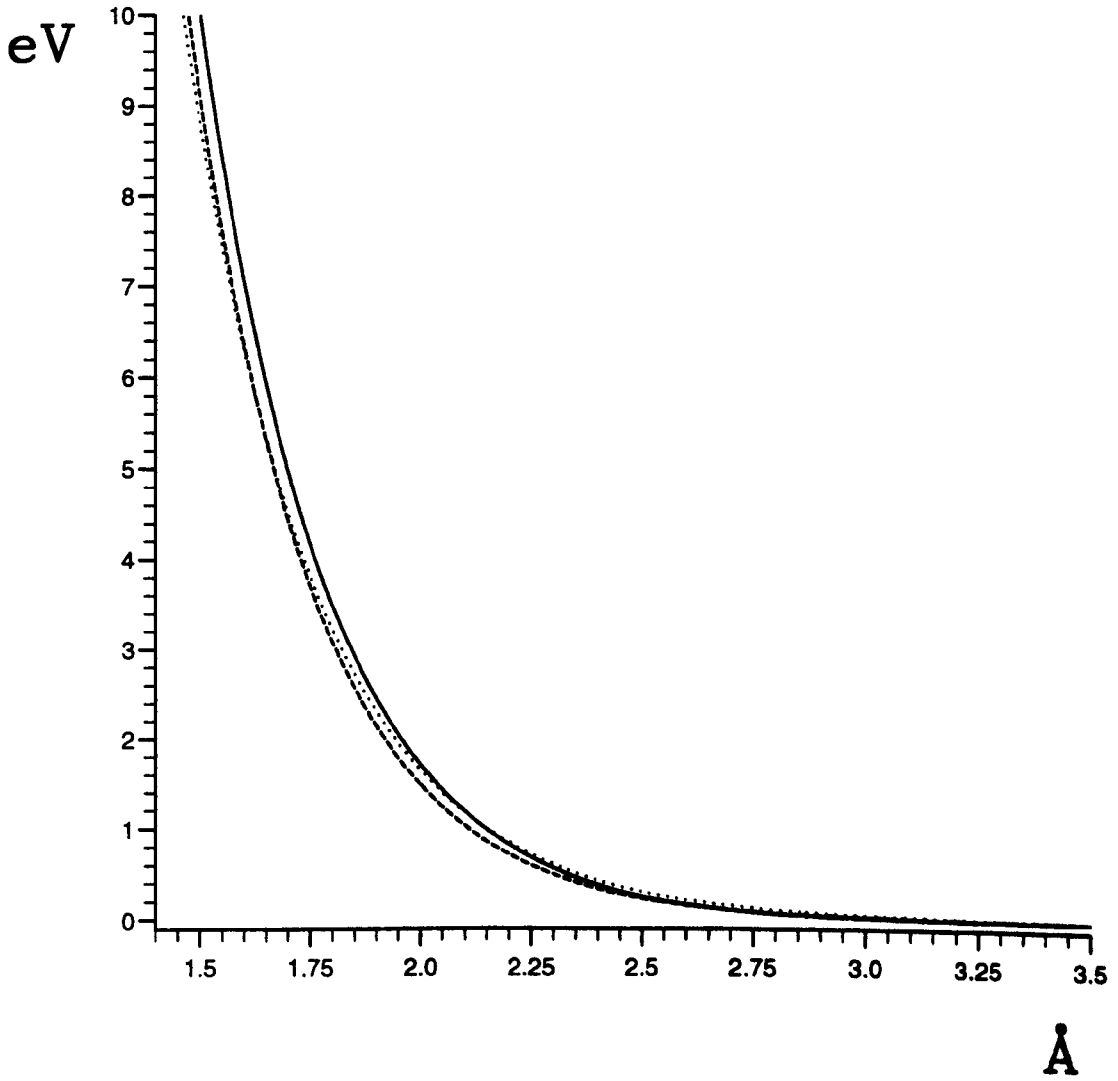

Figure 6.8: Various Mg^{2+}–O^{2-} potentials. Full line=our potential, dotted line=empirical and broken line=electron gas.

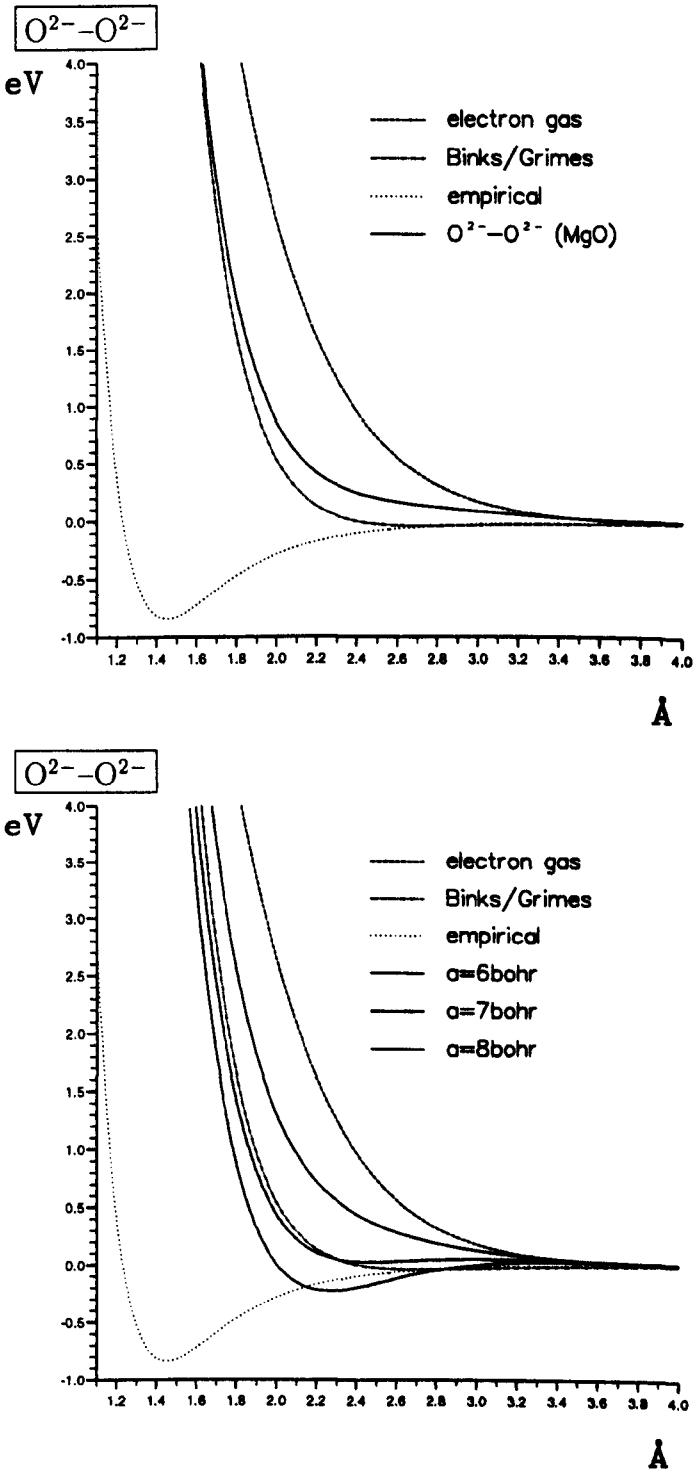

Figure 6.9: Various O^{2-}–O^{2-} potentials.

System	Potentials	r/Å
LiF	M/M & X/X	2.42–2.85, 3.42–4.04
	M/X	1.71–2.02, 2.96–3.50
LiCl	M/M & X/X	2.91–3.48, 4.12–4.93
	M/X	2.06–2.46, 3.56–4.27
NaF	M/M & X/X	2.76–3.29, 3.91–4.66
	M/X	1.96–2.33, 3.39–4.03
NaCl	M/M & X/X	3.48–3.92, 4.93–5.64
	M/X	2.46–2.82, 4.27–4.88
MgO	M/M & X/X	2.66–3.16, 3.76–4.48
	M/X	1.88–2.24, 3.26–3.88
CaO	M/M & X/X	3.04–3.40
	M/X	2.15–2.41, 3.73–4.17

Table 6.2: Ranges of inter-ionic separations (r) sampled for the B1 phase of some of the materials studied at high pressures and elevated temperatures by Allan and coworkers [27].

to leave this variable free—as an empirically determined parameter. This would probably involve fitting the potentials to an assumed function form, in order to obtain analytical expressions for the dependence on the lattice separation. Another problem with the strict comparison of potential curves in figure 6.9 is that the empirical potential is *generic* in the sense that it is not particular to the MgO system. With this in mind we show also the previously used potentials with the new potential calculated for three different point charge lattice separations. It is clear that very good agreement can be obtained by an appropriate choice of a. The small, long-range maximum that remains for all three curves is almost certainly an artifact of neglecting the dispersion contribution to the energy. Attempts at estimating this has not been made in this work, but we refer to appendix 6.a for a strategy that might work well for the case of inter-ionic potentials. Generally the convergence of this term is quite rapid [8], so reliable estimates should be possible with quite reasonable computational effort.

Now we compare calculated potentials with a range of more 'benign' cases for which potentials exist. The work by Allan and coworkers [27] has been based on electron gas potentials. These are probably not very reliable in themselves (cf. figure 6.9, but these were shifted left or right so as to obtain sensible crystal data. No justification for this has previously been offered, but in light of the discussion in section 6.4 this seems very likely to be roughly equivalent to varying the Madelung field experienced by individual ions. This means that the shifted potentials are probably reasonably reliable in the sampled regions.

For Li^+ we employed a (14s9p)/[6s5p] basis set [35], whereas for most of the other ions the basis sets were of triple-zeta-valence (TZV) quality. For O^{2-} and F^- a (10s6p)/[5s3p] basis was used [22], for Na^+, Mg^{2+} and Cl^- a (12s9p)/[6s5p] was used [23], and for K^+ and Ca^{2+} (12s6p)/[7s3p] Huzinaga basis sets were available from the MOLPRO library.

The potentials involving the *smaller* cations can be seen to agree exceptionally well with the electron gas potentials. The K^+–K^+ and Ca^{2+}–Ca^{2+} potentials

diverge for smaller r, but this will occur far below the relevant sampling regions for these cases. Significant deviations tend to occur for the anion–anion potentials, as well as for the cation–anion potentials involving K^+ and Ca^{2+}. In all of these cases the tendency is for our new potentials to predict too repulsive potential curves compared with the more reliable 'shifted' potentials. It therefore seems that further attention must be given to the accurate simulation of the Madelung field.

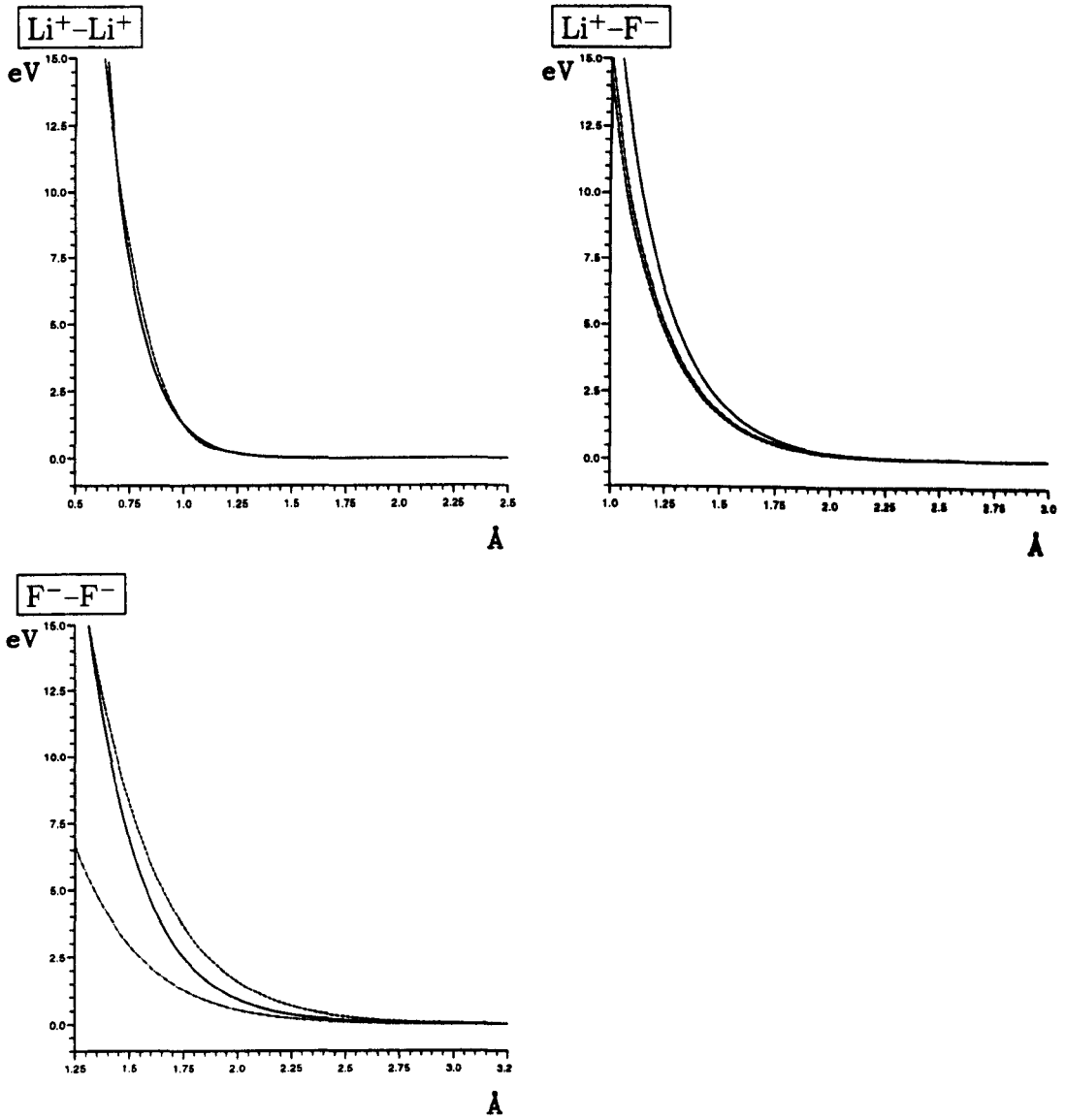

Figure 6.10: Interionic potentials for LiF. Key: full curves – SCF/SCF; dashed curves – electron gas; dot-dash curves – shifted electron gas.

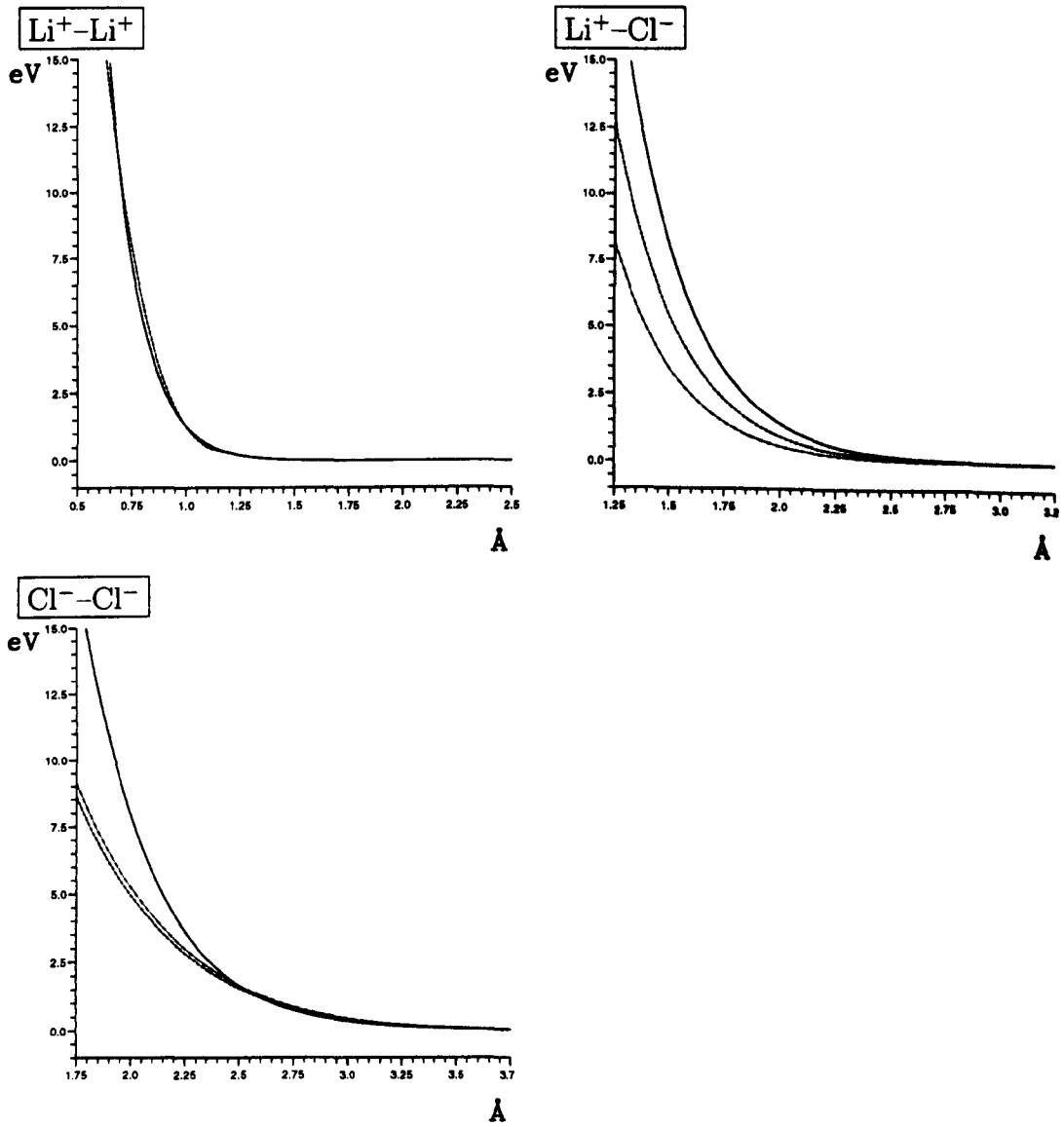

Figure 6.11: Interionic potentials for LiCl. Key: full curves – SCF/SCF; dashed curves – electron gas; dot-dash curves – shifted electron gas.

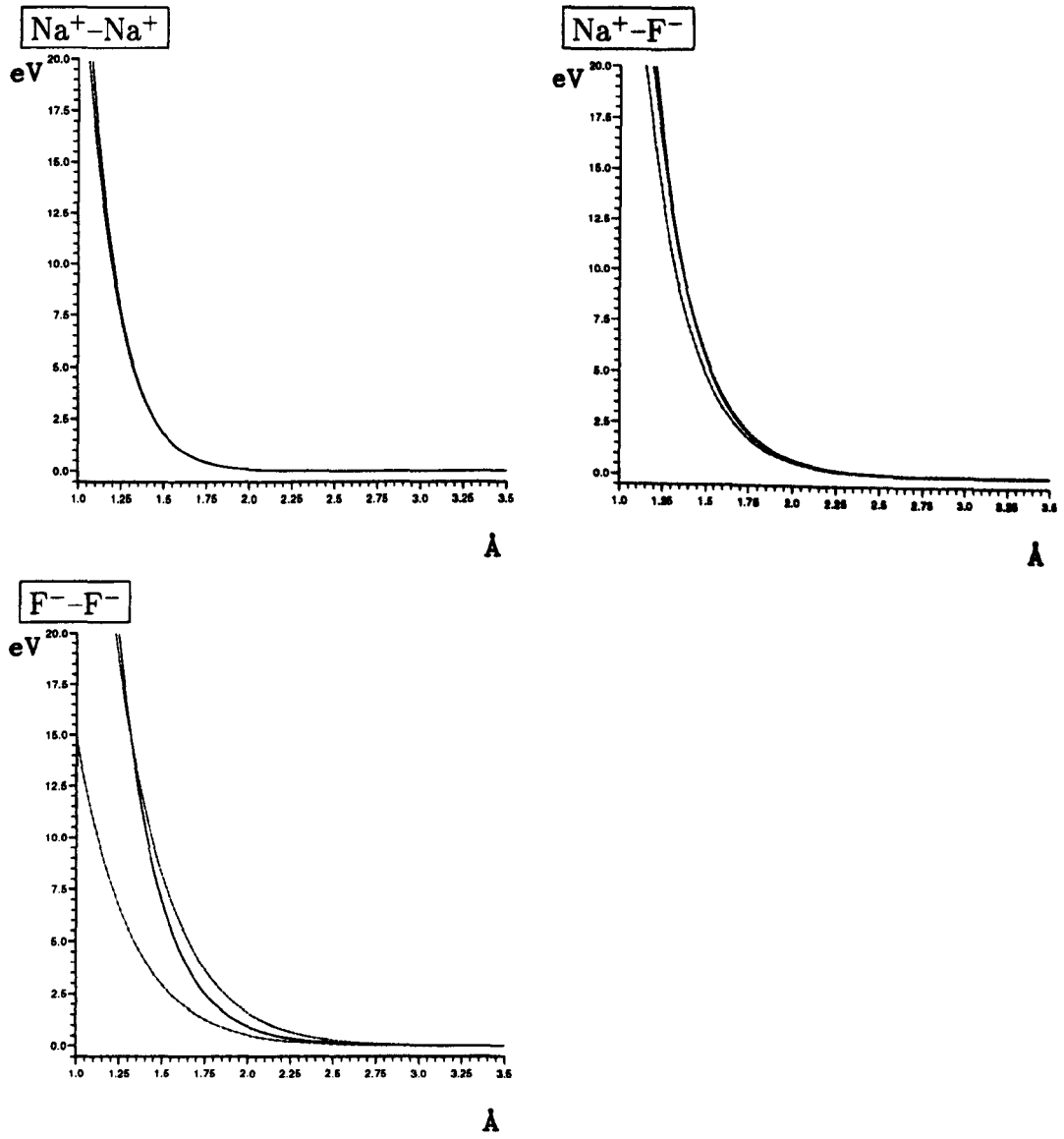

Figure 6.12: Interionic potentials for NaF. Key: full curves – SCF/SCF; dashed curves – electron gas; dot-dash curves – shifted electron gas.

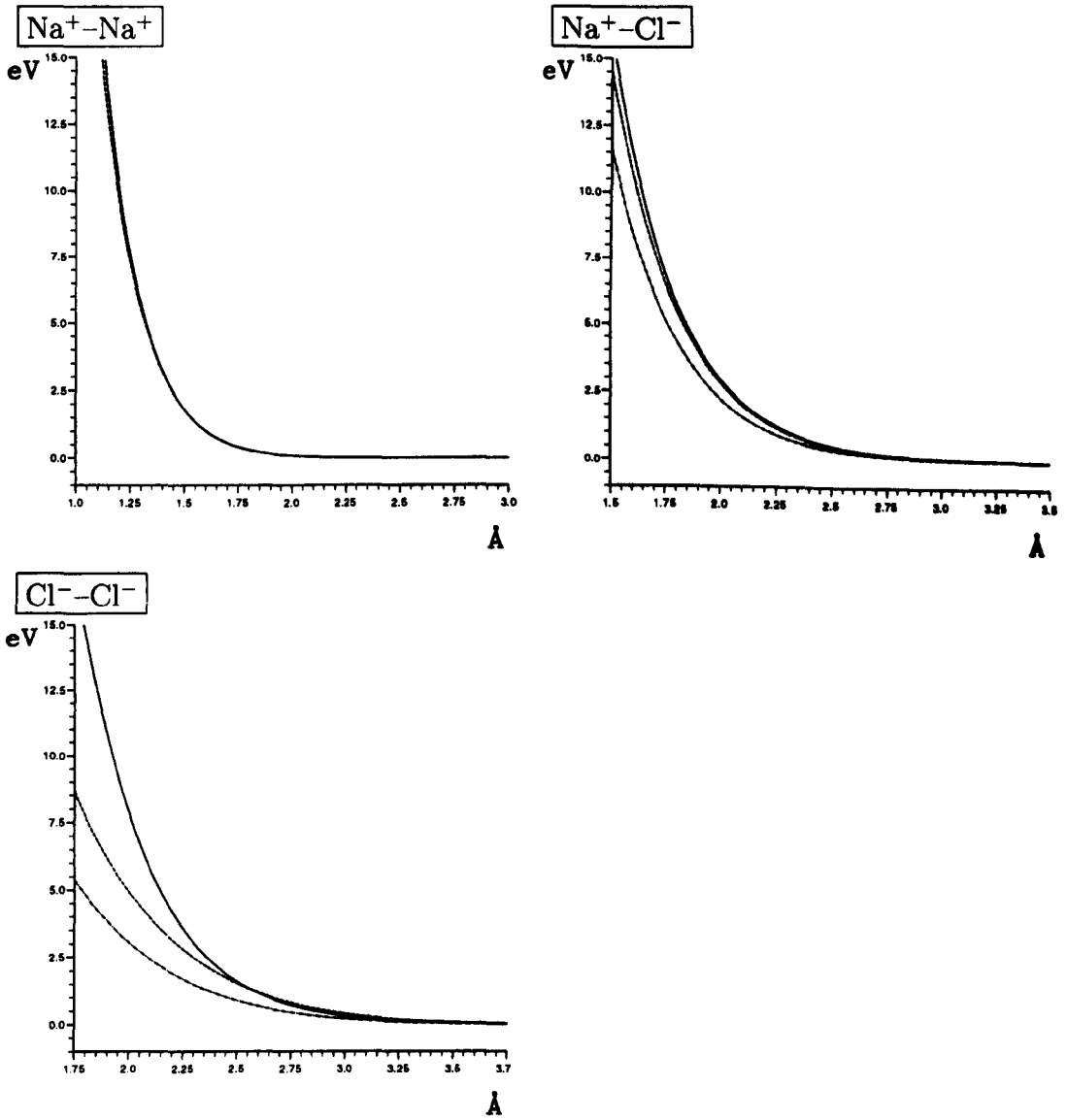

Figure 6.13: Interionic potentials for NaCl. Key: full curves – SCF/SCF; dashed curves – electron gas; dot-dash curves – shifted electron gas.

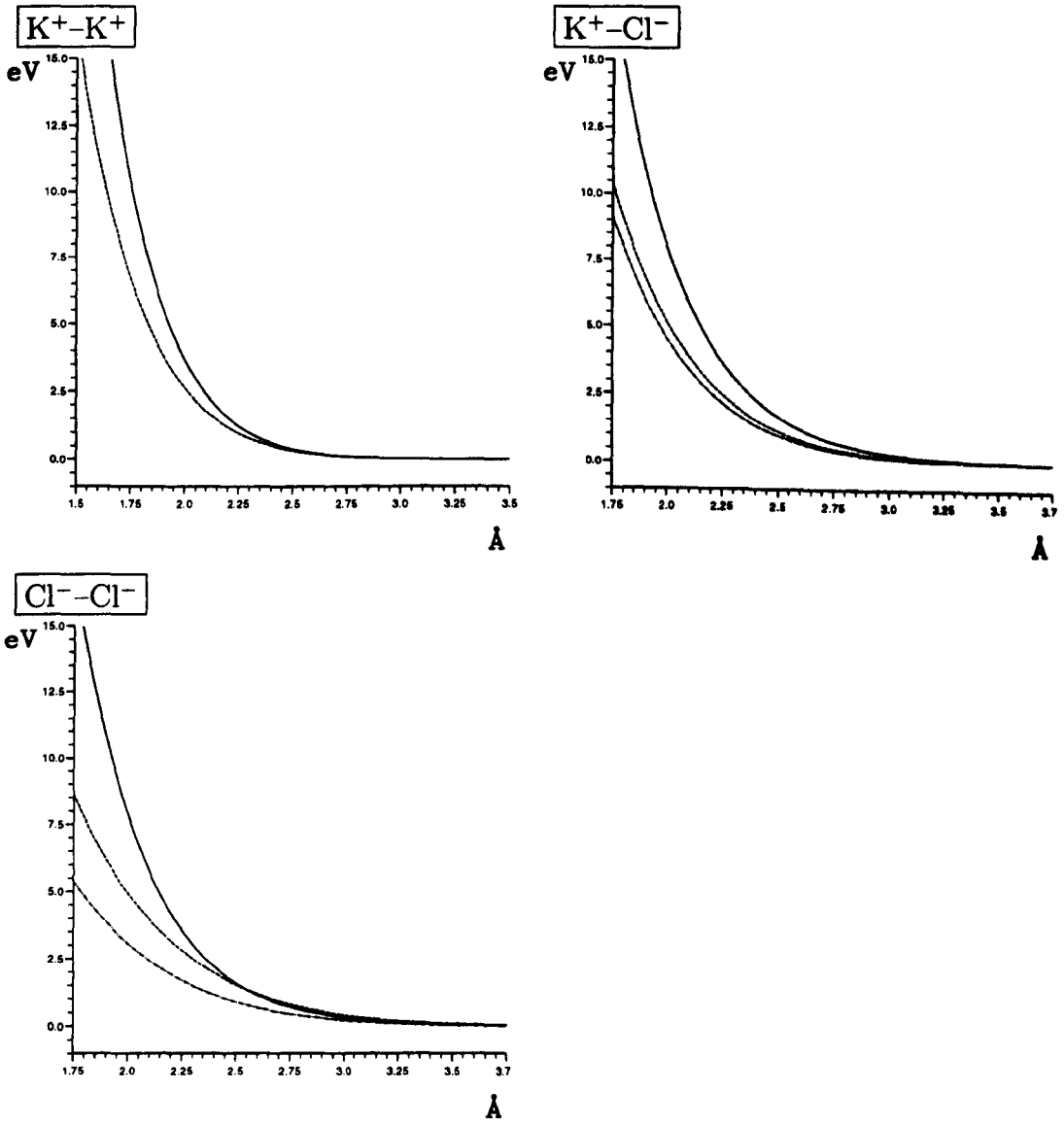

Figure 6.14: Interionic potentials for KCl. Key: full curves – SCF/SCF; dashed curves – electron gas; dot-dash curves – shifted electron gas.

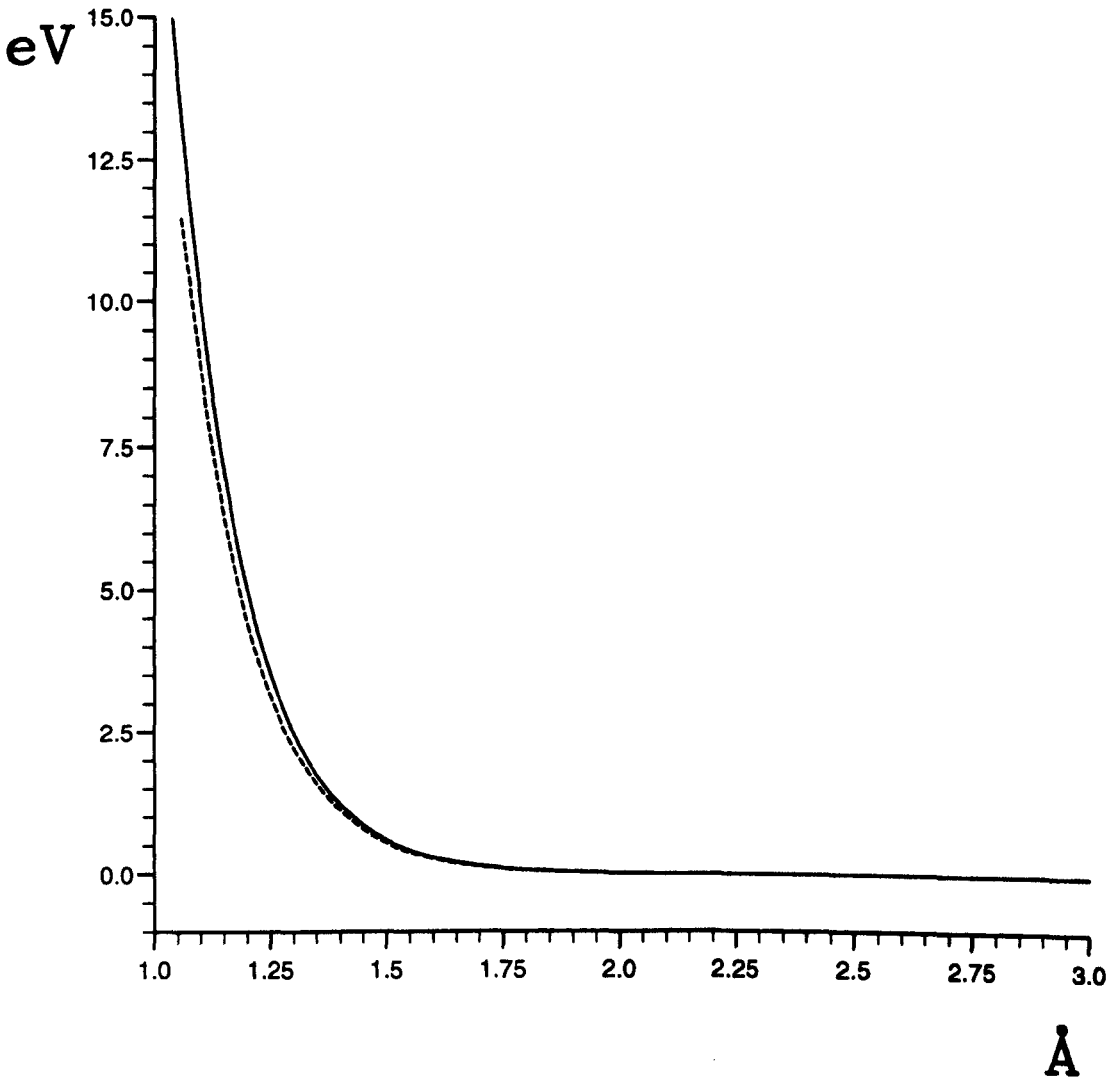

Figure 6.15: Mg^{2+}–Mg^{2+} potential. Key: full curves – SCF/SCF; dashed curves – electron gas.

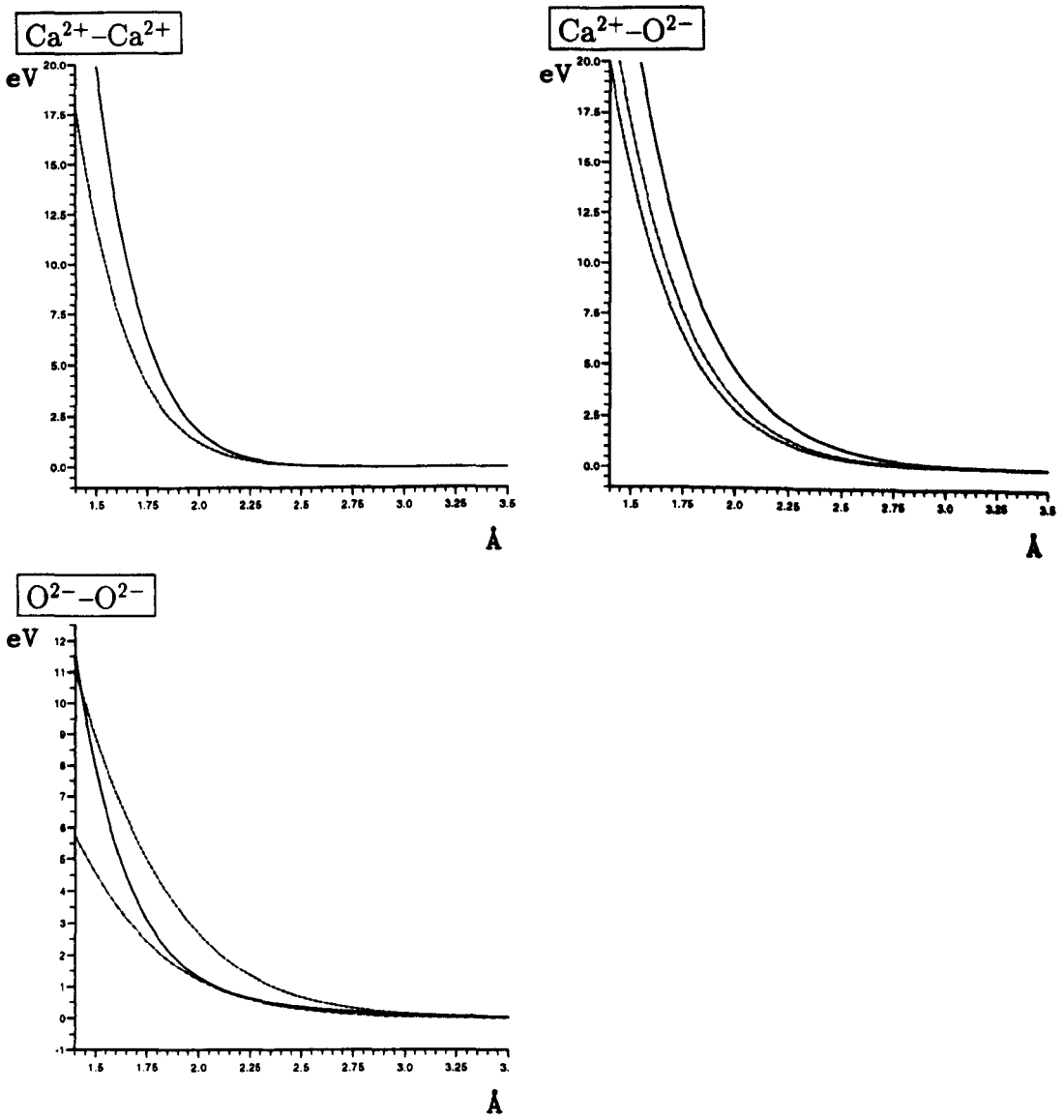

Figure 6.16: Interionic potentials for CaO. Key: full curves – SCF/SCF; dashed curves – electron gas; dot-dash curves – shifted electron gas.

6.6 Discussion

In this work we have showed how the theory of inter-molecular forces may be applied to the calculation of inter-ionic potentials. This extension is in principle quite straightforward, just requiring essentially a core removal for the optimisation of the fragment wavefunction. A similar strategy could be adopted for the super-system for schemes that seek to relax the fragments in the presence of each other (as outlined in appendix 6.a).

The major difficulty in the potential generation has proven to be the realistic description of the Madelung field experienced by the central ions. In comparison the role of correlation and dispersion effects are relatively minor, although the dispersion contribution is probably required particularly for accurate anion–anion potentials. Of course, the potential will normally be employed in conjunction with, say, the shell model to take into account the polarisation energy. The relatively strong dependence of the potentials on the Madelung field suggests that the whole concept of widely transferable potentials is perhaps a dubious one. A single potential is simply not going to be appropriate to widely differing crystal systems, not to mention modelling of interstitials or surface effects where the Madelung potential is likely to vary particularly drastically.

The simplifying factor we have found is that the potentials seem to vary relatively simply according to just the 'strength' of the Madelung field. This suggests two alternative ways forward which would make simulations practical. As mentioned previously, one may leave this as a free variable (as defined for example by the point charge lattice separation, a), i.e., as a parameter to be fitted to experimental data. Another approach would be to obtain a crude estimate of the magnitude of the field, by, during the simulation, taking the immediate surroundings of the respective ions into account. Both strategies would require obtaining suitable functional forms for our potentials, so that interaction energies could be obtained cheaply for essentially arbitrary Madelung fields.

From an electronic structure point of view we believe that most problems

concerned with the potential generation have been essentially solved. Also for the O^{2-}–O^{2-}potential our calculated potentials had the ability to perform quite well—but again only if the 'right' Madelung field was applied. Further work would of course take into account the dispersion contribution. This being said, we have highlighted some fundamental problems which should be addressed by *any* non-empirical approach attempting to use anion–anion potentials in practice.

6.a Appendix: Relaxing the fragments

Reoptimising the wavefunction for the super-system is necessary in order to describe the relaxation of the two ions in the presence of each other. The problem is how to define a point charge lattice consistent with both subsystem calculations. Ideally one would like an A-electron to 'feel' a point charge lattice appropriate to subsystem A, and a B-electron to 'feel' a point charge lattice appropriate to subsystem B.

An iterative scheme in which subsystem A is reoptimised in the field of the A point charge lattice and the B ion (with 'frozen' variational parameters), followed by a similar reoptimisation of ion B, etc., can be envisaged. The convergence characteristics of such a method would most likely depend heavily on the number of structures, but would probably be feasible in the case of two SCF wavefunctions (the crucial point being that due to the weakness of the interaction, the relaxation is fairly small).

More practical than an iterative procedure, would be a simple modification of the one-electron integrals, so that the reoptimisation could be performed in a single step. This suggests modifying the one-electron integrals according to

$$\langle \phi_1^a | \hat{h} | \phi_2^a \rangle \rightarrow \langle \phi_1^a | \hat{h} - \sum_{Q_A} \frac{Z_{Q_A}}{R_{iQ_A}} | \phi_2^a \rangle \tag{6.10}$$

where basis functions ϕ_1^a and ϕ_2^a are centred on system A, the Q_As are the point charges (charge: Z_{Q_A}) appropriate to A, and the distance to electron i is R_{iQ_A}.

Similarly

$$\langle \phi_1^b | \hat{h} | \phi_2^b \rangle \rightarrow \langle \phi_1^b | \hat{h} - \sum_{Q_B} \frac{Z_{Q_B}}{R_{iQ_B}} | \phi_2^b \rangle. \tag{6.11}$$

To be equivalent to the aforementioned iterative procedure, the integrals between orbitals on different centres should be left unchanged. Furthermore in order to be equivalent to the iterative procedure, only direct product spin functions should be included in the super-system calculation. After the reoptimisation step the one-electron energy should, as was the case for the subsystems, be recalculated without the point charge lattices.

The considerations from intermolecular forces regarding size-consistency etc. still hold true, so that if subsystem A is described by N_A structures and the subsystem B is described by N_B structures, all $N_A \times N_B$ structures must be included in the super-molecule calculation.

The effect of charge transfer can, in this type of scheme, be studied in a manner similar to the case of traditional intermolecular forces (this is one additional advantage of modifying the integrals instead of using the iterative scheme directly). Especially for systems containing transition metal ions this is likely to be important. The orbitals used for an ion of a given charge are obtained by an initial optimisation of the subsystem with the same charge. The core orbitals are best kept unchanged in each charged species, since reoptimisation of these would in most cases dramatically enlarge the calculation.

Bibliography

[1] C. R. A. Catlow and W. C. Mackrodt (eds), *Computer Simulations of Solids*, Springer-Verlag, Berlin, FRG (1982).

[2] J. H. Harding and A. H. Harker, UK AERE Harwell Report R-10425 (1982).

[3] J. Kendrick, W. C. Mackrodt, Sol. State Ion. **8**, 247 (1983).

[4] J. H. Harding, Mol. Sim. **4**, 255 (1990).

[5] N. L. Allan, D. L. Cooper and W. C. Mackrodt, Mol. Simul. **4**, 269 (1990).

[6] N. C. Pyper, Mol. Sim. **5**, 23 (1990).

[7] N. C. Pyper, Adv. Solid-State Chem. **2**, 223 (1991)

[8] D. L. Cooper, T. Thorsteinsson, M. Raimondi and J. Gerratt, Phil. Mag. B (in press).

[9] B. G. Dick and A. W. Overhauser, Phys. Rev. **112**, 90 (1958).

[10] G. D. Mahan, Solid State Ionics **1**, 29 (1980).

[11] P. W. Fowler and P. A. Madden, J. Phys. Chem. **89**, 2581 (1985).

[12] P. W. Fowler, Mol. Sim. **4**, 313 (1990).

[13] P. W. Fowler and P. Tole, Rev. Sol. State Sci. **5**, 149 (1991).

[14] A. J. Stone and I. C. Hayes, Faraday Disc. Chem. Soc. **73**, 19 (1984).

[15] G. A. Gallup and J. Gerratt, J. Chem. Phys. **83**, 2316 (1985).

[16] P. E. S.Wormer, T. van Berkel and Ad. van der Avoird, Mol. Phys. **29**, 1181 (1975).

[17] D. L. Cooper, J. Gerratt and M. Raimondi, Adv. Chem. Phys., **69**, 319 (1987).

[18] M. A. Matías, M. Raimondi, E. Tornaghi, D. L. Cooper and J. Gerratt, Mol. Phys. **83**, 87 (1994).

[19] M. Raimondi, M. Sironi, J. Gerratt and D. L. Cooper, (submitted for publication).

[20] R. W. Grimes, Mol. Sim. **5**, 9 (1990).

[21] R. D. Goodenough and V. A. Stenger, in: "*Comprehensive Inorganic Chemistry*", Vol. 1, Pergamon Press, New York (1973).

[22] T. H. Dunning Jr., J. Chem. Phys. **55**, 716 (1971).

[23] A. D. McLean and G. S. Chandler, J. Chem. Phys. **72**, 5639 (1980).

[24] H. M. Evjen, Phys. Rev. **39**, 675 (1932).

[25] MOLPRO is a package of *ab initio* programs written by H.-J. Werner and P. J. Knowles, with contributions of J. Almlöf, R. Amos, S. Elbert, K. Hampel, W. Meyer, K. Peterson, R. Pitzer, and A. Stone.

[26] P. W. M. Jacobs and M. L. Vernon, J. Chem. Soc. Faraday Trans. **86**(8), 1233 (1990).

[27] N. L. Allan, M. Braithwaite, D. L. Cooper, W. C. Mackrodt and S. C. Wright, J. Chem. Phys. **95**, 6792 (1991).

N. L. Allan, M. Braithwaite, D. L. Cooper, W. C. Mackrodt and B. Petch, Mol. Sim. **9**, 161 (1992).

N. L. Allan, M. Braithwaite, D. L. Cooper, B. Petch and W. C. Mackrodt, J. chem. Soc. Faraday Trans. **89**, 4369 (1993).

[28] D. J. Binks and R. W. Grimes, J. Amer. Cer. Soc. **76**, 2370 (1993).

[29] K. Recker, F. Wallrafen and K. Dupre, Naturwiss. **75**, 156 (1988).

[30] A. F. Ievin'sh, M. Straumanis and K. Karlsons, Zeit. Phys. Chem. B **40**, 146 (1938).

[31] T. Barth and G. Lunde, Zent. Min. Geol. A **1927**, 57 (1927).

[32] V. S. Urusov and V. V. Blinov, Izvest. Akad. Nauk SSSR, Ser. Khim. **12**, 278 (1970).

[33] M. Ahtee, Ann. Acad. Scient. Fenn., Ser. A6 **313**, 1 (1969).

[34] W. Primak, H. Kaufman and R. Ward, J. Am. Chem. Soc. **70**, 2043 (1948).

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