The resonating-valence-bond state; the singlet ground state

The resonating-valence-bond state; the singlet ground state

I V Abarenkov†‡

† Cavendish Laboratory, Madingley Road, Cambridge CB3 0HE, UK

Abstract. The resonating-valence-bond (RVB) state with one localized state per site is a state in which correlations are important and therefore for its determination it is necessary to go beyond the one-determinant approximation. Although the number of one-electron space orbitals is equal to the number of electrons, N, the number of necessary determinants increases factorially with N, thus making a rigorous treatment of the correlation effects impracticable even for systems containing relatively small number of electrons. On the other hand, by introducing a many-electron spin function a coupled system of equations for one-electron localized non-orthogonal space orbitals and N-electron spin function can be obtained which can be solved in practice. In this procedure the main difficulty resides in the equation for the N-electron spin function. In this paper it is shown that a perturbation theory based on the decomposition of the spin system into clusters, when taken beyond the first order can, in general, offer a suitable method of mapping the results of numerical calculations of spin clusters onto the total system.

1. Introduction

The resonating-valence-bond state (RVB) with one localized state per site is often found in chemical systems and in solids. Two examples are the state of homogeneous electron gas in the low-density regime (Wigner crystal) [1, 2], and the state of Si(111) surface with dangling bonds. The electronic state of the Cu–O planes in the HTC superconductors is also essentially the RVB state [3].

With a good localization, an electronic system with one localized state per site is in fact a system with open shells, the number of these being equal to the number of electrons. As such systems are out of the scope of the conventional one-determinant Hartree–Fock method, it is evident that the actual wave function should be a superposition of many Slater determinants. Applying the valence-bond version of the spin-extended Hartree–Fock method, one encounters the problem that the number of required determinants increases factorially with the number of electrons.

An alternative approach which avoids the explicit expansion of the wave function into the series of determinants, is based on an N-electron spin function. For this as well as for the required one-electron space functions, a coupled system of equations can be derived. In the case of orthogonal one-electron space functions this can be done exactly and easily. Unfortunately, contrary to the one-determinant Hartree–Fock method, the orthogonality conditions become essential constrains in the valence-bond method and they usually give rise to poor result. To obtain reasonably good quantitative results the orthogonality conditions should be lifted. In this case the said coupled system of equations can be obtained approximately with, in principle, any desired precision. In this approach the complexity of the problem is mostly transferred to the problem of solving the equation for the N-electron spin function. For a comparatively small number of electrons the equation for spin function can be easily solved with the help of standard numerical techniques. For large systems

‡ On leave from Physics Department, St Petersburg University, St Petersburg, Russia.

we propose to consider small subsystems numerically and afterwards map the obtained solutions, in the way to be described below, onto that of the whole system.

In the present paper the perturbation theory based on the decomposition the spin system into clusters is considered as the method of mapping. It is shown that, when taken beyond the first order, the perturbation theory can indeed offer a fruitful method of mapping the results of cluster calculations onto the solution for the whole system.

2. The coupled system of equations

We consider the ground state of N-electron system in the external potential $V(r)$. The wave function of this state, Ψ, is the solution of the Schrödinger equation

$$\hat{H}\,\Psi(x_1, \ldots, x_N) = E\,\Psi(x_1, \ldots, x_N) \tag{2.1}$$

with (we employ the Hartree atomic units throughout)

$$\hat{H} = \sum_{k=1}^{N} \hat{h}(r_k) + \sum_{k<\ell=1}^{N} g(r_k, r_\ell) \tag{2.2}$$

$$\hat{h}(r) = -\frac{1}{2}\Delta + V(r_k) \tag{2.3}$$

$$g(r_k, r_\ell) = \frac{1}{|r_k - r_\ell|}. \tag{2.4}$$

Here x_k stands for the space, r_k, and spin, σ_k, variables of the kth electron. The potential $V(r)$ is the sum

$$V(r) = \sum_{j=1}^{N} V_j(r) \tag{2.5}$$

of the potentials $V_j(r)$ centred at sites R_j. No special symmetry of the sites or similarity of the potentials $V_j(r)$ is supposed. We assume N to be even and the ground state of the system to be a singlet state.

To take account of the correlation effects, we employ the valence-bond (VB) version [4] of the spin-extended Hartree–Fock method for the solution of (2.1). For this let

$$\{\phi_k(r), \ k = 1, 2, \ldots, N\} \tag{2.6}$$

be the set of N space orbitals. These orbitals are normalized, i.e.

$$\int |\phi_k(r)|^2 \, dr = 1 \tag{2.7}$$

but they are not necessarily orthogonal and the overlap integral of the orbitals ϕ_k and ϕ_ℓ we denote by $S_{k,\ell}$,

$$S_{k,\ell} = \int \phi_k^*(r)\phi_\ell(r) \, dr. \tag{2.8}$$

The N-electron space function is taken as a product of the one-electron functions ϕ_i

$$\Phi(r_1, r_2, \ldots, r_N) = \phi_1(r_1)\phi_2(r_2)\cdots\phi_N(r_N) \tag{2.9}$$

To make the N-electron spin function let us bind the electrons in pairs. For this one should first group all sites in pairs $\{i, j\}$, so that every site belongs to one and only one pair. The grouping defines a particular bonding scheme, the number of different bonding schemes for N electrons being equal to $(N-1)!!$. To distinguish different bonding schemes we will use the index g. Next, one should couple the electrons in each pair with the two-electron singlet spin function

$$\chi(\sigma_i, \sigma_j) = \frac{1}{\sqrt{2}}\{\alpha(\sigma_i)\beta(\sigma_j) - \beta(\sigma_i)\alpha(\sigma_j)\} \tag{2.10}$$

The function

$$\Theta_g(\sigma_1, \sigma_2, \ldots, \sigma_N) = \prod_{\{i,j\}} \chi(\sigma_i, \sigma_j) \tag{2.11}$$

in which the product is over all pairs in the gth bond scheme, is the N-electron singlet spin function. The function

$$\Psi_g(x_1, \ldots, x_N) = \frac{1}{\sqrt{N!}}\sum_P \epsilon(P)\hat{P}\Phi(r_1, \ldots, r_N)\Theta_g(\sigma_1, \ldots, \sigma_N) \tag{2.12}$$

in which $\hat{P}$ is a permutation operator, with $\epsilon(P)$ the signature of $\hat{P}$, and the summation is over all possible permutations, is an antisymmetric N-electron singlet function corresponding to a particular bonding scheme g. The functions Θ_g and Ψ_g are normalized to unity, but the functions Θ_g and $\Theta_{g'}$ (and Ψ_g and $\Psi_{g'}$) corresponding to different bonding schemes g and g' are in general not orthogonal. Moreover only $N!/(N/2)!(N/2+1)!$ of spin functions (2.11) are linearly independent [5]. There exist a powerful method to select the linearly independent spin functions called Rumer's diagram technique [4]. Employing linearly independent spin functions in (2.12) one can look for the (approximate) solution of equation (2.1) as the linear combination of Ψ_g

$$\Psi(x_1, \ldots, x_N) = \sum_g C_g \Psi_g(x_1, \ldots, x_N). \tag{2.13}$$

From the Schrödinger equation (2.1) for Ψ, one obtains a system of algebraic equations for the coefficients C_g. Equation (2.13) corresponds to mixing different bonding schemes (hence the name—'resonating valence bonds'). The problem that arises in this way is that the number of basis functions Ψ_g increases with the increasing number of electrons so quickly that calculations become impractical even for a moderate number of electrons. A way out of this problem could be to select the most important basis functions and to neglect all others. There are few exceptional cases where one coefficient C_g is much bigger than all others so that only one particular bonding scheme can be left in equation (2.13) without loss of much accuracy. These are the so-called 'perfect pairing' cases. Usually, many coefficients are of the same order of magnitude and therefore should be left in (2.13). In most problems it is impossible to know the essential coefficients C_k *a priori*.

At the same time, the linear combination in (2.13) is in fact a linear combination of spin functions only, as the spatial part is the same. This linear combination

$$\Theta(\sigma_1, \sigma_2, \ldots, \sigma_N) = \sum_g C_g \Theta_g(\sigma_1, \sigma_2, \ldots, \sigma_N) \tag{2.14}$$

with arbitrary coefficients C_g is the general singlet spin function of N-electron system. We write, therefore, the wavefunction of the system as

$$\Psi(x_1, \ldots, x_N) = \frac{1}{\sqrt{N!}} \sum_P \epsilon(P) \hat{P} \Phi(r_1, \ldots, r_N) \Theta(\sigma_1, \ldots, \sigma_N) \tag{2.15}$$

and instead of deriving the system of linear equations for C_g we derive the equation for the N-electron spin function $\Theta(\sigma_1, \sigma_2, \ldots, \sigma_N)$. To this end it is expedient to employ the variational principle and to look for the minimum of the energy functional

$$W = \langle \Psi | \hat{H} | \Psi \rangle / \langle \Psi | \Psi \rangle. \tag{2.16}$$

As the orbitals $\phi_k(r)$ are not orthogonal, the matrix elements in both the numerator and the denominator of equation (2.16) contain (due to the non-linked clusters) terms which scale as N^2, N^3 etc with increasing N. Therefore direct use of equation (2.16) is impractical even for moderate values of N. To cancel all ambiguous $O(N^2)$, $O(N^3)$, etc., terms one can use the technique developed in [2].

Using the approximation—to the second order in the exchange parameter—introduced in [2] (see equation (23) in [2]) we arrive at the following expression for the energy functional

$$W = \sum_i \langle i | \hat{h} | i \rangle + \frac{1}{2} \sum_{i \neq j} \langle i, j | g | i, j \rangle + \frac{1}{2} \sum_{i \neq j} K_{i,j} \, G_{i,j}. \tag{2.17}$$

The first two terms in this expression comprise the Hartree energy of the system. The last term is the exchange–correlation energy in the adopted approximation. In this term $G_{i,j}$ is the diagonal matrix element of the operator for the transposition of spin variables

$$G_{i,j} = \langle \Theta | \hat{P}_{i,j} | \Theta \rangle, \tag{2.18}$$

and $K_{i,j}$ is a combination of molecular integrals, which can be written in the following form

$$\begin{aligned}
K_{i,j} = {}&-\langle i, j | g | j, i \rangle - S_{i,j} H_{j,i}^{(i,j)} - H_{i,j}^{(i,j)} S_{j,i} \\
&+ S_{i,j} S_{j,i} \left(H_{i,i}^{(i,j)} + H_{j,j}^{(i,j)} + \langle i, j | g | i, j \rangle \right)
\end{aligned} \tag{2.19}$$

with

$$\hat{H}^{(i,j)} = -\frac{1}{2}\Delta + V_i(r) + V_j(r) + \sum_{k \neq i,j} \left(V_k(r) + \int \frac{1}{|r - r'|} |\phi_k(r')|^2 \, dr' \right). \tag{2.20}$$

The latter is the Hamiltonian of an electron moving in the direct potentials V_i and V_j of sites i and j and the screened potential of all other sites.

Applying the variational principle to (2.17), we immediately obtain a system of coupled equations for the one-electron space orbitals ϕ_i and the N-electron spin function Θ,

$$\left(-\frac{1}{2}\Delta + V_i(r) + \sum_{j\neq i}\left(V_j(r) + \int \frac{1}{|r-r'|}|\phi_j(r')|^2\,\mathrm{d}r'\right)\right)\phi_i(r)$$

$$+ \frac{1}{2}\sum_{\ell\neq m}G_{\ell,m}\frac{\delta}{\delta\phi_i}K_{\ell,m} = \lambda_i\,\phi_i(r) \tag{2.21}$$

$$\sum_{i\neq j}K_{i,j}\,\hat{P}_{i,j}\,\Theta = \Lambda\,\Theta. \tag{2.22}$$

In these equations λ_i and Λ are Lagrange multipliers for satisfying the normalisation conditions, respectively, for the one-electron space functions ϕ_i and the N-electron spin function Θ.

Equations (8) are written in a form ready for self-consistent solution. For this we proceed as follows. First we assume an initial approximation $\{\phi_i^{(0)}\}$ for the spatial functions. From this we calculate $K_{i,j}^{(0)}$ through equation (2.19), and insert the obtained values into (2.22), from which $\Theta^{(0)}$ is obtained. Subsequently we calculate $G_{i,j}^{(0)}$ from (2.18) so that through solving equations (2.21) the new approximation $\{\phi_i^{(1)}\}$ is obtained. The same process, as for $\{\phi_i^{(0)}\}$, is continued with $\{\phi_i^{(1)}\}$ and the subsequent functions $\{\phi_i^{(m)}\}$ until convergence is achieved. In this process the equations (2.21) for the set of functions $\{\phi_i\}$ can be solved with the help of the well-developed techniques of quantum-chemistry, which are mainly based upon the expanding ϕ_i into an appropriate basis set and solving for the expansion coefficients from the resulting (non-linear) algebraic equations. In choosing the basis functions it should be taken into account that $\{\phi_i\}$ are well localized functions.

In spite of its simple form, equation (2.22) for the spin function is, in general, more difficult to solve than equations (2.21). Because of the relation

$$4(\hat{s}_i,\hat{s}_j) = 2\,\hat{P}_{i,j} - \hat{I} \tag{2.23}$$

which is valid for spin-1/2 particles, equation (2.22) is in fact the well known Heisenberg equation. Exact solutions (such as one based on the Bethe *ansatz* [6, 7]) exist only for few simple cases. Numerical solutions (see, for example [8–10]) are obtainable only for comparatively small systems.

Table 1. The total energy of chain, $W^{(c)}$, ring, $W^{(r)}$, and their difference as a function of the number of sites N, together with the energy per site of chain, $w^{(c)}$, and ring, $w^{(r)}$; ($w_{\infty} = -0.3863$).

N	$W^{(c)}$	$W^{(r)}$	$W^{(c)} - W^{(r)}$	$w^{(c)}$	$w^{(r)}$
4	−1.7320	−2.0000	0.2680	−0.4330	−0.5000
6	−2.4872	−2.6056	0.1184	−0.4867	−0.4343
8	−3.2499	−3.3024	0.0525	−0.4062	−0.4128
10	−4.0161	−4.0309	0.0148	−0.4016	−0.4031
12	−4.7841	−4.7748	−0.0092	−0.3987	−0.3979
14	−5.5533	−5.5271	−0.0262	−0.3967	−0.3948
16	−6.3236	−6.2846	−0.0392	−0.3952	−0.3928

In table 1 we present some results corresponding to a number of linear chains and rings with even number of sites, for which only the nearest-neighbours interaction was assumed;

$$K_{i,j} = \begin{cases} 1 & \text{if } i, j \text{ are the nearest neighbours} \\ 0 & \text{otherwise.} \end{cases} \tag{2.24}$$

In this table $w^{(c)}$ and $w^{(t)}$ stand for the energy per bond in, respectively, an open chain and a closed ring. In both cases the energy per bond is negative; therefore one would intuitively expect that breaking the bond will always result in increasing the energy of the system. For $N \leqslant 10$ the results are as expected; surprisingly, for $N > 10$ the energy of the chain turns out to be lower than the energy of the ring. This can be interpreted in terms of the 'surface' energy which change sign with increasing number of spins. The described result suggests possible clusterisation of long chains. More peculiar result is that sum of energies of two rings with N sites is lower than the energy of one ring with $2N$ sites by the amount that is comparable with the absolute value of the energy per bond. In this case there is no 'surface', and the number of bonds is the same. This result indicates that the possible clusterisation can have a rather complicated structure.

The numerical calculations are practical for small spin systems (up to approximately 20 spins) as the demand for computing resources grows factorially with the number of spins involved. The quantum Monte Carlo method has been successfully applied [11] to this problem with larger number of spins, but still it is not an inexpensive method. The coupled-cluster method has also been used recently for $s = 1/2$ quantum antiferromagnets [12]. Still for the self-consistent solution of system of equations (8a), (8b) a simpler method that can handle much larger systems is needed, especially for those without symmetry.

The perturbation-theory approach, to be presented below, turns out to be a simple and efficient enough method of solving equation (2.22) for large systems.

3. The perturbation-theory approach

To apply the perturbation theory in solving equation (2.22) we consider a finite system of spins. The Hamiltonian corresponding to (2.22) is

$$\hat{H} = \sum_{i<j} K_{i,j} \hat{P}_{i,j} \tag{3.1}$$

in which, owing to the good localization of ϕ_i, the constants $K_{i,j}$ decrease rapidly with increasing distance between sites so that for practically interesting systems, the sum in (3.1) is over the first, the second, and possibly the third neighbours. The signs of the parameters $K_{i,j}$ are assumed such that the system corresponds to a quantum antiferromagnet. Now considering the spin system as an array of interacting clusters, each cluster a having an even number $N_a = 2M_a$ of spins. The decomposition of the Hamiltonian is straightforward

$$\hat{H} = \sum_{a} \hat{H}_a + \sum_{a<b} \hat{H}_{a,b}. \tag{3.2}$$

Here

$$\hat{H}_a = \sum_{i<j}^{(a)} K_{i,j} \hat{P}_{i,j} \tag{3.3}$$

is the Hamiltonian of the cluster a, and

$$\hat{H}_{a,b} = \sum_{i}^{(a)} \sum_{j}^{(b)} K_{i,j} \hat{P}_{i,j} \tag{3.4}$$

is the interaction between clusters a and b. Looking for the ground state of the system

$$\hat{H} \Theta_g = E_g \Theta_g \tag{3.5}$$

we take

$$\hat{H}_0 = \sum_{a} \hat{H}_a \tag{3.6}$$

as unperturbed operator, and

$$\hat{V} = \sum_{a<b} \hat{H}_{a,b} \tag{3.7}$$

as the perturbation. We denote the eigenfunctions and the corresponding eigenvalues of $\hat{H}_0$ by Φ_k and $\mathcal{E}_k$, respectively

$$\hat{H}_0 \Phi_k = \mathcal{E}_k \Phi_k \tag{3.8}$$

with Φ_0 and $\mathcal{E}_0$ being the wave function and the energy of the ground state of the unperturbed operator. The assumption that the number of spins in every cluster is even has been made to ensure that the ground state of the unperturbed system is non-degenerate. In the case of clusters with an odd number of spins, a different technique than presented here should be used. We employ the intermediate normalization condition

$$\langle \Phi_0 | \Theta_g \rangle = 1 \tag{3.9}$$

and the perturbation equations

$$\begin{aligned} \Delta E &= E_g - E_0 = \langle \Phi_0 | \hat{V} | \Theta_g \rangle \\ \Theta_g &= \Theta_0 + \hat{G}_0 (\Delta E - \hat{V}) \Theta_g, \\ \hat{G}_0 &= \hat{Q} (\hat{H}_0 - \mathcal{E}_0)^{-1} \hat{Q} \end{aligned} \tag{3.10}$$

in which the projection operator $\hat{Q}$ projects out the state Φ_0. Applying $\hat{V} \rightarrow \kappa \hat{V}$ in the above equations, expanding the results in powers of κ, and setting $\kappa = 1$ in the final equations we arrive at

$$\Theta_g = \sum_{k=0}^{\infty} \Theta^{(k)} \qquad E_g = \sum_{k=0}^{\infty} E^{(k)} \tag{3.11}$$

in which

$$\begin{aligned} E^{(0)} &= \mathcal{E}_0 \\ \Theta^{(0)} &= \Theta_0 \\ E^{(k+1)} &= \langle \Theta_0 | \hat{V} | \Theta^{(k)} \rangle \\ \Theta^{(k+1)} &= \hat{G}_0 \left(\sum_{i=0}^{k} E^{(i)} \Theta^{(k-i)} - \hat{V} \Theta^{(k)} \right). \end{aligned} \tag{3.12}$$

In general, for any property $\hat{R}$ the following result can be obtained in the same manner

$$\bar{R} = \frac{\langle \Theta \mid \hat{R} \mid \Theta \rangle}{\langle \Theta \mid \Theta \rangle} = \sum_{k=0}^{\infty} R^{(k)} \tag{3.13}$$

in which

$$R^{(0)} = \langle \Theta^{(0)} \mid \hat{R} \mid \Theta^{(0)} \rangle$$

$$R^{(k)} = \sum_{m=0}^{k} d_m \sum_{n=0}^{m} \langle \Theta^{(n)} \mid \hat{R} \mid \Theta^{(m-n)} \rangle \tag{3.14}$$

the coefficients d_m are obtained through the recurrent relation

$$d_m = -\sum_{\ell=1}^{m} d_{m-\ell} \sum_{n=0}^{\ell} \langle \Theta^{(n)} \mid \Theta^{(\ell-n)} \rangle \qquad d_0 = 1. \tag{3.15}$$

The convergence of the series (3.11) and (3.12) can be either a true convergence (numerical results in some cases point to it), or an asymptotic one (when the deviation from the exact value of the value calculated *up to some order* becomes smaller and smaller with increasing the size of the clusters and correspondingly reducing the relative strength of the perturbation).

To employ perturbation theory we express the Green's function $\hat{G}_0$ of the unperturbed Hamiltonian in terms of projectors on the eigenstates of the unperturbed operator as follows

$$\hat{G}_0 = \sum_{k>0} \mid \Phi_k \rangle \frac{1}{\mathcal{E}_k - \mathcal{E}_0} \langle \Phi_k \mid. \tag{3.16}$$

The eigenstates of $\hat{H}_0$ are the products of clusters eigenstates, and eigenvalues of $\hat{H}_0$ are sums of the clusters eigenvalues

$$\Phi_k = \prod_a \Theta_{a,k_a} \qquad \mathcal{E}_k = \sum_a e_{a,k_a}. \tag{3.17}$$

The Hamiltonian $\hat{H}_a$ of a cluster commutes with the operators for the square of the total spin of the cluster

$$\hat{S}_a^2 = N_a - \frac{N_a^2}{4} + \sum_{i<j}^{(a)} \hat{P}_{i,j} \tag{3.18}$$

and its z-projection $\hat{S}_{a,z}$. Hence, the eigenstates of $\hat{H}_a$ can be the simultaneous eigenstates of $\hat{S}_a^2$ and $\hat{S}_{a,z}$, i.e.

$$\hat{H}_a \, \Theta_{a,n,j,m} = e_{a,n,j} \, \Theta_{a,n,j,m}$$

$$\hat{S}_a^2 \, \Theta_{a,n,j,m} = j(j+1) \, \Theta_{a,n,j,m}$$

$$\hat{S}_{a,z} \, \Theta_{a,n,j,m} = m \, \Theta_{a,n,j,m}$$

where

$$n = 0, 1, 2, \ldots \qquad j = 0, 1, \ldots, M_a \qquad m = -j, \ldots, j. \tag{3.19}$$

The clusters' eigenfunctions are assumed to be normalized to unity and orthogonal within a given cluster

$$\langle \Theta_{a,n,j,m} \mid \Theta_{a,n',j',m'} \rangle = \delta_{n,n'} \, \delta_{j,j'} \, \delta_{m,m'}. \tag{3.20}$$

The perturbation theory is especially suitable for the singlet state, which is believed to be the RVB ground state. In the case of the singlet state many matrix elements of the perturbation occur to be zero due to the symmetry of the spin systems as will be shown in the next section.

4. The matrix elements

The perturbation operator is the linear combination of transposition operators $\hat{P}_{\ell,\ell'}$, with $\ell \in a$, $\ell' \in b$, so let us consider a matrix element of this operator between two eigenstates Φ_k and $\Phi_{k'}$ of the unperturbed Hamiltonian (see (3.17)). These states are products of the wave functions of isolated clusters. Because of (3.20) this matrix element will be different from zero only if

$$k_c = k'_c \qquad c \neq a, b$$

and when this condition is fulfilled we have

$$\langle \Phi_k \mid \hat{P}_{\ell,\ell'} \mid \Phi_{k'} \rangle = \langle \Theta_{a,n1,j1,m1} \, \Theta_{b,n2,j2,m2} \mid \hat{P}_{\ell,\ell'} \mid \Theta_{a,n1',j1',m1'} \, \Theta_{b,n2',j2',m2'} \rangle. \tag{4.1}$$

To calculate this matrix element we decompose each wave function of an isolated cluster as follows

$$\Theta_{g,n,j,m}(\sigma_1, \ldots, \sigma_{N_g}) = F^{(\ell)}_{g,n,j,m,\alpha}(\sigma_1, \ldots, \sigma_{\ell-1}, \sigma_{\ell+1}, \ldots, \sigma_{N_g}) \, \alpha(\sigma_\ell)$$
$$+ F^{(\ell)}_{g,n,j,m,\beta}(\sigma_1, \ldots, \sigma_{\ell-1}, \sigma_{\ell+1}, \ldots, \sigma_{N_g}) \, \beta(\sigma_\ell) \tag{4.2}$$

in which

$$F^{(\ell)}_{g,n,j,m,\alpha}(\sigma_1, \ldots, \sigma_{\ell-1}, \sigma_{\ell+1}, \ldots, \sigma_{N_g}) = \int \alpha(\sigma_\ell) \Theta_{n,j,m}(\sigma_1, \ldots, \sigma_{N_g}) \, d\sigma_\ell$$

$$F^{(\ell)}_{g,n,j,m,\beta}(\sigma_1, \ldots, \sigma_{\ell-1}, \sigma_{\ell+1}, \ldots, \sigma_{N_g}) = \int \beta(\sigma_\ell) \Theta_{n,j,m}(\sigma_1, \ldots, \sigma_{N_g}) \, d\sigma_\ell. \tag{4.3}$$

Consequently we have

$$\begin{aligned}
\langle \Phi_k \mid \hat{P}_{\ell,\ell'} \mid \Phi_{k'} \rangle = {} & \langle F^{(\ell)}_{a,n1,j1,m1,\alpha} \mid F^{(\ell)}_{a,n1',j1',m1',\alpha} \rangle \langle F^{(\ell')}_{b,n2,j2,m2,\alpha} \mid F^{(\ell')}_{b,n2',j2',m2',\alpha} \rangle \\
& + \langle F^{(\ell)}_{a,n1,j1,m1,\alpha} \mid F^{(\ell)}_{a,n1',j1',m1',\beta} \rangle \langle F^{(\ell')}_{b,n2,j2,m2,\beta} \mid F^{(\ell')}_{b,n2',j2',m2',\alpha} \rangle \\
& + \langle F^{(\ell)}_{a,n1,j1,m1,\beta} \mid F^{(\ell)}_{a,n1',j1',m1',\alpha} \rangle \langle F^{(\ell')}_{b,n2,j2,m2,\alpha} \mid F^{(\ell')}_{b,n2',j2',m2',\beta} \rangle \\
& + \langle F^{(\ell)}_{a,n1,j1,m1,\beta} \mid F^{(\ell)}_{a,n1',j1',m1',\beta} \rangle \langle F^{(\ell')}_{b,n2,j2,m2,\beta} \mid F^{(\ell')}_{b,n2',j2'm2',\beta} \rangle.
\end{aligned} \tag{4.4}$$

The scalar products in equation (4.4) are the first-order density matrices of the clusters, and many of them are identically vanishing as can be seen from the fact that functions $F^{(\ell)}_{g,n,j,m,\alpha}$ and $F^{(\ell)}_{g,n,j,m,\beta}$ can be expressed as

$$F^{(\ell)}_{g,n,j,m,\alpha} = a(j, m, \tfrac{1}{2}, -\tfrac{1}{2}) \chi^{(\ell)}_{g,n,j,m;j+\frac{1}{2},m-\frac{1}{2}} + a(j, m, -\tfrac{1}{2}, -\tfrac{1}{2}) \chi^{(\ell)}_{g,n,j,m;j-\frac{1}{2},m-\frac{1}{2}}$$

$$F^{(\ell)}_{g,n,j,m,\beta} = a(j, m, \tfrac{1}{2}, \tfrac{1}{2}) \chi^{(\ell)}_{g,n,j,m;j+\frac{1}{2},m+\frac{1}{2}} + a(j, m, -\tfrac{1}{2}, \tfrac{1}{2}) \chi^{(\ell)}_{g,n,j,m;j-\frac{1}{2},m+\frac{1}{2}} \tag{4.5}$$

in which $\chi^{(\ell)}_{g,n,j,m;j+\mu,m+\nu}$ is a normalized to unity wave function of the cluster $\tilde{g}$ with $N-1$ spins resulting from the cluster g by removing the spin σ_ℓ ($\chi^{(\ell)}$ is not necessarily the eigenfunction of the cluster $\tilde{g}$). This function corresponds to the $j + \mu$ and $m + \nu$ eigenvalues of the square and z-projection operators of the spin of the cluster $\tilde{g}$. The coefficients $a(j, m, \mu, \nu)$ are the Clebsch–Gordan coefficients

$$a(j, m, \mu, \nu) = (j + \mu, \tfrac{1}{2}, j, m \mid j + \mu, \tfrac{1}{2}, m - \nu, \nu) \qquad \mu, \nu = \pm\tfrac{1}{2}. \tag{4.6}$$

From (4.5) it follows that

$$
\begin{aligned}
\langle F_{g,n,j,m,\alpha} \mid F_{g,n',j',m',\alpha} \rangle &= 0 \qquad \text{if } |j - j'| > 1, \text{ or } m' \neq m \\
\langle F_{g,n,j,m,\beta} \mid F_{g,n',j',m',\beta} \rangle &= 0 \qquad \text{if } |j - j'| > 1, \text{ or } m' \neq m \\
\langle F_{n,j,m,\alpha} \mid F_{n',j',m',\beta} \rangle &= 0 \qquad \text{if } |j - j'| > 1, \text{ or } m' \neq m - 1 \\
\langle F_{n,j,m,\beta} \mid F_{n',j',m',\alpha} \rangle &= 0 \qquad \text{if } |j - j'| > 1, \text{ or } m' \neq m + 1.
\end{aligned} \tag{4.7}
$$

The matrix element (4.1) has the simplest form if both states of clusters a and b in Φ_k (or $\Phi_{k'}$) are singlet states. This is due to the fact that in the case of the singlet state equation (4.5) reads

$$
\begin{aligned}
F^{(\ell)}_{g,n,0,0,\alpha}(\sigma_1, \ldots, \sigma_{\ell-1}, \sigma_{\ell+1}, \ldots, \sigma_{N_g}) &= \frac{1}{\sqrt{2}} \chi^{(\ell)}_{g,n,0,0;\frac{1}{2},-\frac{1}{2}}(\sigma_1, \ldots, \sigma_{\ell-1}, \sigma_{\ell+1}, \ldots, \sigma_{N_g}) \\
F^{(\ell)}_{g,n,0,0,\beta}(\sigma_1, \ldots, \sigma_{\ell-1}, \sigma_{\ell+1}, \ldots, \sigma_{N_g}) &= -\frac{1}{\sqrt{2}} \chi^{(\ell)}_{g,n,0,0;\frac{1}{2},\frac{1}{2}}(\sigma_1, \ldots, \sigma_{\ell-1}, \sigma_{\ell+1}, \ldots, \sigma_{N_g}).
\end{aligned}
$$
$$\tag{4.8}$$

From equation (4.7) it follows that

$$
\langle F_{g,n,0,0,\alpha} \mid F_{g,n',0,0,\beta} \rangle = 0. \tag{4.9}
$$

Moreover, there is a simple relation between $F_{g,n,0,0,\alpha}$ and $F_{g,n',0,0,\beta}$. To find this relation, consider the operation $\hat{T}$ which changes the direction of the z axes to the opposite. Under this operation $\alpha(\sigma)$ changes to $\beta(\sigma)$ and $\beta(\sigma)$ to $\alpha(\sigma)$. The spin function $\Theta_{g,n,0,0}$ of the cluster under this operation will either change sign or will remain unchanged, depending on the parity of M_g,

$$
\hat{T} \, \Theta_{n,0,0}(\sigma_1, \ldots, \sigma_{N_g}) = (-1)^{M_g} \Theta_{n,0,0}(\sigma_1, \ldots, \sigma_{N_g}) \tag{4.10}
$$

This follows, for example, from the fact that the singlet function $\Theta_{g,n,0,0}$ can be expressed as a linear combination of valence-bond structures, each valence-bond structure being a product of M_g two-electron singlet spin functions $\chi(\sigma_1, \sigma_2)$. The latter changes sign under the operation $\hat{T}$, hence (4.10). Applying $\hat{T}$ to both sides of equation (4.3) and making use of (4.10) one immediately obtains

$$
\begin{aligned}
\hat{T} \, F^{(l)}_{g,n,0,0,\alpha} &= (-1)^{M_g} F^{(\ell)}_{g,n,0,0,\beta} \\
\hat{T} \, F^{(l)}_{g,n,0,0,\beta} &= (-1)^{M_g} F^{(\ell)}_{g,n,0,0,\alpha}.
\end{aligned} \tag{4.11}
$$

Consequently the orthogonality condition

$$
\langle \Theta_{g,n,0,0} \mid \Theta_{g,n',0,0} \rangle = \delta_{n,n'}. \tag{4.12}
$$

will read

$$
\langle F^{(\ell)}_{g,n,0,0,\alpha} \mid F^{(\ell)}_{g,n',0,0,\alpha} \rangle = \langle F^{(\ell)}_{g,n,0,0,\beta} \mid F^{(\ell)}_{g,n',0,0,\beta} \rangle = \frac{1}{2} \delta_{n,n'}. \tag{4.13}
$$

Combining equations (4.4), (4.8), and (4.13) we obtain

$$
\begin{aligned}
\langle \Theta_{a,n1,j1,m1} \; \Theta_{b,n2,j2,m2} \mid \hat{P}_{\ell,\ell'} \mid \Theta_{a,n1',0,0} \; \Theta_{b,n2',0,0} \rangle \\
= \frac{1}{2} \delta_{n1,n1'} \, \delta_{j1,0} \, \delta_{m1,0} \delta_{n2,n2'} \, \delta_{j2,0} \, \delta_{m2,0} \\
+ 2 \langle F^{(\ell)}_{a,n1,j1,m1,\alpha} \mid F^{(\ell)}_{a,n1',0,0,\alpha} \rangle \langle F^{(\ell')}_{b,n2,j2,m2,\alpha} \mid F^{(\ell')}_{b,n2',0,0,\alpha} \rangle \\
\times \delta_{j1,1} \delta_{j2,1} \left(\delta_{m1,0}\delta_{m2,0} - \delta_{m1,1}\delta_{m2,-1} - \delta_{m1,-1}\delta_{m2,1} \right).
\end{aligned}
\tag{4.15}
$$

Among other things, this result shows that if only singlet states of isolated clusters are taken into account in the perturbation theory, then the matrix of the perturbation operator will be automatically diagonal and every state of the unperturbed system will be simultaneously an eigenstate of the perturbed system. Hence in order to make the perturbation theory effective it is necessary to employ the states of isolated clusters with higher multiplicity. As the perturbation is a linear function of the transposition operators, the second- (and also the third-) order correction is due to the states in which only two interacting clusters are excited into triplet states; these excited states are coupled such as to make the total state a singlet one. The fourth-order correction is due to the states in which two, three, and four clusters are excited, and so on. This makes application of the perturbation theory to the singlet state exceptionally simple.

5. Comparison with variational approach

In the previous section we have shown that the first-order correction to the spin function is due to states in which only two clusters are excited. Therefore it seems appropriate to employ these states in the variational approach in order to improve the ground-state wavefunction. Unfortunately this procedure results in an incorrect dependence of the energy of the system upon the number of sites. To show this explicitly, let us consider a one-dimensional chain, taking into account the nearest-neighbours interaction (2.24) only. We impose periodic boundary condition, and assume all the clusters comprising the chain to be identical. Let the number of sites in each cluster be even, $n_p = 2m_p$, and the number of clusters be M. We take into account only those excited states of the unperturbed chain in which two neighbouring clusters p and $p+1$ are excited to their first excited states, triplet state, and coupled to make a singlet state of the chain. These excited states of the chain we denote by Φ_p, $p = 1, \ldots, M$. These states Φ_p are both orthogonal to each other and to the state Φ_0. Note that both unperturbed and perturbed systems are translationally invariant. Therefore the normalized to unity correction to the wavefunction Φ_0 should be

$$
\chi = \frac{1}{\sqrt{M}} \sum_{p=1}^{M} \Phi_p
\tag{5.1}
$$

and the spin function should have the form

$$
\Theta = C_1 \Phi_0 + C_2 \chi.
\tag{5.2}
$$

The diagonal matrix elements of the Hamiltonian with respect to Φ_0 and χ are

$$
\begin{aligned}
\langle \Phi_0 | \hat{H} | \Phi_0 \rangle &= E_0 = M \left(e_0 + \tfrac{1}{2} \right) \\
\langle \chi | \hat{H} | \chi \rangle &= E_0 + 2\Delta.
\end{aligned}
\tag{5.3}
$$

The excitation energy $2\Delta = 2\Delta_0 + \Delta_1$ consists of $\Delta_0 = e_1 - e_0$, the excitation energy of a single cluster, and $\Delta_1 = \langle\Phi_p|\hat{H}_{p,p+1}|\Phi_p\rangle - \langle\Phi_0|\hat{H}_{p,p+1}|\Phi_0\rangle$, the difference of the interaction energy in the excited and the ground states. Note that Δ_1 does not depend on p and neither Δ_0 nor Δ_1 depend on M.

The non-diagonal matrix element is

$$\langle\Phi_0|\hat{H}|\chi\rangle = \sqrt{M}\,Q. \tag{5.4}$$

The quantity $Q = \langle\Phi_0|\hat{H}|\Phi_p\rangle = \langle\Phi_0|\hat{H}_{p,p+1}|\Phi_p\rangle$ does not depend on neither p nor M.

Straightforward calculations give the following result for the ground state

$$
\begin{aligned}
E &= E_0 + \Delta - \sqrt{\Delta^2 + MQ^2} \\
C_1 &= \frac{Q\sqrt{M}}{\sqrt{2R}} \\
C_2 &= \frac{\Delta - \sqrt{\Delta^2 + MQ^2}}{\sqrt{2R}} \\
R &= \Delta^2 + MQ^2 - \Delta\sqrt{\Delta^2 + MQ^2}.
\end{aligned}
\tag{5.5}
$$

Let us now keep the number of sites n_p in the cluster constant, so that Δ will also be constant, while increasing the number of sites N in the chain through increasing the number of clusters M. From equation (5.5) it can be seen that the energy of the chain will be decreasing, and it will be decreasing as $\sqrt{N}$; the coefficients C_1 and C_2 will tend to $\frac{1}{2}$ and $-\frac{1}{2}$, respectively. At the same time in the system with short-range interaction the energy at large N should be proportional to N, and the calculations on finite spin systems clearly indicate this.

In the perturbation-theory approach in every order only terms proportional to N survive, all other contributions proportional to N^2, N^3, etc., due to the disconnected diagrams, cancel. For the above-considered chain the zero-order energy is $M \times e_0$, the first-order correction is $M \times \frac{1}{2}$, and the second-order correction (only the above-described states being taken into account) amounts to $-M \times Q^2/2\Delta_0$. The factor 2 in the denominator of the last expression is a consequence of two clusters being excited. Combining these results we obtain the following expression for the energy (up to the second order)

$$E = M\left(e_0 + \frac{1}{2} - \frac{Q^2}{2\Delta_0}\right). \tag{5.6}$$

This result is exactly the same as that one would have obtained by expanding the energy expression in equation (5.5) to first order in MQ^2/Δ^2 and replacing Δ by Δ_0.

It is well known from the configuration-interaction method that in order to obtain the proper behavior of the energy with the increasing number of particles one must use in the variational approach the function space that contains all the multiple excitations of a given type (i.e., the complete active space), which makes a much more complicated trial function than the one according to equations (5.1), (5.2). It is therefore obvious that the perturbation-theory approach is preferable for large systems.

6. Results of test calculations on one-dimensional chains and conclusions.

To check the applicability of the perturbation theory to the problem under consideration we chose the simplest system for which the exact results are known. It is the one-dimensional chain with nearest neighbours interaction (2.24) only, and with free ends. First,

we have numerically calculated the energies and the wavefunctions of the finite chains with $N = 4, 6, 8, 10$. Subsequently we have considered these systems as composed of two-sites clusters (dimers) and the perturbation theory equations (3.12), (3.16) were applied up to 20th order, showing good convergence. The results for up to 6th order are shown in table 2. They show also convergence with increasing number of sites in the chain. In the last column the results for an infinite chain are presented. They have been calculated analytically up to fourth order. The exact value for the infinite chain is $1 - 2 \ln 2$, known from the Bethe *ansatz* solution (also presented in the table). These results show that the perturbation calculations converge, and that it is necessary to employ higher than the first order of perturbation for obtaining accurate results.

Table 2. Perturbation theory corrections to the energy per site of chains with N sites, their sum, and the exact values.

Order	$N = 4$	$N = 6$	$N = 8$	$N = 10$	∞
0	−0.5000	−0.5000	−0.5000	−0.5000	−0.5000
1	0.1250	0.1667	0.1875	0.2000	0.2500
2	−0.0469	−0.0625	−0.0703	−0.0750	−0.0937
3	−0.0117	−0.0150	−0.0176	−0.0187	−0.0234
4	−0.0007	−0.0026	−0.0035	−0.0041	−0.0064
5	0.0009	−0.0006	−0.0013	−0.0018	
6	0.0004	−0.0002	−0.0008	−0.0011	
Sum	−0.4330	−0.4149	−0.4061	−0.4007	−0.3735
Exact	−0.4330	−0.4145	−0.4062	−0.4016	−0.3863

The perturbation calculations of correlation coefficients have also been made for finite chains. In table 3 we present the values of the averaged nearest-neighbours correlation coefficient

$$\omega = \frac{1}{N} \sum_{k=1}^{N-1} \frac{\langle \Theta | \hat{P}_{k,k+1} | \Theta \rangle}{\langle \Theta | \Theta \rangle}. \tag{6.1}$$

They have been calculated directly with the help of equations (3.13)–(3.15), without employing the energy expression. These results also show convergence, although with a somewhat slower rate.

The calculations for an infinite chain can be simplified if the diagonal part of perturbation is included into the unperturbed Hamiltonian by means of the transformation

$$\hat{H}_0 \rightarrow \hat{H}_0 + \sum_{k=0}^{\infty} |\Phi_k\rangle \langle \Phi_k| \hat{V} |\Phi_k\rangle \langle \Phi_k|$$

$$\hat{V} \rightarrow \hat{V} - \sum_{k=0}^{\infty} |\Phi_k\rangle \langle \Phi_k| \hat{V} |\Phi_k\rangle \langle \Phi_k|. \tag{6.2}$$

This transformation is not unitary invariant and therefore will produce different results if different sets of functions Φ_k are used for the degenerate unperturbed states. With the help of this transformation the perturbation-theory corrections for the energy per site of an infinite linear chain are

$$w^{(0)} = -\frac{1}{4}, \qquad w^{(1)} = 0, \qquad w^{(2)} = -\frac{1}{8}$$

$$w^{(3)} = 0, \qquad w^{(4)} = -\frac{1}{60}, \qquad w^{(5)} = \frac{1}{240}. \tag{6.3}$$

Table 3. The perturbation-theory corrections of the averaged nearest neighbours correlation coefficient for chains with N sites, their sum, and the exact values.

Order	$N = 4$	$N = 6$	$N = 8$	$N = 10$
0	−0.5000	−0.4000	−0.3571	−0.3333
1	−0.1250	−0.1500	−0.1607	−0.1667
2	0.0156	0.0187	0.0201	0.0208
3	0.0273	0.0250	0.0240	0.0234
4	0.0090	0.0058	0.0044	0.0036
5	−0.0014	0.0011	0.0009	0.0008
6	−0.0025	0.0020	0.0019	0.0019
Sum	−0.5769	−0.4973	−0.4666	−0.4495
Exact	−0.5774	−0.4974	−0.4643	−0.4462

The sum of these amounts to −0.3875, to be compared with the exact result −0.3863.

Calculations for the same infinite chain were also performed with larger clusters ($n_p = 4, 6, 8$) in the unperturbed state. The transformation (6.2) was used in these calculations. The results of calculations of the energy per site (and not the corrections, as in tables 2 and 3) up to the third order are given in table 4. These results show the stability with increasing number of sites in the cluster. Even in the third order the energy is found to be in good agreement with the exact result. At the same time the corrections are almost independent of the size of the clusters, excluding the minimal size. This is evidence that application of perturbation theory, using comparatively small clusters, which can easily be calculated numerically, can offer a powerful method of calculating large systems.

Table 4. The energy per site of the infinite chain, calculated by the perturbation theory with clusters of n sites. The exact value −0.3863.

Order	$n = 2$	$n = 4$	$n = 6$	$n = 8$
0	−0.2500	−0.3080	−0.3312	−0.3437
1	−0.2500	−0.3080	−0.3312	−0.3437
2	−0.3750	−0.3776	−0.3772	−0.3777
3	−0.3750	−0.3867	−0.3875	−0.3872

Application of perturbation theory to more complex than one-dimensional systems is presently in progress.

It should be noted that in the present approach we have only considered the pure singlet spin state of many-electron system. If there are gapless excitations in the system into states with larger spin, other states can have the same energy as the singlet state. In such cases the singlet state is still an exact state and therefore can be dealt with with the method described above. For other states the method needs modifications.

Acknowledgments

References

[1] Moulopoulos K and Ashcroft N W 1993 *Phys. Rev.* B **48** 11646
[2] Abarenkov I V 1993 *J. Phys.: Condens. Matter* **5** 2341
[3] Anderson P W 1987 *Science* **235** 1196
[4] McWeeny R 1989 *Methods of Molecular Quantum Mechanics* (New York: Academic)
[5] Wigner E 1959 *Group Theory and its Applications to the Quantum Mechanics of Atomic Spectra* (New York: Academic)
[6] Bethe H 1931 Z *Phys.* **71** 205
[7] Hulten L 1938 *Ark. Mat. Astron. Fys.* **26** A 11
[8] Borysowicz J, Kaplan T A and Horsh P 1985 *Phys. Rev.* B **31** 1590
[9] Soos Z G, Kuwajima S and Mihalick J E 1985 *Phys. Rev.* B **32** 3124
[10] Gagliano E R, Dagotto E, Moreo A and Alcaraz F C 1986 *Phys. Rev.* B **34** 1677
[11] Trivedi N and Ceperley D M 1990 *Phys. Rev.* B **41** 4552
[12] Harris F E 1993 *Phys. Rev.* B **47** 7903