Journal of Statistical Physics, Vol. 24, No. 4, 1981

Cluster Expansion and Generalized Transfer Matrices for the Statistical Mechanics of Linear Chains

Douglas J. Klein[1,2] and Terry L. Welsher[1,3]

Received June 26, 1980

A cluster expansion of the statistical mechanical density operator for a general linear chain model with nearest-neighbor interactions is made. This expansion is then shown to lead to an expansion of a generalized transfer matrix, whose maximum eigenvalue is the per-site partition function. A number of computational features, as well as some illustrative examples, of this approach are described.

KEY WORDS: Ising model; Heisenberg model; transfer matrix; cluster expansion.

1. INTRODUCTION

We consider linear chain models with nearest-neighbor interactions only. Such models are widely applicable in describing, for example,

(a) magnetic properties of many crystals with linear chains of superexchange-coupled paramagnetic transition metal ions;
(b) magnetic and electric properties of many crystals of aromatic donors and/or acceptors;
(c) conformational distributions in linear polymers;
(d) orientational or substitutional disorder in many linear chain molecular crystals;

Research supported in part by the Robert A. Welch Foundation, Houston, Texas.

[1] Departments of Physics and Chemistry, The University of Texas, Austin, Texas 78712.
[2] Present address: Department of Marine Science, Moody College, Galveston, Texas 77550.
[3] Present address: Bell Laboratories, Whippany, New Jersey 07981.

(e) phonon and/or electronic structure of linear chain polymers;
(f) motion of adsorbed molecules on linear biopolymers, on very anisotropic crystal surfaces, or along "channels" within certain crystals;
(g) properties of linear chain models composed from, say, m parallel and interacting subchains such that their properties can be extrapolated to $m \to \infty$ so as to obtain results for a variety of two-dimensional models, with a variety of applications.

For the more refined versions of these various models in (a)–(g) one typically finds that accurate quantitative thermodynamic properties are difficult to obtain. [Even in cases (c) and (d) above, where generalized but exactly soluble Ising models have almost always been employed so far, more refined models encounter difficulties in solution; this occurs, for instance in (c), when one allows for conformational twisting motions by including nonzero matrix elements between configurations representing different conformations.] In some cases there is difficulty even with the more naively simple models.

Here we describe a generalized transfer matrix method for computing the thermodynamic properties of such linear models. Our transfer matrix is developed in a series expansion such that the $(Q-1)$th term requires numerical (or analytic) results (generally in the form of both eigenvalues and correlation functions) for subchains with up to Q sites. When truncating this expansion at Q sites the present method yields free energies which are accurate through order $[Q/2]$ in terms of a perturbation expansion about *any* zero-order model of the generalized Ising type. Hence our generalized transfer matrix expansion truncates exactly at the first term for any generalized Ising model, and indeed it then takes the form of the well-known transfer matrix which has previously(1–4) been applied to such models. This relation is further emphasized if we note that in general the maximum eigenvalue of our generalized transfer matrix is the "per-site" partition function of an infinite chain. The relation of this generalized transfer matrix to generalized Ising models, its high-temperature convergence properties, and its systematic incorporation of finite subchain calculations suggest that this method is of use in obtaining accurate approximations to a variety of linear chain models.

The characterization of the problem and our fundamental approach is developed in Sections 2, 3, and 4. In Sections 5, 6, 7, 9, and 12 various aspects of the theory are further refined for use in different sorts of applications. Sections 8, 10, 11, and 12 present some illustrative examples. Finally some supplementary material is given in the Appendices.

2. LINEAR CHAIN MODELS

We consider models which are defined on a space with a basis of product kets

$$|r(N)\rangle \equiv |r_1 \cdot r_2 \cdot \cdots \cdot r_N\rangle, \qquad r(N) \equiv r_1 r_2 \cdots r_N \text{ ranging} \tag{2.1}$$

where r_i is a site-state label for site i. Our models are to be defined in terms of operators such as X_i^{st} which shifts the site states on site i

$$X_i^{st}|r(N)\rangle \equiv \delta_{t,r_i}|r_1 \cdot r_2 \cdot \cdots \cdot r_{i-1} \cdot s \cdot r_{i+1} \cdot \cdots \cdot r_N\rangle \tag{2.2}$$

so that

$$\begin{gathered} \left[X_i^{rs}, X_j^{tu}\right] = 0, \qquad i \neq j \\ X_i^{rs} X_i^{tu} = \delta_{st} X_i^{ru} \\ \sum_r X_i^{rr} = 1 \end{gathered} \tag{2.3}$$

Our linear chain Hamiltonian with nearest-neighbor interactions only is

$$H \equiv H_{1 \to N} \equiv \sum_{i=1}^{N-1} H_{ii+1} \tag{2.4}$$

where the H_{ii+1} affect only sites i and $i+1$, and hence are of the form

$$H_{ii+1} = \sum_{rstu} \langle r \cdot s|H_{ii+1}|t \cdot u\rangle X_i^{rt} X_{i+1}^{su} \tag{2.5}$$

Models satisfying this description include the nearest-neighbor linear Heisenberg, Ising, and lattice-gas models. For instance, for the isotropic spin-1/2 Heisenberg model the r_i are α and β spin states and the Hamiltonian may be represented in terms of X-operators as indicated in Section 10. Single-site terms, such as H_i acting only on the site states of site i, may be included in H_{i-1i} or H_{ii+1} or divided up between H_{i-1i} and H_{ii+1}. In Appendix A we also show that second quantized Fermion models can be easily expressed so as to satisfy the conditions of Eq. (2.1) to (2.5). Hence the allowed models also include a variety of Hubbard, PPP, and multiband electronic models.

Clearly the conditions upon H imply

$$\left[H_{ii+1}, H_{jj+1}\right] = 0, \qquad i \neq j \pm 1 \tag{2.6}$$

If (2.6) also holds for $i = j \pm 1$, so that all H_{ii+1} commute, then H is a *generalized Ising* (or lattice-gas) model. The common *Ising-type* models, for which there exists a site-state basis with

$$\langle r \cdot s|H_{ii+1}|t \cdot u\rangle \sim \delta_{rt}\delta_{su} \tag{2.7}$$

are clearly examples of such generalized Ising models. It is well known that the usual(1-4) transfer matrix method solves these Ising-type models exactly.

In the following we shall use the notation

$$H_{m\to n} \equiv \sum_{i=m}^{n-1} H_{ii+1}, \qquad \rho_{m\to n} \equiv e^{-\beta H_{m\to n}}, \qquad m < n \tag{2.8}$$

Further the (unnormalized) density operator $\rho_{1\to N}$ will often be abbreviated to ρ.

3. CLUSTER EXPANSION OF DENSITY OPERATORS

If H is an Ising-type model, then the first step in deriving the transfer matrix (between sites 1 and 2) is to factor ρ to $\rho_{12}\rho_{2\to N}$. However, for a more general H this factorization requires a correction. Most simply this correction might involve a factorization with the first term involving H_{12} and H_{23} while the second term would involve just $H_{3\to N}$. Generally again such a simple correction would itself require an additional "higher" correction, which in turn one might try to factor. Such ideas are suggestive and motivate us to write

$$\rho \equiv \sum_{i=2}^{N-1} \xi_{1\to i}\rho_{i\to N} \tag{3.1}$$

where $\xi_{1\to i}$ is a factor affecting only sites 1 to i. For chain lengths $N = 2, 3, 4, \ldots$ (3.1) yields

$$\begin{aligned} \rho_{12} &= \xi_{12} \\ \rho_{1\to 3} &= \xi_{12}\rho_{23} + \xi_{1\to 3} \\ \rho_{1\to 4} &= \xi_{12}\rho_{2\to 4} + \xi_{1\to 3}\xi_{34} + \xi_{1\to 4} \\ &\text{etc.} \end{aligned} \tag{3.2}$$

and we see that the ξ-factors are consistently and uniquely determined in an iterative manner. Inversion of these equations readily leads to

$$\begin{aligned} \xi_{12} &= \rho_{12} \\ \xi_{1\to 3} &= \rho_{1\to 3} - \rho_{12}\rho_{23} \\ \xi_{1\to 4} &= \rho_{1\to 4} - \rho_{12}\rho_{2\to 4} - \rho_{1\to 3}\rho_{34} + \rho_{12}\rho_{23}\rho_{34} \end{aligned} \tag{3.3}$$

etc., or generally

$$\xi_{1\to n+1} = \rho_{1\to n+1} + \sum_{l=1}^{n-1} (-1)^l \sum_{i_1 < i_2 < \cdots < i_l} \rho_{1\to i_1}\rho_{i_1\to i_2} \cdots \rho_{i_l\to n+1} \tag{3.4}$$

where of course $1 < i_1$ and $i_l < n + 1$. Further we may express the various

ρ's exclusively in terms of the ξ's

$$\begin{aligned}\rho_{12} &= \xi_{12}\\ \rho_{1\to3} &= \xi_{12}\xi_{23} + \xi_{1\to3}\\ \rho_{1\to4} &= \xi_{12}\xi_{23}\xi_{34} + \xi_{12}\xi_{2\to4} + \xi_{1\to3}\xi_{34} + \xi_{1\to4}\end{aligned} \tag{3.5}$$

etc., or generally

$$\rho_{1\to n+1} = \xi_{1\to n+1} + \sum_{l=1}^{n-1} \sum_{i_1<i_2<\cdots<i_l} \xi_{1\to i_1}\xi_{i_1\to i_2}\cdots\xi_{i_l\to n+1} \tag{3.6}$$

where again $i < i_1$ and $i_l < n + 1$. Clearly there is a duality in equations of this section such that $+\rho$ and $-\xi$ may be interchanged. Thus (3.4) and (3.6) are duals, (3.3) and (3.5) are duals, and

$$\xi_{1\to N} = \rho_{1\to N} - \sum_{i=2}^{N-1} \rho_{1\to i}\xi_{i\to N} \tag{3.7}$$

is the dual to (3.1).

In an application we could use (3.3) and (3.4) to determine the ξ's up through some order, say those involving up to Q sites, and then approximate ρ via (3.6) omitting all terms with a ξ involving more than Q sites. We call this a *Q-site decoupling* approximation, and in Appendix B two theorems indicating the accuracy of such approximations are indicated. In particular it is shown that this approximation is accurate to order $Q - 1$ in the high-temperature expansion in powers of β. We also show that if H is expressed as a zero-order generalized Ising model H^0 plus a perturbation λV, then this approximation is accurate through order $[(Q - 1)/2]$ in the perturbation expansion in powers of λ. Since this result is true regardless of the choice of generalized Ising model H^0, it is true of the resolution in which H^0 is "as much like" H as possible.

4. GENERALIZED TRANSFER MATRIX

Here we develop a computationally practical expansion of the partition function

$$Z = \mathrm{Tr}\,\rho = \mathrm{Tr}\,e^{-\beta H} \tag{4.1}$$

through the use of the Q-site decoupling approximation of Section 3. In this development we define

$$\begin{aligned} z_i &\equiv \mathrm{Tr}_{i\to N}\rho_{i\to N}/\mathrm{Tr}_{i+1\to N}\rho_{i+1\to N}\\ c_i^{rs} &\equiv \mathrm{Tr}_{i\to N}(X_i^{sr}\rho_{i\to N})/\mathrm{Tr}_{i\to N}\rho_{i\to N}\end{aligned} \tag{4.2}$$

where $\mathrm{Tr}_{i\to j}$ indicates a trace just over site states for sites $i\to j$, with $i<j$.

Here c_i^{rs} is an expectation value for an end site on a chain of length $N - i + 1$, and z_i is the *per-site* partition function, such that

$$Z = \prod_{i=1}^{N-1} z_i \tag{4.3}$$

Now using the cluster expansion of (3.1) in our definition of c_i^{rs}, we obtain

$$c_i^{rs} = \frac{\mathrm{Tr}_{i\to N}(X_i^{sr}\xi_{ii+1}\rho_{i+1\to N})}{z_i \mathrm{Tr}_{i+1\to N}\rho_{i+1\to N}} + \frac{\mathrm{Tr}_{i\to N}(X_i^{sr}\xi_{i\to i+2}\rho_{i+2\to N})}{z_i z_{i+1}\mathrm{Tr}_{i+2\to N}\rho_{i+2\to N}} + \cdots \tag{4.4}$$

Next we recall the sum of (2.3) and note that

$$\begin{aligned} c_i^{rs} &= \sum_{tu}\left\{ \frac{\mathrm{Tr}_{i\to N}(X_{i+1}^{uu}X_i^{sr}\xi_{ii+1}X_{i+1}^{tt}\rho_{i+1\to N})}{z_i\mathrm{Tr}_{i+1\to N}\rho_{i+1\to N}} \right. \\ &\qquad \left. + \frac{\mathrm{Tr}_{i\to N}(X_{i+2}^{uu}X_i^{sr}\xi_{i\to i+2}X_{i+2}^{tt}\rho_{i+2\to N})}{z_i z_{i+1}\mathrm{Tr}_{i+2\to N}\rho_{i+2\to N}} + \cdots \right\} \\ &= \sum_{tu}\left\{ \frac{1}{z_i}\mathrm{Tr}_{ii+1}(X_i^{sr}\xi_{ii+1}X_{i+1}^{tu})c_{i+1}^{tu} \right. \\ &\qquad \left. + \frac{1}{z_i z_{i+1}}\mathrm{Tr}_{i\to i+2}(X_i^{sr}\xi_{i\to i+2}X_{i+2}^{tu})c_{i+2}^{tu} + \cdots \right\} \end{aligned} \tag{4.5}$$

This is a recursion formula expressing an expectation value c_i^{rs} for a general chain of length $N - i + 1$ in terms of those for shorter chains.

Now we consider an infinite linear chain with translationally equivalent sites and bonds. In this case we may identify limiting values

$$c^{rs} \equiv \lim_{N-i\to\infty} c_i^{rs} \tag{4.6}$$

$$z \equiv \lim_{N\to\to\infty} z_i = \lim_{N\to\infty} Z^{1/N}$$

and we introduce a *generalized transfer matrix* T with (rs, tu)th element

$$\mathsf{T}^{rs,tu} \equiv \sum_{j\geqslant 2} \frac{1}{z^{j-2}} \mathsf{T}_{(j)}^{rs,tu} \tag{4.7}$$

where

$$\mathsf{T}_{(j)}^{rs,tu} \equiv \mathrm{Tr}_{1\to j}(X_1^{sr}\xi_{1\to j}X_j^{tu}) \tag{4.8}$$

Then (4.5) becomes

$$c^{rs} = \sum_{tu}\sum_{j\geqslant 2}\frac{1}{z^{j-1}}\mathsf{T}_{(j)}^{rs,tu}c^{tu} = \frac{1}{z}\sum_{tu}\mathsf{T}^{rs,tu}c^{tu} \tag{4.9}$$

or

$$\mathsf{T}\mathbf{c} = z\mathbf{c} \tag{4.10}$$

where $\mathbf{c}$ is simply the column vector with (rs)th element c^{rs}. Thus the per-site partition function z and the free energy per site

$$f \equiv -\frac{1}{\beta} \lim_{N\to\infty} \frac{1}{N} \ln Z = -\frac{1}{\beta} \ln z \tag{4.11}$$

are obtained via an eigenvalue to our generalized transfer matrix T. The *Q-site decoupling* approximation for T and z is simply implemented by truncating the j-sum of (4.9) beyond $j = Q$.

The components $\mathsf{T}_{(n)}$ of the transfer matrix T are conveniently computed in terms of "correlation" matrices $\mathsf{W}_{(j)}$ with (rs, tu)th element which is a correlation function

$$\mathsf{W}_{(j)}^{rs,tu} \equiv \mathrm{Tr}_{1\to j} X_1^{sr} \rho_{1\to j} X_j^{tu} \tag{4.12}$$

Such unnormalized end-to-end correlation functions are, of course, obtainable for finite chains via standard computational procedures. The desired relation of the $\mathsf{T}_{(n)}$ now follows from (3.7),

$$\mathsf{T}_{(n)} = \mathsf{W}_{(n)} - \sum_{i=1}^{n-2} \mathsf{W}_{(n-i)} \mathsf{T}_{(i+1)}, \qquad n \geqslant 3$$

$$\mathsf{T}_{(2)} = \mathsf{W}_{(2)} \tag{4.13}$$

so that the $\mathsf{T}_{(n)}$ can be built up in a recursive manner.

In order to recognize the desired eigenvalue z to $\mathbf{T}$ one could start at high temperatures, where (as shown in Appendix C) $\mathbf{T}$ becomes independent of z and z is the maximum eigenvalue, which is nondegenerate. Then one could follow this maximum eigenvalue down in temperature, using the z value of the preceding slightly higher temperature as an initial estimate for a self-consistizing iteration scheme for z and $\mathbf{T}$ (since $\mathbf{T}$ itself generally is a function of z). The only ambiguity which might arise in this procedure is the instance in which some other eigenvalue of $\mathbf{T}$ would come up and cross over the one we are following down in temperature. However, this never occurs for the $\mathbf{T}$ matrix of an Ising-type model, since in that case (as shown in Appendix C also) $\mathbf{T}$ essentially reduces to the ordinary type$^{(1-4)}$ of transfer matrix. In the event that the maximum eigenvalue is involved in a crossing, we realize that an (infinitesimally) small perturbation of H should make the crossing forbidden so that the maximum eigenvalue should still be followed after the crossing. Consequently we see that such a crossing would imply a phase transition, which, however, is a phenomenon generally believed not to occur in short-range one-dimensional systems.

5. SUPER TRANSFER MATRIX

In Section 4 the determination of the per-site partition function is presented in a nonlinear framework, since T itself generally depends upon z. However the problem, at least within the Q-site decoupling approximation, may be cast in an entirely linear form via the *super* transfer matrix $\mathfrak{T}$ composed from $(Q-1)^2$ submatrices with the (ij)th submatrix being

$$\mathfrak{T}_{ij} \equiv \begin{cases} \mathsf{T}_{(j)}, & i=2 \\ \mathsf{1}, & i=j+1 \\ \mathsf{0}, & \text{otherwise} \end{cases} \tag{5.1}$$

where the row and column labels $i, j \in \{2, 3, \ldots, Q\}$. Now we consider the eigenvalue problem

$$\mathfrak{T}\mathbf{C} = z\mathbf{C} \tag{5.2}$$

and let $\mathbf{C}_{(2)}, \mathbf{C}_{(3)}, \ldots, \mathbf{C}_{(Q)}$ be the portions of the column eigenvector $\mathbf{C}$ which are associated with the blocking of $\mathfrak{T}$ into submatrices. Then (5.2) is equivalent to the set of equations

$$\begin{gathered} \mathsf{T}_{(2)}\mathbf{C}_{(2)} + \mathsf{T}_{(3)}\mathbf{C}_{(3)} + \cdots + \mathsf{T}_{(Q)}\mathbf{C}_{(Q)} = z\mathbf{C}_{(2)} \\ \mathbf{C}_{(j)} = z\mathbf{C}_{(j+1)}, \qquad j = 2 \text{ to } Q-1 \end{gathered} \tag{5.3}$$

These last $Q-2$ equations are invertible to

$$\mathbf{C}_{(j+1)} = \frac{1}{z^{j-1}}\mathbf{C}_{(2)}, \qquad j = 2 \text{ to } Q-1 \tag{5.4}$$

so that on substituting into the first of the equations of (5.3), we obtain

$$\left\{\mathsf{T}_{(2)} + \frac{1}{z}\mathsf{T}_{(3)} + \cdots + \frac{1}{z^{Q-2}}\mathsf{T}_{(Q)}\right\}\mathbf{C}_{(2)} = z\mathbf{C}_{(2)} \tag{5.5}$$

Since this is recognized to be identical to (4.10), we see that (5.2) is indeed a linear formulation yielding the desired per-site partition function, as an eigenvalue. Further since at high temperatures the $\mathsf{T}_{(i)}$, $i \geqslant 3$, asymptotically vanish, the desired z is the maximum eigenvalue at least in this region. Finally $\mathbf{C}_{(2)}$ may be identified as $\mathbf{c}$, if the correct normalization

$$\sum_r C_{(2)}^{rr} = \sum_r c^{rr} = 1 \tag{5.6}$$

is chosen.

6. THERMODYNAMIC PROPERTIES

Here we consider the thermodynamic properties which are obtained as first and second derivatives of the free energy. Hence we wish to compute derivatives of z, and since z is an eigenvalue to $\mathfrak{T}$, we may do so using a rather straightforward modification (as described in Appendix D) of the standard perturbation theory. Basically, the non-Hermiticity of $\mathfrak{T}$ entails the introduction of a left (row) eigenvector $\mathbf{A}^\dagger$ with the same eigenvalue as the right (column) eigenvector $\mathbf{C}$,

$$\mathbf{A}^\dagger\mathfrak{T} = z\mathbf{A}^\dagger \tag{6.1}$$

Then letting $\mathbf{A}_{(2)}, \mathbf{A}_{(3)}, \ldots, \mathbf{A}_{(Q)}$ be the portions of $\mathbf{A}$ associated with the blocking of $\mathfrak{T}$ into submatrices, we may follow the procedure of (5.5)–(5.7) to show that

$$\mathbf{A}^\dagger_{(j)} = \frac{1}{z}\mathbf{A}^\dagger_{(2)}\left\{\sum_{i=j}^{Q}\frac{1}{z^{i-j}}\mathsf{T}_{(i)}\right\} \tag{6.2}$$

and that $\mathbf{A}_{(2)}$ is an eigenvector to T with eigenvalue z. Now, presuming H_{ii+1} is symmetric with respect to interchange of site labels i and $i+1$, so that T is Hermitian, it follows that $\mathbf{A}_{(2)}$ is the same as $\mathbf{C}_{(2)}$ except possibly for normalization. This we specify via

$$\mathbf{A}^\dagger\mathbf{C} = 1 \tag{6.3}$$

and hence, via (5.4) and (6.2),

$$\mathbf{A}_{(2)} = \left\{\sum_{i=2}^{Q}\frac{i-1}{z^{i-1}}\mathbf{C}^\dagger_{(2)}\mathsf{T}_{(i)}\mathbf{C}_{(2)}\right\}^{-1}\mathbf{C}_{(2)} \tag{6.4}$$

First-order perturbation formulas now give

$$\frac{\partial z}{\partial\lambda} = \mathbf{A}^\dagger\frac{\partial\mathfrak{T}}{\partial\lambda}\mathbf{C} = \sum_{j=2}^{Q}\mathbf{A}^\dagger_{(2)}\frac{\partial\mathsf{T}_{(j)}}{\partial\lambda}\mathbf{C}_{(j)} \tag{6.5}$$

$$\frac{\partial^2 z}{\partial\lambda^2} = \sum_{j=2}^{Q}\left\{\mathbf{A}^\dagger_{(2)}\frac{\partial^2\mathsf{T}_{(j)}}{\partial\lambda^2}\mathbf{C}_{(j)} + 2\mathbf{A}^\dagger_{(2)}\frac{\partial\mathsf{T}_{(j)}}{\partial\lambda}\mathfrak{R}_{j,2}\sum_{i=2}^{Q}\frac{\partial\mathsf{T}_{(i)}}{\partial\lambda}\mathbf{C}_{(i)}\right\} \tag{6.6}$$

where we require the submatrices $\mathfrak{R}_{j,2}$ from the first block column of the generalized inverse $\mathfrak{R}$ of $z-\mathfrak{T}$, as indicated in Appendix D. A formula for these $\mathfrak{R}_{j,2}$ is given in Appendix E. Finally the derivatives of the transfer matrix components $\mathsf{T}_{(j)}$ are calculated using (4.13) from the derivatives of finite chain correlation functions $\mathsf{W}_{(i)}$. For instance,

$$\frac{\partial\mathsf{T}_{(j)}}{\partial\lambda} = \frac{\partial\mathsf{W}_{(j)}}{\partial\lambda} - \sum_{i=1}^{j-2}\left\{\left(\frac{\partial\mathsf{W}_{(j-i)}}{\partial\lambda}\right)\mathsf{T}_{(i+1)} + \mathsf{W}_{(j-i)}\left(\frac{\partial\mathsf{T}_{(i+1)}}{\partial\lambda}\right)\right\} \tag{6.7}$$

The $(\partial\mathsf{W}_{(i)}/\partial\lambda)$, involving derivatives of the $\rho_{1\to i}$, presumably are computable via conventional procedures, as in the case when λ is the temperature or magnetic field, of use in computing internal energies or susceptibilities.

7. LOCAL SYMMETRIES

Here we investigate the consequences arising from *local* (or site) *symmetries*, for site i, with operators

$$\Lambda_i \equiv \sum_{rs} \langle r|\Lambda_i|s\rangle X_i^{rs} \tag{7.1}$$

for which

$$[H_{i-1i}, \Lambda_i] = [H_{ii+1}, \Lambda_i] = 0 \tag{7.2}$$

If the $H_{j\,j+1}$ are Hermitean, it follows that Λ_i also commutes with H_{i-1i} and H_{ii+1} and that the algebra $\mathcal{A}_i$ of local symmetry operators for site i is self-adjoint. Consequently[4] $\mathcal{A}_i$ is "semisimple" and has a (matric) basis of operators

$$\{\Lambda_{i\alpha}^{ab};\quad \alpha \text{ ranging},\quad a,b=1 \quad \text{to} f^{\alpha}\} \tag{7.3}$$

such that

$$\begin{aligned} &\Lambda_{i\alpha}^{ab}\Lambda_{i\beta}^{cd} = \delta_{\alpha\beta}\delta_{bc}\Lambda_{i\alpha}^{ad} \\ &(\Lambda_{i\alpha}^{ab})^{\dagger} = \Lambda_{i\alpha}^{ba} \\ &\sum_{\alpha a}\Lambda_{i\alpha}^{aa} = 1 \end{aligned} \tag{7.4}$$

Since the $\Lambda_{i\alpha}^{aa}$ mutually commute and are orthogonal projection operators, there is a basis of site states for site i such that

$$\Lambda_{i\alpha}^{aa} = \sum_{r}^{\alpha a} X_i^{rr} \tag{7.5}$$

where the r-sum is over a set, say $\mathbb{S}(\alpha a)$, of site-state indices identified with αa. Clearly these $\mathbb{S}(\alpha a)$ are disjoint amd their union yields the whole basis of site states. Further there is a one-to-one correspondence between the elements of $\mathbb{S}(\alpha a)$ and $\mathbb{S}(\alpha b)$, say $r_a \in \mathbb{S}(\alpha a)$ corresponds to $r_b \in \mathbb{S}(\alpha b)$, such that

$$\Lambda_{i\alpha}^{ab} = \sum_{r_a}^{\alpha a} X_i^{r_a r_b} \tag{7.6}$$

Now realizing that the algebra of local symmetry operators has the above-mentioned features, we see that for $r \in \mathbb{S}(\alpha a)$

$$\begin{aligned} \langle r\cdot t|H_{ii+1}|s\cdot u\rangle &= \langle r\cdot t|\Lambda_{i\alpha}^{aa}H_{ii+1}|s\cdot u\rangle \\ &= \langle r\cdot t|H_{ii+1}\Lambda_{i\alpha}^{aa}|s\cdot u\rangle = 0, \qquad s \notin \mathbb{S}(\alpha a) \end{aligned} \tag{7.7}$$

---

[4] Although this is implicit in the theory of finite C^* algebras, the results quoted here are given in the present language in Theorems IV.1 and IV.2 of Ref. 6.

Further if $r_b, s_b \in \mathbb{S}(\alpha a)$ correspond to $r_b, s_b \in \mathbb{S}(\alpha b)$, then

$$\begin{aligned}\langle r_a \cdot t | H_{ii+1} | s_a \cdot u\rangle &= \langle r_a \cdot t | \Lambda_{i\alpha}^{aa} H_{ii+1} | s_a \cdot u\rangle \\ &= \langle r_a \cdot t | \Lambda_{i\alpha}^{ab} \Lambda_{i\alpha}^{ba} H_{ii+1} | s_a \cdot u\rangle \\ &= \langle r_a \cdot t | \left(\Lambda_{i\alpha}^{ba}\right)^{\dagger} H_{ii+1} \Lambda_{i\alpha}^{ba} | s_a \cdot u\rangle = \langle r_b \cdot t | H_{ii+1} | s_b \cdot u\rangle \end{aligned} \tag{7.8}$$

Here (7.7) says H_{ii+1} is block-diagonalized with each block labeled by an αa, while (7.8) identifies different blocks with the same α label as having identical matrix elements. Of course the local symmetry operators for site $i+1$ provide additional block diagonalization and identification. Indeed with site symmetries for all sites, all the $H_{j\,j+1}$ can be blocked up as well as functions (such as the ρ's, W's, and T's) of them.

If all the $\mathbb{S}(\alpha a)$ have just a single site state, then H is an Ising-type model. Moreover, some important simplifications arise if any of the $\Lambda_{i\alpha}^{aa}$ are *primitive* in the sense that $\mathbb{S}(\alpha a)$ has just one site state. For instance, the product ket of (3.1) is an eigenket to H if each r_i is in such a single-member $\mathbb{S}(\alpha a)$. Also

$$\langle r \cdot r_2 \cdot r_3 \cdot \cdots \cdot r_j \cdot u | \xi_{1\to j+1} | s \cdot r_2 \cdot r_3 \cdot \cdots \cdot r_j \cdot t\rangle = 0, \qquad j \geqslant 2 \tag{7.9}$$

if any r_i, $i = 2$ to j, is of a primitive symmetry. This result is seen to hold on using the lemma of Appendix B and noting that the portions of H_{i-1i} and H_{ii+1} surviving in this matrix element of (7.9) commute with one another. As a consequence of (7.9) the computation of the three- and more-site components of T are simplified,

$$\mathsf{T}_{(j+1)}^{rs,\,tu} = \sum_{r_2, r_3, \ldots, r_j}{}' \langle r \cdot r_2 \cdot r_3 \cdot \cdots \cdot r_j \cdot u | \xi_{1\to j+1} | s \cdot r_2 \cdot r_3 \cdot \cdots \cdot r_j \cdot t\rangle \tag{7.10}$$

where the sum here is restricted to go only over those site states of nonprimitive symmetries. If also the ground state is of primitive symmetry, then the generalized transfer matrix approach displays low-temperature convergence properties; an example of this is described in the following section.

8. SOME EXAMPLES

One example is the Heisenberg–Ising model,$^{(7)}$ in which each site consists of a pair of spin-1/2 atoms coupled together by a Heisenberg interaction while there is a classical spin-1/2 Ising-type interaction between the nearer atoms of an adjacent pair of sites. Taking the four site states to be a singlet s and three components $+$, 0, $-$ of a triplet, we obtain the

Hamiltonian in the form

$$H = J \sum_{i=1}^{N} \left(X_i^{++} + X_i^{00} + X_i^{--} - X_i^{ss}\right)$$
$$+ \Delta \sum_{i=1}^{N-1} \left\{(X_i^{++} - X_i^{--})(X_{i+1}^{++} - X_{i+1}^{--}) - \left(X_i^{0s} + X_i^{s0}\right)\left(X_{i+1}^{0s} + X_{i+1}^{s0}\right)\right\} \tag{8.1}$$

In this case the local symmetry algebra $\mathcal{A}_i$ for each site i is of dimension 3 with three orthogonal projection operators

$$\Lambda_{i+} = X_i^{++}, \qquad \Lambda_{i-} = X_i^{--}, \quad \text{and} \quad \Lambda_{i0} = X_i^{00} + X_i^{ss} \tag{8.2}$$

The three associated site-state index sets $\mathbb{S}(\alpha a)$, abbreviated to $\mathbb{S}(\alpha)$ since $f^\alpha = 1$, are

$$\mathbb{S}(+) = \{+\}, \qquad \mathbb{S}(-) = \{-\}, \quad \text{and} \quad \mathbb{S}(0) = \{0, s\} \tag{8.3}$$

Since there is only one index set $\mathbb{S}(0)$ here with more than one index, the computation of the $\mathsf{T}_{(j)}$ and $\mathsf{W}_{(j)}$ reduces to a relatively simply problem with just two states per site (except possibly the end sites of the finite subchain). Indeed this latter problem is solved exactly,$^{(7)}$ so that z can be obtained exactly by our present method.

Another model of a similar nature is described by Mattis and Nam.$^{(8)}$ Indeed their transformed Hamiltonian (in terms of their P's and J's) is readily seen to yield a three-dimensional local symmetry algebra much like that of (8.2) here. Again the relevant two-state per site model has been solved, so that z can be obtained exactly.

A third model we mention is that of Zawadowski and Cohen,$^{(9)}$ where again each site consists of two "atoms," between which there is shared a single unpaired electron. The electron may be transferred between the two atoms of a site, while two electrons on two adjacent atoms of two adjacent sites undergo Coulomb repulsion and exchange coupling. Thus, denoting the four site states by $a\alpha$, $a\beta$, $b\alpha$, $b\beta$ depending on whether the a or b atom of a site is occupied by an α or β spin electron, we obtain the Hamiltonian as

$$H = T \sum_{i=1}^{N} \sum_{\sigma} \left(X_i^{a\sigma, b\sigma} + X_i^{b\sigma, a\sigma}\right)$$
$$+ U_1 \sum_{i=1}^{N-1} \sum_{\sigma,\tau} X_i^{b\sigma, b\sigma} X_{i+1}^{a\tau, a\tau} + J \sum_{i=1}^{N-1} \sum_{\sigma,\tau} X_i^{b\sigma, b\tau} X_{i+1}^{a\tau, a\sigma} \tag{8.4}$$

Here there does not seem to be any local symmetry. But it is of interest to note that if $T = 0$, we obtain a simple example (which in our present nomenclature) is a generalized Ising model although not an Ising-type model.

9. GLOBAL SYMMETRIES

In this section we investigate the consequences of a *global* symmetry group $\mathcal{G}$ such that

$$\begin{aligned} &\left[G_{i\to j}, H_{i\to j} \right] \equiv 0 \\ &G_{i\to j} \equiv G_i G_{i+1} \cdots G_j \in \mathcal{G}_{i\to j} \\ &G_k \equiv \sum_{rs} \Gamma_{rs}(G) X_k^{rs} \end{aligned} \tag{9.1}$$

with the matrices $\Gamma(G)$ being unitary. Further we presume that the ordinary group orthogonality properties for the irreducible representations of $\mathcal{G}$ apply

$$\frac{1}{v} \underset{G\in\mathcal{G}}{\mathsf{S}} \Gamma^{\alpha}_{a\alpha}(G) \Gamma^{\beta}_{b'b}(G^{-1}) = \frac{1}{f^{\alpha}} \delta_{\alpha\beta} \delta_{ab} \delta_{a'b'} \tag{9.2}$$

where $\Gamma^{\alpha}(G)$ is the αth irreducible representation matrix of $G \in \mathcal{G}$, f^{α} is the dimension of α, S indicates the appropriate sum and/or integration over group elements, and v is the group volume. The group $\mathcal{G}$ might be realized as a Lie group obtained from a Lie algebra with generators Λ such that

$$\begin{aligned} &\left[\Lambda_{i\to j}, H_{i\to j} \right] = 0 \\ &\Lambda_{i\to j} = \Lambda_i + \Lambda_{i+1} + \cdots + \Lambda_j \\ &\Lambda_k = \sum_{rs} \Gamma_{rs}(\Lambda) X_k^{rs} \end{aligned} \tag{9.3}$$

or $\mathcal{G}$ might represent a discrete set of rotations about the chain axis of our linear model. Finally we also define the matrices G with (rs, tu)th element

$$\mathsf{G}^{rs,tu} \equiv \Gamma_{rt}(G) \Gamma_{us}(G^{-1}), \quad G \in \mathcal{G} \tag{9.4}$$

The set $\mathcal{G}$ of these matrices is readily verified to be a homomorphic image of (i.e., representation) $\mathcal{G}$.

Now

$$\begin{aligned} \sum_{tu} \mathsf{G}^{rs,tu} \mathsf{T}^{tu,vw}_{(j)} &= \sum_{tu} \Gamma_{rt}(G) \Gamma_{us}(G^{-1}) \mathrm{Tr}_{1\to j} X_1^{ut} \xi_{1\to j} X_j^{vw} \\ &= \mathrm{Tr}_{1\to j} G_1^{-1} X_1^{sr} G_1 \xi_{1\to j} X_j^{vw} \\ &= \mathrm{Tr}_{1\to j} G_{1\to j}^{-1} X_1^{sr} G_{1\to j} \xi_{1\to j} X_j^{vw} \\ &= \mathrm{Tr}_{1\to j} X_1^{sr} G_{1\to j} \xi_{1\to j} X_j^{vw} G_{1\to j}^{-1} \end{aligned} \tag{9.5}$$

But since $G_{1\to j}$ commutes with all $H_{i'\to j'}$ for $1 \leqslant i' < j' \leqslant j$, we see that $G_{1\to j}$ commutes with $\xi_{1\to j}$ and

$$\begin{aligned}\sum_{tu} \mathsf{G}^{rs,tu}\mathsf{T}^{tu,vw}_{(j)} &= \mathrm{Tr}_{1\to j} X_1^{sr}\xi_{1\to j} G_{1\to j} X_j^{vw} G_{1\to j}^{-1} \\ &= \mathrm{Tr}_{1\to j} X_1^{sr}\xi_{1\to j} G_j X_j^{vw} G_j^{-1} \\ &= \sum_{tu} \Gamma_{tv}(G)\Gamma_{wu}(G^{-1})\mathrm{Tr}_{1\to j} X_1^{sr}\xi_{1\to j} X_j^{tu} \\ &= \sum_{tu} \mathsf{T}^{rs,tu}_{(j)}\mathsf{G}^{tu,vw} \qquad (9.6)\end{aligned}$$

Consequently

$$[\mathsf{G},\mathsf{T}] = 0, \qquad \mathsf{G} \in \mathcal{G} \qquad (9.7)$$

a result which obviously persists even in the Q-site approximation for the transfer matrix. Thus T (and $\mathfrak{T}$ also) may be block-diagonalized with different blocks of T labeled by different irreducible representations of $\mathcal{G}$ (or $\mathcal{G}$). Now the transformation properties of the eigenvector $\mathbf{c}$ with respect to the $\mathsf{G} \in \mathcal{G}$ can be studied in a similar manner,

$$\begin{aligned}\sum_{tu} \mathsf{G}^{rs,tu} c^{tu} &= \lim_{N\to\infty} \sum_{tu} \frac{\Gamma_{rt}(G)\Gamma_{us}(G^{-1})\mathrm{Tr}_{1\to N} X_1^{ut}\rho_{1\to N}}{\mathrm{Tr}_{1\to N}\rho_{1\to N}} \\ &= \lim_{N\to\infty} \frac{\mathrm{Tr}_{1\to N} G_1^{-1} X_1^{sr} G_1 \rho_{1\to N}}{\mathrm{Tr}_{1\to N}\rho_{1\to N}} \\ &= \lim_{N\to\infty} \frac{\mathrm{Tr}_{1\to N} X_1^{sr} G_{1\to N}\rho_{1\to N} G_{1\to N}^{-1}}{\mathrm{Tr}_{1\to N}\rho_{1\to N}} \\ &= c^{rs} \qquad (9.8)\end{aligned}$$

Hence

$$\mathsf{G}\mathbf{c} = \mathbf{c}, \qquad \mathsf{G} \in \mathcal{G} \qquad (9.9)$$

and $\mathbf{c}$ is clearly associated with the block of T labeled by the identity irreducible representation.

This desired block of T is clearly identified as the nonzero block of

$$\frac{1}{v} \underset{G\in\mathcal{G}}{\mathsf{S}} \mathsf{GT} = \frac{1}{v} \underset{G\in\mathcal{G}}{\mathsf{S}} \mathsf{TG} \qquad (9.10)$$

An explicit formula for this nonzero block is obtained if we choose the site states for site i to be symmetry adapted to $\mathcal{G}_i$, so that a typical site state is $|i\rho\alpha a\rangle$ such that

$$\Gamma_{\rho\alpha a,\sigma\beta b}(G) = \delta_{\rho\sigma}\delta_{\alpha\beta}\Gamma^{\alpha}_{ab}(G) \qquad (9.11)$$

$$G|i\rho\alpha a\rangle = \sum_{a'} \Gamma^{\alpha}_{a'a}(G)|i\rho\alpha a'\rangle$$

Then the $(\rho\alpha a\rho'\alpha' a', vw)$ element of (9.10) is

$$\begin{aligned}
&\frac{1}{v} \underset{G\in\mathcal{G}}{\mathsf{S}} \sum_{\sigma\beta b} \sum_{\sigma'\beta' b'} \mathsf{G}^{\rho\alpha a\rho'\alpha' a',\, \sigma\beta b\sigma'\beta' b'} \mathsf{T}^{\sigma\beta b\sigma'\beta' b',\, vw} \\
&\quad = \frac{1}{v} \underset{G\in\mathcal{G}}{\mathsf{S}} \sum_{\sigma\beta b} \sum_{\sigma'\beta' b'} \delta_{\rho\sigma}\delta_{\alpha\beta}\Gamma^{\alpha}_{ab}(G)\delta_{\rho'\sigma'}\delta_{\alpha'\beta'}\Gamma^{\alpha'}_{a'b'}(G^{-1})\mathsf{T}^{\sigma\beta b\sigma'\beta' b',\, vw} \\
&\quad = \sum_{bb'} \frac{1}{v} \underset{G\in\mathcal{G}}{\mathsf{S}} \Gamma^{\alpha}_{ab}(G)\Gamma^{\alpha'}_{b'a'}(G^{-1})\mathsf{T}^{\rho\alpha b\rho'\alpha' b', vw} \\
&\quad = \delta_{\alpha\alpha'}\delta_{aa'} \frac{1}{f^{\alpha}} \sum_{b} \mathsf{T}^{\rho\alpha b\rho'\alpha b, vw}
\end{aligned} \tag{9.12}$$

But a similar result applies using the right-hand side of (9.10), so that the $(\rho\alpha a\rho'\alpha' a', \sigma\beta b\sigma'\beta' b')$ element is in fact just

$$\delta_{\alpha\alpha'}\delta_{aa'}\delta_{\beta\beta'}\delta_{bb'} \frac{1}{\left(f^{\alpha}f^{\beta}\right)^{1/2}} \hat{\mathsf{T}}^{\rho\alpha\rho', \sigma\beta\sigma'} \tag{9.13}$$

where

$$\hat{\mathsf{T}}^{\rho\alpha\rho', \sigma\beta\sigma'} \equiv \frac{1}{\left(f^{\alpha}f^{\beta}\right)^{1/2}} \sum_{cd} \mathsf{T}^{\rho\alpha c\rho'\alpha c, \sigma\beta d\sigma'\beta d} \tag{9.14}$$

Now the (invariant) space of vectors upon which the matrix with elements as in (9.13) acts to give a generally nonzero result consists of those vectors $\mathbf{v}$ with components

$$v^{\sigma\beta b, \sigma'\beta' b'} = \delta_{\beta\beta'}\delta_{bb'}\hat{v}^{\sigma\beta\sigma'} \tag{9.15}$$

Hence the identity irreducible representation block of T is just the matrix $\hat{\mathsf{T}}$ of (9.14), acting on a space of vectors $\hat{\mathbf{v}}$ with components $\hat{v}^{\sigma\beta\sigma'}$. Often $\hat{\mathsf{T}}$ yields a much smaller matrix to diagonalize. Finally by appending a subscript (j) in (9.14) we can define matrices $\hat{\mathsf{T}}_{(j)}$, and by appending carats in (5.1) obtain the block $\hat{\mathfrak{T}}$ of $\mathfrak{T}$ which is of present interest.

10. ISOTROPIC HEISENBERG MODELS

In this section we consider the spin-s Heisenberg model

$$H \equiv \frac{J}{2s^2} \sum_{i=1}^{N-1} \left\{ \frac{1}{2}(s_i^+ s_{i+1}^- + s_i^- s_{i+1}^+) + s_i^z s_{i+1}^z \right\} \tag{10.1}$$

for a linear chain. Here the site states for a site i are labeled by their eigenvalues $m = -s, -s+1, \ldots, +s$ to the s_i^z operator. If we wish we can express the s_i^{μ} appearing in H in terms of our X-operators via

$$s_i^{\mu} = \sum_{m=-s}^{+s} \lambda_{\mu}(s,m) X_i^{m+\mu,\,m}$$

$$\lambda_{\mu}(s,m) \equiv \begin{cases} \left[s(s+1) - m(m+1)\right]^{1/2}, & \mu = + \\ m, & \mu = z \\ \left[s(s+1) - m(m-1)\right]^{1/2}, & \mu = - \end{cases} \tag{10.2}$$

For this model H commutes with the global symmetry operators s^{+}, s^{z}, and s^{-} which are generators for the Lie algebra of the familiar $\mathcal{SU}(2)$ group.

Since all the site states of a given site belong to a simple irreducible representation of $\mathcal{SU}(2)$, the desired block $\hat{\mathsf{T}}$ of the transfer matrix is one dimensional, and

$$z = \hat{\mathsf{T}} = \frac{1}{2s+1} \sum_{mm'} \mathsf{T}^{mm,\,m'm'} \tag{10.3}$$

Now

$$\begin{aligned} \hat{\mathsf{T}}_{(j)} &\equiv \frac{1}{2s+1} \sum_{mm'} \hat{\mathsf{T}}_{(j)}^{mm,\,m'm'} \\ &= \frac{1}{2s+1} \sum_{mm'} \mathrm{Tr}_{1\to j} X_1^{mm} \xi_{1\to j} X_j^{m'm'} = \frac{1}{2s+1} \mathrm{Tr}_{1\to j} \xi_{1\to j} \end{aligned} \tag{10.4}$$

and in analogy to (4.12) and (4.13) we have

$$\begin{aligned} \hat{\mathsf{T}}_{(n)} &= \hat{\mathsf{W}}_{(n)} - \sum_{i=1}^{n-2} \hat{\mathsf{W}}_{(n-1)} \hat{\mathsf{T}}_{(i+1)}, \qquad n \geqslant 3 \\ \hat{\mathsf{T}}_{(2)} &= \hat{\mathsf{W}}_{(2)} \\ \hat{\mathsf{W}}_{(j)} &= \frac{1}{2s+1} \mathrm{Tr}_{1\to j} \rho_{1\to j} \end{aligned} \tag{10.5}$$

Thus for the isotropic Heisenberg model all we require to compute our eigenvalue z is the sequence of finite chain partition functions. We also note that the present treatment exactly solves the infinite-spin Heisenberg model since

$$\lim_{s\to\infty} \left[\frac{s_i^{\mu}}{s}, \frac{s_i^{\nu}}{s} \right] = 0 \tag{10.6}$$

The two-site Heisenberg model is readily solved, with the eigenkets being the states of total spin $s_{12} = 0, 1, \ldots, 2s$. Thus

$$\begin{aligned} \hat{\mathsf{T}}_{(2)} = \hat{\mathsf{W}}_{(2)} = \frac{1}{2s+1} \sum_{s_{12}=0}^{2s} (2s_{12}+1)\exp \\ - \frac{\beta J}{4s^2} \left\{ s_{12}(s_{12}+1) - 2s(s+1) \right\} \end{aligned} \tag{10.7}$$

giving the two-site approximation for z. For three sites the eigenkets to (10.1) are obtained by first coupling $\mathbf{s}_1$ and $\mathbf{s}_3$ to a spin $s_{13} = 0, 1, \ldots, 2s$, then second coupling $\mathbf{s}_{13}$ and $\mathbf{s}_2$ to a total spin $s_{123} = |s_{13} - s|, |s_{13} - s| + i, \ldots, s + s_{13}$. Thus

$$\mathsf{W}_{(3)} = \frac{1}{2s+1} \sum_{s_{13}=0}^{s} \sum_{s_{123}=|s_{13}-s|}^{s+s_{13}} (2s_{123}+1)\exp$$
$$- \frac{\beta J}{4s^2} \{ s_{123}(s_{123}+1) - s_{13}(s_{13}+1) - s(s+1) \} \qquad (10.8)$$

and the three-site approximation for the symmetry-reduced super transfer matrix is

$$\hat{\mathfrak{T}} = \begin{pmatrix} \hat{\mathsf{W}}_{(2)} & \hat{\mathsf{W}}_{(3)} - \hat{\mathsf{W}}_{(2)}^2 \\ 1 & 0 \end{pmatrix} \qquad (10.9)$$

and the three-site approximation for z, obtained as an eigenvalue to $\hat{\mathfrak{T}}$, is

$$z = \tfrac{1}{2}\hat{\mathsf{W}}_{(2)} + \left(\hat{\mathsf{W}}_{(3)} - \tfrac{3}{4}\hat{\mathsf{W}}_{(2)}^2 \right)^{1/2} \qquad (10.10)$$

To obtain the internal energy and specific heat in the three-site approximation one could differentiate (10.9) directly. In general, however, such closed form expressions for z are not readily available, in which case the formulas of Section 6 become useful. The application of these formulas to the three-site isotropic Heisenberg may be effected if we note that for $Q = 3$

$$\frac{dz}{\partial\lambda} = A_{(2)}C_{(2)} \left\{ \hat{\mathsf{T}}'_{(2)} + \frac{1}{z}\hat{\mathsf{T}}'_{(3)} \right\} \qquad (10.11)$$

$$\frac{d^2z}{\partial\lambda^2} = 2A_{(2)}C_{(2)} \left\{ \frac{1}{2}\left[\hat{\mathsf{T}}''_{(2)} + \frac{1}{z}\hat{\mathsf{T}}''_{(3)} \right] \right.$$
$$\left. + \left[\mathcal{R}_{22}\hat{\mathsf{T}}'_{(2)} + \mathcal{R}_{32}\hat{\mathsf{T}}'_{(3)} \right] \left[\hat{\mathsf{T}}'_{(2)} + \frac{1}{z}\hat{\mathsf{T}}'_{(3)} \right] \right\}$$

where all the primes indicate derivatives with respect to λ and we have noted that all matrices involved are one dimensional. By (6.4) we have

$$A_{(2)}C_{(2)} = \left(1 + \frac{1}{z^2}\hat{\mathsf{T}}_{(3)} \right)^{-1} \qquad (10.12)$$

and the resolvent matrix elements are given by the formulas in Appendix E as

$$\mathcal{R}_{22} = \frac{1}{z^3}\hat{\mathsf{T}}_{(3)} \left(1 + \frac{1}{z^2}\hat{\mathsf{T}}_{(3)} \right)^{-2} \qquad (10.13)$$

$$\mathcal{R}_{32} = \frac{1}{z}\mathcal{R}_{22} - \frac{1}{z^2}\left(1 + \frac{1}{z^2}\hat{\mathsf{T}}_{(3)} \right)^{-1}$$

In the present case the formulas for the derivatives of z may also be obtained by direct differentration of (10.10). If we identify λ as temperature the above formulas lead to the internal energy and specific heat. Magnetic properties (with λ the field strength) require the use of 2×2 transfer matrices $\mathsf{T}_{(j)}$.

11. SOME NUMERICAL RESULTS

The free energy per site F/JN was computed in the $Q = 2$ and $Q = 3$ approximations for the Heisenberg with antiferromagnetically signed $J > 0$ models of Section 10. Results for spins $s = 1/2, 1, 3/2, 2, 5/2$ are shown in Fig. 1. The ($Q = 3$)-site approximations are terminated at a temperature $1/\beta_c J$, where the approximation fails, as is evidenced by the radical in (10.10) becoming imaginary. This temperature for the failure of the $Q = 3$ approximation is plotted in Fig. 2 as a function of $1/s$, whence it is seen that the approximation avoids failure to lower temperatures as s increases.

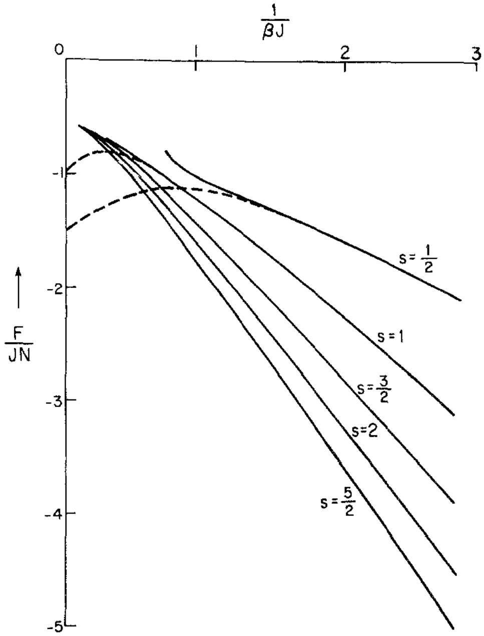

Fig. 1. The ($Q = 2$)-site approximations to the free energy for spin-s linear chain Heisenberg models are shown in the solid curves, and the ($Q = 3$)-site approximations (which are the same to within the accuracy of this plot at higher temperatures, and higher spins) are shown in dashed curves.

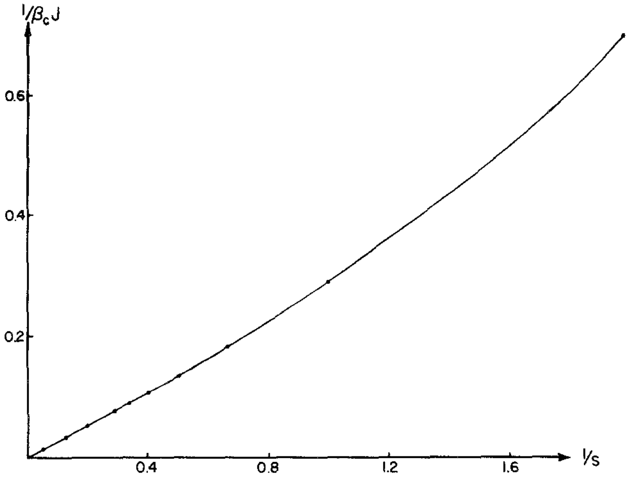

Fig. 2. The (reduced) temperature $1/\beta_c J$ at which the ($Q = 2$)-site approximation for the spin-s linear chain Heisenberg model fails.

Further the $Q = 2$ and $Q = 3$ approximations agree more and more closely to lower temperatures as s increases. Thus for the larger spins $s \geqslant 3/2$ it appears the $Q = 2$ and $Q = 3$ approximations are both rather accurate down to the temperatures given in Fig. 2.

For the $s = 1/2$ case we carried out numerical computations up through the ($Q = 9$)-site approximation. The reduced transfer matrix components $\hat{\mathsf{T}}_{(j)}$ converge rapidly at higher temperatures. But at lower temperatures (around $1/\beta J = 1$) the $\hat{\mathsf{T}}_{(j)}$ increase rapidly with j, numerical problems arise for larger j ($j \geqslant 9$ in our present program), and finally at a lower temperature the method fails, much as in the $Q = 3$ case. The failure temperature becomes lower as Q increases, though the shift is rather slow, so that for $Q = 8$ the failure temperature is $1/\beta_c J \cong 0.53$ as compared to $1/\beta_c J \cong 0.69$ for $Q = 3$. Evidently the longer-range (large-j) $\hat{\mathsf{T}}_{(j)}$ are rather important for this Heisenberg model at these lower temperatures. We note that although the $\hat{\mathsf{T}}_{(j)}$ become very large at lower temperatures, their ratios $\hat{\mathsf{T}}_{(j)}/\hat{\mathsf{T}}_{(j-1)}$ appear to converge better, so that one might attempt to extrapolate the higher values of $\hat{\mathsf{T}}_{(j)}$ and use them in an extended approximation.

The specific heat for the $s = 1/2$ model is shown in Fig. 3, where it is seen that it seems to converge (to within the accuracy of the plot) down to temperatures $1/J\beta \cong 0.66$, whence the different approximations split apart and soon thereafter diverge. For the higher temperatures, $1/\beta J \gtrsim 0.7$, the convergence appears to be very good, in comparison to straightforward finite chain calculations[10] of the specific heat. For instance, the maximum

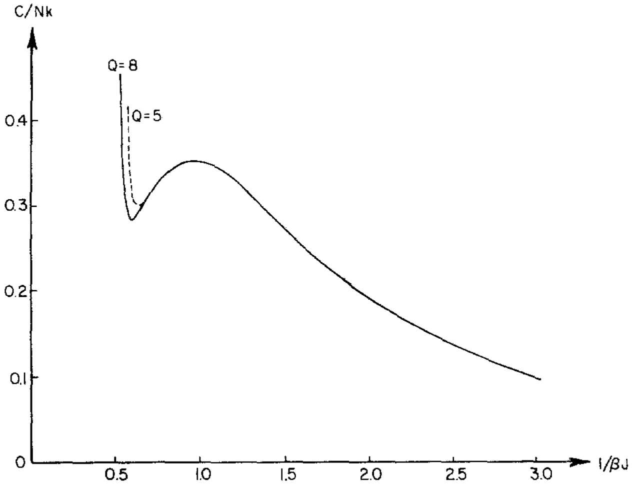

Fig. 3. Some higher Q-site approximations to the specific heat of the spin $s = 1/2$ linear chain Heisenberg model. The divergent artifact of the approximation appears to move rather slowly to lower temperatures as the "order" of the approximation is increased.

of the specific heat is located to be $C/Nk = 0.3499$ at $1/\beta J = 0.961$, where the error in C/Nk is estimated as 0.0003, the difference in the values computed in the $Q = 7$ and 8 approximations.

12. HEISENBERG MODELS FOR DIMERS

A type of Heisenberg model which is especially accurately soluble via the present techniques is that describing a collection[5] of weakly exchange-coupled dimers, the monomer units of which are more strongly exchange-coupled. Here we shall assume a linear chain of dimers composed from doublet monomer units. The dimers are identified as sites so that there are four site states: a singlet, labeled by S, and three triplets, labeled by their eigenvalues $+, 0, -$ to the s_i^z operator for that site. The Hamiltonian we consider then is

$$H = J\sum_{i=1}^{N}\left(2\mathbf{s}_{ia}\cdot\mathbf{s}_{ib}+\frac{1}{2}\right)+K\sum_{i=1}^{N}\sum_{c}^{a,b}\sum_{c'}^{a,b}2\mathbf{s}_{ic}\cdot\mathbf{s}_{i+1c'} \qquad (12.1)$$

where a, b label the two monomer units of a dimer. In terms of the X

---

[5] See, e.g., Refs. 11–15.

operators this is

$$H = \sum_{i=1}^{N-1} H_{ii+1}$$

$$\begin{aligned} H_{ii+1} = \frac{J}{2} \{ & X_i^{++} + X_i^{00} + X_i^{--} - X_i^{ss} + X_{i+1}^{++} + X_{i+1}^{00} + X_{i+1}^{--} - X_{i+1}^{ss} \} \\ & + 2K \{ (X_i^{+0} + X_i^{0-})(X_{i+1}^{0+} + X_i^{0-}) + (X_i^{0-} + X_i^{-0})(X_{i+1}^{+0} + X_{i+1}^{0-}) \\ & + (X_i^{++} - X_i^{--})(X_{i+1}^{++} - X_{i+1}^{--}) \} \end{aligned} \tag{12.2}$$

where the changes involved on sites 1 and N are of negligible importance if N is large.

Evidently H has local symmetry algebras $\mathcal{A}_i$ with basis

$$\Lambda_{is} = X_i^{ss}, \qquad \Lambda_{iT} = X_i^{++} + X_i^{00} + X_i^{--} \tag{12.3}$$

Since Λ_{is} is primitive the computation of the $\mathsf{T}_{(j)}$ with $j \geqslant 3$ involves only the triplet (T) site states in the $j-2$ sites between the two end sites of $\mathsf{T}_{(j)}$. Further, since the coupling between an s state on one site and a T state on an adjacent state is the same for each of the T states, the portion of H_{ii+1} surviving when site i is in an S state commutes with $H_{i-1\,i}$ and $H_{i+1\,i+2}$, so that (via the lemma of Appendix B) the end sites of $\mathsf{T}_{(j)}$, $j \geqslant 3$, are necessarily in a T state in order to obtain a nonzero contribution. Also H displays a global $\mathcal{S}\mathcal{U}(2)$ symmetry so that the generalized transfer matrix is reduced to a 2×2 form

$$\begin{aligned} \hat{\mathsf{T}} &= \begin{bmatrix} \hat{\mathsf{T}}_{(2)}^{SS,SS} & \hat{\mathsf{T}}_{(2)}^{SS,TT} \\ \hat{\mathsf{T}}_{(2)}^{TT,SS} & \hat{\mathsf{T}}^{TT,TT} \end{bmatrix} \\ \hat{\mathsf{T}}_{(2)}^{SS,SS} &= \langle SS | e^{-\beta H_{12}} | SS \rangle = e^{\beta J} \\ \hat{\mathsf{T}}_{(2)}^{SS,TT} &= \hat{\mathsf{T}}_{(2)}^{TT,SS} = \frac{1}{\sqrt{3}} \sum_{\mu}^{+,0,-} \langle S\mu | e^{-\beta H_{12}} | S\mu \rangle = \sqrt{3} \\ \hat{\mathsf{T}}^{TT,TT} &= \sum_{j \geqslant 2} \frac{1}{z^{j-2}} \hat{\mathsf{T}}_{(j)}^{TT,TT} \end{aligned} \tag{12.4}$$

Here the $\hat{\mathsf{T}}_{(j)}^{TT,TT}$ are given in terms partition functions $Z_{1\to i}^{T}$ for a regular Heisenberg chain of spin-1 sites

$$\hat{\mathsf{T}}_{(j)}^{TT,TT} = \frac{1}{3} \left\{ Z_{1\to j}^{T} - \sum_{i=1}^{j-2} Z_{1\to j-i}^{T} \hat{\mathsf{T}}_{(i+1)}^{TT,TT} \right\} \tag{12.5}$$

$$Z_{1\to i}^{T} = \mathrm{Tr}_{1\to i}^{(T)} \exp - 2\beta k \sum_{l=1}^{i-1} \mathbf{s}_l \cdot \mathbf{s}_{l+1}$$

The per-site partition function is then the maximum eigenvalue to $\hat{\mathsf{T}}$,

$$z = \tfrac{1}{2}\left(e^{\beta J} + \hat{\mathsf{T}}^{TT,TT}\right) + \tfrac{1}{2}\left[\left(e^{\beta J} + \hat{\mathsf{T}}^{TT,TT}\right)^2 + 12\right]^{1/2} \tag{12.6}$$

a result which is exact.

In addition to the standard high-temperature convergence of the series for $\hat{\mathsf{T}}^{TT,TT}$, it displays a rapid low-temperature convergence for strong antiferromagnetically signed intrasite exchange J. Indeed for $J > 2|K|$ the ground state of the system is the simple product state of local singlet (S) symmetries, and because these local singlet symmetries are primitive, they are treated exactly even by the leading ($j = 2$) term in the expansion for $\hat{\mathsf{T}}$. In fact the J dependence of $\hat{\mathsf{T}}_{(j)}^{TT,TT}$ is entirely due to a factor $e^{-\beta(j-1)J}$, which tends to quench the effects of these higher-order terms when J is large and positive. In detail we find

$$\begin{aligned}
\hat{\mathsf{T}}_{(2)}^{TT,TT} &= \tfrac{1}{3} Z_{12}^{T} \\
\hat{\mathsf{T}}_{(3)}^{TT,TT} &= \tfrac{1}{3} Z_{1\to 3}^{T} - \tfrac{1}{9}\left(Z_{12}^{T}\right)^2 \\
Z_{12}^{T} &= e^{-\beta J}\left\{e^{4\beta K} + 3e^{2\beta K} + 5e^{-2\beta K}\right\} \\
Z_{1\to 3}^{T} &= e^{-2\beta J}\left\{3e^{6\beta K} + e^{4\beta K} + 8e^{2\beta K} + 3 + 5e^{-2\beta K} + 7e^{-4\beta K}\right\}
\end{aligned} \tag{12.7}$$

which in conjunction with (12.5) and (12.6) yield a quite accurate [($Q = 3$)-site] approximation when $J > 2|k|$, which is the regime of interest if we are to reasonably speak of dimers.

13. RANDOM CHAINS

The present cluster expansion approach can also be applied to random linear chains in which each

$$H_{ii+1} \equiv H_{ii+1}(\boldsymbol{\alpha}_i) \tag{13.1}$$

depends upon independent variables $\boldsymbol{\alpha}_i \equiv (\alpha_{i1}, \ldots, \alpha_{iM})$ varying in accordance with a probability distribution $p_i(\boldsymbol{\alpha}_i)$,

$$\begin{gathered}
\int d\boldsymbol{\alpha}_i p_i(\boldsymbol{\alpha}_i) = 1 \\
p_i(\boldsymbol{\alpha}_i) \geqslant 0
\end{gathered} \tag{13.2}$$

For a scalar function φ of the Hamiltonian $H_{i\to j}$ we introduce the notation

$$\langle \varphi(H_{i\to j}) \rangle \equiv \int d\boldsymbol{\alpha}_i \int d\boldsymbol{\alpha}_{i+1} \cdots \int d\boldsymbol{\alpha}_{j-1} \varphi(H_{i\to j}) \prod_{l=i}^{j-1} p_l(\boldsymbol{\alpha}_l) \tag{13.3}$$

Then for two functions φ and θ of $H_{1\to i}$ and $H_{i\to N}$, respectively, we have

$$\langle \varphi(H_{1\to i})\theta(H_{i\to N}) \rangle = \langle \varphi(H_{1\to i}) \rangle \langle \theta(H_{i\to N}) \rangle \tag{13.4}$$

Now defining

$$\begin{aligned} c_i^{rs} &\equiv \langle \mathrm{Tr}_{i\to N} X_i^{sr} \rho_{i\to N} \rangle / \langle \mathrm{Tr}_{i\to N} \rho_{i\to N} \rangle \\ z_i &\equiv \langle \mathrm{Tr}_{i\to N} \rho_{i\to N} \rangle / \langle \mathrm{Tr}_{i+1\to N} \rho_{i+1\to N} \rangle \end{aligned} \tag{13.5}$$

and using (13.4) it follows that we can obtain equations analogous to those of (4.4) and (4.5).

Next we consider an infinite linear chain with an average translational symmetry in the sense that the probability distributions $p_i \equiv p$ are independent of i. Then defining c^{rs} and z as in (4.6) and the transfer matrix T with elements

$$\mathsf{T}^{rs,tu} \equiv \sum_{j \geqslant 2} \frac{1}{z^{j-2}} \langle \mathrm{Tr}_{1\to j} X_1^{sr} \xi_{1\to j} X_j^{tu} \rangle \tag{13.6}$$

we still obtain

$$\mathbf{Tc} = z\mathbf{c} \tag{13.7}$$

Further much of the interpretation and developments following from this equation are similar to those of Sections 4–9. Hence we have a systematic method for solving not only random (generalized) Ising models exactly but also for solving various other (random) models approximately.

14. CONCLUSION

The generalized transfer matrix method which we have introduced for solving linear models is seen to exhibit many of the same features the usual transfer matrix method exhibits in application to Ising-type models. For general linear models we find the present method becomes approximate in that the elements of the generalized transfer matrix become difficult to determine. However, even in truncating the expansion of these matrix elements in our Q-site decoupling approximation, results accurate through given orders in high temperature and perturbation expansions are retained. Indeed it seems that the method should be especially accurate for linear models differing only slightly from generalized Ising models, or more generally from models in which the low-lying states are eigenstates to site symmetries identified with single site states (as discussed in Section 7). In even more general circumstances we might view the present method of making Bonner- and Fisher-type extrapolations, using now not only the eigenvalue spectra of finite chains but also certain end-to-end correlation functions. However, for general models low-temperature convergence is not necessarily easily achieved, as is found in Section 11. Nevertheless accurate low-temperature results are obtained for some models which are not generalized Ising models, an example of which is found in Section 12.

ACKNOWLEDGMENTS

The authors acknowledge encouragement from Professor F. A. Matsen and support from The Robert A. Welch Foundation of Houston, Texas.

APPENDIX A: LINEAR CHAIN FERMION MODELS

We consider an electronic Hamiltonian H defined on an antisymmetric space. If H is expressed in terms of ordinary Fermion creation and annihilation operators one can carry out a Jordan–Wigner type of transformation(16) to new Pauli-type raising and lowering operators, which in turn may be used rather directly but tediously to construct the desired X-operators. Here we follow a line of development in terms of a more general *site representation*.(17–19) In this representation we presume antisymmetric sites states such that the whole space for H is spanned by kets

$$|r_1 \wedge r_2 \wedge \cdots \wedge r_N\rangle \tag{A.1}$$

where r_i labels the state on site i and the $\wedge$-symbol indicates an antisymmetrized (or Grassman) product. Then H can be expressed as

$$H = \sum_i \sum_{rs} \epsilon_i^{rs} A_{ir}^+ A_{is} + \sum_{i<j} \sum_{rstu} v_{ij}^{rstu} A_{ir}^+ A_{js}^+ A_{ju} A_{it} + \cdots \tag{A.2}$$

where the A_{ir}^+ and A_{ir} are site-state creation and annihilation operators for state r on site i. These operators satisfy

$$\begin{aligned}
&\left.\begin{aligned}
A_{ir}^+ A_{js}^+ &= (-1)^{n_{ir}n_{js}} A_{js}^+ A_{ir}^+ \\
A_{ir}^+ A_{js} &= (-1)^{n_{ir}n_{js}} A_{js} A_{ir}^+ \\
A_{ir} A_{js} &= (-1)^{n_{ir}n_{js}} A_{js} A_{ir}
\end{aligned}\right\} \quad i \neq j \\
&A_{ir}^+ A_{is} = \delta_{r0} A_{is} \\
&A_{ir} A_{is}^+ = \delta_{rs} A_{ir} A_{ir}^+ \\
&\left.\begin{aligned}
A_{ir}^+ A_{is}^+ &= 0 \\
A_{is} A_{ir} &= 0
\end{aligned}\right\} \quad s \neq 0
\end{aligned} \tag{A.3}$$

where n_{ir} denotes the number of electrons in the state r on site i.

The only restrictions placed upon H are that it be electron conserving and that it involve interactions between no more than nearest-neighbor sites. (It may be noted that one-site terms can be chosen to appear like nearest-neighbor pair interactions through the use of the identity $\sum_r A_{ir}^+ A_{ir} = 1$.) Now we define a type of Jordan–Wigner transformed(17–19) operator

$$X_i^{rs} \equiv (-1)^{(n_{ir} - n_{is})N_{1\to i-1}} A_{ir}^+ A_{is} \tag{A.4}$$

with $N_{1\to i-1}$ being the number operator for the electrons on sites 1 through

$i - 1$. Then for $i < j$

$$\begin{aligned} X_i^{rs} X_j^{tu} &= (-1)^{(n_{ir} - n_{is}) N_{1\to i-1}} A_{ir}^+ A_{is} (-1)^{(n_{jt} - n_{ju}) N_{1\to j-i}} A_{jt}^+ A_{ju} \\ &= (-1)^{(n_{ir} - n_{is}) N_{1\to i-1}} (-1)^{(n_{jt} - n_{ju})(N_{1\to j-1} + n_{ir} - n_{is})} A_{ir}^+ A_{is} A_{jt}^+ A_{ju} \\ &= (-1)^{(n_{ir} - n_{is}) N_{1\to i-1}} (-1)^{(n_{jt} - n_{ju}) N_{1\to j-1}} A_{jt}^+ A_{ju} A_{ir}^+ A_{is} \\ &= X_j^{tu} X_i^{rs} \end{aligned} \tag{A.5}$$

and we readily see that these X-operators satisfy (2.3) for all i and j. Thus the set of operators

$$\{ X_i^{rr}; \quad ir \text{ ranging} \} \tag{A.6}$$

forms a complete commuting set for our space, which hence has a basis of kets of the form given in (2.7). Finally on comparing (A.2) to (2.4) and (2.5), we see that

$$\begin{aligned} H_{ii+1} &= \sum_{rstu} \Big\{ \tfrac{1}{2}(1 + \delta_{i,1}) \epsilon_i^{rt} \delta_{su} + \tfrac{1}{2}(1 + \delta_{i,N-1}) \delta_{rt} \epsilon_{i+1}^{su} \\ &\qquad + v_{ii+1}^{rstu} (-1)^{n_{it}(n_{i+1s} - n_{i+1c})} \Big\} A_{ir}^+ A_{it} A_{i+1s}^+ A_{i+1u} \\ &= \sum_{rstu} \langle r \cdot s | H_{ii+1} | t \cdot u \rangle X_i^{rt} X_{i+1}^{su} \end{aligned} \tag{A.7}$$

with

$$\begin{aligned} \langle r \cdot s | H_{ii+1} | t \cdot u \rangle &= \tfrac{1}{2}(1 + \delta_{i,1}) \epsilon_i^{rt} \delta_{su} + \tfrac{1}{2}(1 + \delta_{i,N-1}) \delta_{rt} \epsilon_i^{su} \\ &\qquad + (-1)^{n_{ir}(n_{i+1s} - n_{i+1u})} v_{ii+1}^{rstu} \end{aligned} \tag{A.8}$$

If nonnearest-neighbor interactions were allowed, the present Jordan–Wigner transformation would yield more complex interactions in terms of the X-operators.

APPENDIX B: PROPERTIES OF $\xi_{1\to n+1}$

We first wish to establish a working lemma:

Lemma. If $[H_{i-1i}, H_{ii+1}] = 0$ for some $i \in \{2, 3, \ldots, n\}$ with $n \geqslant 2$, then $\xi_{1\to n+1} = 0$.

Proof. Assuming the hypothesis of this theorem, we consider the first ($l = 0$) term of (3.4)

$$\rho_{1\to n+1} = \rho_{1\to i} \rho_{i\to n+1}$$

so that this $l = 0$ term cancels with the $l = 1$ term for which $i = i_1$. Now consider the other $l = 1$ terms

$$\rho_{1\to i_1} \rho_{i_1\to n+1} = \begin{cases} \rho_{1\to i_1} \rho_{i_1\to i} \rho_{i\to n+1}, & i_1 < i \\ \rho_{1\to i} \rho_{i\to i_1} \rho_{i_1\to n+1}, & i_1 > i \end{cases} \tag{B.1}$$

so that these other $l = 1$ terms cancel with the $l = 2$ terms for which $i \in \{i_1, i_2\}$. Similarly, the other $l = 2$ terms (with $i \notin \{i_1, i_2\}$) cancel with the $l = 3$ terms for which $i \in \{i_1, i_2, i_3\}$, etc. Finally the $l = n - 2$ term with $j \notin \{i_1, \ldots, i_{n-2}\}$ cancels with the single $l = n - 1$ term, and the theorem is proved.

This lemma leads directly to some useful results:

Theorem B.1. If $\xi_{1\to n+1}$ is expanded in powers of β, then the first nonvanishing term is of order n, involving β^n.

Proof. We let $H_{ii+1} \equiv \lambda_{ii+1}\tilde{H}_{ii+1}$ and consider first powers of these λ's which may arise in $\xi_{1\to n+1}$. Thus if $H_{ii+1} \to 0$ for some $i = 1$ to n, then from the lemma we see that $\xi_{1\to n+1} \to 0$ so that $\xi_{1\to n+i}$ is at least linear in λ_{ii+1}. Hence $\xi_{1\to n+1}$ is proportional to $\prod_{i=1}^{n}\lambda_{ii+1}$, and higher powers. But each of these λ_{ii+1} appears only along with a β, so that $\xi_{1\to n+1}$ is proportional to β^n, and higher powers.

Theorem B.2. If $H = H^0 + \lambda V$ with a generalized Ising model and if $\xi_{1\to n+1}$ is expanded in powers of λ, then the first nonvanishing term is of order $[n/2]$, involving $\lambda^{[n/2]}$, where $[n/2]$ is $(n-1)/2$ or $n/2$ as n is odd or even.

Proof. Here we let $H_{ii+1} = H^0_{ii+1} + \lambda V_{ii+1}$ with the H^0_{ii+1} the terms of the generalized Ising model. Now since all these H^0_{ii+1} mutually commute, our lemma implies that a term in the expansion in powers of these λ's will vanish unless it is, at least, first-order in λ_{i-1i} or λ_{ii+1}, with $i = 2$ to n. The lowest-order way in which this "nonvanishing" criterion can be achieved involves $\prod_{i=1}^{[n/2]}\lambda_{2i-1\,2i}$. Setting all these λ_{jj+1} equal to just λ, we then obtain the theorem.

APPENDIX C: PROPERTIES OF T AND z

In the theorems of this appendix it will sometimes be of use to consider that the generalized transfer matrix T is defined by

$$\mathsf{T}^{rs,tu} \equiv \sum_{vw} \sum_{j \geqslant 2} \mathrm{Tr}_{1\to j}(X_1^{sr}\xi_{1\to j}X_j^{vw})\left[\mathsf{T}^{j-2}\right]^{vw,tu} \tag{C.1}$$

instead of by (4.7).

Theorem C.1. At sufficiently high temperatures z is the nondegenerate maximum eigenvalue to T.

Proof. Because of the high-temperature properties of the ξ's, as discussed in Appendix B, we see from (3.6) that

$$Z \cong \mathrm{Tr}\{\xi_{12}\xi_{23}\xi_{34}\cdots\xi_{N-1N}\}, \qquad \beta \cong 0 \tag{C.2}$$

and from (4.7), or (C.1), that

$$\mathsf{T}^{rs,tu} \cong \mathrm{Tr}_{12}(X_1^{sr}\xi_{12}X_2^{tu}) \tag{C.3}$$

Thus combining these two results we obtain

$$Z \cong \sum_{st} \left[\mathsf{T}^{N-1}\right]^{ss,tt}, \qquad \beta \cong 0 \tag{C.4}$$

Then (since $N \to \infty$)

$$Z \cong \sum_{st} c^{ss} z^{N-1} c^{tt}, \qquad \beta \cong 0 \tag{C.5}$$

where z is the maximum eigenvalue to T for which the corresponding eigenket $\mathbf{c}$ has

$$\sum_{s} c^{ss} \neq 0 \tag{C.6}$$

But from (C.3) we see that

$$\mathsf{T}^{rs,tu} \cong \langle r \cdot u | e^{-\beta H_{12}} | s \cdot t \rangle \cong \delta_{rs}\delta_{tu} - \beta \langle r \cdot u | H_{12} | s \cdot t \rangle, \qquad \beta \cong 0 \tag{C.7}$$

so that on using a perturbation expansion of the T matrix about $\beta = 0$, we find a single eigenvalue near 1 with all other eigenvalues near 0. Further, the eigenket associated with this nondegenerate maximum eigenvalue, near 1, satisfies (C.6). Thus we identify this eigenvalue and eigenket, z and $\mathbf{c}$, as in (4.6).

Theorem C.2. If H is an Ising-type model, then the elements of T are 0 except for a diagonal block identical to the conventional transfer matrix, and z is the nondegenerate maximum eigenvalue to T at all temperatures.

Proof. Using the site-states effecting the diagonalization, we see that

$$\mathsf{T}^{rs,tu} = \mathrm{Tr}_{12}\left(X_1^{sr} e^{-\beta H_{12}} X_2^{tu}\right) = \delta_{rs}\delta_{tu} e^{-\beta\langle r \cdot t | H_{12} | r \cdot t \rangle} \tag{C.8}$$

and we recognize the nonzero block of T as the usual transfer matrix for Ising-type models. Now the maximum eigenvalue to T is already known(1-4) to be the per-site partition function and to be[6] nondegenerate at all $\beta > 0$.

Theorem C.3. If both T and z are obtained in the Q-site decoupling approximation, then their expansions in powers of β are both accurate up through order Q.

Proof. In the Q-site approximation only the terms in (C.1) through $j = Q$ are retained, and the lowest-order errors are contributed by the first

---

[6] See, for instance, Ref. 20.

omitted term $\xi_{1\to Q+1}$. Since Theorem B.1 implies $\xi_{1\to Q+1}$ is at least of order Q, we immediately see that T and z are accurate up through order $Q-1$. Now using the lemma of Section B we see, as in the proof to Theorem B.1, that the portion of $\xi_{1\to Q+1}$ which is of order Q involves a linear combination of products with each of the interactions H_{ii+1}, $i=1$ to Q, raised to the first power. Thus the lowest-order portion of the lowest-order term $\mathrm{Tr}_{1\to Q+1}(X_1^{sr}\xi_{1\to Q+i}X_2^{tu})$, which is deleted from T, is a linear combination of

$$\mathrm{Tr}_{1\to Q+1}\{X_1^{sr}H_{12}H_{23}\cdots H_{QQ_1}X_{Q+1}^{tu}\} \tag{C.9}$$

and similar terms merely with different orderings for the interactions. However, using the cyclic trace property ($\mathrm{Tr}\,AB=\mathrm{Tr}\,BA$) we may reorder each of these terms to the form in (C.9). For instance,

$$\begin{aligned}\mathrm{Tr}_{1\to 4}\{X_1^{sr}H_{23}H_{12}H_{34}X_4^{tu}\} &= \mathrm{Tr}_{1\to 4}\{X_1^{sr}H_{23}H_{34}X_4^{tu}H_{12}\}\\ &= \mathrm{Tr}_{1\to 4}\{H_{12}X_1^{sr}H_{23}H_{34}X_4^{tu}\}\\ &= \mathrm{Tr}_{1\to 4}\{X_1^{sr}H_{12}H_{23}H_{34}X_4^{tu}\}\end{aligned} \tag{C.10}$$

Then since commutation of the various interactions would do no more than to achieve these same reorderings, we see that all these Qth-order terms cancel, and the only errors in T and z are of order $Q+1$, and higher.

Theorem C.4. If $H=H^0+\lambda V$ with H^0 a generalized Ising model and if T and z are obtained in the Q-site decoupling approximation, then the expansions of T and z in powers of λ are both accurate up through order $[(Q-1)/2]+1$.

Proof. The proof of accuracy up through order $[(Q-1)/2]$ follows readily from Theorem B.2, while the result for the last order again follows using the cyclic trace property, much as in Theorem C.3.

APPENDIX D: NON-HERMITIAN PERTURBATION THEORY

Here we briefly consider the perturbation formulas for a real nondegenerate isolated eigenvalue z of a (possibly) non-Hermitian matrix or operator $\mathfrak{T}$. Since it is nondegenerate there is a similarity transformation which isolates z in its own 1×1 block and there are both right and left eigenvectors, $|\psi\rangle$ and $(\psi|$, for z:

$$\begin{aligned}\mathfrak{T}|\psi\rangle &= z|\psi\rangle\\ (\psi|\mathfrak{T} &= z(\psi|\end{aligned} \tag{D.1}$$

Next we expand $\mathfrak{T}$, z, $|\psi\rangle$, and $(\psi|$ in powers of a perturbation parameter λ with $\mathfrak{T}^{(n)}, z^{(n)}, |n\rangle$, and $(n|$ being the nth-order terms, i.e., the coefficients of

λ^n in the corresponding expansion. Then the basic perturbation equations are

$$\begin{gathered}(\mathcal{T}^{(0)}-z^{(0)})|0\rangle=0\\(\mathcal{T}^{(1)}-z^{(1)})|0\rangle+(\mathcal{T}^{(0)}-z^{(0)})|1\rangle=0\\(\mathcal{T}^{(2)}-z^{(2)})|0\rangle+(\mathcal{T}^{(1)}-z^{(1)})|1\rangle+(\mathcal{T}^{(0)}-z^{(0)})|2\rangle=0\\\text{etc.}\end{gathered}\tag{D.2}$$

Now choosing the intermediate normalization

$$(0|0\rangle=(0|\psi\rangle=1\tag{D.3}$$

we have

$$(0|n\rangle=0,\qquad n\geqslant 1\tag{D.4}$$

Then applying $(0|$ to the equations (D.2), one obtains

$$\begin{aligned}z^{(1)}&=(0|\mathcal{T}^{(1)}|0\rangle\\z^{(2)}&=(0|\mathcal{T}^{(2)}|0\rangle+(0|\mathcal{T}^{(1)}|1\rangle\end{aligned}\tag{D.5}$$

In characterizing the perturbation vectors $|n\rangle$ we find it convenient to introduce complimentary (non-Hermitian) projectors

$$P\equiv|0\rangle(0|,\qquad Q\equiv 1-P\tag{D.6}$$

and the zero-order resolvent

$$\mathcal{R}\equiv Q\{z^{(0)}-\mathcal{T}^{(0)}+\alpha P\}^{-1}Q,\qquad \alpha\neq 0\tag{D.7}$$

[In terms of matrices the zero-order resolvent may be computed by first carrying out the similarity transformation block-diagonalizing the $M\times M$ matrix $z^{(0)}-\mathcal{T}^{(0)}$ so that the zero eigenvalue appears in its own individual block, second inverting the complimentary (and nonsingular) $(M-1)\times(M-1)$ block, and third performing the reverse similarity transformation.] Now noting that Q is orthogonal to P and commutes with $\mathcal{T}^{(0)}$, we note

$$\begin{aligned}\mathcal{R}(z^{(0)}-\mathcal{T}^{(0)})&=Q\{z^{(0)}-\mathcal{T}^{(0)}+\alpha P\}^{-1}Q(z^{(0)}-\mathcal{T}^{(0)}+\alpha P)\\&=Q\{z^{(0)}-\mathcal{T}^{(0)}+\alpha P\}^{-1}\{z^{(0)}-\mathcal{T}^{(0)}+\alpha P\}Q=Q\end{aligned}\tag{D.8}$$

Thus applying the zero-order resolvent to the second of the equations of (D.2), we obtain

$$|1\rangle=Q|1\rangle=\mathcal{R}(\mathcal{T}^{(1)}-z^{(1)})|0\rangle=\mathcal{R}\mathcal{T}^{(1)}|0\rangle\tag{D.9}$$

and hence

$$z^{(2)}=(0|\{\mathcal{T}^{(2)}+\mathcal{T}^{(1)}\mathcal{R}\mathcal{T}^{(1)}\}|0\rangle\tag{D.10}$$

Clearly, with the present formulation, all these results bear a close resemblance to the familiar case with $\mathfrak{T}^{(0)}$ Hermitian and $(0|$ simply the adjoint of $|0\rangle$. Similar remarks apply to the higher-order formulas, and those of double perturbation expansions also.

APPENDIX E: FORMULA FOR THE ZERO-ORDER RESOLVENT

A basic computation $\mathfrak{R}$ satisfies is

$$(z - \mathfrak{T})\mathfrak{R} = 1 - \mathbf{C}\mathbf{A}^{\dagger} \tag{E.1}$$

From the $(i, 2)$-subblock of this matrix, $i \geqslant 3$, we find

$$\sum_{j=2}^{Q} (z\delta_{ij} - \mathfrak{T}i_j)\mathfrak{R}_{j,2} = -\mathbf{C}_{(i)}\mathbf{A}^{\dagger}_{(2)} \tag{E.2}$$

$$\mathfrak{R}_{i,2} = \frac{1}{z}\mathfrak{R}_{i-1,2} - \frac{1}{z^{i-1}}\mathbf{C}_{(2)}\mathbf{A}^{\dagger}_{(2)} = \frac{1}{z^{i-2}}\mathfrak{R}_{2,2} - \frac{i-2}{z^{i-1}}\mathbf{C}_{(2)}\mathbf{A}^{\dagger}_{(2)}$$

so that all the blocks of $\mathfrak{R}$ appearing in (6.6) are expressed in terms of the $(2, 2)$-subblock. Using the $(2, 2)$-block of (E.1) and then the final result of (E.2), we obtain

$$z\mathfrak{R}_{2,2} - \sum_{j=2}^{Q} \mathsf{T}_{(j)}\mathfrak{R}_{j,2} = 1 - \mathbf{C}_{(2)}\mathbf{A}^{\dagger}_{(2)}$$

$$(z - \mathsf{T})\mathfrak{R}_{2,2} = 1 - \mathbf{C}_{(2)}\mathbf{A}^{\dagger}_{(2)} - \sum_{j=3}^{Q} \frac{j-2}{z^{j-1}}\mathsf{T}_{(j)}\mathbf{C}_{(2)}\mathbf{A}^{\dagger}_{(2)} \tag{E.3}$$

$$\mathfrak{R}_{2,2} = \mathsf{R}\left\{1 - \sum_{j=3}^{Q} \frac{j-2}{z^{j-1}}\mathsf{T}_{(j)}\mathbf{C}_{(2)}\mathbf{A}^{\dagger}_{(2)}\right\} + \zeta\mathbf{C}_{(2)}\mathbf{A}^{\dagger}_{(2)}$$

where R is the generalized inverse to $z - \mathsf{T}$. Since R is defined to be zero on the null space of $z - \mathsf{T}$, these equations do not determine ζ. To find ζ we can use the condition that $\mathbf{A}^{\dagger}$ is in the left null-space of $\mathfrak{R}$, followed by the results of (E.2) and (E.3), to obtain

$$\sum_{j=2}^{Q} \mathbf{A}^{\dagger}_{(j)}\mathfrak{R}_{j,2} = 0$$

$$\begin{aligned}
\zeta \sum_{j=2}^{Q} \frac{1}{z^{j-2}}\mathbf{A}^{\dagger}_{(j)}\mathbf{C}_{(2)}\mathbf{A}^{\dagger}_{(2)} \\
= -\sum_{j=2}^{Q} \frac{1}{z^{j-2}}\mathbf{A}^{\dagger}_{(j)}\mathsf{R}\left\{1 - \sum_{i=3}^{Q} \frac{i-2}{z^{i-1}}\mathsf{T}_{(i)}\mathbf{C}_{(2)}\mathbf{A}^{\dagger}_{(2)}\right\} \\
+ \sum_{j=3}^{Q} \frac{j-2}{z^{j-1}}\mathbf{A}^{\dagger}_{(j)}\mathbf{C}_{(2)}\mathbf{A}^{\dagger}_{(2)}
\end{aligned} \tag{E.4}$$

Next recalling (5.4), we find

$$\sum_{j=2}^{Q} \frac{1}{z^{j-2}} \mathbf{A}^{\dagger}_{(j)} \mathbf{C}_{(2)} = \sum_{j=2}^{Q} \mathbf{A}^{\dagger}_{(j)} \mathbf{C}_{(j)} = \mathbf{A}^{\dagger}\mathbf{C} = 1 \tag{E.5}$$

Now applying $\mathbf{A}_{(2)}$ to the right of (E.4) and dividing by $\mathbf{A}^{\dagger}_{(2)}\mathbf{A}_{(2)}$, we obtain

$$\zeta = \sum_{j=3}^{Q} \frac{j-2}{z^{j-1}} \mathbf{A}^{\dagger}_{(j)} \mathbf{C}_{(2)} + \sum_{j=2}^{Q} \sum_{i=3}^{Q} \frac{i-2}{z^{i+j-3}} \mathbf{A}_{(j)} \mathsf{RT}_{(i)} \mathbf{C}_{(2)}$$

our final result.

APPENDIX F: LOWER BOUNDS TO THE GROUND STATE ENERGY

Here we consider the ground state to H of (12.2), for certain values of the exchange parameters J and K. The product of dimer singlets is evidently an eigenvector with eigenvalue $-J(N-1)$ for any values of J and K. Now we consider Bopp-type lower bounds[7] to the ground state of H. The energy of a normalized ket $|\Psi\rangle$ is

$$E_{\Psi} = \langle\Psi|H|\Psi\rangle = \sum_{i=1}^{N-1} \mathrm{Tr}_{ii+1}(H_{ii+1}\rho_{ii+1}) \tag{F.1}$$

where ρ_{ii+1} is the reduced density matrix obtained by a (partial) trace over the indices of all dimers other than i and $i+1$,

$$\rho_{ii+1} \equiv \mathrm{Tr}_{1\to i-1}\mathrm{Tr}_{i+2\to N}|\Psi\rangle\langle\Psi| \tag{F.2}$$

Clearly ρ_{ii+1} is positive semidefinite, Hermitean, and normalized (to 1). Now if we minimize the pair energies,

$$\epsilon_{ii+1} \equiv \mathrm{Tr}_{ii+1}(H_{ii+1}\hat{\rho}_{ii+1}) \tag{F.3}$$

by varying the $\hat{\rho}_{ii+1}$ independently subject to the above-mentioned constraints, a lower bound to the true pair energies (computed with the true ρ_{ii+1}) results. Hence a lower bound to E_{Ψ} also results. Since the eigenvalues to H_{ii+1} are $-J$, 0, $J+2K$, $J-2K$, and $J-4k$, we see that $-J$ is the lowest eigenvalue for J greater than 0 and greater than $2K$. Thus the product of dimer singlets has an energy equal to this Bopp-type lower bound for $J > 2K, 0$. Further the ground state is nondegenerate, for these parameter values, since any other state would entail reduced density matrices with some local triplet component, so that a lower bound for such a state would be strictly greater than $-J(N-1)$. Similar remarks apply to the cyclic Hamiltonian of (12.1).

---

[7] See, for instance, Refs. 21–24.

REFERENCES

1. E. W. Montroll, *J. Chem. Phys.* **9**:706 (1941).
2. H. A. Kramers and G. H. Wannier, *Phys. Rev.* **60**:252 (1941).
3. E. N. Lassettre and J. P. Howe, *J. Chem. Phys.* **9**:747 (1941).
4. R. Kudo, *Busseiron Kenkyu* **1** (1943).
5. L. D. Landau and E. M. Lifschitz, *Statistical Physics* (Pergamon Press, London, 1958), p. 482.
6. D. J. Klein, in *Group Theory and Its Applications*, Vol. III, E. M. Loebl, ed., (Academic Press, New York, 1975).
7. E. H. Lieb, T. D. Schultz, and D. C. Mattis, *Ann. Phys. (N.Y.)* **16**:407 (1961).
8. D. C. Mattis and S. B. Nam, *J. Math. Phys.* **13**:1185 (1972).
9. A. Zawadowski and M. H. Cohen, *Solid State Comm.*, preprint.
10. J. C. Bonner and M. E. Fisher, *Phys. Rev.* **135**:A640 (1964).
11. P. L. Nordio, Z. G. Soos, and H. M. McConnell, *Ann. Rev. Phys. Chem.* **17**:237 (1966).
12. I. F. Shchegolev, *Phys. Status Solidi A* **12**:4 (1972).
13. G. F. Kokoszka and R. W. Duerst, *Coord. Chem. Rev.* **5**:209 (1970).
14. A. P. Ginsberg, *Inorg. Chim. Acta Rev.* **5**:45 (1971).
15. L. J. DeJongh and A. R. Miedema, *Adv. Phys.* **23**:1 (1974).
16. P. Jordan and E. P. Wigner, *Z. Phys.* **47**:631 (1928).
17. J. Hubbard, *Proc. R. Soc. London Ser. A* **285**:542 (1965).
18. D. J. Klein and Z. G. Soos, *Mol. Phys.* **20**:1013 (1971).
19. Z. G. Soos and D. J. Klein in *Molecular Association*, R. Foster, ed., (Academic Press, New York, 1975).
20. D. Ruelle, *Statistical Mechanics, Rigorous Results*, Theorem 5.6.2 (W. A. Benjamin, New York, 1969).
21. F. Bopp, *Z. Phys.* **156**:348 (1959).
22. A. J. Coleman, *Rev. Mod. Phys.* **35**:668 (1963).
23. L. J. Kijewski and J. K. Percus, *Phys. Rev.* **164**:228 (1967).
24. L. J. Kijewski and J. K. Percus, *Phys. Rev.* **179**:45 (1969).