Cluster-expansion representation

M A Garcia-Bach† and D J Klein‡

† Departomento de Física Fonamental, Facultat de Física, Universitat de Barcelona, Diagonal 647, 08028 Barcelona, Catalunya, Spain
‡ Texas A&M University at Galveston, Galveston, TX 775531–675, USA

Abstract. We describe how starting from a type of nearest-neighbour localized wavefunction cluster expansion one may conveniently build up a locally correlated orthonormal 'many-body' basis for a linear-chain system. The initial correlated cluster expansion approximant to the ground state plays the role of a 'vacuum' while the other members of the basis are single or multiple excitations above this. The scheme is illustrated for the case of two states per site, in application to the linear-chain Heisenberg model.

1. Introduction

Wavefunction cluster expansions have long been recognized as a general 'size consistent' means by which to introduce many-body correlation effects (see, e.g., [1–5]). Development of cluster expansions within a real-space localized viewpoint is natural in dealing with strong correlation and should relate directly to classical chemical bonding ideas. But there are several conceptually different manners of implementation of a localized-viewpoint cluster expansion, even aside from different sorts of schemes for the evaluation of matrix elements, and even in application to so simple a system as the antiferromagnetically-signed spin-$\frac{1}{2}$ Heisenberg model. As illustrated for this last model there are three general cluster-expansion developments:

(a) one that we view as focusing on the expansion of the wave-operator to be applied to a reference state [6–8] (the Neél state for the Heisenberg-model example);

(b) one that focuses on an expansion of the configuration coefficients [9–11] (involving a 'correlative' factorization of these coefficients); and

(c) a resonating valence-bond (VB) cluster expansion [11, 12] focusing on electron singlet-spin pairing patterns.

Much the same categorization of methods applies to other models as well, and indeed such (e.g. fermionic) models have more frequently been the focus of consideration. The computational techniques are understood conceptually since each type of approach relates directly to (what are better known as) non-trivial statistical-mechanical combinatoric problems—involving:

(a′) novel graphico-combinatoric problems of a statistical-mechanical flavour for the wave-operator expansion method;

(b′) Ising (or lattice gas) model problems for the configuration-coefficient expansion; and

(c′) dimer-covering related problems for the resonating VB ansatz.

Therefore such schemes are treated without any matrix-element approximation most readily for quasi-one-dimensional (or polymer) systems, hence the transfer-matrix scheme then is uniformly applicable, as 'emphasized' in the present context in [13], though the general technique in application to statistical mechanical problems dates back to the 1940s (see, e.g., [14]). Again, successful utilization of these wavefunction might be made using other (often non-variational) approaches, such as coupled-cluster [7] or Monte Carlo [10, 11] methods.

Here we investigate a novel feature of the Ising-model related configuration-coefficient cluster expansion. Namely, within the transfer-matrix framework, we show how from such a low-order (nearest-neighbour) cluster expansion for the ground state one may build up a set of excitations and multiple excitations all of which turn out to be orthonormal and to form a new (still largely 'local') basis, which in some cases is essentially complete. This approach differs from much previous work usually focusing solely on the ground state. With the proposed cluster-expansion basis one may reformulate the Hamiltonian in terms of it and then go on to an approximate solution (e.g., many-body perturbation expansion or perhaps a second cluster expansion), noting that just the single 'vacuum' basis state is already the result of the initial (correlated) cluster expansion. We illustrate the basis with application to the linear-chain spin-$\frac{1}{2}$ antiferromagnetically-signed Heisenberg model.

Despite such a correlated picture it is cautioned that there remains the question of the adequacy of the 'vacuum' and single-excitation above it as a description of the actual ground state and low-lying excitations above it—in particular, although the 'vacuum' is variationally optimized, it may in some cases fall into a different 'universality class' than the actual ground state.

2. Background

First the background framework and notation should be set. We presume a linear chain of *sites*, $i = 1 \to N$, each of which may have complex internal structure as encoded in their orthogonal *site states* $|ir\rangle$, r ranging. They combine into *system* product-like states

$$|r(N \to 1)\rangle \equiv |r_N, \ldots, r_2, r_1\rangle \tag{2.1}$$

that form an orthonormal basis for the system space

$$\langle r(N \to 1)|s(N \to 1)\rangle = \delta_{N \to 1}^{r,s} \equiv \prod_{i=1}^{N} \delta(r_i, s_i). \tag{2.2}$$

The states of (2.1) may be conventional Kronecker products of site states, but they could be antisymmetrized products, or the correlated cluster-expansion states of section 6. All we require is (2.2).

There is a natural set of generators of operators on the system space. These are the single-site operators

$$X_i(r, s) \equiv \sum_{r(N \to 1)}^{r_i=r} \sum_{s(N \to 1)}^{s_i=s} \delta_{N \to i+1}^{r,s} \delta_{i-1 \to 1}^{r,s} |r(N \to 1)\rangle \langle s(N \to 1)| \tag{2.3}$$

where the summations are restricted to fixed r_i and s_i. These operators act on the basis vectors thus

$$X_i(t, s)|r(N \to 1)\rangle = \delta(r_i, s)|r'(N \to 1)\rangle \tag{2.4}$$

where $|r'(N \to 1)\rangle$ is the same as $|r(N \to 1)\rangle$ except (perhaps) on site i which is in site state $r'_i \equiv t$. Further, these operators satisfy commutation relations

$$[X_i(r, s), X_j(t, u)] = \delta_{ij}\{\delta(s, t)X_i(r, u) - \delta(u, r)X_i(t, s)\} \tag{2.5}$$

and any other operators on the system may be generated from them. Typical operators of interest will be expressed in terms of summations (on i) of products over only a few $X_i(r, s)$. Here we assume all the sites are of the same type, with the site states on different sites having the same number M, and we assume these site states on different sites are translatable one into another. That is, each site has the same labels r.

3. Cluster expansion

The nearest-neighbour cluster-expanded wavefunctions are to be defined in terms of (variational) parameters each associated with neighbour pairs of sites, or with single sites at the chain ends (N and 1). The site-pair parameters are viewed as elements $(r|\mathbf{x}_{i+1i}|s)$ of a matrix $\mathbf{x}_{i+1i}$, $i = N-1 \to 1$, while the end-site parameters $(r|x_i)$ are viewed as components of vectors $|x_i\rangle$, $i = N$ or 1. Then the cluster-expanded basis states are introduced as

$$|\Phi[x(N \to 1)]\rangle = \sum_{r(N \to 1)} (x_N|r_N) \prod_i^{N-1 \to 1} (r_{i+1}|\mathbf{x}_{i+1i}|r_i)(r_1|x_1)|r(N \to 1)\rangle . \tag{3.1}$$

The overlap matrix element between two such wavefunctions (with possibly different x- and y-sets of parameters) is

$$\langle \Phi[x(N \to 1)]|\Phi[y(N \to 1)]\rangle = (x_N^* \cdot y_N| \prod_i^{N-1 \to 1} (\mathbf{x}_{i+1i}^* \cdot \mathbf{y}_{i+1i})|x_1^* \cdot y_1) \tag{3.2}$$

where we have recalled the orthonormality of (2.2) and have utilized the so-called 'Hadamard' product notation: a vector $|x \cdot y\rangle$ is the Hadamard product of $|x\rangle$ and $|y\rangle$, or a matrix $\mathbf{x} \cdot \mathbf{y}$ is the Hadamard product of $\mathbf{x}$ and $\mathbf{y}$, when

$$\begin{aligned}
(r|x \cdot y) &\equiv (r|x)(r|y) \\
(r|\mathbf{x} \cdot \mathbf{y}|s) &\equiv (r|\mathbf{x}|s)(r|\mathbf{y}|s) .
\end{aligned} \tag{3.3}$$

Thus with M site states per site, all has been reduced in (3.2) to the manipulation of $M \times M$ matrices

$$\mathbf{T}_{i+1i} \equiv \mathbf{x}_{i+1i}^* \cdot \mathbf{y}_{i+1i} \tag{3.4}$$

called *transfer* matrices.

Local matrix elements may be dealt with upon making a few additional refinements. For a one-site matrix element over $X_i(s, s')$ we introduce a *connection* matrix $\mathbf{C}_{i+1 \to i-1}(s, s')$ with elements

$$(r|\mathbf{C}_{i+1 \to i-1}(s, s')|t) = (r|\mathbf{x}_{i+1i}^*|s)(s|\mathbf{x}_{ii-1}^*|t)(r|\mathbf{y}_{i+1i}|s')(s'|\mathbf{y}_{ii-1}|t) . \tag{3.5}$$

Of course, one may be interested in a suitable linear combination $\mathcal{H}_i$ of $X_i(s, s')$—that is, $\mathcal{H}_i$ involves a summation over the site-state indices s, s'. Therefore one may introduce the same linear combination $\mathbf{C}_{i+1 \to i-1}$ of the $\mathbf{C}_{i+1 \to i-1}(s; s')$. Now, making an expansion parallelling that for the overlap, one obtains

$$\begin{aligned}
\langle \Phi[x(N \to 1)]|\mathcal{H}_i|\Phi[y(N \to 1)]\rangle \\
= (x_N^* \cdot y_N|\mathbf{T}_{NN-1} \cdots \mathbf{T}_{i+3i+2}\mathbf{C}_{i+1 \to i-1}\mathbf{T}_{i-1i-2} \cdots \mathbf{T}_{21}|x_1^* \cdot y_1) .
\end{aligned} \tag{3.6}$$

So all is reduced to M by M matrices.

For a two-site matrix element over $X_{i+1}(s, s')X_i(t, t')$ one introduces a ('longer') connection matrix $\mathbf{C}_{i+2 \to i-1}(s, t; s', t')$ with elements

$$(r|\mathbf{C}_{i+2 \to i-1}(s, t; s', t')|u) = (r|\mathbf{x}^*_{i+2\,i+1}|s)(s|\mathbf{x}^*_{i+1\,i}|t)(t|\mathbf{x}^*_{i\,i-1}|u)$$
$$\times (r|\mathbf{y}_{i+2\,i+1}|s')(s'|\mathbf{y}_{i+1\,i}|t')(t'|\mathbf{y}_{i\,i-1}|u) . \tag{3.7}$$

Again, for linear combinations $\mathcal{H}_{i+1\,i}$ of two-site operators, corresponding linear-combination connection matrices $\mathbf{C}_{i+2 \to i-1}$ may be introduced, and an expression obtained like (3.6) but now with the matrix $\mathbf{C}_{i+2 \to i-1}$ while $\mathbf{T}_{i+2\,i+1}$ is missing.

4. The ground state

Frequently, the ground state (away from the ends) approaches translational invariance (as $N \to \infty$), in which case this ground state might be approximated by a single cluster-expanded wavefunction with a single set of (variational) pair parameters $x_{i+1\,i}$. Such a state might be denoted by $\Psi(x_1, x_N)$ in place of $\Phi[x(N \to 1)]$. Also, the transfer matrices all become identical so that their site labels are dropped. Solution of the eigenproblem

$$\mathbf{T}|\lambda) = \lambda|\lambda) \tag{4.1}$$

allows a useful analysis of the overlap

$$\langle \Psi(x_1, x_N)|\Psi(x_1, x_N)\rangle = \sum_\lambda (x^*_N \cdot x_N|\lambda)\lambda^{N-1}(\lambda|x^*_1 \cdot x_1) \tag{4.2}$$

assuming that the eigenvectors λ are normalized and $\mathbf{T}$ is symmetric. (If $\mathbf{T}$ is not symmetric, then an analogous result follows when one pays attention to distinguish left and right eigenvectors, and biorthonormalizes them.) Now $\mathbf{T}$ is seen from (3.4) to have non-negative elements, hence the Frobenius–Perron theorem [15] implies that there is a maximum-magnitude eigenvalue Λ which is real and positive, and whose eigenvector $|\Lambda)$ has all components $(r|\Lambda)$ of the same phase, so that one may choose all $(r|\Lambda)$ positive. Further, if a sufficient set of elements of (our often symmetric) $\mathbf{T}$ are non-zero, then $|\Lambda)$ is non-degenerate, not only in value but magnitude as well, so that for long chains this eigenvalue dominates over the others in both the overlap (4.2) and the analogously developed formula for matrix elements as (3.6).

But an analogous formula can be obtained exactly without resort to $N \to \infty$ asymptotics if we choose the end parameters x_1 and x_N judiciously

$$(r|x_1) = (r|x_N) = (r|\Lambda)^{1/2} . \tag{4.3}$$

Then $|x^*_i \cdot x_i) = |\Lambda)$, $i = N$ and 1, so that (now also dropping the x_1, x_N indices on Ψ)

$$\langle \Psi|\Psi\rangle = \Lambda^{N-1} . \tag{4.4}$$

Likewise

$$\langle \mathcal{H}_{i+1\,i}\rangle \equiv \frac{\langle \Psi|\mathcal{H}_{i+1\,i}|\Psi\rangle}{\langle \Psi|\Psi\rangle} = \frac{(\Lambda|\mathbf{C}|\Lambda)}{\Lambda^3} \tag{4.5}$$

at least so long as neither i nor $i + 1$ are end sites.

The $(r|\mathbf{x}|s)$ may be treated as variational parameters. The minimization of the Hamiltonian expectation with respect to these $(r|\mathbf{x}|s)$ is a non-linear problem but is numerically feasible for modest M, this involving just repeated manipulations of $M \times M$ matrices $\mathbf{T}$.

If the variational parameters $(r|\mathbf{x}_{i+1\,i}|s)$ do not well approach translationally fixed values, they should still often exhibit some small period, say p. Then a p-fold product of the p different asymptotic transfer matrices, enables a somewhat similar treatment to be made, though now with up to a power of p times as many variational parameters.

5. Bond excitations

Now, even for somewhat more general cluster expansions results analogous to those in the preceding section (at least in an asymptotic sense) can be obtained and even extended [16] to define and deal with 'single' local excitations. But the present cluster expansion (at least sometimes) allows much more: a 'complete' set of 'multiple' excitations may be developed and a new representation obtained. To ultimately accomplish this we seek M different pair-excitation matrices $\mathbf{x}^\xi$, first with $\xi = 0$ and $|\Lambda) \equiv |0)$ corresponding to the ground-state case of the preceding section and second with

$$(\mathbf{x}^{\xi^*} \cdot \mathbf{x}^\zeta)|0) = \delta_{\xi\zeta}|0) \,. \tag{5.1}$$

Here it is understood that the introduction of a factor Λ_ξ on the right of (5.1) would not effectively relax these conditions, since if satisfied with this factor, then rescaling of the so-found $\mathbf{x}^\Lambda$ by $\Lambda_\xi^{-1/2}$ would yield the result as written. The Hadamard products on the left of (5.1) are evidently transfer matrices $\mathbf{T}^{(\xi,\zeta)}$, with $\mathbf{T}^{(0,0)} = \mathbf{T}$ and $\mathbf{x}^0 = \mathbf{x}$ of the preceding section rescaled so that $\Lambda = 1$. The matter of generation of *canonical* bond excitations satisfying (5.1) is delayed until section 7.

Assuming M matrices $\mathbf{x}^\xi$ satisfying (5.1), we define many-body cluster-expanded basis states thus

$$|\Psi[\xi(N - 1 \to 1)]\rangle \equiv \sum_{r(N \to 1)} (0|r_N)^{1/2} \prod_{i=1}^{N-1} (r_{i+1}|\mathbf{x}^{\xi_i}|r_i)(r_1|0)^{1/2}|r(N \to 1)\rangle \tag{5.2}$$

where, of course, $\xi(N - 1 \to 1) \equiv \xi_{N-1}, \ldots, \xi_2, \xi_1$. That is, the states are just $|\Phi[\mathbf{x}(N \to 1)]\rangle$ as in (3.1) but with all the variational parameters chosen in a special way. The key point is the development of the consequent overlaps in terms of transfer matrices $\mathbf{T}^{(\xi,\zeta)}$, followed by the use of (5.1),

$$\langle\Psi[\xi(N - 1 \to 1)]|\Psi[\zeta(N - 1 \to 1)]\rangle = (0|\prod_{i=1}^{N-1} \mathbf{T}^{(\xi_i,\zeta_i)}|0)$$

$$\equiv \prod_{i=1}^{N-1} \delta(\xi_i, \zeta_i) \equiv \delta_{N-1\to 1}^{\xi\zeta} \,. \tag{5.3}$$

The cluster-expanded basis states are orthonormal, and they are essentially complete. (There is some sort of minor deficiency in connection with the chain ends—there being just $N - 1$ bonds on which to make excitations, though there were N sites.)

6. Cluster-expansion representation

In general, our essentially orthonormal cluster-expansion basis may be used to develop a new representation. Of key importance, of course, is the representation of the Hamiltonian.

Since the bond excitations are somewhat local, Hamiltonian interactions originally local in the representation of section 2 remain somewhat so. The neighbour-interaction matrix elements are

$$\langle\Psi[\xi(N - 1 \to 1)]|\mathcal{H}_{i+1\,i}|\Psi[\zeta(N - 1 \to 1)]\rangle = \delta_{N-1\to i+2}^{\xi\zeta}\delta_{i-1\to 1}^{\xi\zeta}(0|\mathbf{C}_{i+2\to i-1}^{(\xi,\zeta)}|0) \tag{6.1}$$

where we have abbreviated two Kronecker delta functions (as in (2.2) or (5.3)), and the $\mathbf{C}$-matrix here is as described in (3.7) but now for special parameter values, with the $\mathbf{x}_{j+1\,j}^* = \mathbf{x}_{j+1\,j}^{\xi^*}$ and the $\mathbf{y}_{j+1\,j} = \mathbf{x}_{j+1\,j}^\zeta$. For i or $i + 1$ at a chain end there would be a

modification. Notably, if the original interactions extend only over n nearest neighbour sites, then the new renormalized interactions extend only over $n + 1$ neighbouring bonds.

To effect the new representation introduce local operators, analogous to the local $X_i(r, s)$ of section 2, thus

$$Y_i^{\xi\zeta} = \sum_{\xi(N-1\to 1)}^{\xi_i=\xi} \sum_{\zeta(N-1\to 1)}^{\zeta_i=\zeta} \delta_{N\to i+1}^{\xi\zeta} \delta_{i-1\to 1}^{\xi\zeta} |\Psi[\xi[N-1\to 1]]\rangle\langle\Psi[\zeta(N-1\to 1)]| . \tag{6.2}$$

Evidently, bonds now play the role that sites did in section 2. Presuming the original Hamiltonian to have just a nearest-neighbour interaction, we have its new representation

$$\mathcal{H} = \sum_i \sum_{\xi_3\xi_2\xi_1} \sum_{\zeta_3\zeta_2\zeta_1} (0|\mathbf{C}_{4-\text{site}}^{(\xi,\zeta)}|0) Y_{i+2}^{\xi_3\zeta_3} Y_{i+1}^{\xi_2\zeta_2} Y_i^{\xi_1\zeta_1} \tag{6.3}$$

where again there would be corrections at the ends of the chain.

7. Satisfaction of canonicalness

The point now is to indicate that at least sometimes it is possible that there are M solutions $\mathbf{x}^\xi$ to the $\sim M^2$ matrix conditions of (5.1), and hence that we have a complete set of bond excitations. For simplicity consider the common case where the $\mathbf{x}^\xi$ are symmetric and real (so that there are just $M(M + 1)/2$ conditions (5.1)). Again, we imagine that $\mathbf{x}^0 = \mathbf{x}$ for the $\xi = 0$ ground state is already available. For the $\xi \neq \zeta$ case the conditions (5.1) merely say that the transfer matrix $(\mathbf{x}^\xi \cdot \mathbf{x}^\zeta)$ has $|0)$ as an eigenvector with eigenvalue 0, so that $\mathbf{x}^\xi \cdot \mathbf{x}^\zeta$ is confined to the $(M - 1)$-dimensional subspace spanned by all the eigenvectors of $\mathbf{T} = \mathbf{T}^{0,0}$ other than $|0)$. When $\xi = \zeta$ evidently $|0)$ still is to be an eigenvector to $(\mathbf{x}^\xi \cdot \mathbf{x}^\zeta)$ but now with eigenvalue unity. Thus

$$(\mathbf{x}^\xi \cdot \mathbf{x}^\zeta) = \delta_{\xi\zeta}|0)(0| + \sum_{\mu,\nu}^{\neq 0} |\mu)\mathbf{t}_{\mu\nu}^{\xi\zeta}(\nu| . \tag{7.1}$$

But not all these $(M - 1) \times (M - 1)$ matrices $\mathbf{t}^{\xi\zeta}$ can be chosen independently. As a step toward seeing this more precisely, abbreviate those with $\zeta = 0$ (and $\xi \neq 0$) thus $\mathbf{t}^{\xi 0} = \mathbf{t}^\xi$, hence from (7.1) one has

$$(r|\mathbf{x}^\xi|s) = \frac{1}{(r|\mathbf{x}|s)} \sum_{\mu,\nu}^{\neq 0} (r|\mu)\mathbf{t}_{\mu\nu}^\xi(\nu|s) . \tag{7.2}$$

But now the remaining $\mathbf{t}^{\xi\zeta}$ with both subscripts $\xi, \zeta \neq 0$ are to be related via

$$\frac{\sum_{\mu,\nu}^{\neq 0}(r|\mu)\mathbf{t}_{\mu\nu}^\xi(\nu|s) \sum_{\mu,\nu}^{\neq 0}(r|\mu)\mathbf{t}_{\mu\nu}^\zeta(\nu|s)}{(r|\mathbf{x}|s)^2} = \delta_{\xi\zeta}(r|0)(0|s) + \sum_{\mu,\nu}^{\neq 0}(r|\mu)\mathbf{t}_{\mu\nu}^{\xi\zeta}(\nu|s) . \tag{7.3}$$

Here, we see that the number of scalar equations (7.3), letting ξ, ζ, r, s range and noting the symmetries $\mathbf{x}^{\xi\dagger} = \mathbf{x}^\xi$, $\mathbf{t}^{\xi\dagger} = \mathbf{t}^\xi$, $\mathbf{t}^{\xi\zeta\dagger} = \mathbf{t}^{\xi\zeta} = \mathbf{t}^{\zeta\xi}$, is

$$\#_{\text{eq}} = \tfrac{1}{2}(M - 1)M \tfrac{1}{2}M(M + 1) . \tag{7.4}$$

We also see that the number of different scalar elements in the matrices $\mathbf{t}^\xi$, $\mathbf{t}^{\xi\zeta}$ is

$$\#_{\text{t}}(M) = \tfrac{1}{2}(M - 1)M\left[(M - 1) + \tfrac{1}{2}(M - 1)M\right] . \tag{7.5}$$

Here, the number $\#_{\text{eq}}(M)$ of conditions to be satisfied exceeds the number $\#_{\text{t}}(M)$ of parameters by $M(M - 1)/2$, so one might ordinarily expect that these conditions cannot always be met.

However, for many interesting models a symmetry is manifested which effectively reduces the number of conditions. For instance, for $M = 2$ the spin-$\frac{1}{2}$ Heisenberg model (in the absence of a magnetic field) there is a symmetry interchanging the site states labelled by spin-up α and spin-down β, so that the $r = s = \alpha$ and $r = s = \beta$ conditions of (7.3) become identical, hence all can be satisfied, as illustrated explicitly in section 8. For $M = 3$, the spin-1 Heisenberg model and the tJ-model each have site states one might label $+, 0, -$ with a symmetry interchanging $+$ and $-$. For $M = 4$ the spin-$\frac{3}{2}$ Heisenberg model and the Hubbard model also have site states exhibiting an interchange symmetry under spin-flip. But regardless of whether all M^2 of the conditions of (5.1) can be met we have at hand a renormalization generally leaving at least $M - 1$ lower local states for each bond.

8. Two states per site

The ideas of the preceding sections may be illustrated by their application to the simplest non-trivial linear-chain case—that with $M = 2$ states per site. The best known quantum model for $M = 2$ is the spin-$\frac{1}{2}$ Heisenberg model, as well as various modifications to it. If we label the site states $r = \alpha, \beta$, then the ground-state ($\xi = 0$) transfer matrix appears as

$$\mathbf{T} = \begin{pmatrix} (\alpha|\mathbf{x}|\alpha)^2 & (\alpha|\mathbf{x}|\beta)^2 \\ (\beta|\mathbf{x}|\alpha)^2 & (\beta|\mathbf{x}|\beta)^2 \end{pmatrix}. \tag{8.1}$$

This form may be used to treat ground states for quite general models—even models which in some sense have $1 < M < 2$, as in [17]. But if α and β are equivalent or 'degenerate', as now assumed, then it is reasonable to take not only $(\alpha|\mathbf{x}|\beta) = (\beta|\mathbf{x}|\alpha)$ but also $(\beta|\mathbf{x}|\beta) = (\alpha|\mathbf{x}|\alpha)$. Therefore the maximum-eigenvalue eigenvector to $\mathbf{T}$ becomes

$$|0) = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \tag{8.2}$$

independently of the value of $(\alpha|\mathbf{x}|\alpha)$ and $(\alpha|\mathbf{x}|\beta)$. Further, because of the chosen normalization (that the maximum eigenvalue Λ of $\mathbf{T}$ be 1), only one of the parameters actually remains independent. That is, we may express this in terms of one (new) parameter x,

$$\mathbf{x} = \frac{1}{\sqrt{1 + x^2}} \begin{pmatrix} x & 1 \\ 1 & x \end{pmatrix} \tag{8.3}$$

so that the transfer matrix $\mathbf{x} \cdot \mathbf{x}$ has $\Lambda = 1$. For the antiferromagnetically-signed Heisenberg model one expects $|x| < 1$.

To obtain a full cluster-expansion representation we develop the canonical excitations following section 7. Letting μ be the second (i.e. smaller) eigenvalue to $\mathbf{T}$, the parameters for the excitation $\xi = \star$ are to be given by

$$(r|\mathbf{x}^\star|s) = \frac{(r|\mu)\mathbf{t}^\star_{\mu\mu}(\mu|s)}{(r|\mathbf{x}|s)} \tag{8.4}$$

where $\mathbf{t}^\star_{\mu\mu}$ and a second parameter $\mathbf{t}^{\star\star}_{\mu\mu}$ are to satisfy the conditions

$$\frac{(r|\mu)\mathbf{t}^\star_{\mu\mu}(\mu|s)}{(r|\mathbf{x}|s)^2} = (r|0)(0|s) + (r|\mu)\mathbf{t}^{\star\star}_{\mu\mu}(\mu|s). \tag{8.5}$$

But again with the symmetric case this reduces to just two distinct conditions, for $(r, s) = (\alpha, \alpha)$ and $(r, s) = (\alpha, \beta)$. Recalling (8.3) these two conditions are

$$
\left\{ \frac{\sqrt{x^2 + 1}}{x} \mathbf{t}^\star_{\mu\mu \frac{1}{2}} \right\}^2 = \tfrac{1}{2} + \tfrac{1}{2} \mathbf{t}^{\star\star}_{\mu\mu}
$$

$$
\left\{ \sqrt{x^2 + 1}\, \mathbf{t}^\star_{\mu\mu}\!\left(-\tfrac{1}{2}\right) \right\}^2 = \tfrac{1}{2} + \left(-\tfrac{1}{2}\right) \mathbf{t}^{\star\star}_{\mu\mu}
$$

$$(8.6)$$

whence we obtain

$$
\mathbf{t}^\star_{\mu\mu} = \frac{2x}{x^2 + 1} \qquad \mathbf{t}^{\star\star}_{\mu\mu} = \frac{1 - x^2}{1 + x^2}.
$$

$$(8.7)$$

Thus

$$
\mathbf{x}^\star = \frac{1}{\sqrt{1 + x^2}} \begin{pmatrix} 1 & -x \\ -x & 1 \end{pmatrix}
$$

$$(8.8)$$

and a cluster-expansion basis as in section 7 has been realized.

9. The linear-chain Heisenberg model

The isotropic spin-$\frac{1}{2}$ Heisenberg model for a linear chain is

$$
\mathcal{H}_0 = 2J \sum_i \mathbf{s}_{i+1} \cdot \mathbf{s}_i
$$

$$(9.1)$$

where J is the 'exchange' parameter and $\mathbf{s}_j$ is the spin operator for site j. It is also the Pauling–Wheland valence-bond model for a conjugated chain (e.g. a polyene or polyacetylene). For convenience we choose the initial basis as products of site states with a sign

$$
|\sigma(N \to 1)\rangle = \pm \sigma_1(1) \times \sigma_2(2) \times \cdots \times \sigma_N(N) \qquad \sigma_i = \alpha, \beta \qquad i = 1 \to N
$$

$$(9.2)$$

where the $\pm$ sign is simply the parity of the number of β-spins on odd sites. The advantage of this choice [18] is that the ground state is 'nodeless' on this basis, i.e. the ground-state configuration–interaction coefficients are then all of the same sign (which is positive, with our choice of phase).

The ground-state cluster expansion is set up as indicated in section 8, with a variational parameter and transfer matrix as in (8.3). The connection matrix is conveniently broken into two parts, the first $\mathbf{C}_d$ being that for the 'diagonal' part of the interaction

$$
(\sigma|\mathbf{C}_d|\tau) = (\sigma|\mathbf{x}|\tau)\langle \sigma, \tau | 2 s_1^z s_2^z | \sigma, \tau \rangle (\sigma|\mathbf{x}|\tau)
$$

$$(9.3)$$

so that

$$
\mathbf{C}_d = \frac{1}{x^2 + 1} \begin{pmatrix} x^2/2 & -1/2 \\ -1/2 & x^2/2 \end{pmatrix}.
$$

$$(9.4)$$

The second part $\mathbf{C}_{od}$ for the 'off-diagonal' part involves four sites

$$
(\rho|\mathbf{C}_{od}|\omega) = -\sum_{\sigma\sigma'} \sum_{\tau\tau'} (\rho|\mathbf{x}|\sigma)(\sigma|\mathbf{x}|\tau)(\tau|\mathbf{x}|\omega)
$$
$$
\times \langle \sigma, \tau | s_1^+ s_2^- + s_1^- s_2^+ | \sigma', \tau' \rangle (\rho|\mathbf{x}|\sigma')(\sigma'|\mathbf{x}|\tau')(\tau'|\mathbf{x}|\omega)
$$

$$(9.5)$$

and leads to

$$
\mathbf{C}_{od} = -\frac{2x^2}{(x^2 + 1)^3} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}.
$$

$$(9.6)$$

Thus via (4.5) the ground-state energy expectation per site is

$$\varepsilon_0 = J \langle 0 | \left[\mathbf{C}_d + \mathbf{C}_{od} \right] | 0 \rangle = J \left\{ \frac{x^2 - 1}{2(x^2 + 1)} - \frac{4x^2}{(x^2 + 1)^3} \right\} . \tag{9.7}$$

This is sufficiently simple that it may be analytically optimized. The expression is cubic in the (positive) variable $z \equiv 1/(x^2 + 1)$, and the derivative of this leads to a quadratic which leads us back to optimal values

$$x = \left\{ 2\sqrt{7} - 5 \right\}^{1/2} \approx +0.539\,909\,83$$
$$\varepsilon_0/J = -\tfrac{7}{27} \left\{ \sqrt{7} + \tfrac{1}{2} \right\} \approx -0.815\,565\,15 . \tag{9.8}$$

Indeed, this is just the result of Kasteleyn [9], although here obtained in a somewhat different manner. Table 1 compares it with a few other results, including the exact solution [19] near $-0.886\,28$ (which does not, however, always readily extend to other related linear-chain models). There are a number of other (even variational) results which give more accurate estimates of the exact result, but not all yield so complete a presentation of the excitation spectrum as follows with the present approach, which in addition yields a whole new still local renormalized representation.

Table 1. Ground-state energy estimates per site.

Method	Energy	References
Néel state	−0.5000	
Kekülé state	−0.7500	
Two-site CI-coefficient cluster expansion	−0.8156	[9], here
RHF (after Jordan–Wigner transform)	−0.8383	[20]
Two-bond resonating-VB cluster expansion	−0.8538	[11, 12]
Two-site wave-operator cluster expansion	−0.8558	[6, 8]
SCF antiferromagnetic spin-wave	−0.8634	[21]
Exact	−0.8863	[19]
2nd order Néel-state perturbation expansion	−1.0000	[22]

10. Renormalized Heisenberg model

The 'bond' excitations for the new basis are indicated near the end of section 8. For the excitation spectrum (and newly represented Hamiltonian) we also need to develop the bond-excitation labelled connection matrices, as in section 6. For the present case we break the interactions into two parts, as in (9.3) and (9.5) for the ground state. The first part gives matrix elements

$$(\sigma | \mathbf{C}_d^{(\xi,\zeta)} | \tau) = (\sigma | \mathbf{x}^\xi | \tau)\langle \sigma, \tau | 2s_1^z s_2^z | \sigma, \tau \rangle (\sigma | \mathbf{x}^\zeta | \tau) \tag{10.1}$$

so that

$$\mathbf{C}_d^{\xi,\zeta} = \tfrac{1}{2} \mathbf{x}^\xi \cdot \mathbf{x}^\zeta \cdot \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} . \tag{10.2}$$

Similarly, the second part leads to

$$\mathbf{C}_{od}^{(\xi(3\to1);\zeta(3\to1))} = - \left(\mathbf{x}^{\xi_3} \cdot \mathbf{x}_{\zeta_3} \right) \begin{pmatrix} 0 & (\alpha|(\mathbf{x}^{\xi_2} \cdot \mathbf{x}^{\zeta_2})|\beta) \\ (\alpha|(\mathbf{x}^{\xi_2} \cdot \mathbf{x}^{\zeta_2})|\beta) & 0 \end{pmatrix} \left(\mathbf{x}^{\xi_1} \cdot \mathbf{x}_{\zeta_1} \right) \tag{10.3}$$

where a matrix $\mathbf{x}_\zeta$ is obtained from $\mathbf{x}^\zeta$ by interchanging its rows. Following the ideas in section 6, we then find the newly represented Hamiltonian

$$\mathcal{H}_0 = J \sum_i \left\{ \frac{1}{x^2+1} \Delta_i - \frac{1}{(x^2+1)^3} A_{i+1} \nabla_i A_{i-1} \right\} \tag{10.4}$$

in terms of three types of one-bond operators

$$\Delta \equiv \frac{1-x^2}{2}(-Y^{00}+Y^{\star\star}) + x(Y^{0\star}+Y^{\star 0})$$
$$A \equiv 2x(Y^{00}-Y^{\star\star}) + (1-x^2)(Y^{0\star}+Y^{\star 0}) \tag{10.5}$$
$$\nabla \equiv Y^{00}+x^2 Y^{\star\star} - x(Y^{0\star}+Y^{\star 0}).$$

Again, there are some corrections in the operator representation at the chain ends, but we neglect these here. It should be emphasized that bond operators for different bonds commute, even if the bonds are adjacent (and viewed to share a site, in terms of which the original representation of section 2 was built).

11. Next-nearest-neighbour Heisenberg model

Beyond the nearest-neighbour Heisenberg model which has already been solved exactly [19], one may apply the current methodology to other models which have not been solved exactly. Such is the next-nearest-neighbour Heisenberg model

$$\mathcal{H}_\gamma = 2J \sum_i (s_{i+1} \cdot s_i + \gamma s_{i+1} \cdot s_{i-1}). \tag{11.1}$$

The results of section 8 still apply, but new connection matrices for the next-nearest-neighbour interaction need to be constructed, much as done for the nearest-neighbour interaction in sections 9 and 10. We find the consequent new cluster-expansion representation to be

$$\mathcal{H}_\gamma = \mathcal{H}_0 + \gamma J \sum_i \left\{ \frac{2}{(x^2+1)^2} \Delta_{i+1}\Delta_i + \frac{1}{(x^2+1)^4} A_{i+2}(R^+_{i+1}R^-_i + R^-_{i+1}R^+_i)A_{i-1} \right\}$$
$$\tag{11.2}$$

where the Δ and A operators are as before in (10.5) and R^+, R^- are a conjugate pair

$$R^+ \equiv x(Y^{00}-Y^{\star\star}) - x^2 Y^{0\star} + Y^{\star 0}$$
$$R^- \equiv (R^+)^\dagger \tag{11.3}$$

and actually appear in a hidden form in (10.4) since $R^+ + R^- = A$.

The ground-state expectation per site for $\mathcal{H}_\gamma$ is readily obtained from this general representation, as

$$\varepsilon_\gamma = \varepsilon_0 + \gamma \left\{ \frac{(x^2-1)^2}{2(x^2+1)^2} + \frac{8x^4}{(x^2+1)^4} \right\}. \tag{11.4}$$

Again this is a function of $z = 1/(1+x^2)$, now quartic in this variable, so that optimization leads to a cubic polynomial and the result of figure 1. Here too, earlier results [8, 12, 23] are better, though again we have a new representation (in equation (11.2)).
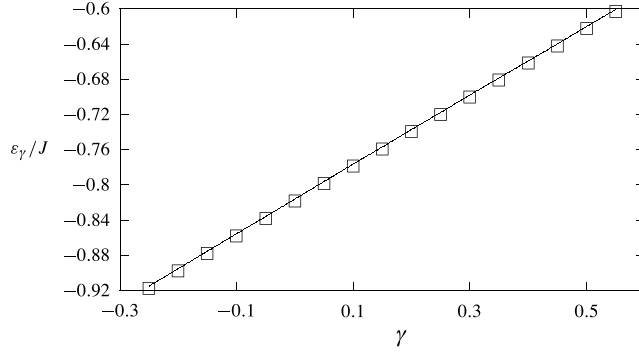
Figure 1. Energy (in units of J) per site for the next-nearest Heisenberg Hamiltonian.

12. Discussion

Thus we have proposed a novel type of basis, which although it has the mathematical features of a simple orthonormal product of site states, also has much of the physical correlation included. The ground 'vacuum' is itself a correlated cluster expansion with (possibly optimized) variational parameters. The single excitations involve localized (correlated changes) from the ground 'vacuum', and may be chosen such that they do not directly couple to the ground 'vacuum'. The multiple excitations appear in some sense product-like as independent single excitations, retaining mutual orthogonality though as represented in the original basis they are not simple products. Moreover, we have indicated a scheme to realize all this for a certain class of linear-chain systems. Even in cases where the transformed basis states are not complete (i.e. span only a proper subspace) the approach may offer a useful (approximate) renormalization technique.

The present outline raises several questions. Can the scheme to achieve canonicalness of bond excitations be made more elegantly constructive? Can the scheme be formulated without so intimate a connection to the transfer-matrix formulation for treating matrix elements? Can the scheme be extended to other dimensionalities or types of cluster expansions?

Another question concerns the degree of completeness of our present basis. Since with N sites there are only $N - 1$ bonds in an (open) linear chain, there are just M^{N-1} basis states of the type in (5.2) while there are M^N original basis states in (2.1). This discrepancy (as it concerns boundary conditions) should be asymptotically negligible in many respects: the ground-state energy per site should be accurately included to $\sim 1/N$ and the bulk of the single-excitation eigenspectrum should be similarly accurate. In fact, much is elucidated if we modify the states of (5.2) thus

$$|\Psi_{\mu 0}[\xi(N-1 \to 1)]\rangle = \sum_{r(n \to 1)} \frac{(\mu|r_N)}{(0|r_N)^{1/2}} \left\{ \prod_i^{N-1 \to 1} (r_{i+1}|\mathbf{x}^{\xi_i}|r_i) \right\} (r_1|0)^{1/2} |r(n \to 1)\rangle \quad (12.1)$$

and also

$$|\Psi_{0\mu}[\xi(N-1 \to 1)]\rangle = \sum_{r(n \to 1)} (0|r_N) \left\{ \prod_i^{N-1 \to 1} (r_{i+1}|\mathbf{x}^{\xi_i}|r_i) \right\} \frac{(\mu|r_1)}{(0|r_1)^{1/2}} |(r(N \to 1)\rangle. \quad (12.2)$$

For $\mu \neq \Lambda$ one finds that the sets of (12.1) and (12.2) are both orthogonal to that of (5.2). Further, (12.1) and (12.2) tend to be biorthonormal to one another (for $M = 2$, or in general

if there are not too many excitations) so that with $M - 1$ eigenvalues $\mu \neq \Lambda$ we seem to have cluster-expansion basis states for the remnant $M - 1$ M^{N-1}-dimensional space. Matrix element formulae even follow in much the same manner though μ and $|\mu\rangle$ replace Λ and $|\Lambda\rangle$ in much of the formulation. We do not pursue this here.

There is also a further question of how readily one might make an accurate solution to the transformed model, as in (6.3) or (10.4). Of course, the whole exercise has been based upon a simple cluster expansion, foremost for the ground state, but then also for excitations above it. Thus even the single basis vector $\Psi[0, 0, \ldots, 0]\rangle \equiv |\Psi\rangle$ should hopefully yield a reasonable ground-state estimate, though in the case of the Heisenberg model in section 9 we found this estimate to be only of rather modest accuracy. Ordinary (non-degenerate) many-body perturbation theory might be tried. Or a second cluster expansion might be tried. Indeed, such a next ground-state cluster expansion is considered in a later paper with application to the Heisenberg model.

Overall the cluster-expansion basis seems novel, possibly opening new paths in treating quasi-one-dimensional systems.

Acknowledgments

References

[1]   Bethe H 1955 *Phys. Rev.* **103** 1353–90
[2]   Hubbard J 1957 *Proc. R. Soc.* A **240** 539–60
[3]   Brout R 1958 *Phys. Rev.* **111** 1324–33
[4]   Primas H 1965 *Modern Quantum Chemistry* ed O Sinanoglu (New York: Academic) pp 45–74
[5]   Cizek J 1966 *J. Chem. Phys.* **45** 4256–66
[6]   Vroelant C and Daudel R 1949 *Bull. Soc. Chim.* **16** 36–43
      Nebanzahl I 1969 *Phys. Rev.* **177** 1001–5
      Bartowski R R 1972 *Phys. Rev.* B **5** 4536–41
[7]   Bishop R F, Parkinson J B and Xian Y 1991 *Phys. Rev.* B **44** 9425–43
[8]   Garcia-Bach M A and Klein D J 1977 *Int. J. Quantum Chem.* **12** 273–89
[9]   Kasteleyn P W 1952 *Physica* **28** 104–13
[10] Huse D A and Elser V 1988 *Phys. Rev. Lett.* **60** 2531–34
[11] Liang S, Doucot B and Anderson P W 1988 *Phys. Rev. Lett.* **61** 365–8
[12] Klein D J 1979 *Phys. Rev.* B **19** 870–6
      Klein D J and Garcia-Bach M A 1979 *Phys. Rev.* B **19** 877–86
[13] Klein D J, Hite G E and Schmalz T G 1986 *J. Comput. Chem.* **7** 443–56
[14] Biggs N L 1977 *Interaction Models* (Cambridge: Cambridge University Press)
[15] Gantmacher F P 1959 *The Theory of Matrices* vol II (New York: Chelsea) ch 13
[16] Klein D J, Garcia-Bach M A and Seitz W A 1989 *J. Mol. Struct.* **185** 275–85
[17] Klein D J and Schmalz T G 1989 *Int. J. Quantum Chem.* **35** 373–83
      Garcia-Bach M A, Valenti R and Klein D J 1989 *J. Mol. Spectrosc.* **185** 287–96
[18] Marshall W 1955 *Proc. R. Soc.* A 232 48–68
[19] Hulthén L 1938 *Ark. Mat. Astron. Fys.* A **26** A 11 188
[20] Rodriguez S 1959 *Phys. Rev.* **116** 1474–7
[21] Anderson P W 1952 *Phys. Rev.* **86** 694–701
      Kubo R 1952 *Phys. Rev.* **87** 568–80
[22] Hartmann H 1947 *Z. Naturf.* A **2** 259
[23] Neimeijer Th 1971 *J. Math. Phys.* **12** 1487–91