

## Valence bond theory. A tribute to the pioneers of 1927–1935

Roy McWeeny

Dipartimento di Chimica e Chimica Industriale, Università di Pisa, Italy

**Abstract.** The account presented illustrates that the ideas and intuitions of the pioneers of the early 'thirties were fundamentally correct and that, in one form or another, the theoretical methods they developed are here to stay.

### INTRODUCTION

Sixty years ago, what is now called “quantum chemistry” was just beginning: Heitler and London had demonstrated in their hydrogen molecule calculation (ref.1) the validity and feasibility of the new “quantum mechanics”, in whose development both were deeply involved, as a means of attacking chemical problems. And every month saw new applications, to more complicated molecules, of the theoretical approach they had established. The key ideas had a deceptive simplicity, involving the parallel or anti-parallel “vector coupling” of the spins of unpaired valence electrons, which masked a formidable mathematical apparatus connected with the construction of antisymmetric wavefunctions. The general formulation, to which Rumer and Weyl made huge contributions, was largely dependent on group theory – which at that time was not much loved even by theoretical physicists – but qualitative applications of the new “valence bond (VB) theory”, notably by Pauling and Wheland, continued; and the idea of “resonance” was immediately embraced by organic chemists. There can be no better testimony to the value of this approach, and its general acceptance by chemists, than the continuing popularity of Pauling's magnificent book “The Nature of the Chemical Bond” (ref.2).

Much of this was soon to be swept away. When Slater (ref.3) introduced his determinants in 1929 (claiming to have killed “the group pest”) a rival approach to the theory of the chemical bond, based on an independent-particle model (IPM) in which every electron occupied its own “molecular orbital”, began to gain popularity: MO theory involved no very sophisticated mathematics, it apparently performed almost equally well, and it was particularly successful – especially in the hands of Robert Mulliken – in the interpretation of molecular spectra. It was this theory that, with the advent of the computer, was to be almost universally adopted as the means of making completely *ab initio* calculations of molecular electronic structures. Nowadays, VB theory is hardly mentioned in many quantum chemistry textbooks.

After 50 years in the centre of the stage, with all the successes of self-consistent field (SCF) techniques and large-scale configuration interaction (CI), the limitations of the MO approach are by now well known. And new forms of VB theory are emerging. Modern VB theory rests heavily on the work of the pioneers of 1927–1935: it validates most of the qualitative ideas they propounded. The object of this review is to present the main issues involved and to show how some of the mathematical difficulties once considered insuperable are now being overcome.

### The hydrogen molecule calculation

Heitler and London's 1927 paper on the hydrogen molecule has been described (ref. 4) as “the greatest single contribution to the clarification of the chemist's conception of valence

that has been made since G.N.Lewis's suggestion in 1916 that the chemical bond between two atoms consists of a pair of electrons held jointly by the two atoms". In retrospect their idea was simple: if  $\chi_A(\mathbf{r}_1)$  is the wavefunction for an electron (1) of hydrogen atom A, and  $\chi_B(\mathbf{r}_2)$  is that for an electron (2) of atom B, then

$$\Phi_1(\mathbf{r}_1, \mathbf{r}_2) = \chi_A(\mathbf{r}_1)\chi_B(\mathbf{r}_2) \quad (1)$$

is an 'exact' wavefunction when A and B are remote, and will be a reasonable *approximate* wavefunction when they approach. But when A and B are in close proximity\*

$$\Phi_2(\mathbf{r}_1, \mathbf{r}_2) = \chi_B(\mathbf{r}_1)\chi_A(\mathbf{r}_2) \quad (2)$$

is equally acceptable (indistinguishability) and, taking account of spatial symmetry, the appropriate linear combinations to use will be \*

$$\Psi_S = \Phi_1 + \Phi_2, \quad \Psi_A = \Phi_1 - \Phi_2 \quad (3)$$

These symmetric and antisymmetric combinations lead to the famous results

$$E_S = \frac{Q + K}{1 + S^2}, \quad E_A = \frac{Q - K}{1 - S^2} \quad (4)$$

in which  $Q$  (the coulomb integral) represents the energy of the two atoms interacting as if they were 'rigid' charge distributions, while  $K$  (the exchange integral) reduces, with familiar notation, to

$$K = 2S \langle \chi_A | \hat{h} | \chi_B \rangle + \langle \chi_A \chi_B | g | \chi_B \chi_A \rangle . \quad (5)$$

At the equilibrium internuclear distance  $K$  is large and negative and accounts for about 90% of the deep minimum in the plot of  $E_S$  against distance. The negative value of  $K$ , however, arises not from the 2-electron term (which although it does involve exchange of electronic variables is actually positive) but rather from the *one*-electron term which contains a negative interaction energy representing the coulomb attraction between the nuclei and an *overlap density*  $\chi_A(\mathbf{r})\chi_B(\mathbf{r})$ . Nevertheless, the wavefunction  $\Psi_S (= \chi_A\chi_B + \chi_B\chi_A)$  is the prototype function for describing a covalent bond A-B; and the bonding has come to be associated with the 'exchange energy'  $K$ .

Of course, spin has been ignored so far. But provided relativistic terms in the Hamiltonian are ignored this is permissible: for with only two spin functions,  $\alpha(s)$  and  $\beta(s)$ , the ground state including spin will be represented correctly as

$$\Psi = (\chi_A\chi_B + \chi_B\chi_A)(\alpha\beta - \beta\alpha) \quad (6)$$

which satisfies the Pauli (antisymmetry) principle; and in computing the energy expectation value  $E = \langle \Psi | \hat{H} | \Psi \rangle / \langle \Psi | \Psi \rangle$  the spin factor will disappear in the spin integration! As van Vleck and Sherman (ref.5) first noted, it is the *symmetry of the spatial factor* that counts - to which that of the spin factor is only indirectly linked by the need for overall antisymmetry. As they put it, the spin is merely an 'indicator'.

The idea of linking the bonding and antibonding spatial factors in (3) with the coupling of spins is attractive: antisymmetric coupling of the spins ('pairing') produces a spin eigenfunc-

\* For convenience, electronic variables in a product function will always be in the order 1,2,...

\* Throughout this paper, trivial normalizing factors will be omitted.

tion with  $S=0$  and an energy  $E_S$  corresponding to bonding; symmetric ('parallel') coupling produces an antibonding triplet state with  $S = 1$ . Indeed, it is easy to show that, if the overlap term  $S^2$  in (4) is neglected, then the energies of the two states are exactly those of a *model* system with Hamiltonian

$$\hat{H}_s = Q' - 2K\hat{S}(1)\cdot\hat{S}(2) \quad (7)$$

—which contains only *spin operators* and numerical parameters,  $Q' (= Q - \frac{1}{2}K)$  and  $K$ . In other words, a *model* consisting of two interacting spins with the 'spin Hamiltonian' (7) has eigenstates whose energies coincide (in this VB approximation) with those of the lowest singlet and triplet states of the *real* system. Not surprisingly, this 'vector model' (generalized by Dirac, van Vleck and others) achieved instant popularity as a means of rationalizing the discussion of bonding in more complicated molecules.

Let us now recall briefly the IPM description of  $H_2$ . With the same two basis functions,  $\chi_A, \chi_B$ , the symmetry-adapted MOs are

$$\phi_1 = \chi_A + \chi_B, \quad \phi_2 = \chi_A - \chi_B, \quad (8)$$

the first being bonding, the second anti-bonding. The ground state is then represented by

$$\Psi = \phi_1\phi_2(\alpha\beta - \beta\alpha), \quad (9)$$

with both electrons in the bonding MO, instead of by (6). For the equilibrium bond length, the corresponding variational energy is not quite as good as in VB theory: but as the internuclear distance increases a much more dramatic failure occurs. The MO and VB curves are compared in Fig.1. The conclusion is obvious — and general. In all situations where covalent bonds are broken or formed (and thus in most chemical reactions) a simple MO approximation is of little value.

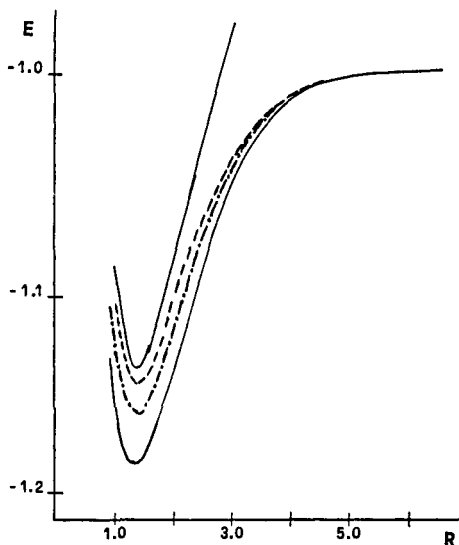


Figure 1. Energy curves for the hydrogen molecule. Upper and lower solid curves show, respectively, the energy from a simple MO approximation and the experimental energy. The curve --- is the corresponding VB approximation, and curve -.- is the "full-CI" approximation (3 structures).

The defects of the simple MO approximation can of course be corrected by admitting configuration interaction: in the  $H_2$  example the electrons (one or both) are promoted into the anti-bonding MO and the resultant configurational functions (CFs) are allowed to mix, with coefficients which are optimized by solving a set of secular equations. The Heitler-London function can likewise be improved by admitting CFs which correspond to electron transfer  $\chi_A \leftrightarrow \chi_B$  and thus to 'ionic' situations  $A^+ \dots B^-$  or  $A^- \dots B^+$ . When all possible configurations are admitted (in general the case of 'full CI') we obtain the 'basis-set limit', determined only by the limitation to the two functions  $\chi_A, \chi_B$  and it does not matter which approach (MO or VB) we use.

The 'full-CI' curve is also shown in Fig.1, but it is marked 'CF' because it is mathematically equivalent to the Coulson-Fischer approximation (ref.6), which is of *Heitler-London* form, namely

$$\Psi = (\tilde{\chi}_A \tilde{\chi}_B + \tilde{\chi}_B \tilde{\chi}_A)(\alpha\beta - \beta\alpha). \quad (10)$$

The spherical 1s orbitals ( $\chi_A, \chi_B$ ) have been replaced by

$$\tilde{\chi}_A = \chi_A + \lambda\chi_B, \quad \tilde{\chi}_B = \chi_B + \lambda\chi_A, \quad (\lambda \text{ positive}), \quad (11)$$

which are egg-shaped and have a much larger overlap than the unmodified AOs: they are *overlap-enhanced orbitals* (OEOs). With OEOs the admission of 'ionic' configurations is unnecessary: in this simple example one covalent structure reaches the full-CI limit.

### Generalizations

Some aspects of the calculations just discussed are of great generality. In particular, away from equilibrium geometry, the MO formulation which is usually adopted as the basis for *ab initio* electronic structure calculations cannot give reliable information on potential energy surfaces without including CI; and even for small molecules this may become very extensive. Fig.2 shows the situation for the fluorine molecule. The 1-configuration MO curve is completely false and the standard methods of admitting extensive CI (by many-body perturbation theory) are very slowly convergent. Indeed, in calculations by Handy *et al* (ref.7) on the water molecule, many thousands of CFs were admitted and MBPT estimates showed unsatisfactory convergence even in 48th order! Only when *orbital forms* are optimized (e.g. by MC SCF theory) is a realistic energy curve obtained; and the optimized orbitals are then *localized* on the separate fragments – as they would be in a VB calculation.

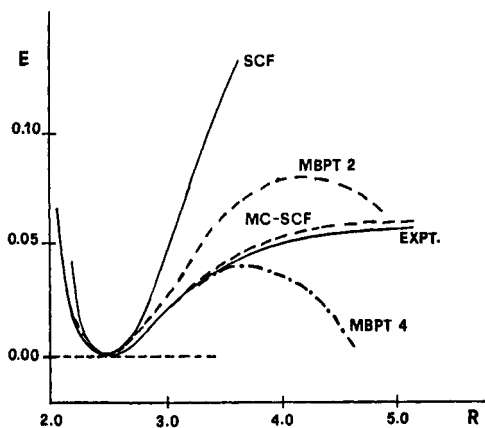


Figure 2. Energy curves for the fluorine molecule.

At this point it is clear that *in principle* the Heitler-London approach might well offer a more accurate and more compact description of molecular electronic structure as a function of geometry. The reasons for the slow development of VB techniques, at the *ab initio* level, are in fact purely technical: the non-orthogonality of the atomic orbitals (AOs or OEOs) leads to great computational problems when  $N$ , the number of electrons, increases; for in one way or another the matrix elements of the Hamiltonian must involve  $N!$  permutations, whereas with determinants of orthogonal orbitals they follow from trivial algorithms (Slater's rules). On the other hand, non-orthogonality is a key feature of all forms of VB theory: if overlap is eliminated (e.g. by orthogonalizing the AOs) exchange integrals such as (5) become positive and the 'covalent structures', with spins paired as in (4), predict no bonding (refs.7,8)! Indeed, the efficiency of the OEOs in (11) in describing the covalent bond simply reflects a well known 'principle of maximum overlap', used extensively in the 'thirties – in choosing the orbitals to be 'paired', in a polyatomic molecule, we look for those that overlap the most, combining them if necessary to form 'hybrids' with an enhanced overlap.

Fifty years ago, the technical difficulties just indicated were computationally insuperable: but the underlying theoretical structure of the VB approach was sound and at a qualitative level much progress was made. Two simple examples will serve to recall the generalization of Heitler and London's method to polyatomic molecules.

Example (i) The hydrogen molecule dimer,  $H_4$ .

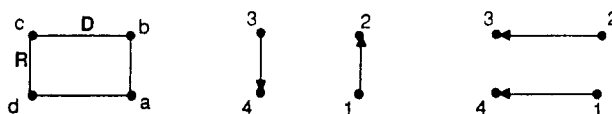


Figure 3. Rectangular conformation of the  $H_2$  dimer, with two VB structures.

For the  $H_2$  dimer in a rectangular conformation (Fig.3), the four 1s AOs ( $a, b, c, d$ ) may be used to construct a wavefunction by associating the orbital product

$$\Omega = a(\mathbf{r}_1)b(\mathbf{r}_2)c(\mathbf{r}_3)d(\mathbf{r}_4) \quad (12)$$

with a spin eigenfunction and antisymmetrizing. The spin functions introduced and extensively used by Rumer and Weyl (refs.10,11) contain pair-factors

$$\theta(s_i, s_j) = \alpha(s_i)\beta(s_j) - \beta(s_i)\alpha(s_j) \quad (13)$$

followed by, for non-singlet states, a parallel-coupled product such as  $\alpha(s_k)\alpha(s_l)\dots\alpha(s_p)$ . With  $g$  pairs, such a function is an eigenfunction of  $\hat{S}^2$  with eigenvalue  $S(S+1)$  and  $S = \frac{1}{2}(N-2g)$ . For two  $H_2$  molecules ( $D/R$  large in Fig.3) a plausible singlet function would thus be (introducing the usual antisymmetrizer  $\hat{A}$ )

$$\Phi_1 = \hat{A}[\Omega(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3, \mathbf{r}_4)\theta(s_1, s_2)\theta(s_3, s_4)] \quad (14)$$

which is equivalent to an antisymmetrized product of two Heitler-London functions. But another possibility would be

$$\Phi_2 = \hat{A}[\Omega(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3, \mathbf{r}_4)\theta(s_2, s_3)\theta(s_4, s_1)] \quad (15)$$

in which, by analogy, the 'bonds' are  $b \rightarrow c$  and  $d \rightarrow a$ . As Rumer first noted, there are no other independent possibilities:  $\Phi_1$  and  $\Phi_2$  are independent 'structures' and a structure with crossed links is merely a linear combination of the two. A general singlet CF for this electron configuration is

$$\Psi = c_1 \Phi_1 + c_2 \Phi_2 \quad (16)$$

— usually known as a 'resonance mixture' of the two structures — and the coefficients may be optimized variationally in the usual way. For  $D/R$  large,  $\Psi \rightarrow \Phi_1$  and the one structure uniquely describes the two separate hydrogen molecules: this is the famous *perfect pairing approximation*, used so extensively for saturated molecules where the assignment of the bonds may seem unambiguous.

All the necessary matrix element formulae for such calculations were derived in convenient pictorial form, notably by Rumer and by Pauling, during the early years of VB theory. Unfortunately, however, such algorithms depended on neglect of products of overlap integrals, an approximation now known to be untenable.

Example (ii) The water molecule,  $H_2O$ .

Here the 'valence configuration' of the oxygen atom may be regarded as  $O[1s^2 2s^2 2p_x^2 2p_y]$ , in which the  $2p_x$  and  $2p_y$  orbitals point roughly towards the hydrogen  $1s$  orbitals ( $H_1, H_2$  say) in order to maximize the overlaps. An electron-pairing scheme for the valence electrons may then be symbolized by

$$\Phi_1 = (ss)(p_x p_x)(p_x H_1)(p_y H_2),$$

which represents a perfect pairing approximation with two O-H electron pairs and two 'lone pairs'. But an alternative scheme would be

$$\Phi_2 = (ss)(p_x p_x)(p_x p_y)(H_1 H_2),$$

which shows a bond between the hydrogen atoms and an *internal* pairing on the oxygen. Other possibilities would be

$$\Phi_3 = (p_x p_x)(p_x p_x)(s H_1)(p_y H_2),$$

in which the s-type lone pair has been 'opened' and replaced by one of p-type; and

$$\Phi_4 = (ss)(p_x p_x)(p_x p_x)(p_y H_2),$$

which represents an ionic structure  $H^+ O^- - H$ . The full-CI function, with this limited orbital basis, will be a resonance mixture of 175 such structures in which  $\Phi_1$  (the perfect pairing structure) has the greatest weight.

The maximum overlap criterion suggests at once how the performance of the single-term approximation  $\Phi_1$  might be improved. Hückel and Pauling, both in 1931, introduced the concept of *hybridization*, essentially as a means of improving the perfect-pairing approximation. By using three in-plane mixtures of  $2s, 2p_x, 2p_y$ , namely  $h_1, h_2$  (pointing towards the hydrogens) and  $h_0$  (pointing along the axis of symmetry), the perfect-pairing function becomes

$$\Phi_1 = (p_x p_x)(h_0 h_0)(h_1 H_1)(h_2 H_2)$$

where the lone pair ( $h_0 h_0$ ) now points to the rear of the oxygen and the two bond pairs, with

enhanced overlap, give an excellent description of the bonding in the two O–H regions. In particular, as we shall presently confirm, improving the overlap almost eliminates the need to consider ionic structures.

Before passing to present-day developments, it must be repeated that all these fundamental ideas were in everyday use more than 50 years ago; and that much of the mathematical machinery for implementing them had already been set up by Heitler, London, Rumer and Weyl, and developed by Serber, Pauling, and many others.

### ***Ab initio* valence bond theory**

There are several ways of dealing with the non-orthogonality difficulty. The most direct is simply to expand every structure,

$$\Phi_{\kappa} = \hat{A}[\Omega\Theta_{\kappa}] \quad (17)$$

say, in which  $\Omega$  is an orbital product  $\phi_1\phi_2\dots\phi_N$  and  $\Theta_{\kappa}$  is a particular spin-coupled function, in terms of Slater determinants of non-orthogonal orbitals; and then to use the Löwdin rules (ref.12) for matrix element evaluation. Although feasible for small molecules (in the forms used by Karplus and Balint-Kurti (ref.13), Simonetta *et al* (ref.14), and others) this method is indirect, computationally heavy, and makes no use of existing VB techniques.

Other methods, initiated largely by Goddard (ref.15) and Gerratt (refs. 16,17), employ group theoretical techniques and are somewhat closer to traditional VB theory. With functions such as (17), matrix elements of the Hamiltonian may be reduced by writing the antisymmetrizer in the form

$$\hat{A} = (N!)^{-1} \sum_P \epsilon_P \hat{P} \quad (18)$$

where  $\epsilon_P = \pm 1$  (a parity factor) and the sum is over all  $N!$  electron index permutations. It follows readily that

$$H_{\kappa\lambda} = \langle \Phi_{\kappa} | \hat{H} | \Phi_{\lambda} \rangle = (N!)^{-1} \sum_P \epsilon_P H_P D_S(\hat{P})_{\kappa\lambda}, \quad (19)$$

where

$$H_P = \langle \Omega | \hat{H} | \hat{P}\Omega \rangle, \quad (20)$$

involving only *orbital products*, while  $D_S(\hat{P})_{\kappa\lambda}$  is the  $\kappa\lambda$ -element of the matrix associated with permutation  $\hat{P}$  in the irreducible representation  $D_S$  (corresponding to total spin quantum number  $S$ ) of the permutation group. Thus, when the spin eigenfunctions are orthogonal,

$$D_S(\hat{P})_{\kappa\lambda} = \langle \Theta_{\kappa} | \hat{P} | \Theta_{\lambda} \rangle. \quad (21)$$

Unfortunately, the Weyl-Rumer functions, constructed by spin pairing, are *not* orthogonal and the matrices defined by (21) do not provide a representation. The standard irreducible representations ('irreps') of the symmetric group  $S_N$  of  $N!$  permutations are carried by the *branching diagram* functions (refs.15,16), which are set up by coupling the spins  $s_1, s_2, \dots, s_N$  sequentially instead of in pairs. Every coupling sequence (up-down-up..., up-up-down..., etc.) is represented by a path in the branching diagram (Fig.4) and the number of distinct paths leading to a resultant spin  $S$  is Wigner's number

$$f_S^N = \binom{N}{\frac{1}{2}N - S} - \binom{N}{\frac{1}{2}N - S - 1}, \quad (22)$$

which is the number of linearly independent spin eigenfunctions of given  $S, M$  (quantum

numbers) and determines the dimension of  $D_S$ .

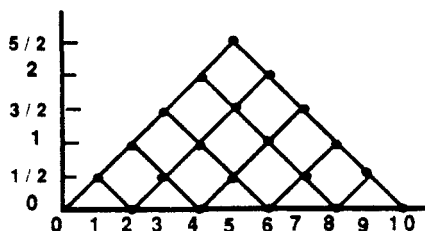


Figure 4. The branching diagram.

To determine the coefficients (21) from group theory one may consider the *generators*  $\hat{P}_{n-1,n}$  of the permutation group: these operators exchange only adjacent indices and the corresponding matrices are extremely sparse, with elements which are easily determined using an algorithm due to Young (see refs.19,20). All other matrices follow by matrix multiplication. By using this approach, Gerratt and coworkers (refs.16,17,21) have shown that the energy expectation value for a resonance mixture

$$\Psi = \sum_{\kappa} c_{\kappa} \Phi_{\kappa} \quad (23)$$

assumes the form

$$E = \sum_{\kappa, \lambda} c_{\kappa} [D_{ji}^{\kappa \lambda} \langle \phi_i | \hat{h} | \phi_j \rangle + \frac{1}{2} D_{ki,ij}^{\kappa \lambda} \langle \phi_i \phi_j | g | \phi_k \phi_l \rangle] c_{\lambda}^* \quad (24)$$

in which the  $D$ 's (which are density matrix elements) can be generated efficiently by recursion.

Since (23) is a *one-configuration* approximation (not including, for example, ionic configurations) it is important to optimize not only the expansion coefficients  $c_{\kappa}$  but also the forms of the orbitals  $\phi_1, \phi_2, \dots, \phi_N$  - which at this stage have been left unspecified. The best mixing coefficients satisfy the usual secular equations. But the best orbitals are much more difficult to obtain: they satisfy a complicated set of coupled pseudo-eigenvalue equations

$$\hat{F}_i \phi_i = \sum_k \phi_k \epsilon_{ki} \quad (\text{all } i) \quad (25)$$

in which there is one (exceedingly complicated) operator  $\hat{F}_i$  for each orbital. Iterative solution is of course necessary.

The results of such calculations, notably by Gerratt, Raimondi, and collaborators, are by now well known: they appear to give a spectacular confirmation of many of the conclusions reached by qualitative methods in the early 'thirties. To mention only one, a calculation on the benzene  $\pi$ -electron system (ref.22) shows that the optimized orbitals *do* localize mainly around the six conjugated carbon atoms; and furthermore that, on transforming from branching-diagram functions to the Weyl-Rumer basis, *two Kekulé structures* lead to almost the same basis-set limit as a full-CI calculation with 175 structures! This suggests that it might be preferable to work from the beginning in terms of the Weyl-Rumer coupling-scheme of classical VB theory. To examine this possibility it is convenient to use a "spin-free" approach - which again goes back to 1928.

### Spin-free methods

Heitler first showed that *exact* eigenfunctions of a spinless Hamiltonian could be combined



with spin factors to give an antisymmetric wavefunction

$$\Psi = \sum_{\kappa} F_{\kappa} \Theta_{\kappa} \quad (26)$$

where the spin eigenfunctions  $\{ \Theta_{\kappa} \}$  carry one irrep  $D_S$ , while the (degenerate) spatial eigenfunctions  $\{ F_{\kappa} \}$  carry another,  $D_{\tilde{S}}$  say, the two irreps being “associate” \* The exact energy corresponding to (26) is then

$$E = \frac{\langle \Psi | \hat{H} | \Psi \rangle}{\langle \Psi | \Psi \rangle} = \frac{\langle F_{\kappa} | \hat{H} | F_{\kappa} \rangle}{\langle F_{\kappa} | F_{\kappa} \rangle} \quad (27)$$

(numerator and denominator reducing to sums of identical terms) and the result is valid for *any* choice of  $\kappa$ :  $F_{\kappa}$  may be any one of the degenerate functions which carry the irrep  $D_{\tilde{S}}$ . This shows very clearly the significance of the statement (ref.5) that the spin serves only as an *indicator of symmetry* of the spatial wavefunction with respect to permutation of the electronic variables. Instead of explicitly considering spin it is sufficient to require that the spatial function belongs to a particular symmetry species, this observation being the basis of ‘spin-free quantum chemistry’ (ref.23).

Let us now choose a particular basis vector of the irrep  $D_{\tilde{S}}$ , that for  $\kappa = 1$  say, and set up a Wigner operator

$$\rho = (f_{\tilde{S}}^N / N!) \sum_P D_{\tilde{S}}(\hat{P})_{11} \hat{P} \quad (28)$$

The effect of this operator on *any orbital product*  $\Omega_K = k_1 k_2 \dots k_N$ , where  $k_1, k_2, \dots, k_N$  (collectively denoted by  $K$ ) indicates a selection of orbitals from the set  $\{ \phi_k \}$ , will be to produce a function

$$F_K = \rho \Omega_K \quad (29)$$

of the desired symmetry species.\*\* It will then be possible to expand an *exact* wavefunction of this symmetry species in the form

$$F = \sum_K c_K F_K. \quad (30)$$

This expansion is the spin-free analogue of CI and thus goes beyond the single-configuration approximation used in the last Section: it permits the calculation of wavefunctions up to the full-CI limit.

The matrix elements which enter the usual secular equations now reduce to

$$H_{KL} = \sum_P \epsilon_P D_S(\hat{P})_{11} \langle \Omega_K | \hat{H} | \hat{P} \Omega_L \rangle \quad (31)$$

and the sole technical problem is the rapid evaluation of such elements. There are two extreme possibilities, depending on which branching-diagram function is chosen as “the first” ( $\kappa = 1$ ). Gallup (ref.24) has, in effect, taken the first function when the standard ordering is used (i.e. the ‘last-letter’ sequence of the corresponding Young tableau): but this function, which corresponds to the uppermost path in the branching diagram, does not lead to VB theory in its classical form – for it is a linear combination of *all* the Weyl-Rumer, spin-paired, functions. Another possibility (ref.25) is to take the *last* function of the standard set i.e. the first function using *reverse* last-letter sequence. This second choice has many

\*\* From now on the species labels ( $\tilde{S}, 1$  in (28)) are redundant and will be omitted.

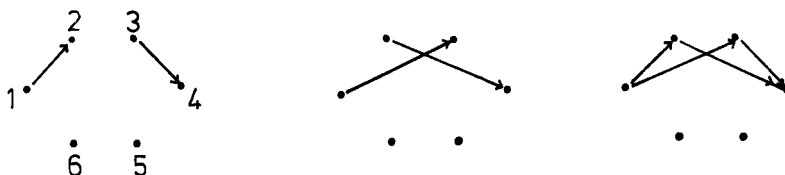
\* When the matrices are real orthogonal,  $D_{\tilde{S}}(\hat{P}) = \epsilon_P D_S(\hat{P})$ .

advantages. In particular, the corresponding spin eigenfunction is the single Weyl-Rumer function with couplings 1-2, 3-4,.... and when the orbitals in  $\Omega_K$  have indices  $k_1, k_2, k_3, \dots$  the projected function (29) is in one-to-one correspondence with the VB structure with links  $k_1 \rightarrow k_2, k_3 \rightarrow k_4, \dots$ etc.. In this way it is possible to retain close contact with the classical VB approach and the development (30) corresponds exactly to an expansion in classical structures. Moreover, this is a multi-configuration approach in which ionic structures are admitted simply by allowing repeated orbitals. There are other interesting connections not only with the work of the 'thirties but also with very recent progress. For example, the choice of index sets  $K$  needed in order to reach the full-CI limit is obtained by allocating the available indices to a two-column Young shape so that they increase on going down the columns but are non-decreasing in reading along the rows: the resultant *standard Weyl tableaux* are the ones adopted in current unitary group approaches to CI. Such aspects are considered in detail elsewhere (e.g. ref.26, Ch.10 and references therein).

Finally, let us consider the actual evaluation of the matrix elements in (31). The factor  $\langle \Omega_K | \hat{H} | \hat{P} \Omega_L \rangle$ , which involves only an orbital product and a permuted product, is trivial: it yields only sums of 1- and 2-electron integrals multiplied by chains of overlaps. The coefficient

$$D_S(\hat{P})_{11} = \langle \Theta_1 | \hat{P} \Theta_1 \rangle = \langle \Theta_1 | \Theta'_1 \rangle \quad (32)$$

is also simple when  $\Theta_1$  is of Rumer type, for the permutation merely produces a similar function  $\Theta'_1$  with a new set of links and the scalar product in (32) may be evaluated at once by inspection of a superposition pattern and use of the Rumer-Pauling rules. In dealing with a 6-electron triplet state, for example, with  $\Theta_1 = \theta(s_1, s_2)\theta(s_3, s_4)\alpha(s_5)\alpha(s_6)$ , the operator  $\hat{P}_{23}$  gives  $\Theta'_1 = \theta(s_1, s_3)\theta(s_2, s_4)\alpha(s_5)\alpha(s_6)$ : the two Rumer diagrams and their superposition pattern are then



and the required scalar product is (ref.26, p.218)

$$D_S(\hat{P}_{23})_{11} = (-1)^\nu \delta_E 2^{n-g} = \frac{1}{2},$$

$\nu$  being the number of arrow reversals to achieve head-to-head, tail-to-tail matching,  $g$  the number of electron pairs, and  $n$  the number of closed 'islands'. Generally, the factor  $\delta_E$  is 1 when there are no 'E chains' (open chains) connecting an even number of points) and is zero otherwise.

To perform complete *ab initio* VB calculations, with all non-orthogonality effects included, it is thus only necessary to generate permutations systematically, evaluate coefficients efficiently using the Rumer-Pauling algorithm, and accumulate the matrix element contributions in (31).

### Some applications

To test the practical feasibility of the approach outlined above, *ab initio* calculations have been made on the systems referred to in earlier Sections – the hydrogen-molecule dimer,

the water molecule, and the benzene  $\pi$ -electron system. Some preliminary results have been reported already (refs.25,27,28) and other applications are in progress. A detailed discussion will be published elsewhere.

(i) *The  $H_2$  dimer.* For four hydrogen atoms in plane rectangular geometry, with a minimal basis set of four 1s orbitals, there are two linearly independent non-polar singlet structures (Fig.3) and three triplets. Calculations were performed at three levels: (a) with unmodified AOs; (b) with OEOs of the form  $a' = \lambda(S_{ab}b + S_{ad}d)$  etc., with one variational parameter  $\lambda$ ; and (c) at the full-CI limit (20 singlet structures, 15 triplet). The results, for a range of geometries, are given in Table 1. Even though there is no attempt at full orbital optimization, two singlet structures or three triplets are sufficient to give close approach to the full-CI limit.

Table 1. Energies of lowest singlet and triplet states of  $H_2 \dots H_2$ .

D/R	1.0	1.5	2.0	2.5	3.0
<i>Singlet</i>					
2 Kek.(AO)	-1.636	-2.020	-2.176	-2.226	-2.242
2 Kek.(OEO)	-1.641	-2.052	-2.205	-2.254	-2.269
Full CI	-1.641	-2.058	-2.208	-2.255	-2.269
<i>Triplet</i>					
3 Kek.(AO)	-1.574	-1.693	-1.683	-1.672	-1.668
3 Kek.(OEO)	-1.645	-1.772	-1.732	-1.698	-1.684
Full CI	-1.647	-1.775	-1.737	-1.700	-1.685

(ii) *The  $H_2O$  molecule.* With axes and orbitals chosen as in Sect.3 Example (ii), ground-state calculations have been performed using a minimal basis (STO-6G). The 1s and 2p-z oxygen AOs (the latter being the only function of  $\pi$  symmetry) conveniently define a closed-shell core and a corresponding effective Hamiltonian for the 6 remaining valence electrons. The VB perfect-pairing approximation is then symbolized by

$$\Phi_1 = (ss)(p_x H_1)(p_y H_2)$$

in which the doubly occupied 2s orbital represents a first approximation to the oxygen  $\sigma$ -type lone pair. Energy calculations using the unmodified AOs yield rather poor results, much inferior at the equilibrium geometry) to those obtained by standard SCF methods (Fig. 5); and the situation is not much improved by admitting hybridization in which  $s, p_x, p_y \rightarrow h_0, h_1, h_2$  (with or without an orthogonality constraint on the hybrids). On allowing a small amount of mixing between the bond hybrids and corresponding hydrogen AOs ( $H_1, H_2$ ), to obtain OEOs of the form

$$\tilde{h}_i = h_i + \lambda H_i, \quad \tilde{H}_i = H_i + \mu h_i \quad (i = 1, 2)$$

there is a dramatic improvement (Fig.5). The perfect-pairing approximation gives a ground-state valence electron energy of  $E = -14.0887E_h$ , compared with the SCF value  $-14.0528E_h$  and the full-CI (frozen-core) limit  $-14.0981E_h$ . With two more structures, namely  $(h_0 h_0)(h_1 h_2)(H_1 H_2)$  and  $(h_1 h_1)(h_2 h_2)(H_1 H_2)$ , the energy falls to  $-14.0905E_h$  – recovering about 83% of the conventional ‘correlation’ energy.

(iii) *The benzene  $\pi$ -electron system.* In this case the  $\sigma$ -bonded framework is regarded as the core and the VB calculation refers to the remaining  $\pi$  electrons in the corresponding

effective field. The core orbitals (MOs) are obtained from a conventional SCF calculation and the effective Hamiltonian follows in the usual way.

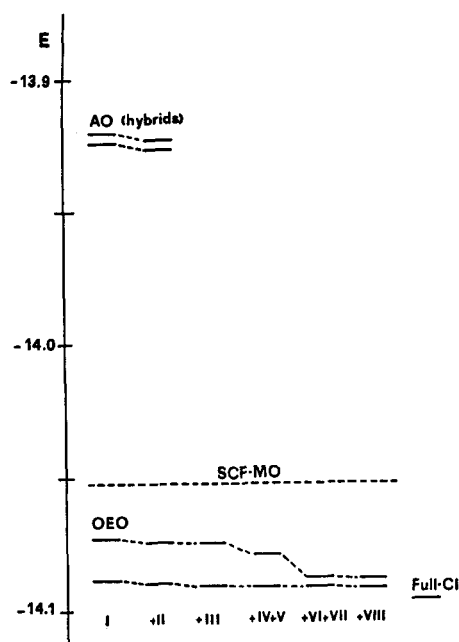


Figure 5. Some energy approximations for  $H_2O$ . The upper and lower levels, top left, result from the use of hybrid AOs with and without the usual orthogonality constraint. The corresponding levels at the bottom of the diagram arise when free mixing with all other AOs is admitted. Structures I-VIII are, respectively,  $(h_0h_0)(h_1H_1)(h_2H_2)$ ,  $(h_0h_0)(h_1h_2)(H_1H_2)$ ,  $(h_1h_1)(h_2h_2)(H_1H_2)$ ,  $(h_0H_1)(h_1h_1)(h_2H_2)$ ,  $(h_0H_2)(h_2h_2)(h_1H_1)$ ,  $(h_0h_0)(h_1h_1)(H_1H_2)$ ,  $(h_0h_0)(h_2h_2)(H_1H_2)$ ,  $(h_0h_0)(h_1h_1)(h_2h_2)$

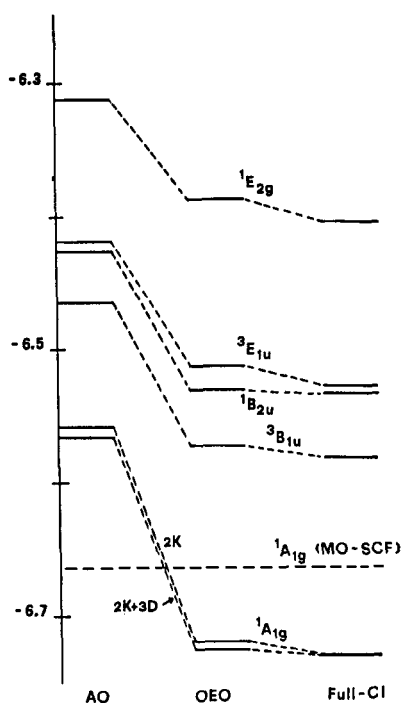


Figure 6. Energy approximations for ground and excited states of the benzene  $\pi$  system. H and K indicate the use of structures of Kekulé or Dewar (i.e. "long-bonded") type.

The calculations were performed at three levels, exactly as in (i), with OEOs which admit only nearest-neighbour mixing. The ground-state results (Fig.6) confirm those of Cooper et al (ref.22), two Kekulé-type structures based on OEOs giving a remarkably close approximation to the full-CI result \* which employs 175 singlet structures. The first few excited-state energies, obtained using the same optimized orbitals, also approach quite closely the full-CI results.

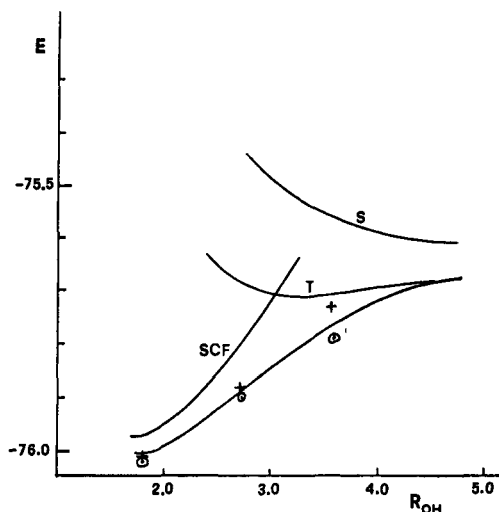


Figure 7. Energy curves for symmetric-stretch dissociation of  $H_2O$ . The ground-state curve was calculated using two Kekulé-type structures. Curves marked S and T are for first-excited singlet and triplet states. The points indicated by circles and crosses show the results of full-CI (245025 determinants) and MBPT3 calculations (ref. 7).

(iv) *Dissociation of  $H_2O$* . Since one of the main strengths of VB theory is its capacity of following the energy to a correct dissociation limit, it is of special interest to make a test calculation on such a process. Fig.7 shows the results of a minimal basis (STO-6G) calculation on the symmetric stretch dissociation of  $H_2O$ . The oxygen 1s and  $2p_x$  orbitals define a 'frozen' core, as in (ii), and the number of 'active' electrons is thus reduced to 6 (2 for the remaining  $\sigma$ -type lone pair and 4 for the two O-H bonds). The solid curve in Fig.7 was obtained using the structures  $(h_0 h_0)(h_1 H_1)(h_2 H_2)$  and  $(h_0 h_0)(h_1 h_1)(H_1 H_2)$ , with full optimization of the valence orbitals for every point. At the equilibrium distance the contribution of the second structure is very small but towards the dissociation limit it becomes important: there is an avoided crossing with the first excited singlet curve, after which the ground-state wavefunction begins to describe *triplet* spin couplings on the oxygen atom and in the H...H region (the spin eigenfunction implied then being that of the 'top' path in the branching diagram). The first triplet energy surface is also shown.

The points enclosed in open circles were obtained by Handy et al (ref.7) in calculations designed to test the efficiency of many-body perturbation theory (MBPT) as a means of correcting the usual closed-shell SCF reference function: they correspond to *complete* CI (involving the processing, at each point, of 245,025 Slater determinants) with a 6-31G basis\*. The MBPT approximations, even taken to order 48, failed to give satisfactory convergence for bond lengths more than twice the equilibrium value, showing clearly the need to employ

\* The full-CI results for all states were obtained by Dr.F.E.Jorge, using the same basis but a completely independent approach (see ref.29 for the method).

\* Slightly inferior in quality to the one used in the present work.

a multi-configuration reference function. By contrast, the two-structure VB approximation gives a completely satisfactory account of both ground and first excited singlet surfaces at all distances. The first excited triplet needs only three structures. And at long range the oxygen atom correctly goes towards its triplet ground state.

### CONCLUSION

It is clear from the results described in the last Section that there are no insuperable difficulties in performing VB calculations of the kind envisaged and discussed in the 'thirties – but now in a fully *ab initio* form. It is true that the famous "N! problem" still remains; but this is also the case with conventional MO calculations, whenever the CI expansion converges slowly, because the number of configurations available is also dependent on factorials. In all cases where a serious attempt is made to approach the full-CI limit, the most important criterion for success is the rapid convergence of the expansion – and this is very strongly dependent on the nature of the orbitals employed. It should also be noted that  $N$  generally denotes the number of 'active' electrons: normally there is a core of 'passive' electrons (assigned in an MO-based approach to a common set of doubly-occupied orbitals) whose effect can always be rigorously absorbed into an effective Hamiltonian for the rest. It is normally assumed, with considerable experimental and theoretical confirmation, that processes such as electronic excitation, or bond breaking as a result of geometry change, primarily involve the active electrons in an effective field provided by the rest of the system. If this is indeed the case, then a vast area of chemistry now comes within the scope of modern valence bond theory. Some indication of the potential of the method is evident from other recent reviews (refs.30-32).

The main features of the spin-free formulation described in Section 5 may be summarized as follows:

- (i) For up to at least 10 electrons outside a closed shell there are no technical difficulties in a 'permutation-driven' approach. Permutations and corresponding coefficients in the Wigner operator can be generated rapidly and efficiently using only logical operations and integer arithmetic.
  - (ii) With the relatively small non-orthogonal basis sets needed, a complete calculation can be done in the fast memory of a small computer \*\*, no storage and repeated processing of intermediate data (e.g. permutation matrices) being required.
  - (iii) By using localized orbitals, which pass smoothly into AOs on separating the atoms, correct dissociation is ensured for all possible bond-breaking geometries. In other words there are no 'size-consistency' difficulties and calculations of entire reaction surfaces are feasible.
  - (iv) Perhaps most important of all: wavefunctions of excellent quality can be obtained using only a small number of classical VB structures, provided overlap-enhanced orbitals are admitted, and every structure has an immediate significance for any chemist. All the intuitions built up during the last 150 years of chemistry – concerning for example the importance of Kekulé-type structures or the role of hybridization – can thus be fully exploited.
- These conclusions strongly suggest that the ideas and intuitions of the pioneers of the early

---

\*\* All the calculations reported in this paper were in fact performed on a personal computer with only 640 Kb of fast store.

'thirties were fundamentally correct and that, in one form or another, the theoretical methods they developed are here to stay.

## REFERENCES

- 1) W.Heitler and F.London, Z.Phys. 44(1927)455
- 2) *The Nature of the Chemical Bond*, L.Pauling, Cornell University Press, Ithaca, 1948
- 3) J.C.Slater, Phys.Rev. 34(1929)1293
- 4) *Introduction to Quantum Mechanics*, L.Pauling and E.B.Wilson, McGraw-Hill, New York, 1935
- 5) J.H.van Vleck and A.Sherman, Rev.Mod.Phys. 7(1935)167
- 6) C.A.Coulson and I.Fischer, Phil.Mag. 40(1949)306
- 7) N.C.Handy, P.J.Knowles and K.Somarsundran, Theor.Chim.Acta 68(1985)87
- 8) J.C.Slater, J.Chem.Phys. 19(1950)220
- 9) R.McWeeny, Proc.Roy.Soc. A227(1955)288
- 10) G.Rumer, Göttingen Nachr.(1932)377
- 11) *The Theory of Groups and Quantum Mechanics*, H.Weyl, Dover, New York, 1956 (Translation of the German Edition of 1931)
- 12) P.-O.Löwdin, Phys.Rev. 97(1955)1474
- 13) G.Balint-Kurti and M.Karplus, J.Chem.Phys. 50(1968)478
- 14) M.Simonetta, E.Gianinetti and I.Vandone, J.Chem.Phys. 48(1968)1579
- 15) W.A.Goddard Jr., Phys.Rev. 157(1967)73, 81
- 16) J.Gerratt, Adv.in At.Mol.Phys. 7(1971)141
- 17) N.C.Pyper and J.Gerratt, Proc.Roy.Soc.(Lond.) A355(1977)402
- 18) *Tables of Molecular Integrals*, M.Kotani, A.Ameniya, E.Ishiguro and T.Kimura, Maruzen, Tokyo, 1955
- 19) *Spin Eigenfunctions: Construction and Use*, R.Pauncz, Plenum, New York, 1979
- 20) *Substitutional Analysis*, D.E.Rutherford, Edinburgh University Press, Edinburgh, 1948
- 21) J.Gerratt and M.Raimondi, Proc.Roy.Soc.(Lond.) A371(1980)525
- 22) D.L.Cooper, J.Gerratt and M.Raimondi, Nature 323(1980)699
- 23) F.A.Matsen, Adv.Quantum Chem. 1(1964)50
- 24) G.A.Gallup, Adv.Quantum Chem. 16(1973)229
- 25) R.McWeeny, Int.J.Quantum.Chem. 23(1983)405
- 26) *Methods of Molecular Quantum Mechanics 2nd Ed.*, R.McWeeny, Academic Press, London, 1988
- 27) R.McWeeny, Theoretica Chim.Acta 73(1988)115
- 28) R.McWeeny and F.E.Jorge, J.Mol.Struct.(Theochem) 169(1988)459
- 29) J.M.Norbeck and R.McWeeny, Chem.Phys.Letters 34(1975)206
- 30) G.A.Gallup, R.L.Vance, J.R.Collins, J.M.Norbeck, Adv. Quantum Chem.16(1982)229
- 31) M.Raimondi, M.Simonetta and G.F.Tantardini (1985) Comput. Phys.Repts. 2(1985)171
- 32) D.L.Cooper, J.Gerratt and M.Raimondi, Adv.Chem.Phys. 69(1987)319