CAMBRIDGE

Cambridge University Press
0521803926 - Valence Bond Methods: Theory and Applications
Gordon A. Gallup
Excerpt
More information

Part I

Theory and two-electron systems

1

Introduction

1.1 History

In physics and chemistry making a direct calculation to determine the structure or properties of a system is frequently very difficult. Rather, one assumes at the outset an ideal or asymptotic form and then applies adjustments and corrections to make the calculation adhere to what is believed to be a more realistic picture of nature. The practice is no different in molecular structure calculation, but there has developed, in this field, two different "ideals" and two different approaches that proceed from them.

The approach used first, historically, and the one this book is about, is called the valence bond (VB) method today. Heitler and London[8], in their treatment of the H_2 molecule, used a trial wave function that was appropriate for two H atoms at long distances and proceeded to use it for all distances. The ideal here is called the "separated atom limit". The results were qualitatively correct, but did not give a particularly accurate value for the dissociation energy of the H−H bond. After the initial work, others made adjustments and corrections that improved the accuracy. This is discussed fully in Chapter 2. A crucial characteristic of the VB method is that the orbitals of different atoms must be considered as nonorthogonal.

The other approach, proposed slightly later by Hund[9] and further developed by Mulliken[10] is usually called the molecular orbital (MO) method. Basically, it views a molecule, particularly a diatomic molecule, in terms of its "united atom limit". That is, H_2 is a He atom (not a real one with neutrons in the nucleus) in which the two positive charges are moved from coinciding to the correct distance for the molecule.[1] HF could be viewed as a Ne atom with one proton moved from the nucleus out to the molecular distance, etc. As in the VB case, further adjustments and corrections may be applied to improve accuracy. Although the united atom limit is not often mentioned in work today, its heritage exists in that MOs are universally

---

[1] Although this is impossible to do in practice, we can certainly calculate the process on paper.

considered to be mutually orthogonal. We touch only occasionally upon MO theory in this book.

As formulated by Heitler and London, the original VB method, which was easily extendible to other diatomic molecules, supposed that the atoms making up the molecule were in (high-spin) S states. Heitler and Rumer later extended the theory to polyatomic molecules, but the atomic S state restriction was still, with a few exceptions, imposed. It is in this latter work that the famous Rumer[11] diagrams were introduced. Chemists continue to be intrigued with the possibility of correlating the Rumer diagrams with bonding structures, such as the familiar Kekulé and Dewar bonding pictures for benzene.

Slater and Pauling introduced the idea of using whole atomic configurations rather than S states, although, for carbon, the difference is rather subtle. This, in turn, led to the introduction of hybridization and the maximum overlap criterion for bond formation[1].

Serber[12] and Van Vleck and Sherman[13] continued the analysis and introduced symmetric group arguments to aid in dealing with spin. About the same time the Japanese school involving Yamanouchi and Kotani[14] published analyses of the problem using symmetric group methods.

All of the foregoing work was of necessity fairly qualitative, and only the smallest of molecular systems could be handled. After WWII digital computers became available, and it was possible to test many of the qualitative ideas quantitatively.

In 1949 Coulson and Fisher[15] introduced the idea of nonlocalized orbitals to the VB world. Since that time, suggested schemes have proliferated, all with some connection to the original VB idea. As these ideas developed, the importance of the spin degeneracy problem emerged, and VB methods frequently were described and implemented in this context. We discuss this more fully later.

As this is being written at the beginning of the twenty-first century, even small computers have developed to the point where *ab initio* VB calculations that required "supercomputers" earlier can be carried out in a few minutes or at most a few hours. The development of parallel "supercomputers", made up of many inexpensive personal computer units is only one of the developments that may allow one to carry out ever more extensive *ab initio* VB calculations to look at and interpret molecular structure and reactivity from that unique viewpoint.

1.2 Mathematical background

Data on individual atomic systems provided most of the clues physicists used for constructing quantum mechanics. The high spherical symmetry in these cases allows significant simplifications that were of considerable usefulness during times when procedural uncertainties were explored and debated. When the time came

Cambridge University Press
0521803926 - Valence Bond Methods: Theory and Applications
Gordon A. Gallup
Excerpt
More information

to examine the implications of quantum mechanics for molecular structure, it was immediately clear that the lower symmetry, even in diatomic molecules, causes significantly greater difficulties than those for atoms, and nonlinear polyatomic molecules are considerably more difficult still. The mathematical reasons for this are well understood, but it is beyond the scope of this book to pursue these questions. The interested reader may investigate many of the standard works detailing the properties of Lie groups and their applications to physics. There are many useful analytic tools this theory provides for aiding in the solution of partial differential equations, which is the basic mathematical problem we have before us.

1.2.1 Schrödinger's equation

Schrödinger's space equation, which is the starting point of most discussions of molecular structure, is the partial differential equation mentioned above that we must deal with. Again, it is beyond the scope of this book to give even a review of the foundations of quantum mechanics, therefore, we assume Schrödinger's space equation as our starting point. Insofar as we ignore relativistic effects, it describes the energies and interactions that predominate in determining molecular structure. It describes in quantum mechanical terms the kinetic and potential energies of the particles, how they influence the wave function, and how that wave function, in turn, affects the energies. We take up the potential energy term first.

Coulomb's law

Molecules consist of electrons and nuclei; the principal difference between a molecule and an atom is that the latter has only one particle of the nuclear sort. Classical potential theory, which in this case works for quantum mechanics, says that Coulomb's law operates between charged particles. This asserts that the potential energy of a pair of spherical, charged objects is

$$V(|\vec{r}_1 - \vec{r}_2|) = \frac{q_1 q_2}{|\vec{r}_1 - \vec{r}_2|} = \frac{q_1 q_2}{r_{12}}, \tag{1.1}$$

where q_1 and q_2 are the charges on the two particles, and r_{12} is the scalar distance between them.

Units

A short digression on units is perhaps appropriate here. We shall use either Gaussian units in this book or, much more frequently, Hartree's atomic units. Gaussian units, as far as we are concerned, are identical with the old cgs system of units with the added proviso that charges are measured in unnamed *electrostatic units*, esu. The value of $|e|$ is thus $4.803206808 \times 10^{-10}$ esu. Keeping this number at hand is all that will be required to use Gaussian units in this book.

Hartree's atomic units are usually all we will need. These are obtained by assigning mass, length, and time units so that the mass of the electron, $m_e = 1$, the electronic charge, $|e| = 1$, and Planck's constant, $\hbar = 1$. An upshot of this is that the Bohr radius is also 1. If one needs to compare energies that are calculated in atomic units (hartrees) with measured quantities it is convenient to know that 1 hartree is 27.211396 eV, 6.27508×10^5 cal/mole, or 2.6254935×10^6 joule/mole. The reader should be cautioned that one of the most common pitfalls of using atomic units is to forget that the charge on the electron is -1. Since equations written in atomic units have no m_es, es, or $\hbar$s in them explicitly, their being all equal to 1, it is easy to lose track of the signs of terms involving the electronic charge. For the moment, however, we continue discussing the potential energy expression in Gaussian units.

The full potential energy

One of the remarkable features of Coulomb's law when applied to nuclei and electrons is its additivity. The potential energy of an assemblage of particles is just the sum of all the pairwise interactions in the form given in Eq. (1.1). Thus, consider a system with K nuclei, $\alpha = 1, 2, \ldots, K$ having atomic numbers Z_α. We also consider the molecule to have N electrons. If the molecule is uncharged as a whole, then $\sum Z_\alpha = N$. We will use lower case Latin letters, $i, j, k, \ldots$, to label electrons and lower case Greek letters, $\alpha, \beta, \gamma, \ldots$, to label nuclei. The full potential energy may then be written

$$V = \sum_{\alpha<\beta} \frac{e^2 Z_\alpha Z_\beta}{r_{\alpha\beta}} - \sum_{i\alpha} \frac{e^2 Z_\alpha}{r_{i\alpha}} + \sum_{i<j} \frac{e^2}{r_{ij}}. \tag{1.2}$$

Many investigations have shown that any deviations from this expression that occur in reality are many orders of magnitude smaller than the sizes of energies we need be concerned with.[2] Thus, we consider this expression to represent exactly that part of the potential energy due to the charges on the particles.

The kinetic energy

The kinetic energy in the Schrödinger equation is a rather different sort of quantity, being, in fact, a differential operator. In one sense, it is significantly simpler than the potential energy, since the kinetic energy of a particle depends only upon what it is doing, and not on what the other particles are doing. This may be contrasted with the potential energy, which depends not only on the position of the particle in question, but on the positions of all of the other particles, also. For our molecular

---

[2] The first correction to this expression arises because the transmission of the electric field from one particle to another is not instantaneous, but must occur at the speed of light. In electrodynamics this phenomenon is called a *retarded potential*. Casimir and Polder[16] have investigated the consequences of this for quantum mechanics. The effect within distances around 10^{-7} cm is completely negligible.

system the kinetic energy operator is

$$T = -\sum_{\alpha} \frac{\hbar^2}{2M_\alpha} \nabla_\alpha^2 - \sum_i \frac{\hbar^2}{2m_e} \nabla_i^2, \tag{1.3}$$

where M_α is the mass of the α^{th} nucleus.

The differential equation

The Schrödinger equation may now be written symbolically as

$$(T + V)\Psi = E\Psi, \tag{1.4}$$

where E is the numerical value of the total energy, and Ψ is the wave function. When Eq. (1.4) is solved with the various constraints required by the rules of quantum mechanics, one obtains the total energy and the wave function for the molecule. Other quantities of interest concerning the molecule may subsequently be determined from the wave function.

It is essentially this equation about which Dirac[17] made the famous (or infamous, depending upon your point of view) statement that all of chemistry is reduced to physics by it:

> The general theory of quantum mechanics is now almost complete, the imperfections that still remain being in connection with the exact fitting in of the theory with relativity ideas. These give rise to difficulties only when high-speed particles are involved, and are therefore of no importance in the consideration of atomic and molecular structure and ordinary chemical reactions .... The underlying physical laws necessary for the mathematical theory of a large part of physics and the whole of chemistry are thus completely known, and the difficulty is only that the exact application of these laws leads to equations much too complicated to be soluble ....

To some, with what we might call a practical turn of mind, this seems silly. Our mathematical and computational abilities are not even close to being able to give useful general solutions to it. To those with a more philosophical outlook, it seems significant that, at our present level of understanding, Dirac's statement is apparently true. Therefore, progress made in methods of solving Eq. (1.4) is improving our ability at making predictions from this equation that are useful for answering chemical questions.

The Born–Oppenheimer approximation

In the early days of quantum mechanics Born and Oppenheimer[18] showed that the energy and motion of the nuclei and electrons could be separated approximately. This was accomplished using a perturbation treatment in which the perturbation parameter is $(m_e/M)^{1/4}$. In actuality, the term "Born–Oppenheimer approximation" is frequently ambiguous. It can refer to two somewhat different theories. The first is the reference above and the other one is found in an appendix of the book by Born and Huang on crystal structure[19]. In the latter treatment, it is assumed, based upon physical arguments, that the wave function of Eq. (1.4) may be written as the product of two other functions

$$\Psi(\vec{r}_i, \vec{r}_\alpha) = \phi(\vec{r}_\alpha)\psi(\vec{r}_i, \vec{r}_\alpha), \tag{1.5}$$

where the nuclear positions $\vec{r}_\alpha$ given in ψ are parameters rather than variables in the normal sense. The ϕ is the actual wave function for nuclear motion and will not concern us at all in this book. If Eq. (1.5) is substituted into Eq. (1.4), various terms are collected, and small quantities dropped, we obtain what is frequently called the Schrödinger equation for the electrons using the Born–Oppenheimer approximation

$$-\frac{\hbar^2}{2m_e}\sum_i \nabla_i^2\psi + V\psi = E(\vec{r}_\alpha)\psi, \tag{1.6}$$

where we have explicitly observed the dependence of the energy on the nuclear positions by writing it as $E(\vec{r}_\alpha)$. Equation (1.6) might better be termed the Schrödinger equation for the electrons using the *adiabatic* approximation[20]. Of course, the only difference between this and Eq. (1.4) is the presence of the nuclear kinetic energy in the latter. A heuristic way of looking at Eq. (1.6) is to observe that it would arise if the masses of the nuclei all passed to infinity, i.e., the nuclei become stationary. Although a physically useful viewpoint, the actual validity of such a procedure requires some discussion, which we, however, do not give.

We now go farther, introducing atomic units and rearranging Eq. (1.6) slightly,

$$-\frac{1}{2}\sum_i \nabla_i^2\psi - \sum_{i\alpha}\frac{Z_\alpha}{r_{i\alpha}}\psi + \sum_{i<j}\frac{1}{r_{ij}}\psi + \sum_{\alpha<\beta}\frac{Z_\alpha Z_\beta}{r_{\alpha\beta}}\psi = E_e\psi. \tag{1.7}$$

This is the equation with which we must deal. We will refer to it so frequently, it will be convenient to have a brief name for it. It is the *electronic Schrödinger equation*, and we refer to it as the ESE. Solutions to it of varying accuracy have been calculated since the early days of quantum mechanics. Today, there exist computer programs both commercial and in the public domain that will carry out calculations to produce approximate solutions to the ESE. Indeed, a program of this sort is available from the author through the Internet.[3] Although not as large as some of the others available, it will do many of the things the bigger programs will do, as well as a couple of things they do not: in particular, this program will do VB calculations of the sort we discuss in this book.

---

[3] The CRUNCH program, `http://phy-ggallup.unl.edu/crunch/`

Cambridge University Press
0521803926 - Valence Bond Methods: Theory and Applications
Gordon A. Gallup
Excerpt
More information

1.3 The variation theorem

1.3.1 General variation functions

If we write the sum of the kinetic and potential energy operators as the Hamiltonian operator $T + V = H$, the ESE may be written as

$$H\Psi = E\Psi. \tag{1.8}$$

One of the remarkable results of quantum mechanics is the variation theorem, which states that

$$W = \frac{\langle\Psi|H|\Psi\rangle}{\langle\Psi|\Psi\rangle} \geq E_0, \tag{1.9}$$

where E_0 is the lowest allowed eigenvalue for the system. The fraction in Eq. (1.9) is frequently called the *Rayleigh quotient*. The basic use of this result is quite simple. One uses arguments based on similarity, intuition, guess-work, or whatever, to devise a suitable function for Ψ. Using Eq. (1.9) then necessarily gives us an upper bound to the true lowest energy, and, if we have been clever or lucky, the upper bound is a good approximation to the lowest energy. The most common way we use this is to construct a trial function, Ψ, that has a number of parameters in it. The quantity, W, in Eq. (1.9) is then a function of these parameters, and a minimization of W with respect to the parameters gives the best result possible within the limitations of the choice for Ψ. We will use this scheme in a number of discussions throughout the book.

1.3.2 Linear variation functions

A trial variation function that has linear variation parameters only is an important special case, since it allows an analysis giving a systematic improvement on the lowest upper bound as well as upper bounds for excited states. We shall assume that $\phi_1, \phi_2, \ldots,$ represents a complete, normalized (but not necessarily orthogonal) set of functions for expanding the exact eigensolutions to the ESE. Thus we write

$$\Psi = \sum_{i=1}^{\infty} \phi_i C_i, \tag{1.10}$$

where the C_i are the variation parameters. Substituting into Eq. (1.9) we obtain

$$W = \frac{\sum_{ij} H_{ij} C_i^* C_j}{\sum_{ij} S_{ij} C_i^* C_j}, \tag{1.11}$$

where

$$H_{ij} = \langle\phi_i|H|\phi_j\rangle, \tag{1.12}$$

$$S_{ij} = \langle\phi_i|\phi_j\rangle. \tag{1.13}$$

We differentiate W with respect to the C_i^*s and set the results to zero to find the minimum, obtaining an equation for each C_i^*,

$$\sum_j (H_{ij} - WS_{ij})C_j = 0; \qquad i = 1, 2, \ldots. \tag{1.14}$$

In deriving this we have used the properties of the integrals $H_{ij} = H_{ji}^*$ and a similar result for S_{ij}. Equation (1.14) is discussed in all elementary textbooks wherein it is shown that a $C_j \neq 0$ solution exists only if the W has a specific set of values. It is sometimes called the *generalized eigenvalue problem* to distinguish from the case when S is the identity matrix. We wish to pursue further information about the Ws here.

Let us consider a variation function where we have chosen n of the functions, ϕ_i. We will then show that the eigenvalues of the n-function problem divide, i.e., occur between, the eigenvalues of the $(n+1)$-function problem. In making this analysis we use an extension of the methods given by Brillouin[21] and MacDonald[22].

Having chosen n of the ϕ functions to start, we obtain an equation like Eq. (1.14), but with only $n \times n$ matrices and n terms,

$$\sum_{j=1}^{n} \left(H_{ij} - W^{(n)} S_{ij}\right) C_j^{(n)} = 0; \qquad i = 1, 2, \ldots, n. \tag{1.15}$$

It is well known that sets of linear equations like Eq. (1.15) will possess nonzero solutions for the $C_j^{(n)}$s only if the matrix of coefficients has a rank less than n. This is another way of saying that the determinant of the matrix is zero, so we have

$$\left|H - W^{(n)} S\right| = 0. \tag{1.16}$$

When expanded out, the determinant is a polynomial of degree n in the variable $W^{(n)}$, and it has n real roots if H and S are both Hermitian matrices, and S is positive definite. Indeed, if S were not positive definite, this would signal that the basis functions were not all linearly independent, and that the basis was defective. If $W^{(n)}$ takes on one of the roots of Eq. (1.16) the matrix $H - W^{(n)} S$ is of rank $n - 1$ or less, and its rows are linearly dependent. There is thus at least one more nonzero vector with components $C_j^{(n)}$ that can be orthogonal to all of the rows. This is the solution we want.

It is useful to give a matrix solution to this problem. We affix a superscript $^{(n)}$ to emphasize that we are discussing a matrix solution for n basis functions. Since $S^{(n)}$ is Hermitian, it may be diagonalized by a unitary matrix, $T = (T^\dagger)^{-1}$

$$T^\dagger S^{(n)} T = s^{(n)} = \mathrm{diag}\left(s_1^{(n)}, s_2^{(n)}, \ldots, s_n^{(n)}\right), \tag{1.17}$$

where the diagonal elements of $s^{(n)}$ are all real and positive, because of the Hermitian and positive definite character of the overlap matrix. We may construct the inverse square root of $s^{(n)}$, and, clearly, we obtain

$$\left[T\left(s^{(n)}\right)^{-1/2}\right]^{\dagger} S^{(n)} T\left(s^{(n)}\right)^{-1/2} = I. \tag{1.18}$$

We subject $H^{(n)}$ to the same transformation and obtain

$$\left[T\left(s^{(n)}\right)^{-1/2}\right]^{\dagger} H^{(n)} T\left(s^{(n)}\right)^{-1/2} = \bar{H}^{(n)}, \tag{1.19}$$

which is also Hermitian and may be diagonalized by a unitary matrix, U. Combining the various transformations, we obtain

$$V^{\dagger} H^{(n)} V = h^{(n)} = \mathrm{diag}\left(h_1^{(n)}, h_2^{(n)}, \ldots, h_n^{(n)}\right), \tag{1.20}$$

$$V^{\dagger} S^{(n)} V = I, \tag{1.21}$$

$$V = T\left(s^{(n)}\right)^{-1/2} U. \tag{1.22}$$

We may now combine these matrices to obtain the null matrix

$$V^{\dagger} H^{(n)} V - V^{\dagger} S^{(n)} V h^{(n)} = 0, \tag{1.23}$$

and multiplying this on the left by $(V^{\dagger})^{-1} = U(s^{(n)})^{1/2} T$ gives

$$H^{(n)} V - S^{(n)} V h^{(n)} = 0. \tag{1.24}$$

If we write out the k^{th} column of this last equation, we have

$$\sum_{j=1}^{n} \left(H_{ij}^{(n)} - h_k^{(n)} S_{ij}^{(n)}\right) V_{jk} = 0; \qquad i = 1, 2, \ldots, n. \tag{1.25}$$

When this is compared with Eq. (1.15) we see that we have solved our problem, if $C^{(n)}$ is the k^{th} column of V and $W^{(n)}$ is the k^{th} diagonal element of $h^{(n)}$. Thus the diagonal elements of $h^{(n)}$ are the roots of the determinantal equation Eq. (1.16).

Now consider the variation problem with $n + 1$ functions where we have added another of the basis functions to the set. We now have the matrices $H^{(n+1)}$ and $S^{(n+1)}$, and the new determinantal equation

$$\left|H^{(n+1)} - W^{(n+1)} S^{(n+1)}\right| = 0. \tag{1.26}$$

We may subject this to a transformation by the $(n + 1) \times (n + 1)$ matrix

$$\bar{V} = \begin{bmatrix} V & 0 \\ 0 & 1 \end{bmatrix}, \tag{1.27}$$