Journal of Statistical Physics, Vol. 57, Nos. 1/2, 1989

Superintegrable Chiral Potts Model: Thermodynamic Properties, an "Inverse" Model, and a Simple Associated Hamiltonian

R. J. Baxter[1,2]

Received June 27, 1989

The partition function of the N-state superintegrable chiral Potts model is obtained exactly and explicitly (if not completely rigorously) for a finite lattice with particular boundary conditions. This yields the bulk and surface free energies, and horizontal and vertical correlation lengths and interfacial tensions. The critical exponents are $\alpha = 1 - 2/N$, $\mu_{\text{hor}} = \nu_{\text{hor}} = 2/N$, and $\mu_{\text{vert}} = \nu_{\text{vert}} = 1$, and the finite-size corrections are obtained at criticality. The eigenvalue spectrum of the column-to-column transfer matrix is that of a direct product of N by N matrices. Inverting this matrix gives a related solvable model which is a generalization of the free-fermion model. The associated Hamiltobian has a very simple form, suggesting there may be a more direct algebraic method (perhaps a generalized Clifford algebra) for obtaining its eigenvalues.

KEY WORDS: Statistical mechanics; exact solution; chiral Potts model; anisotropic scaling; wetting; conformal invariance; Clifford algebra.

1. INTRODUCTION

Recently, a new class of two-dimensional lattice models have been discovered[1] that satisfy the star–triangle (or Yang–Baxter) relation. On each site i of a lattice there is a spin σ_i, taking the values $0,..., N-1$. There are particular two-spin interactions between adjacent spins. Like the Potts model, the interactions are Z_N-symmetric, but they are chiral, in the sense that they are not reflection-symmetric. Exact (if unwieldy) equations for the free energy have been obtained in the infinite-lattice limit.[2]

An intriguing special case occurs when the model is "superintegrable."

[1] Centre for Mathematical Analysis, Australian National University, GPO Box 4, Canberra ACT 2601, Australia.

[2] Permanent Address: Mathematics and Theoretical Physics, IAS, Australian National University, Canberra ACT 2601, Australia.

0022-4715/89/1000-0001$06.00/0

It has been shown[3,4] that the row-to-row transfer matrix $\mathbf{T}_{\text{row}}$ then has a set S of eigenvalues, which have a particularly simple form. Provided the largest eigenvalue of $\mathbf{T}_{\text{row}}$ is contained in S, this makes it easy to obtain the bulk free energy explicity as an integral of elementary functions.

Albertini *et al.*[5] have very recently found, for toroidal boundary conditions and close to criticality, that the largest eigenvalue is *not* contained in the set S. However, as was remarked in ref. 4, if we impose suitable fixed-spin boundary conditions on the top and/or bottom rows of the lattice, then only those eigenvalues in S can contribute to the partition function Z.

Here we consider a model with such fixed-spin boundary conditions at both the top and bottom of the lattice, and cyclic (cylindrical) boundary conditions from right to left. In Sections 2 and 3 we obtain explicit expressions for Z for a *finite* lattice of arbitrary size. These results have been reported earlier.[6] They do depend on the assumption that the matrix $\mathbf{P}$ in Section 2 is independent of the variable k': this is unproven, but is plausible and numerically tested (see Appendix B).

In Section 4 the results are used to deduce the full eigenvalue spectrum of the column-to-column transfer matrix $\mathbf{T}_{\text{col}}$. This turns out to be extremely simple: basically just a direct product of N by N diagonal matrices. In Section 5 we discuss the duality relation, and show that it maps the model to one with free boundary conditions on the top and bottom row, and shifted cyclic boundary conditions from right to left.

We consider the large-lattice limit in Section 6. We show that these results yield not only the bulk free energy, but also a surface free energy due to the fixed (or free) spin boundary condition, and horizontal and vertical correlation lengths and interfacial tensions. (The vertical properties differ from those usually defined: in particular, for $N=2$ they reduce to the known Ising results only when $\mathbf{T}_{\text{row}}$ is Hermitian. Even, so they may have the same critical exponents.) The corresponding critical exponents [defined in (6.27)] are

$$\begin{aligned} &\alpha = 1 - 2/N, \qquad \alpha_s = 2 - 2/N \\ &\mu_{\text{hor}} = \nu_{\text{hor}} = 2/N \\ &\mu_{\text{vert}} = \nu_{\text{vert}} = 1 \end{aligned} \tag{1.1}$$

These results have been reported earlier,[3,4] but without a clear distinction between the horizontal and vertical properties. They satisfy the relations

$$\begin{aligned} &\alpha_s = \alpha + \nu_{\text{vert}} \\ &\nu_{\text{hor}} + \nu_{\text{vert}} = 2 - \alpha \\ &\mu_{\text{hor}} = \nu_{\text{hor}}, \qquad \mu_{\text{vert}} = \nu_{\text{vert}} \end{aligned} \tag{1.2}$$

The first of these follows from finite-size scaling [Eqs. (3.18a) and (3.22) of ref. 7]. The others can be thought of as variants of the usual scaling relations $\mu + \nu = 2 - \alpha = d\nu$ for an intrinsically anisotropic two-dimensional system.[8] From this point of view it is critical isotropy that is violated by this model, rather than scaling. This is also apparent from the finite-size corrections for the critical case, which are shown to differ from the usual inverse power form[9,10] predicted by conformal invariance.

Another phenomenon exhibited by the model is that of interfacial "wetting" (§8.1 of ref. 11). For sufficiently large N it can be energetically favorable for two ordered phases to be separated by one or more other phases.

In Section 7 we return to the exact results for the finite lattice and show that they can be extended to an inhomogeneous system, where the Boltzmann weight parameters (including k') vary from row to row. (This provides an alternative way of going from fixed to free boundary conditions.) We remark that the direct product eigenvalue spectrum of $\mathbf{T}_{\mathrm{col}}$ is very similar to that of the Ising model, which is the $N = 2$ case of this superintegrable chiral Potts model. Indeed, the underlying equations are identical (being independent of N): given the eigenvalue spectrum for the Ising model case, one can deduce at once the spectrum for general N.

This implies a close connection between the Ising model and the superintegrable chiral Potts model, so we have started to investigate this further. The matrix $\mathbf{T}_{\mathrm{col}}$ is a product of local transfer matrices, corresponding to individual edges of the lattices. It is therefore easy to invert, and its inverse is a rather simple sparse matrix, corresponding to a model in which either horizontally adjacent spins are equal or the left-hand spin is one greater than the right-hand spin.

This new "inverse" model is defined in Section 8 and is a generalization of the free-fermion model. We obtain its partition function (for the finite lattice) explicitly.

We also obtain the eigenvalues of an associated non-Hermitian Hamiltonian. This Hamiltonian has a simple structure, and is an N-state generalization of the one-dimensional Ising model in a transverse field. All its coefficients are abitrary, yet its eigenvalues are always simple direct sums. This is intriguing: it suggests that there may be a much more direct way of obtaining the eigenvalues, perhaps a generalization of the Clifford algebra (i.e., fermion operator) method that solves the Ising model.

2. DEFINITION OF THE MODEL

Consider the square lattice, drawn diagonally as in Fig. 1, with $M + 1$ rows, each of L sites. Impose cylindrical boundary conditions connecting

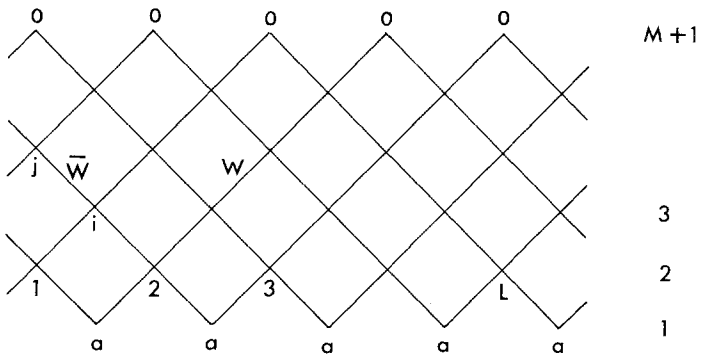

Fig. 1. The square lattice $\mathscr{L}$, with periodic boundary conditions at the sides, and fixed spin boundary conditions at top and bottom (with values 0 and a).

the left and right boundaries. At each site i place a spin σ_i, taking the values $0,\ldots, N-1$. Fix the spins in the top (bottom) row to have value $0\,(a)$, as indicated. Adjacent spins σ_i, σ_j on southwest to northeast edges (with i below j) interact with Boltzmann weight $W(\sigma_i-\sigma_j)$; those on southeast to northwest edges with weight $\bar{W}(\sigma_i-\sigma_j)$. The functions W, $\bar{W}$ are given by

$$\begin{aligned} W(n) &= W(N+n) = \mu^n \prod_{j=1}^{n} (1-y\omega^j)/(1-x\omega^j) \\ \bar{W}(n) &= \bar{W}(N+n) = \mu^{-n} \prod_{j=1}^{n} (\omega - x\omega^j)/(1-y\omega^j) \end{aligned} \tag{2.1}$$

where $\omega = e^{2\pi i/N}$, x and y are parameters (in general complex) that are at our disposal, and

$$\mu^N = (x^N-1)/(y^N-1) \tag{2.2}$$

The partition function, which depends on a, is

$$Z_a = \sum_{\sigma} \prod_{\langle i,j\rangle} E(\sigma_i-\sigma_j) \tag{2.3}$$

where the outer sum is over all values of all the spins, the product is over all edges $\langle i, j\rangle$ of the lattice, and the function E is either W or $\bar{W}$, depending on the orientation of the edge. We take x, y, μ to be the same for all edges in the same row, but possibly different in different rows.

The row-to-row transfer matrix is defined in the obvious way in ref. 4. It can be thought of as a function $\mathbf{T}_{\text{row}}(x, y)$ of x and y. Strictly, we should

distinguish, as in ref. 4, between transfer matrices for odd and even rows. However, since they differ only by cyclic shifts of the column indices 1,..., L in one of the rows, and we shall only consider the "zero-momentum" space of vectors invariant such shifts, the distinction here is immaterial.

Because of the star–triangle relation,[1] two matrices $\mathbf{T}_{\text{row}}(x, y)$ (with different x and different y) commute provided they have the same value of

$$k' = (x^N - 1)(y^N - 1)/(y^N - x^N) \tag{2.4}$$

Taking the limit $x, y, \mu \to 1$, we find that they also commute with the Hamiltonian

$$\mathscr{H} = -2 \sum_{j=1}^{L} \sum_{n=1}^{N-1} (Z_j^n Z_{j+1}^{-n} + k' X_j^n)/(1 - \omega^{-n}) \tag{2.5}$$

This is the Hamiltonian considered in refs. 3 and 5, and the operators X_j, Z_j are defined therein (our k', L are there replaced by λ, $\mathcal{N}$). All we need note is that

$$\mathscr{H} = \mathscr{H}_0 + k' \mathscr{H}_1 \tag{2.6}$$

where $\mathscr{H}_0$ and $\mathscr{H}_1$ are independent even of k', depending only on N and L.

2.1. Nonpositivity of the Boltzmann Weights

This model is a special case of the solvable chiral Potts model in which the vertical rapidity p is fixed so that (in the notation of ref. 1) $a_p = b_p = 1$, $c_p = d_p = \eta$, $x = \eta a_q / d_q$, $y = c_q / \eta b_q$, where $\eta = [(1 + k')/(1 - k')]^{1/2N}$. In the notation of ref. 2, this implies $v_p = -\pi/2$, $u_p = -\pi/2 + i \cos h^{-1}(1/k)$. We are therefore outside the physical range discussed in ref. 2, where k', u_p, u_q are real, $k' > 0$, $0 < u_q - u_p < \pi$, and the weight functions $W(n)$, $\bar{W}(n)$ are positive real. Instead, here we shall focus (in Section 5) on the case when k', x, y are positive real and $1 < x < y < \infty$. For $N = 2$ this is the ferromagnetic case of the Ising model, but for $N > 2$ this does imply that $W(n)$, $\bar{W}(n)$ are complex. This is unphysical, but we feel the model is still interesting for various reasons. Its solution yields eigenvalues of the Hamiltonian H: this is Hermitian and physically interesting. The partition function satisfies $Z_a^* = Z_{-a} = Z_{N-a}$, so Z_0 is real and in fact positive, as are the Fourier coefficients of Z_a, which occur naturally and are defined in (3.2). Many equivalences have been discovered between lattice models: it is possible that this model, with complex weights, is in some sense equivalent to another model with real, positive weights. (Just as, for example, the staggered six-vertex model is equivalent to the usual q-state Potts model: cf. §§12.3, 12.4 of ref. 12, where λ is pure imaginary if $q < 4$.)

2.2. Eigenvalues of $\mathbf{T}_{\text{row}}$

The state of a row of L spins $\sigma_1,..., \sigma_L$ is specified by $\sigma = \{\sigma_1,..., \sigma_L\}$. There are N^L such states. Let $\mathbf{u}_j$ (for $j = 0,..., N-1$) be the N^L-dimensional vector with elements

$$(\mathbf{u}_j)_\sigma = \delta(\sigma_1, j) \cdots \delta(\sigma_L, j) \tag{2.7}$$

(i.e., its elements are zero unless $\sigma_1 = \sigma_2 = \cdots = \sigma_L = j$). Also, for $Q = 0, 1,..., N-1$, let

$$\mathbf{v}_Q = N^{-1/2} \sum_{j=0}^{N-1} \omega^{-Qj} \mathbf{u}_j \tag{2.8}$$

Successively premultiplying $\mathbf{v}_Q$ by $\mathbf{T}_{\text{row}}(x, y)$ (in general with different values of x and y, but the same value of k') generates a set of vectors which we can take to be the basis of a vector space V_Q. This will be a subspace of the full N^L-dimensional space on which $\mathbf{T}_{\text{row}}$ acts.

Define

$$G = -1 + k'(x^N + 1)/(x^N - 1) = (x^N y^N - 1)/(y^N - x^N) \tag{2.9}$$

$$\rho = (x-1)/(x^N - 1)^{1/N} \tag{2.10}$$

and, for $Q = 0, 1,..., N-1$, consider the expression

$$P(z^N) = z^{-Q} \sum_{n=0}^{N-1} \omega^{(Q+L)n} \{(z^N - 1)/(z - \omega^n)\}^L \tag{2.11}$$

The rhs of (2.11) is a polynomial in z^N, of degree

$$m = \text{integer part of } [(NL - L - Q)/N] \tag{2.12}$$

so $P(x)$ is a polynomial of this degree. Let its zeros be $z_1^N, z_2^N,..., z_m^N$ and define, for $j = 1,..., m$ and $Q = 0, 1,..., N-1$,

$$\cos\theta_j = (1 + z_j^N)/(1 - z_j^N) \tag{2.13}$$

$$C_Q = N^L x^{-Q} / \{(2k')^m (1 - x^{-N})^{L-m-L/N}\} \tag{2.14}$$

In ref. 4 it was shown that the eigenvalues $\Lambda_{\text{row}}(x, y)$ of $\mathbf{T}_{\text{row}}(x, y)$ with eigenvectors in V_Q are

$$\Lambda_{\text{row}}(x, y) = C_Q \rho^L \prod_{j=1}^{m} [G \pm (1 + k'^2 - 2k' \cos\theta_j)^{1/2}] \tag{2.15}$$

for $Q=0, 1,..., N-1$ and all 2^m independent choices of the signs of the square roots. This is the set of eigenvalues referred to above as S. (One moderately substantial step in the working was not explicitly given in ref. 4: for completeness we include it here in Appendix A.)

The functional method used in ref. 4 rigorously predicts that any eigenvalue in S must be of the form (2.15), but it tells us nothing about the eigenvectors (except that they depend on x and y only via k' and are the same for $\mathbf{T}_{\text{row}}$ as for $\mathscr{H}$), nor does it predict how many times any particular eigenvalue occurs, if at all. However, from small-L and numerical calculations on $\mathscr{H}$ (discussed in Appendix B), it appears that each eigenvalue occurs just once, and that V_Q is of dimension 2^m. Further, these calculations imply that V_Q is closed not only under multiplication by $\mathscr{H}$, but under multiplication by $\mathscr{H}_0$ and $\mathscr{H}_1$ separately; and that V_Q is the smallest such subspace. (In fact it was generated numerically by such multiplications.) This means that V_Q is independent of k', depending only on N and L.

The operators $\mathbf{T}_{\text{row}}$, $\mathscr{H}$, $\mathscr{H}_0$, and $\mathscr{H}_1$ can therefore be restricted to the subspace V_Q, wherein they are represented by 2^m by 2^m matrices, which we call $\hat{T}_{\text{row}}$, $\hat{\mathscr{H}}$, $\hat{\mathscr{H}}_0$, and $\hat{\mathscr{H}}_1$, respectively. From (2.5), $\mathscr{H}$ is Hermitian: provided we choose a k'-independent orthogonal basis for V_Q, $\hat{\mathscr{H}}$ must also be linear in k' and Hermitian.

From (2.15), the $\Lambda_{\text{row}}(x, y)$ are the eigenvalues of the 2^m by 2^m matrix

$$\mathscr{T}_Q = C_Q \rho^L \prod_{j=1}^{m} \{GI + (1 - k' \cos\theta_j)\, \sigma_j^z - k' \sin\theta_j \sigma_j^x\} \tag{2.16}$$

where σ_j^z, σ_j^x are the usual Pauli matrices, direct products of m two-by-two matrices:

$$\begin{aligned} \sigma_j^z &= e \otimes e \otimes \cdots \otimes \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \otimes e \otimes \cdots \otimes e \\ \sigma_j^x &= e \otimes e \otimes \cdots \otimes \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \otimes e \otimes \cdots \otimes e \end{aligned} \tag{2.17}$$

e is the identity two-by-two matrices and the nonidentity matrices in position j, for $j=1,..., m$. Hence $\hat{\mathbf{T}}_{\text{row}}$ is related to $\mathscr{T}_Q$ by a similarity transformation:

$$\hat{\mathbf{T}}_{\text{row}} = \mathbf{P} \mathscr{T}_Q \mathbf{P}^{-1} \tag{2.18}$$

Because of the commutation relations, the matrix $\mathbf{P}$ can be chosen to depend on x, y only via k'. Taking the Hamiltonian limit x, y, $\mu \to 1$ (keeping k' fixed), it follows that

$$\hat{\mathscr{H}} = \mathbf{P} H \mathbf{P}^{-1} \tag{2.19}$$

where

$$H = [2k'Q + (1+k')(mN - LN + L)]\, I$$
$$- N \sum_{j=1}^{m} [(1 - k' \cos\theta_j)\, \sigma_j^z - k' \sin\theta_j\, \sigma_j^x] \tag{2.20}$$

Since $\hat{\mathscr{H}}$ is Hermitian for all real k, its eigenvalues must be real, and hence $\theta_1,..., \theta_m$ must be real. Thus, H is also Hermitian, and $\mathbf{P}$ can be chosen unitary.

Note that both $\hat{\mathscr{H}}$ and H are linear in k' (remember that $\theta_1,..., \theta_m$ are independent of k'). We want to assert that $\mathbf{P}$ must be independent of k'. Certainly this is the most obvious way of ensuring that (2.19) is satisfied for all k', and it was a basic and consistent feature of our numerical calculations (see Appendix B). We do not yet have a proof, but shall assume that it is so.

2.3. Ground States of $\mathscr{H}_0$ and H_0

From (2.5) and (2.6), and the definitions in ref. 3 of Z_j, $\mathscr{H}_0$ is a diagonal matrix whose diagonal entry corresponding to the row spin-state $\sigma = \{\sigma_1,..., \sigma_L\}$ is

$$f(\sigma_1 - \sigma_2) + f(\sigma_2 - \sigma_3) + \cdots + f(\sigma_L - \sigma_1)$$

where

$$f(\sigma) = -2 \sum_{n=1}^{N-1} \omega^{nj}/(1 - \omega^{-n}) = 1 - N + 2 \bmod(j, N) \tag{2.21}$$

$\bmod(j, N)$ is the value of j modulo N: $0 \leqslant \bmod(j, N) \leqslant N-1$.

It follows that the lowest eigenvalue (ground state) of $\mathscr{H}_0$ is $-L(N-1)$, the corresponding eigenvectors having nonzero elements only for the row spin states $\sigma_1 = \sigma_2 = \cdots = \sigma_L$. These are the vectors $\mathbf{u}_0,..., \mathbf{u}_{N-1}$, or (taking independent linear combinations) $\mathbf{v}_0,..., \mathbf{v}_{N-1}$. In fact, v_Q is the *unique* ground-state eigenvector of $\mathscr{H}_0$ in the subspace V_Q.

Similarly, taking H_0 to be the k'-independent component of H, we see at once from (2.20) and (2.17) that this is also diagonal, with unique lowest eigenvalue $-L(N-1)$ [in agreement with the fact that H_0 is related to $\mathscr{H}_0$ by the similarity transformation (2.19)]. The corresponding ground-state eigenvector is

$$\xi = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \otimes \begin{pmatrix} 1 \\ 0 \end{pmatrix} \otimes \cdots \otimes \begin{pmatrix} 1 \\ 0 \end{pmatrix} \tag{2.22}$$

(a direct product of m vectors, each two-dimensional).

Since H is related to $\hat{\mathscr{H}}$ (the restriction of $\mathscr{H}$ to the subspace V_Q) by the similarity transformation (2.19), it follows that when $k'=0$, ξ must be correspondingly related to V_Q. In fact

$$\xi = \mathbf{P}^{-1}\hat{v}_Q \tag{2.23}$$

$\hat{v}_Q$ is the coordinate vector of v_Q with respect to the orthogonal basis of V_Q.

Note that ξ, V_Q, the basis of V_Q, and hence $\hat{v}_Q$, are independent of k'. If, as we assume, $\mathbf{P}$ is also independent of k', then (2.23) is true not only for $k'=0$, but for all values of k'. We shall use this relation in the next section.

3. PARTITION FUNCTION

Using (2.7), we find the partition function Z_a of the model on the lattice of Fig. 1,

$$Z_a = \mathbf{u}_a^{\dagger}[\mathbf{T}_{\text{row}}(x, y)]^M \mathbf{u}_0 \tag{3.1}$$

(For the moment we consider a homogeneous model, where x, y, k' are the same for all edges.) From (2.8) it follows that

$$Z_a = N^{-1}\sum_{Q=0}^{N-1} \omega^{-Qa}\tilde{Z}_Q \tag{3.2}$$

where $\tilde{Z}_Q$, the Fourier transform of Z_a, is

$$\tilde{Z}_Q = N^{1/2}\mathbf{v}_Q^{\dagger}[\mathbf{T}_{\text{row}}(x, y)]^M \mathbf{u}_0 = \sum_{Q'=0}^{N-1} \mathbf{v}_Q^{\dagger}[\mathbf{T}_{\text{row}}(x, y)]^M \mathbf{v}_{Q'} \tag{3.3}$$

The vector $[\mathbf{T}_{\text{row}}(x, y)]^M v_{Q'}$ lies in V_Q. All such vectors are eigenvectors of the spin translation operator R (that increases all spins by one, modulo N), with eigenvalue ω^Q. Hence this vector is orthogonal to v_Q unless $Q'=Q$, so

$$\tilde{Z}_Q = \mathbf{v}_Q^{\dagger}[\mathbf{T}_{\text{row}}(x, y)]^M \mathbf{v}_Q \tag{3.4}$$

We can replace $\mathbf{T}_{\text{row}}(x, y)$ and $\mathbf{v}_Q$ by their 2^m-dimensional representatives $\hat{\mathbf{T}}_{\text{row}}(x, y)$ and $\hat{\mathbf{v}}_Q$ with respect to an orthogonal basis of V_Q. Making the similarity transformation (2.18), using our result (2.23), it follows that

$$\tilde{Z}_Q = \xi^{\dagger}\mathscr{T}_Q^M \xi \tag{3.5}$$

The matrix $\mathscr{T}_Q$ is given by (2.16), ξ by (2.22). The rhs of (3.5) therefore factors into a product of m terms, each involving only two-dimensional vectors and matrices. Define

$$g = \frac{N(1-x^{-1})}{1-x^{-N}} \tag{3.6}$$

$$B(x, y, \theta) = \frac{1-x^{-N}}{2k'} \begin{pmatrix} G+1-k'\cos\theta & -k'\sin\theta \\ -k'\sin\theta & G-1+k'\cos\theta \end{pmatrix} \tag{3.7}$$

$$D(\cos\theta) = (1, 0)\, B(x, y, \theta)^M \begin{pmatrix} 1 \\ 0 \end{pmatrix} \tag{3.8}$$

[remember from (2.4) that k' is a function of x and y]. Then (3.5) yields, after using the definitions of ρ and C_Q,

$$\tilde{Z}_Q = g^{LM} x^{-QM} D(\cos\theta_1)\, D(\cos\theta_2) \cdots D(\cos\theta_m) \tag{3.9}$$

We can evaluate the matrix B to the power M in (3.8). Let $\Delta(c)$ be the function

$$\Delta(c) = (1 + k'^2 - 2k'c)^{1/2} \tag{3.10}$$

Then

$$\begin{aligned} D(c) = (1-x^{-N})^M \{ &[G+\Delta(c)]^M\, [\Delta(c)+1-k'c] \\ &+ [G-\Delta(c)]^M\, [\Delta(c)-1+k'c] \}/[2^{M+1} k'^M \Delta(c)] \end{aligned} \tag{3.11}$$

The expression (3.9) is thus an explicit and tractable expression for $\tilde{Z}_Q$, the Fourier transform of the partition function Z_a of the finite lattice of Fig. 1. It gives Eq. (55) of ref. 6.

3.1. An Alternative Form

The lattice has L columns and M rows. We have used the row-to-row transfer matrix to obtain (3.9). For that reason, M enters only as an explicit exponent in (3.9) and (3.11), arising from the product of the M transfer matrices. On the other hand, L enters implicitly and at a more basic level via the definitions (2.11)–(2.13) of $\theta_1, \ldots, \theta_m$.

We can transform the result so as to interchange these roles of L and M. The rhs of (3.11) is an even polynomial in $\Delta(c)$, so from (3.10), $D(c)$ is a polynomial of degree

$$r = \text{integer part of } [(M+1)/2] \tag{3.12}$$

Let $c_1,..., c_r$ be the zeros of $D(c)$. Then $\exists$ a constant γ such that

$$D(c) \equiv \gamma(c_1 - c)(c_2 - c) \cdots (c_r - c) \tag{3.13}$$

Substituting this into (3.9), we get

$$\tilde{Z}_Q = g^{LM} x^{-QM} \gamma^m \prod_{i=1}^{r} \prod_{j=1}^{m} (\cos \theta_j - c_i) \tag{3.14}$$

This is a "double product" expression for the Fourier transform of the partition function, rather like the Pfaffian results for the Ising model [Eq. (2.23), p. 85 of ref. 13; see also ref. 14].

From (3.11),

$$D(1) = x^{-MN}, \qquad D(-1) = 1 \tag{3.15}$$

Taking ratios and using (3.13), we obtain

$$\prod_{i=1}^{r} (c_i - 1)/(c_i + 1) = D(1)/D(-1) = x^{-MN} \tag{3.16}$$

We define

$$\zeta_i = [(c_i - 1)/(c_i + 1)]^{1/N} \tag{3.17}$$

Then from (3.16) we can choose the Nth roots so that

$$\zeta_1 \zeta_2 \cdots \zeta_r = x^{-M} \tag{3.18}$$

Using (2.13) and these results, we can now write (3.14) as

$$\tilde{Z}_Q = g^{LM} x^{-QM} \prod_{i=1}^{r} \prod_{j=1}^{m} \frac{\zeta_i^N - z_j^N}{1 - z_j^N} \tag{3.19}$$

Remembering that $z_1^N,..., z_m^N$ are the zeros of the polynomial $P(z)$, the j product is $P(\zeta_i^N)/P(1)$. Using (2.11) and (3.18), it follows that

$$\tilde{Z}_Q = g^{LM} \prod_{i=1}^{r} N^{-L} \sum_{n=0}^{N-1} \omega^{(Q+L)n} \left[\frac{\zeta_i^N - 1}{\zeta_i - \omega^n} \right]^L \tag{3.20}$$

Together with (3.2), this is the result (60) quoted in ref. 6.

To within simple normalization factors, $\tilde{Z}_Q$ is the resultant (§3.2. of ref. 15) of the polynomials $P(u)$ and $(1-u)^r D[(1+u)/(1+u)]$, of degrees m and r, respectively.

4. EIGENVALUES OF $\mathbf{T}_{\text{col}}$

Expanding the product in (3.20), we can write the rhs as a multiple sum over $n_1,..., n_r$, Q entering only via a factor $\omega^{Q(n_1+\cdots+n_r)}$. Using (3.2), if follows that

$$Z_a = g^{LM} \sum (\lambda_{1,n_1}\lambda_{2,n_2}\cdots\lambda_{r,n_r})^L \tag{4.1}$$

where

$$\lambda_{i,n} = N^{-1}(1-\zeta_i^N)/(1-\zeta_i\omega^{-n}) \tag{4.2}$$

and the sum is over all integers $n_1,..., n_r$ such that $0 \leqslant n_i < N$ and

$$n_1 + n_2 + \cdots + n_r = a, \quad \text{mod } N \tag{4.3}$$

Let $\mathbf{n} = \{n_1,..., n_r\}$ be the set of all such integers, and define

$$\Lambda_{\mathbf{n}} = g^M \lambda_{1,n_1}\lambda_{2,n_2}\cdots\lambda_{r,n_r} \tag{4.4}$$

Then (4.1) can be written quite simply as

$$Z_a = \sum_{\mathbf{n}} (\Lambda_{\mathbf{n}})^L \tag{4.5}$$

and $\Lambda_{\mathbf{n}}$ is independent of L.

Let $\mathbf{T}_{\text{col}}(a)$ be the column-to-column transfer matrix of this model. It is independent of L, of dimension N^{r-1}, and because of the periodic boundary condition linking column L to column 1,

$$Z_a = \text{Trace}[\mathbf{T}_{\text{col}}(a)]^L \tag{4.6}$$

Comparing (4.5) and (4.6), it is obvious that the only way they can both be true for all positive integers L is for the $\Lambda_{\mathbf{n}}$ to be the eigenvalues of $\mathbf{T}_{\text{col}}(a)$.

It is helpful to define an N^r-by-N^r matrix $\mathscr{T}_{\text{col}}$ which is block-diagonal, its N diagonal blocks being $\mathbf{T}_{\text{col}}(0),..., \mathbf{T}_{\text{col}}(N-1)$. This can be thought of as the column-to-column transfer matrix when the spins in the bottom row of Fig. 1 are fixed to be equal, without assigning their value. The eigenvalues of this matrix are given by (4.4) for all values 0,..., $N-1$ of $n_1, n_2,..., n_r$, *without* the restriction (4.3). Thus, the diagonal form of $\mathscr{T}_{\text{col}}$ is a direct product of r diagonal matrices, each N by N:

$$\mathscr{T}_{\text{col}} = g^M \begin{pmatrix} \lambda_{10} & & & \\ & \lambda_{11} & & \\ & & \ddots & \\ & & & \lambda_{1,N-1} \end{pmatrix} \otimes \cdots \otimes \begin{pmatrix} \lambda_{r0} & & & \\ & \lambda_{r1} & & \\ & & \ddots & \\ & & & \lambda_{r,N-1} \end{pmatrix} \tag{4.7}$$

This is a fascinating simplification and is very reminiscent of the Ising and free-fermion models, whose transfer matrices involve direct products of two-by-two matrices. (We shall return to this point toward the end of Section 7.) Note that here we are considering the full eigenvalue spectrum of $\mathscr{T}_{\text{col}}$: it is also true that the set of eigenvalues (2.15) of $\mathbf{T}_{\text{row}}$ is that of a direct product of m matrices (each two-by-two), but these are the eigenvalues only for the subspace V_Q.

5. DUALITY

Let $W_f(n)$, $\bar{W}_f(n)$ be the Fourier transforms of $W(n)$, $\bar{W}(n)$, defined slightly asymmetrically by

$$\begin{aligned} W_f(j) &= N^{-1/2} \sum_{n=0}^{N-1} \omega^{-jn} W(n) \\ \bar{W}_f(j) &= N^{-1/2} \sum_{n=0}^{N-1} \omega^{jn} \bar{W}(n) \end{aligned} \tag{5.1}$$

From (2.1) we may note that

$$\mu(1 - y\omega^n)\, \bar{W}(n) = (\omega - x\omega^n)\, \bar{W}(n-1) \tag{5.2}$$

Multiplying this equation by $N^{-1/2}\omega^{jn}$ and summing over n, we find

$$\bar{W}_f(n) = \bar{W}_f(0) \prod_{j=1}^{n} \frac{\mu - \omega^j}{\mu y - x\omega^j} \tag{5.3a}$$

and similarly

$$W_f(n) = W_f(0) \prod_{j=1}^{n} \frac{\omega\mu y - x\omega^j}{\mu - \omega^j} \tag{5.3b}$$

Comparing these equations with (2.1), we see that if we replace x, y, μ by their dual values

$$x_d = x/(\mu y), \qquad y_d = \mu^{-1}, \qquad \mu_d = y^{-1} \tag{5.4}$$

then (2.2) remains satisfied, while (to within n-independent normalization factors) $W(n)$ is interchanged with $\bar{W}_f(n)$, and $\bar{W}(n)$ with $W_f(n)$.

This is a duality transformation. From (2.4) and (2.9), the duals of k' and G are

$$k'_d = 1/k', \qquad G_d = G/k' \tag{5.5}$$

Thus this transformation interchanges the high-temperature disordered phase ($k' > 1$) with the low-temperature ordered phase ($k' < 1$).

We can obtain its effect on the partition function Z (including boundary conditions) by using standard graphical methods (§2 of ref. 16). Let $Z_a^{(d)}$ be the value of Z_a as defined by (2.1)–(2.3), after replacing x, y, μ by x_d, y_d, μ_d. The effect of this is to replace $W(n)$ by $\bar{W}_f(n)/\bar{W}_f(0)$, $\bar{W}(n)$ by $W_f(n)/W_f(0)$, so

$$Z_a^{(d)} = \Omega^{-LM} Z_a^{(f)}, \tag{5.6}$$

where

$$\Omega = W_f(0)/\bar{W}_f(0) \tag{5.7}$$

and $Z_a^{(f)}$ is defined by (2.3), but with E being the function $\bar{W}_f(W_f)$ on SW–NE (SE–NW) edges.

From (5.1)

$$\sum_{j=0}^{N-1} W_f(j)\, \bar{W}_f(j) = \sum_{n=0}^{N-1} W(n)\, \bar{W}(n) \tag{5.8}$$

while from (2.1), (5.3), and (5.7), we have the simple relations

$$\begin{aligned} W(n)\, \bar{W}(n) &= \frac{x-1}{x-\omega^{-n}}, \\ W_f(j)\, \bar{W}_f(j) &= \Omega \frac{x-\mu y}{x-\mu y \omega^{-j}} \end{aligned} \tag{5.9}$$

for $n, j = 0, \ldots, N-1$. Substituting these expressions into (5.8), we get

$$\Omega = \frac{x-1}{x-\mu y} \frac{x^N - \mu^N y^N}{x^N - 1} \tag{5.10}$$

Let $\mathscr{L}$ be the lattice of Fig. 1, and $\mathscr{L}_D$ its dual, as shown in Fig. 2. We now formulate $Z_a^{(f)}$ in terms of "edge spins" on $\mathscr{L}_D$ as follows. With each edge λ of $\mathscr{L}$, with endspins b, c, (b below c), associate an "edgespin" $e_\lambda = b - c$ (modulo N) on the corresponding edge of $\mathscr{L}_D$.

Then these edgespins e_λ satisfy the constraints

$$e_1 + e_4 = e_2 + e_3, \quad \text{mod } N \tag{5.11a}$$

at each internal vertex of $\mathscr{L}_D$, where $e_1, \ldots, e_4$ are arranged as in Fig. 2. At the boundary vertices (of valence 2), with adjacent edge spins e_1, e_2, this

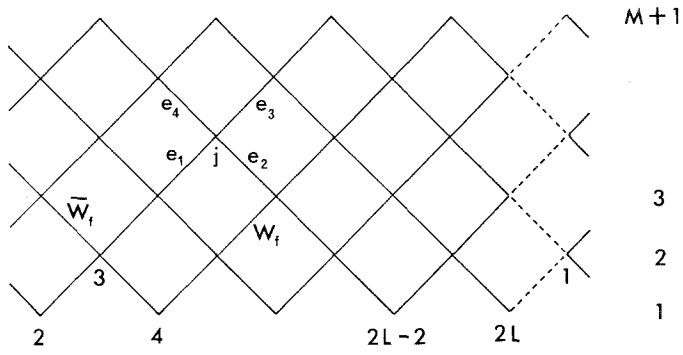

Fig. 2. The dual lattice $\mathscr{L}_D$, showing a typical arrangement of edge spins $e_1,...,e_4$ around a vertex j. The broken edges link column $2L$ to column 1, and form a vertical seam through $\mathscr{L}_D$.

constraint simplifies to $e_1 = e_2$. Also, for a vertical seam of M edges (e.g., the edges denoted by broken lines in Fig. 2), we must have

$$e_1 + e_2 + \cdots + e_M = -a, \quad \text{mod } N \tag{5.11b}$$

$e_1, ..., e_M$ are the spins on those edges.

Because of (5.11a), if (5.11b) is true for any seam, then it is true for all. The top and bottom edges of $\mathscr{L}_D$ each correspond to two edges of $\mathscr{L}$, but theen site-spin configurations on $\mathscr{L}$ and allowed edge-spin configurations on $\mathscr{L}_D$, so

$$Z_a^{(f)} = \sum_e \prod_i E(e_i) \tag{5.12}$$

the sum being over all allowed edgespin configurations $e = \{e_1, e_2,...\}$ on $\mathscr{L}_D$, the product being over all edges i of $\mathscr{L}_D$, $E(e_i)$ being $\bar{W}_f(e_i)$ or $W_f(e_i)$.

First consider the constraint (5.11b). This is most easily handled by introducing the Fourier transform $\tilde{Z}_Q^{(f)}$ of $Z_a^{(f)}$ analogously to (3.2). Then,

$$\tilde{Z}_Q^{(f)} = \sum_{a=0}^{N-1} \omega^{Qa} Z_a^{(f)} \tag{5.13}$$

If we use (5.11b) to replace Q_a by $-Q(e_1 + \cdots + e_m)$, then the effect of the sum over a is to remove the constraint (5.11b), giving

$$\tilde{Z}_Q^{(f)} = \sum_e \omega^{-Q(e_1 + \cdots + e_m)} \prod_i E(e_i) \tag{5.14}$$

Here the sum and product have the same meaning as in (5.12), except that the constraint (5.11b) is removed.

The next step is to expand each of the vertex and seam Kronecker deltas into a Fourier series, e.g.,

$$\delta(e_1+e_4-e_2-e_3)=N^{-1}\sum_{b=0}^{N-1}\omega^{b(e_2+e_3-e_1-e_4)} \tag{5.15}$$

Doing this, and expanding, we find that the summand in (5.14) itself becomes a sum of N^{LM+L} terms involving $b_1,\ldots,b_{LM+L}$, each b_i corresponding to a vertex of $\mathscr{L}_D$. For each such term, the outer e_λ sum in (5.14) is trivial. For instance, for a SW–NE edge λ not in the vertical seam, the corresponding e_λ enters the summand only in a factor

$$\omega^{(b-c)e_\lambda}W_f(e_\lambda) \tag{5.16}$$

b (c) being the b_i associated with the vertex below (above) λ. Summing this over e_λ gives, from (5.1), $N^{1/2}W(b-c)$. Similarly, each SE–NW edge (not on the seam) gives a contribution $N^{1/2}\bar{W}(b-c)$.

For the edges on the seam, there is an extra factor ω^{-Qe_λ}, causing the contributions to be modified to $N^{1/2}\bar{W}(b-c+Q)$.

We still have to sum over the b_i. There are $LM+L$ such b_i, one for each vertex of $\mathscr{L}_D$. We can obviously regard them as new "spins": they take the values $0,\ldots,N-1$. Each vertex gives a factor N^{-1}, while each of the $2LM$ edges gives a factor $N^{1/2}$. Altogether, it follows that

$$\tilde{Z}_Q^{(f)}=N^{-L}\sum_b\prod_{\langle i,j\rangle}E(b_i-b_j) \tag{5.17}$$

where the sum is over all values of all the spins $b=\{b_1,b_2,\ldots\}$, the product is over all edges $\langle i,j\rangle$ of $\mathscr{L}_D$, $E(n)$ is $W(n)$ on SW–NE edges, while on SE–NW it is $\bar{W}(n)$. For edges on the seam, $W(n)$ and $\bar{W}(n)$ are to be replaced by $W(n-Q)$, $\bar{W}(n+Q)$, respectively.

From Fig. 2, $\mathscr{L}_D$ has $2L$ individual columns, of alternating type. If we take the seam to be between columns $2L$ and 1, the seam modification is equivalent to using "Q-cyclic" boundary conditions, where the spins σ_{2L+1} in column $2L+1$ (the one to the right of L) are related to the corresponding spins σ_1 in column 1 (the far left column) by

$$\sigma_{2L+1}=\sigma_1+Q \tag{5.18}$$

The sum in (5.17) is thus very like our original partition function. To be precise, it is the partition function Z_Q^{free} defined by (2.1)–(2.3), but with free boundary conditions on the top and bottom rows, and with the Q-cyclic boundary condition (5.18) linking the last and first columns.

Using (3.2) and (5.6), it follows that $\tilde{Z}_Q$, with x, y, μ replaced by x_d, y_d, μ_d, is

$$\tilde{Z}_Q^{(d)} = N^{-L}\Omega^{-LM} Z_Q^{\text{free}} \tag{5.19}$$

Thus we can go from the fixed-spin top and bottom boundary conditions of this paper to free-spin boundary conditions by applying this duality transformation. We remark in Section 7 that we can also do this by "turning off" the interactions in the top and bottom rows of edges, so we have a useful consistency check on our results. We shall also use the duality transformation to obtin the vertical interfacial tension.

6. LARGE-LATTICE LIMIT

Here we take k', x, y to be real, $k' > 0$, $1 < x < y < \infty$. For $N = 2$ this is the ferromagnetic case of the Ising model. The system is ferromagnetically ordered if $k' < 1$, disordered if $k' < 1$, so we can think of k' as a "temperature." The variable G is positive,

$$G > 1 + k' \tag{6.1}$$

and we shall use the quantity

$$w = \frac{G - |1 - k'|}{G + |1 - k'|} \tag{6.2}$$

Provided M is sufficiently large, the zeros $c_1, \ldots, c_r$ of the polynomial $D(c)$ are real. For $0 < k' < 1$ they all lie in

$$\tfrac{1}{2}(k' + k'^{-1}) < c_j < \infty \tag{6.3}$$

while for $k' > 1$ all but one of them lie in this range, the exceptional zero being (for M large)

$$c_1 = 1 + 2(k' - 1)^2 \, w^M / k'^2 + \text{smaller terms} \tag{6.4}$$

Thus, c_1 tends to 1 as $M \to \infty$.

Using (3.15), it follows that $D(c)$ is real and positive for $-1 \leqslant c \leqslant 1$, so from (3.9) and (3.2), $\tilde{Z}_Q$ and Z_0 are positive real, even though the weight functions $W(n)$, $\bar{W}(n)$ are complex. We remarked on this in Section 2.

In (2.11) let

$$z = e^u \tag{6.5}$$

and consider the function

$$\Psi(u) = P(e^{Nu}) = e^{-Qu} \sum_{n=0}^{N-1} \omega^{Qn} \left\{ \frac{1 - e^{Nu}}{1 - \omega^{-n} e^{u}} \right\}^{L} \tag{6.6}$$

This has simple zeros at $u_1, \ldots, u_m$, where $z_j = \exp(u_j)$, and we can choose $u_1, \ldots, u_m$ all to have imaginary part π/N. Thus, they lie on the line AB in Fig. 3. From (2.13), we can write (3.9) exactly as

$$\ln \tilde{Z}_Q = LM \ln g - QM \ln x + \frac{1}{2\pi i} \oint_C \ln D(\cos\theta) [\Psi'(u)/\Psi(u)] \, du \tag{6.7}$$

where θ, u are related by

$$\cos\theta = (1 + e^{Nu})/(1 - e^{Nu}) \tag{6.8}$$

and C is the contour C_1 shown in Fig. 3. It just encloses the line AB.

The function $\ln D(\cos\theta)$ is analytic on AB, and it follows that if we ignore (for M large) terms of relative order w^M, then we can take

$$D(c) = (1 - x^{-N})^M [G + \Delta(c)]^M [\Delta(c) + 1 - k'c]/[2^{M+1} k'^M \Delta(c)] \tag{6.9}$$

6.1. The Case $k' < 1$

First consider the case $k' < 1$. The integrand of (6.7) is then a periodic function of u, of period $2\pi i/N$. It is analytic in the domain $-\pi/N < \mathrm{Im}\, u < 2\pi/N$, apart from simple zeros on AB and a branch cut

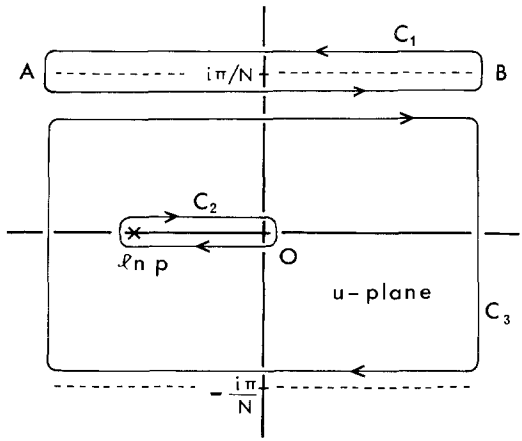

Fig. 3. The contours C_1, C_2, C_3 in the complex u plane. C_1 and C_3 extend to infinity to the left and right.

[where $\Delta(\cos\theta)$ is pure imaginary] on the negative real axis, along the segment

$$\ln p < u < 0 \tag{6.10}$$

where

$$p = \left|\frac{1-k'}{1+k'}\right|^{2/N} \tag{6.11}$$

From (2.9) we can verify that $D(-1)=1$, from which it follows that the integrand of (6.7) tends to zero as $\mathrm{Re}(u)\to+\infty$. Also, $\Psi'(u)/\Psi(u)\sim e^{Nu}$ as $u\to-\infty$, and hence the integrand also tends to zero as $\mathrm{Re}(u)\to-\infty$.

It follows that C_1 and C_2 in Fig. 3 are equivalent contours, so we can take C in (6.7) to be C_2, where C_2 goes clockwise around the branch cut.

Now consider $\Psi(u)$ on C_2. For L large it is dominated by the $n=0$ term in (6.6), the other terms being at most of relative order $\omega^{Qn}[(1-p)/(1-p\omega^{-n})]^L$. Ignoring these terms, we can take

$$\Psi(u) = e^{-Qu}[(1-e^{Nu})/(1-e^{u})]^L$$

Finally, we expand $C=C_2$ to become the closed contour C_3, enclosing the rectangle $-\pi/N<\mathrm{Im}\,u<\pi/N$. Taking account of the contribution from $\mathrm{Re}\,u=-\infty$, (6.7) gives, for $0<k'<1$,

$$\ln \tilde{Z}_Q = -LMf - QMh - 4Lf_s \tag{6.12}$$

where

$$\begin{aligned} -f &= \ln(N\rho) + \frac{1}{\pi}\int_0^{\pi(1-N^{-1})} \ln\left[\frac{G+\Delta(\cos\theta)}{2k'}\right] d\phi \\ h &= \frac{1}{N}\ln\frac{G+|1-k'|}{G+1-k'} \\ -f_s &= \frac{1}{4\pi}\int_0^{\pi(1-N^{-1})} \ln\left[\frac{1-k'\cos\theta+\Delta(\cos\theta)}{2\Delta(\cos\theta)}\right] d\phi \end{aligned} \tag{6.13}$$

Here ϕ and θ are related by

$$\tan\frac{\theta}{2} = \left[\frac{\sin\phi}{\sin(\phi+\pi/N)}\right]^{N/2} \tag{6.14}$$

and θ increases from 0 to π as ϕ increases from 0 to $\pi(1-N^{-1})$.

6.2. The Case $k' > 1$

This case is more complicated, as the factor $\Delta(c) + 1 - k'c$ vanishes when $\mathrm{Re}\, u \to -\infty$. The integrand of (6.7) has an extra branch cut from $u = -\infty$ to $u = \ln p$. We can integrate around this as necessary, using the fact that across this cut $\ln D(\cos\theta)$ merely jumps by $2\pi i$ (from above to below). One also has to be careful with convergence when $u \to -\infty$. The end result is that $\ln \tilde{Z}_Q$ is still given by (6.12)–(6.14), except that there is an additional contribution

$$\ln P(0) - \frac{2Q}{N} \ln\left(\frac{k'}{k'-1}\right) \tag{6.15}$$

to $\ln \tilde{Z}_Q$. Here $P(0)$ is the value of (2.11) when $z \to 0$: equal to $N\,L!/[Q!\,(L-Q)!]$.

6.3. Partition Function Z_a

We can now start to calculate the partition function Z_a, using (3.2). For $k' < 1$, h is zero and the rhs of (5.12) is independent of Q. Thus, we get at this stage

$$Z_a = \delta_{a,0} e^{-LMf - 4Lf_s} \tag{6.16}$$

For $k' > 1$, $h > 0$ and the summand in (3.2) is dominated (for M large) by the $Q = 0$ term. The $Q = 1$ term gives an exponentially small correction

$$Z_a = e^{-LMf - 4Lf_s}\left\{1 + L\omega^{-a}\left(\frac{k'-1}{k'}\right)^{2/N} e^{-Mh}\right\} \tag{6.17}$$

Of course, $Z_1, \ldots, Z_{N-1}$ are not precisely zero for $k' < 1$, merely much smaller than Z_0. We can determine their ratios to Z_0 from (4.1)–(4.5). For L large, the sums in (4.1) and (4.5) are dominated by the largest eigenvalue $\Lambda_{\mathbf{n}}$, where $n_1, \ldots, n_r$ must satisfy the restriction (4.3).

For $a = 0$ the largest eigenvalue occurs when $n_1, \ldots, n_r$ are zero (because $|\lambda_{i,n}| < |\lambda_{i,0}|$ for $n = 1, \ldots, N-1$). For $a \neq 0$ at least one of $n_1, \ldots, n_r$ must be nonzero, and the maximum eigenvalue is achieved when the corresponding ζ_i is as small as possible. For $k' < 1$ and M large, this is when the corresponding c_i, related to ζ_i by (3.17), is at the lower end of the range (6.3), i.e., $c_i = (k' + 1/k')/2$ and $\zeta_i = p$, using (6.11). Thus,

$$\frac{Z_a}{Z_0} = \tau_M^{l_a}\left\{\prod_n \frac{1-p}{1-p\omega^{-n}}\right\}^L \tag{6.18}$$

Here τ_M is a factor arising from the number of choices of the ζ_i that are close to p: we expect it to be proportional to M for M large. The product is over a set of nonzero integers n whose sum is a (modulo N): this set is chosen to maximize the modulus of the product in (6.18). The integer l_a is the number of terms in the set, so $l_0 = 0$ and for $a = 1,..., N-1$ we must have $0 < l_a \leqslant \min(a, N-a)$. Often (but not always) there is just one term in the product, in which case $l_a = 1$, $n = a$, and (for $k' < 1$ and $a \neq 0$)

$$Z_a = \tau_M \left(\frac{1-p}{1-p\omega^{-a}} \right)^L e^{-LMf - 4Lf_s} \tag{6.19}$$

The situation is different for $k' > 1$. Then the smallest ζ_i corresponds to the exceptional c_i (labeled c_1) given by (6.4). This means that ζ_1 is exponentially small for M large and (to relative order ζ_1)

$$\begin{aligned} Z_a/Z_0 &= 1 + L\zeta_r(\omega^{-a} - 1) \\ &= 1 + L(\omega^{-a} - 1)[(k'-1)^2 w^M/k'^2]^{1/N} \end{aligned} \tag{6.20}$$

in agreement with (6.17).

6.4. Correlation Lengths and Interfacial Tensions

There are other exponentially small corrections to $\ln Z_a$. For $k' < 1$ we do not have the correction given by (6.20), but we do have corrections to (3.9) arising from including the second component of $D(c)$ in (3.11): the component proportional to $[G - \Delta(c)]^M$. These are largest when $\cos\theta_j$ is close to -1, giving corrections proportional to Lw^M.

Other corrections, exponentially small in L, come from including the next largest eigenvalues in the sum in (4.5). Because of the constraint (4.3), these are obtained by changing *two* of $n_1,..., n_r$ (say replacing n_1, n_2 by $n_1 + 1, n_2 - 1$). To minimize the effect of this change, the corresponding two ζ_i (ζ_1 and ζ_2) should be the smallest possible. For $k' < 1$ this means that both are effectively equal to p. For $k' > 1$ one is equal to p, while the other is the exceptional ζ_i (namely ζ_1), which is effectively zero.

For $a \neq 0$ and $k' < 1$ the maximum eigenvalue $\Lambda_{\mathbf{n}}$ is but one of a band, and the question of excitations (next lower eigenvalues) is complicated. If we ignore this case, the maximum eigenvalue has $n_1,..., n_r = a, 0,..., 0$, while the next maximum has $n_1,..., n_r = a \pm 1, \mp 1, 0,..., 0$. If we define

$$v_\pm = (1-p)/(1-p\omega^{\pm 1}) \tag{6.21}$$

then for $k' < 1$ the correction terms are proportional to $M^2(v_+ v_-)^L$, while for $k' > 1$ they are proportional to $M(v_+)^L$ and $M(v_-)^L$. Altogether, we get

$$\ln Z_a = -LMf - 2L(2f_s + s_a) + l_a \ln \tau_M + O(Le^{-M/\xi_{\text{vert}}}) + O(M^{\varepsilon(k')} e^{-2L/\xi_{\text{hor}}}) \tag{6.22}$$

where f and f_s are given by (6.13), the integer l_a is defined above, τ_M is proportional to M (for M large), and for $k' < 1$

$$s_a = -\frac{1}{2} \ln \frac{1-p}{1-p\omega^{-a}}, \qquad \varepsilon(k') = 2$$

$$e^{-1/\xi_{\text{vert}}} = w = \frac{G - |1-k'|}{G + |1-k'|} \tag{6.23a}$$

$$e^{-1/\xi_{\text{hor}}} = (v_+ v_-)^{1/2}$$

while for $k' > 1$

$$s_a = \ln \tau_M = 0, \qquad \varepsilon(k') = 1$$

$$e^{-1/\xi_{\text{vert}}} = w^{1/N} = \left[\frac{G - |1-k'|}{G + |1-k'|}\right]^{1/N} \tag{6.23b}$$

$$e^{-1/\xi_{\text{hor}}} = (v_\pm)^{1/2}$$

This f is the bulk free energy per site, f_s is the surface free energy per unit length, due to the fixed-spin top and bottom boundary conditions. (We can use the duality mapping $x, y, \mu \to x_d, y_d, \mu_d$ of Section 5 to transform f_s to the surface free energy due to free boundaries.) We take the unit length to be a single row or column spacing, so in these units the lattice is $2L$ wide, M high.

The quantity ξ_{hor} is the horizontal correlation length (see Section 7.10 of ref. 10); s_a is the tension energy per unit length of a horizontal interface between ferromagnetically ordered domains in phases 0 and a. If the product in (6.18) contains more than one term, then s_a should be replaced by the corresponding sum $\sum s_n$, where $\sum n = a$: this will mean that it is energetically favorable for other phases to intervene between 0 and a, i.e., for "wetting"[11] to occur. This happens when N is sufficiently large.

When $x^N = 1 + k'$ and $y = \infty$ the transfer matrix $\mathbf{T}_{\text{row}}$ is Hermitian. In ref. 4 it is asserted that in this case ξ_{vert} is the vertical correlation length, implying the spin–spin correlation length. This is true for $N = 2$, but is not necessarily true for larger N: we have not evaluated all the eigenvalues of $\mathbf{T}_{\text{row}}$ (only those with eigenvectors in $V_0, \ldots, V_{N-1}$), nor can we argue that the vertical spin–spin correlation length should depend only on k' and N

(not on x). We no longer have the "difference property,"[1] so vertical correlations can depend on the intervening horizontal rapidities, i.e., on x.

Even so, ξ_{vert} does measure the length over which the top and bottom surfaces are correlated, so in this sense we can still think of if as a correlation length.

We can also obtain a measure of the tension of a vertical interface by using (6.12), together with the duality transformation (5.19). This shows that there is a Q-dependent contribution $-Ms'_Q$ to $\ln Z_Q^{\text{free}}$ [coming from the $-QMh$ term in (6.12)], where

$$s'_Q = \frac{Q}{N} \ln \frac{G + |1 - k'|}{G + k' - 1} \tag{6.24}$$

and $Q = 0, 1, \ldots, N-1$.

Since the effect of the Q-cyclic boundary condition (5.18) is to force at least one vertical interface separating phases that differ by Q, it seems plausible to identify s'_Q with the interfacial tension (per unit length) between the domains. For $k' > 0$, (6.24) correctly gives $s'_Q = 0$, while for $k' < 1$,

$$s'_Q = -\frac{Q \ln w}{N} = \frac{Q}{N\xi_{\text{vert}}} \tag{6.25}$$

This result, like ξ_{vert}, depends on x, so it cannot (at least for $N = 2$) be the usual vertical interfacial tension. Presumably the difference is related to the use of free (rather than cyclic) boundary conditions at the top and bottom of the lattice.

For the Hermitian case ($y = \infty$, $x^N = G = 1 + k'$), with $N = 2$, these results simplify to

$$\xi_{\text{vert}} = \xi_{\text{hor}} = (2s_a)^{-1} = (2s'_1)^{-1} = -1/\ln k' \qquad \text{for} \quad k' < 1 \tag{6.26}$$

$$\xi_{\text{vert}} = \xi_{\text{hor}} = 2/\ln k' \qquad \text{for} \quad k' > 1$$

These agree with Eqs. (7.10.18), (7.10.43), and (7.11.4) of ref. 10, provided we note that k therein is our k', and $s/k_B T$ therein is our $2s_a$ or $2s'_1$ (this 2 comes from a difference in the choice of length scale).

6.5. Critical Exponents

Let f_{sing} be the singular part of the free energy at the critical point $k' = 1$. Similarly, let $(f_s)_{\text{sing}}$ be the singular part of the surface free energy f_s. Then near $k' = 1$ we expect

$$
\begin{aligned}
f_{\text{sing}} &\propto |k'-1|^{2-\alpha}, \qquad (f_s)_{\text{sing}} \propto |k'-1|^{2-\alpha_s} \\
s_a &\propto (1-k')^{\mu_{\text{hor}}}, \qquad s'_Q \propto (1-k')^{\mu_{\text{vert}}} \\
\xi_{\text{hor}} &\propto |k'-1|^{-\nu_{\text{hor}}}, \qquad \xi_{\text{vert}} \propto |k'-1|^{\nu_{\text{vert}}}
\end{aligned}
\tag{6.27}
$$

where $\alpha,\ldots, \nu_{\text{vert}}$ are critical exponents. We can find that this is the case, the exponents having the values given in (1.1).

In particular, as $k' \to 1$ from below, to leading order

$$
\begin{aligned}
\xi_{\text{vert}} &= G/[2(1-k')] \\
\xi_{\text{hor}} &= [2/(1-k')]^{2/N}/[2\sin^2(\pi/N)]
\end{aligned}
\tag{6.28}
$$

Eliminating $1-k'$ between these two expressions, we see that the limit

$$
\Gamma = \lim_{k' \to 1^-} \xi_{\text{hor}}^{N/2}/\xi_{\text{vert}} \tag{6.29a}
$$

exists, with value

$$
\Gamma = 4G^{-1}[2\sin^2(\pi/N)]^{-N/2} \tag{6.29b}
$$

(If we take the limit $k' \to 1^+$, we get a slightly different expression, but still inversely proportional to G.)

6.6. Finite-Size Corrections at Criticality

In the previous subsection we took the results for $k' \neq 1$, evaluated in the large L, M limits, and then let $k' \to 1$. Now let us reverse the order of limits and first set $k' = 1$, then take M and then L (or L and then M) large.

First taking M large, we can repeat the derivation of (6.12) above, but now the approximation (valid for large L) of replacing (6.6) by (6.11) means that for finite L we should add to (6.12) a correction term

$$
\frac{1}{2\pi i}\oint_C \ln D(\cos\theta)\, d\ln\left\{1 + \sum_{n=1}^{N-1}\left(\frac{1-e^u}{1-\omega^{-n}e^u}\right)^L\right\} \tag{6.30}
$$

[We have taken $Q=0$, as in the limit $M \to \infty$ the sum in (3.2) is dominated by the contribution from $Q=0$: this gives the largest relevant eigenvalue of $\mathbf{T}_{\text{row}}$.]

Here $D(c)$ is given by the truncated result (6.9), keeping only the terms exponential in M, and C is the contour C_2, which now surrounds the negative real axis in the u plane.

For L large, the integrand in (6.30) is dominated by small values of $z = e^u$, in fact by values of order L^{-1}. Then, from (3.10) and (6.8),

$$
\Delta(\cos\theta) \sim \mp 2ie^{Nu/2} \tag{6.31}
$$

the upper (lower) sign applying above (below) the real axis. Hence (6.30) becomes

$$\frac{M}{2\pi i}\int_{-\infty}^{0} \ln\left(\frac{G-2iz^{N/2}}{G+2iz^{N/2}}\right) d\ln\left\{1+\sum_{n=1}^{N-1}\left(\frac{1-z}{1-\omega^{-n}z}\right)^{L}\right\} \tag{6.32}$$

the integral being from $u=-\infty$ to $u=0$, i.e., from $z=0$ to $z=1$.

Changing variables from u through z to v, where

$$z = v/L \tag{6.33}$$

and then letting $L\to\infty$, we find that (6.32) in turn becomes

$$2Mc_N/(\pi G L^{N/2}) \tag{6.34}$$

where

$$\begin{aligned} c_N &= -\int_0^\infty v^{N/2}\, d\ln\left[1+\sum_{n=1}^{N-1} e^{(\omega^{-n}-1)v}\right] \\ &= \frac{N}{2}\int_0^\infty dv\, v^{(N-2)/2}\ln\left[1+\sum_{n=1}^{N-1} e^{(\omega^{-n}-1)v}\right] \end{aligned} \tag{6.35}$$

Now suppose we first take L large [having already set $k'=1$ in the derivation of (6.12)]. For M finite, the working goes through as before, except that we get an extra contribution to $\ln \tilde{Z}_Q$ of

$$\frac{L}{\pi}\int_0^{\pi(1-N^{-1})} \ln\left\{1+\left[\frac{G-\Delta(\cos\theta)}{G+\Delta(\cos\theta)}\right]^M \frac{\Delta(\cos\theta)-1+\cos\theta}{\Delta(\cos\theta)+1-\cos\theta}\right\} d\phi \tag{6.36}$$

This is proportional to L. As before, θ and ϕ are related by (6.14), and now $\Delta(\cos\theta)=2\sin(\theta/2)$.

If we now let M become large, the integral (6.36) is dominated by its values for small θ, in fact for $\theta\sim M^{-1}$. Setting here $\theta = Gv/(2M)$, we obtain that (6.36) becomes

$$\left(\frac{2L}{N\pi}\right)\left(\frac{G}{4M}\right)^{2/N} d_N \sin\frac{\pi}{N} \tag{6.37}$$

where

$$d_N = \int_0^\infty v^{(2-N)/N}\ln(1+e^{-v})\, dv \tag{6.38}$$

Both the corrections (6.34) and (6.37) involve L and M only via the ratio $M/L^{N/2}$. Obviously this model is intrinsically anisotropic, having

different critical exponents in the horizontal and vertical directions. To compensate for this, for $k' \neq 1$ it would be natural to measure M in units of ξ_{vert}, $2L$ in units of ξ_{hor}. Doing this, then taking the limit $k' \to 1$, suggests that instead of $M/L^{N/2}$ we should express the results in terms of $\Gamma M/(2L)^{N/2}$. We therefore write the above results in the form

$$-\lim_{M\to\infty}\frac{1}{M}\ln Z_a = -fL + \Delta_N^{(\text{per})}\Gamma/(2L)^{N/2} \tag{6.39a}$$

$$-\lim_{L\to\infty}\frac{1}{2L}\ln Z_a = \frac{1}{2}fM + 2f_s + \Delta_N^{(\text{fix})}/(\Gamma M)^{2/N} \tag{6.39b}$$

where

$$\begin{aligned}\Delta_N^{(\text{per})} &= -\left(2\sin\frac{\pi}{N}\right)^N \frac{c_N}{2\pi} \\ \Delta_N^{(\text{fix})} &= -\frac{d_N}{2N\pi\sin(\pi/N)}\end{aligned} \tag{6.40}$$

The first result, (6.39a), is the free energy per unit height of the infinitely high lattice, with width $2L$ and periodic boundary conditions from side to side. The second, (6.39b), is for the infinitely wide lattice, with height M and fixed-spin boundary conditions. For the Ising case, when $N=2$, we get $\Delta_2^{(\text{per})} = -\pi/12$ and $\Delta_2^{(\text{fix})} = -\pi/48$, so the results (6.39) agree with the predictions of conformal invariance[9,10] with central charge $c = 1/2$. For $N > 2$ they differ from these predictions, notably in that the large-L, M corrections are no longer just proportional to L^{-1}, M^{-1}. This is presumably due to the intrinsic anisotropy of the model.

7. ROW-INHOMOGENEOUS MODEL

Because of the star–triangle relation,[1] transfer matrices $\mathbf{T}_{\text{row}}$ with the same value of k, but different values of x, commute. This was an essential step in our derivation[4] of the eigenvalues of $\mathbf{T}_{\text{row}}$ in the spaces $V_0, \ldots, V_{N-1}$. It makes it easy to generalize the above results to a model where x has different values $x_1, \ldots, x_M$ in the M rows: all one has to do is replace Mth powers by the corresponding product; e.g., x^{-M} in (3.18) becomes $(x_1 x_2 \cdots x_M)^{-1}$, while $[G + \Delta(c)]^M$ in (3.11) becomes $\prod_{i=1}^{M}[G_i + \Delta(c)]$. This is because the eigenvectors of $\mathbf{T}_{\text{row}}$ depend on k', but not on x.

However, we can do better than this. In Section 2 we assumed (we hope plausibly and correctly) that the matrix $\mathbf{P}$ is independent of both x *and* k'. Thus, even if k also varies from row to row, with values $k'_1, \ldots, k'_M$,

one can still simultaneously apply the similarity transformation (2.18) (in the space V_Q) to transform each transfer matrix to the form (2.16), i.e., to a direct product of m matrices, each only two-by-two.

Thus, the product of the M transfer matrices is also transformed to a direct product of two-by-two matrices, each being a product of M matrices $B(x_i, y_i, \cos\theta_j)$, for $i = 1,..., M$.

Here y_i is the value of y for each row, related to x_i, k_i' by (2.4). Similarly, G is defined by (2.9) for each row, and takes the values $G_1,..., G_M$.

Replacing $\mathscr{T}_Q^M$ in (3.5) by the product of the $\mathscr{T}_Q$ for each row, we get the analogs of (3.8) and (3.9):

$$D(\cos\theta) = (1, 0) \prod_{i=1}^{M} B(x_i, y_i, \theta) \begin{pmatrix} 1 \\ 0 \end{pmatrix} \tag{7.1}$$

$$\tilde{Z}_Q = (g_1 \cdots g_M)^L (x_1 \cdots x_M)^{-Q} D(\cos\theta_1) \cdots D(\cos\theta_m) \tag{7.2}$$

The two-by-two matrix $B(x_i, y_i, \theta)$ is defined by (3.7), with x, k', G replaced by their values x_i, k_i', G_i for row i.

We can no longer reduce (7.1) to an explicit form like (3.11), but it is possible to use (7.1) directly, and to note that $D(c)$ is still a polynomial of degree r, and to verify (3.15). Thus all the equations (3.12)–(4.7) remain valid, provided only that we replace x^M, g^M therein by $x_1 \cdots x_M$, $g_1 \cdots g_M$.

7.1. Free Boundary Conditions

Such generalizations are not necessarily interesting from a physical viewpoint, but they can be useful in examining the equivalences and mathematical structures of lattice models. They also mean that we can readily change from fixed to free boundary conditions on the bottom row, simply by taking $x_1 = y_1 = \infty$, in which case $W(n) = \bar{W}(n) = 1$ between spin rows 1 and 2, hence there are no intervening interactions. This is equivalent to removing row 1 and imposing free boundary conditions on row 2. We can similarly remove the top row, in which case Z_a (for all a) becomes the partition function of a lattice with free boundary conditions (with M reduced by 2).

We have explicitly verified that the result is the same as that given by the duality relation (5.19) for $Q = 0$. A key step is the similarity relation $B(x_d, y_d, \theta) = X^{-1} B(x, y, \theta) X$, where $X = (1 - \cos\theta)\,\sigma^z - \sin\theta\,\sigma^x$.

7.2. Relation to the Ising Model

One intriguing point that is brought out particularly clearly in the inhomogeneous case is that (apart from a normalization) the eigenvalue

spectrum (4.7) of $\mathbf{T}_{\rm col}$ depends only on N and $\zeta_1,\ldots,\zeta_r$, or equivalently [using (3.17)] on N and the zeros $c_1,\ldots,c_r$ of $D(c)$. From (7.1) and (3.7), we see that $c_1,\ldots,c_r$ are determined completely by M, $G_1,\ldots,G_M$, $k'_1,\ldots,k'_M$. They have no other dependence on N, so are the same for general N as for the Ising model. We know we can solve the Ising model by Pfaffians or Fermian operators,[13,14,17,18] and with our boundary conditions we are indeed led to the expression (7.1). This suggests that there should be some combinatorial or algebraic method of diagonalizing $\mathbf{T}_{\rm col}$ which closely resembles these Ising model methods, having the same underlying equations leading to the same $\zeta_1,\ldots,\zeta_r$, but yielding the general-N form of (4.7).

All the remarks of the next section generalize readily to the row-inhomogeneous mod, though for brevity we do not usually exhibit the row indices.

8. INVERSE SOS MODEL

Consider the column-to-column transfer matrix $\mathbf{T}_{\rm col}$, and for convenience take M to be odd, so $M=2r-1$. [The case of M even can be derived from this by taking $x_{2r-1}=1$, so that in the top row $\bar{W}(n)=\delta_{n,0}$.]

In this section we shall write a_B for the value of the fixed spins in the bottom row of the lattice.

A typical column (or double column) of the lattice is shown in Fig. 4 for $M=5$, $r=3$. There are $r-1$ free spins on the right (center, left), labeled $\sigma'_1,\ldots,\sigma'_{r-1}(\sigma''_1,\ldots,\sigma''_{r-1};\sigma_1,\ldots,\sigma_{r-1})$. The transfer matrix $\mathbf{T}_{\rm col}$ acts on the

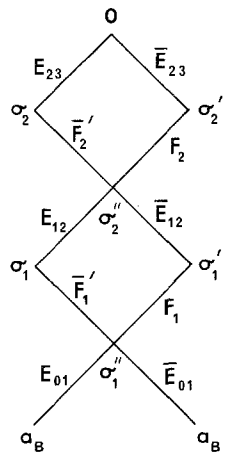

Fig. 4. A typical double column of $\mathscr{L}$, showing the free spins $\sigma_1,\ldots,\sigma''_2$; the fixed spins a_B, 0; and the E, F matrices associated with the edges.

N^{r-1}-dimensional space of vectors ψ with elements $\psi(\sigma_1,..., \sigma_{r-1})$, and has elements

$$(\mathbf{T}_{\mathrm{col}})_{\sigma\sigma'} = \sum_{\sigma''} \prod_{j=1}^{r} W(\sigma_{j-1} - \sigma_j'')\, \bar{W}(\sigma_{j-1}' - \sigma_j'') \times \prod_{j=1}^{r-1} \bar{W}(\sigma_j'' - \sigma_j)\, W(\sigma_j'' - \sigma_j') \tag{8.1}$$

taking $\sigma_0 = \sigma_0' = a_B$, $\sigma_r'' = 0$. Here σ denotes the set of spins $\sigma_1,..., \sigma_{r-1}$; and similarly for σ', σ''.

Define other N^{r-1}-by-N^{r-1} matrices:

$$\begin{aligned} (E_{i,i+1})_{\sigma\sigma'} &= W(\sigma_i - \sigma_{i+1}) \prod_{j=1}^{r-1} \delta(\sigma_j, \sigma_j') \\ (F_i)_{\sigma\sigma'} &= W(\sigma_i - \sigma_i') \prod_{\substack{j=1 \\ \neq i}}^{r-1} \delta(\sigma_j, \sigma_j') \end{aligned} \tag{8.2}$$

Thus $E_{i,i+1}$ is diagonal, F_i is not. Similarly, define $\bar{E}_{i,i+1}, \bar{F}_i$ with W replaced by $\bar{W}$. Let $\bar{F}_i'$ be the transpose of $\bar{F}_i$. Then for $r = 3$ we can write (8.1) as

$$\mathbf{T}_{\mathrm{col}} = E_{23} \bar{F}_2' E_{12} \bar{F}_1' E_{01} \bar{E}_{01} F_1 \bar{E}_{12} F_2 \bar{E}_{23} \tag{8.3}$$

Each of these E or F matrices corresponds to a single edge, as indicated in Fig. 4: they are local "edge transfer matrices" that add a single edge to the lattice.

From (4.2), $\lambda_{i,n}^{-1}$ has a simpler dependence on n than $\lambda_{i,n}$, being just linear in ω^{-n}. From (4.7), this suggests that the inverse of $\mathscr{T}_{\mathrm{col}}$, and hence $\mathbf{T}_{\mathrm{col}}$, may be in some ways simpler than the original matrix.

We shall find that this is so. The E, F matrices are invertible, so from (8.3) it is easy to invert $\mathbf{T}_{\mathrm{col}}$. In fact, the inverse is simply given by interchanging $\bar{E}_{i,i+1}$ with $E_{i,i+1}^{-1}$, and $\bar{F}_i'$ with F_i^{-1}.

The matrices $E_{i,i+1}$, being diagonal, are easily inverted; the matrices F_i are not much harder: one uses the Fourier transforms (5.1)–(5.3). This gives

$$\begin{aligned} (E_{i,i+1}^{-1})_{\sigma\sigma'} &= X(\sigma_i - \sigma_{i+1}) \prod_{j=1}^{r} \delta(\sigma_j, \sigma_j') \\ (F_i^{-1})_{\sigma\sigma'} &= Y(\sigma_i - \sigma_i') \prod_{\substack{j=1 \\ \neq i}}^{r-1} \delta(\sigma_j, \sigma_j') \end{aligned} \tag{8.4}$$

where $X(n) = 1/W(n)$ and

$$Y(n) = N^{-3/2} \sum_{j=0}^{N-1} \omega^{jn}/W_f(j) \tag{8.5}$$

The matrices $\bar{E}_{i,i+1}^{-1}$, $\bar{F}_i^{-1}$ are given similarly, with W, X, W_f, and Y replaced by $\bar{W}$, $\bar{X}$, $\bar{W}_f$, and $\bar{Y}$, and ω^{jn} by ω^{-jn}.

We thus obtain

$$(\mathbf{T}_{\text{col}}^{-1})_{\sigma\sigma'} = \sum_{\sigma''} \prod_{j=1}^{r} \bar{X}(\sigma_{j-1} - \sigma_j'') \, X(\sigma_{j-1}' - \sigma_j'') \times \prod_{j=1}^{r-1} Y(\sigma_j - \sigma_j'') \, \bar{Y}(\sigma_j' - \sigma_j'') \tag{8.6}$$

Summing over $\sigma_1'', \ldots, \sigma_{r-1}''$, we get

$$(\mathbf{T}_{\text{col}}^{-1}) = \bar{X}(\sigma_{r-1}) \, X(\sigma_{r-1}') \prod_{j=1}^{r-1} U(\sigma_{j-1}, \sigma_{j-1}', \sigma_j', \sigma_j) \tag{8.7}$$

where

$$U(a, b, c, d) = \sum_{s=0}^{N-1} \bar{X}(a-s) \, X(b-s) \, \bar{Y}(c-s) \, Y(d-s) \tag{8.8}$$

and a in this section is not necessarily equal to the boundary spin value a_B.

This function $U(a, b, c, d)$ is the weight of a four-pointed star, with center spin s and outer spins a, b, c, d (ordered anticlockwise from the bottom left). Consider first the case when $b = a$ and note that $X(n) \, \bar{X}(n) = [W(n) \, \bar{W}(n)]^{-1}$. From (5.9), this simplifies, giving

$$U(a, a, c, d) = \frac{1}{x-1} \sum_{s=0}^{N-1} (x - \omega^{s-a}) \, \bar{Y}(c-s) \, Y(d-s) \tag{8.9}$$

Using (8.5) and its analog for $\bar{Y}$, this gives

$$U(a, a, c, d) = \frac{1}{N^2(x-1)} \sum_{j=0}^{N-1} \frac{\omega^{j(d-c)}}{W_f(j)} \left\{ \frac{x}{\bar{W}_f(j)} - \frac{\omega^{c-a}}{\bar{W}_f(j-1)} \right\} \tag{8.10}$$

We have already used the unexpectedly simple result (5.9) for $W(n) \, \bar{W}(n)$. Now we note from (5.3) and (5.9) that $W_f(j) \, \bar{W}_f(j)$ and $W_f(j) \, \bar{W}_f(j-1)$ also simplify dramatically. Substituting the results into (8.10) and using (5.7), we obtain

$$U(a, a, c, d) = \sum_{j=0}^{N-1} \frac{\omega^{j(d-c)} \{ x(x - \mu y \omega^{-j}) - \omega^{c-a}(1 - \mu\omega^{-j}) \}}{N^2(x-1)(x-\mu y) \, \Omega} \tag{8.11}$$

The j summation can now be performed:

$$U(a, a, c, d) = \frac{(x^2 - \omega^{c-a})\,\delta_{d,c} + \mu(\omega^{c-a} - xy)\,\delta_{d,c+1}}{N(x-1)(x-\mu y)\,\Omega} \tag{8.12a}$$

where $\delta_{a,b} = 1$ if $b = a$, modulo N; else $\delta_{a,b} = 0$.

Similarly, one can establish that

$$U(a, a-1, c, d) = \frac{\mu(xy - \omega^{c-a})\,\delta_{d,c} + \mu^2(\omega^{c-a} - y^2)\,\delta_{d,c+1}}{N(x-1)(x-\mu y)\,\Omega} \tag{8.12b}$$

These results can be combined into one formula, true for $a - b = 0$ or 1 (mod N):

$$U(a, b, c, d) = \sum_{k=0}^{1} \omega^{k(c-a)} f_k(a-b)\, g_k(d-c) \tag{8.13}$$

where, using (5.10),

$$\begin{aligned} f_0(n) &= (x\delta_{n,0} + \mu y \delta_{n,1})/(x-1) \\ f_1(n) &= (\delta_{n,0} + \mu\delta_{n,1})/(x-1) \\ g_0(n) &= (x^N - 1)(x\delta_{n,0} - \mu y \delta_{n,1})/[N(x-1)(x^N - \mu^N y^N)] \\ g_1(n) &= (x^N - 1)(-\delta_{n,0} + \mu\delta_{n,1})/[N(x-1)(x^N - \mu^N y^N)] \end{aligned} \tag{8.14}$$

Now note that the σ_j, σ'_j in (8.7), as in (8.1), satisfy the bottom-row boundary condition $\sigma_0 = \sigma'_0 = a_B$. Hence, for the $j = 1$ term in the product in (8.7) we do have $U(a, b, c, d)$ with $b = a$. Hence the product vanishes unless $\sigma_1 - \sigma'_1 = 0$ or 1 (mod N). Thus, the $j = 2$ term in the product is given by (8.13), and vanishes unless $\sigma_2 - \sigma'_2 = 0$ or 1.

Continuing this argument, it follows that the elements of $\mathbf{T}_{\text{col}}$ are given by (8.7), with $U(a, b, c, d)$ defined for *all* integers a, b, c, d by (8.13), (8.14). The elements vanish unless $\sigma_j - \sigma'_j = 0$ or 1 for $j = 1, .., r-1$.

We can regard the integer k in (8.13) as a new "spin," living at the center of the star a, b, c, d, and taking only the two values, 0 and 1. Then

$$\omega^{-ka} f_k(a-b), \qquad \omega^{kc} g_k(d-c) \tag{8.15}$$

are the weights of the spin triplets (a, b, k), (d, c, k), respectively.

From this viewpoint, the rhs of (8.7) is the column-to-column transfer matrix of a new model on the lattice of Fig. 5. This lattice has the same number $L(M+1)$ of sites as the original lattice $\mathscr{L}$, but now they are divided into two sublattices, shown by open circles and filled circles. On

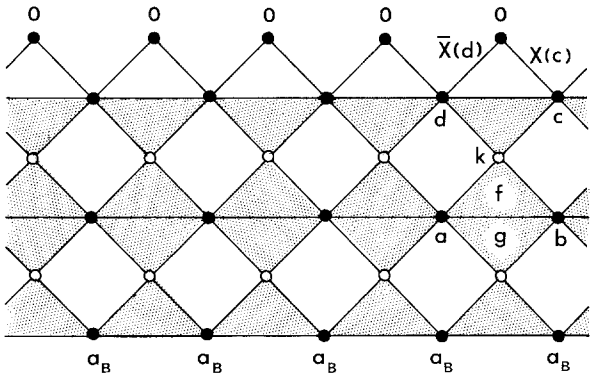

Fig. 5. The lattice on which the "inverse" model is defined. Sites denoted by open (filled) circles contain two-valued (N-valued) spins. On the shaded triangles there are three-spin interactions given by (8.15). The top edges have weights X, $\bar{X}$ as indicated, where $X(n)\,\bar{X}(n) = (x-\omega^{-n})/(x-1)$.

the filled circles we still have the original spins, each taking N values (except for the spins at the top and bottom, which are still fixed at 0 and a_B, respectively). On the open circles there are new spins, with values 0 and 1. Each up-pointing triangle a, b, k (with a, b, k arranged as in Fig. 5) has as weight the first expression (8.15); down-pointing triangles d, c, k have the second expression as weights. The top edges contribute weights $X(c)$, $\bar{X}(d)$, as indicated.

Either horizontally adjacent spins on filled circles are equal, or the left spin is one greater than the right spin (modulo N). In this respect the model is a "solid-on-solid" (SOS) model.

Let $k_1,\ldots, k_L$ be a row of k-spins (on open circles), and $k'_1,\ldots, k'_L$ the row above. From (8.15), incrementing all the intermediate σ-spins (on filled circles) by j gives an extra factor

$$\omega^{j(k_1+\cdots+k_L-k'_1-\cdots-k'_L)}$$

The partition function contains a sum of such increments, which is a Kronecker delta. We can therefore restrict the partition function to values of k-spins such that $k_1+\cdots+k_L$ is the same (modulo N) for each row of the lattice. If we form $\tilde{Z}_Q$ similarly to (3.2), only terms with $k_1+\cdots+k_L = Q \pmod N$ contribute.

Since the column-to-column transfer matrix of this model is just the inverse of $\mathbf{T}_{\rm col}$, it also has a simple direct-product eigenvalue spectrum, given by inverting (4.7). The partition function is, from (4.5) and (4.4),

$$\sum_{\mathbf{n}} (\Lambda_{\mathbf{n}})^{-L} = g^{-LM} \sum_{\mathbf{n}} (\lambda_{1,n_1}\lambda_{2,n_2}\cdots\lambda_{r,n_r})^{-L} \tag{8.16}$$

the summation being over all integers $n_1,..., n_r$ between 0 and $N-1$ that satisfy (4.3), i.e., $n_1 + \cdots + n_r = a_B \pmod N$.

This working generalizes readily to the inhomogeneous case, when x and y vary arbitrarily from row to row (but not column to column). As in Section 7, the ζ_i are then given by (3.17), where the c_i are the zeros of the function $D(c)$ defined by (7.1). The triangle weight functions $f_k(n)$, $g_k(n)$ are then defined by (8.14), using the values of x, y, μ appropriate to the row in which the triangle lies.

There are various trivial gauge transformations that can be made; e.g., multiplying $f_k(n)$, $g_k(n)$ by s_k, s_k^{-1} (for any function s_k) obviously leaves (8.13) and $\mathbf{T}_{\rm col}$ unchanged. By using these, we can probably extend these results to arbitrary functions $f_k(n)$, $g_k(n)$, defined for k, $n = 0, 1$.

It is interesting to note that these inhomogeneous models are solvable, but as k' can vary from row to row, they are *not* Z-invariant.[19]

8.1. Associated Hamiltonian

The Hamiltonian $\mathscr{H}$ in (2.5) is obtained from the transfer matrix $\mathbf{T}_{\rm row}$ of the original homogeneous model by perturbing about the case $x = y = \mu = 1$, when $\mathbf{T}_{\rm row}$ is a simple shift operator: two operations by $\mathbf{T}_{\rm row}$ on a vector ψ, with elements $\psi(\sigma_1,..., \sigma_L)$, produce a vector with elements $\psi(\sigma_2,..., \sigma_L, \sigma_1)$.

Because of the fixed-spin boundary conditions, $\mathbf{T}_{\rm col}$ is not in this case invertible, and it does not seem to be helpful to try to correspondingly expand $\mathbf{T}_{\rm col}$.

However, what we can do is use the freedom of the row-inhomogeneous model to expand about the case when (for ε small)

$$\begin{aligned} &x_j,\ y_j = O(\varepsilon^{-1}), \qquad \mu_j \sim 1, \qquad j \text{ odd} \\ &x_j = 1 + O(\varepsilon^N), \qquad y_j \sim 1, \qquad \mu_j = O(\varepsilon), \qquad j \text{ even} \end{aligned} \tag{8.17}$$

Here, as in Section 7, x_j, y_j, μ_j are related by (2.2), being the values of x, y, μ in row j (between original spin rows j and $j+1$), and $j = 1,..., M$. To leading order we take $\mu_j = x_j/y_j$ for j odd.

From (2.1), when $\varepsilon = 0$ the weight functions $W(n)$, $\bar{W}(n)$ are both equal to 1 for odd rows, $\delta_{n,0}$ for even rows. It follows at once that in this case $\mathbf{T}_{\rm col} = I$.

Expanding to first order in ε, taking M odd, and using the results (8.7), (8.13), (8.14) for $\mathbf{T}_{\rm col}^{-1}$, we obtain

$$\mathbf{T}_{\rm col}^{-1} = \frac{N^r}{g_1 g_2 \cdots g_M} \{I + \varepsilon \mathscr{H}_{\rm col}\} \tag{8.18}$$

where, if the limit is taken so that the quantities

$$\alpha_j = \varepsilon^{-1}/(x_d)_{2j} = \varepsilon^{-1}\mu_{2j}\, y_{2j}/x_{2j}, \qquad \gamma_j = \varepsilon^{-1}/x_{2j-1} \tag{8.19}$$

are held fixed, then

$$\mathscr{H}_{\mathrm{col}} = -\sum_{j=1}^{r-1} \alpha_j X_j - \sum_{j=1}^{r} \gamma_j Z_j Z_{j-1}^{-1} \tag{8.20}$$

Here X_j, Z_j are the N^{r-1}-dimensional operators used in refs. 3 and 5:

$$\begin{aligned} (X_j)_{\sigma\sigma'} &= \delta(\sigma_j, \sigma_j' + 1) \prod_{\substack{k=1 \\ \neq j}}^{r-1} \delta(\sigma_k, \sigma_k') \\ (Z_j)_{\sigma\sigma'} &= \omega^{\sigma_j} \prod_{k=1}^{r-1} \delta(\sigma_k, \sigma_k') \end{aligned} \tag{8.21}$$

For the boundary cases, $Z_0 = \omega^{aB} I$ and $Z_r = I$.

From (8.16), (4.2), and (7.1), the eigenvalues of $\mathscr{H}_{\mathrm{col}}$ are

$$E = -\sum_{j=1}^{r} s_j \omega^{-n_j} \tag{8.22}$$

where $n_1, \ldots, n_r$ are again any integers between 0 and $N-1$ satisfying (4.3), i.e., $n_1 + \cdots + n_r = a_B \pmod N$.

The $s_1, \ldots, s_r$ must satisfy

$$s_1 s_2 \cdots s_r = \gamma_1 \gamma_2 \cdots \gamma_r \tag{8.23}$$

and s_1^N, $s_2^N, \ldots, s_r^N$ are the r zeros of the polynomial $\bar{D}(z)$, where

$$\bar{D}(z) = (1, 0)\, B_1(z)\, B_2 B_3(z) \cdots B_{M-1} B_M(z) \begin{pmatrix} 1 \\ 0 \end{pmatrix} \tag{8.24}$$

and

$$\begin{aligned} B_{2j-1}(z) &= \begin{pmatrix} z - \gamma_j^N & iz^{1/2} \\ iz^{1/2} & -1 \end{pmatrix} \\ B_{2j} &= \begin{pmatrix} 1 & 0 \\ 0 & \alpha_j^N \end{pmatrix} \end{aligned} \tag{8.25}$$

[The original ζ_j are equal to εs_j. The polynomial $\bar{D}(z)$ is proportional to the $D(\cos\theta)$ of (7.1), with $\cos\theta = (1 + \varepsilon^N z)/(1 - \varepsilon^N z)$, evaluated in the desired limit $\varepsilon \to 0$, keeping z, α_j, γ_j fixed.]

Define $q_1, q_2, \ldots, q_M$ by

$$q_{2j-1} = \gamma_j; \qquad q_{2j} = \alpha_j \tag{8.26}$$

and define a set of two-dimensional matrices

$$\bar{B}_j(z) = \begin{pmatrix} z^{1/2} & iq_j \\ i & 0 \end{pmatrix} \tag{8.27}$$

Then we can rewrite (8.24) as

$$\bar{D}(z) = (1, 0)\, \bar{B}_1(z)\, \bar{B}_2(z) \cdots \bar{B}_M(z) \begin{pmatrix} z^{1/2} \\ i \end{pmatrix} \tag{8.28}$$

From this one can verify that

$$\bar{D}(z) = \det(A - z^{1/2} I) \tag{8.29}$$

where A is the $M+1$ by $M+1$ bidiagonal matrix

$$A = \begin{pmatrix} 0 & q_1^{N/2} & 0 & & & \\ q_1^{N/2} & 0 & q_2^{N/2} & & & \\ 0 & q_2^{N/2} & 0 & & & \\ & & \ddots & & & \\ & & & 0 & q_M^{N/2} \\ & & & q_M^{N/2} & 0 \end{pmatrix} \tag{8.30}$$

To summarize this last result: the Hamiltonian $\mathscr{H}_{\text{col}}$ is given by (8.20), with *arbitrary* coefficients $\alpha_1, \ldots, \alpha_{r-1}, \gamma_1, \ldots, \gamma_r$. All its N^{r-1} eigenvalues are given by (8.22), with $s_1^N, \ldots, s_r^N$ being the zeros of $\bar{D}(z)$.

This is an amazingly simple result. For $N = 2$, $\mathscr{H}_{\text{col}}$ is the usual Ising chain operator, and the result can be obtained from Clifford or free-fermion algebras.[18,20]

It is natural to ask whether there is some appropriate generalization of this algebraic approach that will as readily yield (8.22). Further, can it be extended to periodic boundary conditions on the chain (from top to bottom of the lattice)? (Such an extension does not seem as straightforward as one might hope.)

Note that the family of matrices $\mathbf{T}_{\text{col}}$ (with different x, y) do *not* commute, and hence do not commute with $\mathscr{H}_{\text{col}}$. To obtain commutation, we must go back to the full solvable chiral Potts model[1,2] with periodic boundary conditions.

9. SUMMARY

We have obtained exact and explicit expressions for the partition function of the superintegrable chiral Potts model, with fixed (or free)

boundary conditions at top and bottom, periodic from right to left. We have also obtained some of the eigenvalues of $\mathbf{T}_{\text{row}}$ and all of the eigenvalues of $\mathbf{T}_{\text{col}}$. In Section 6 we have considered the large-lattice limit, obtaining the bulk and surface free energies, and the horizontal and vertical interfacial tensions and correlation lengths, and studied their critical properties. The model is intrinsically anisotropic, with different exponents in the horizontal and vertical directions. This is reflected in the fact that at criticality the finite-size corrections do not have the usual form predicted by conformal invariance.[9,10]

The eigenvalues of $\mathbf{T}_{\text{col}}$ have a simple direct product structure, and $\mathbf{T}_{\text{col}}^{-1}$ is a sparse matrix that can be thought of as the transfer matrix of an "inverse" SOS model. $\mathbf{T}_{\text{col}}$ has these simple properties even for a row-inhomogeneous model, where x and y vary arbitrarily from row to row. Since k' can vary, this model is not necessarily Z-invariant:[19] this in itself sets this model apart from most of the other exactly solved models in two-dimensional statistical mechanics (except of course the Ising model, which is the $N=2$ case of this model).

The inverse model contains two types of spin: two-valued and N-valued. It is natural to associate the N-valued spins with the horizontal lines in Fig. 5, and the two-valued spins with corresponding vertical lines (through the open circles of Fig. 5). Perhaps this gives a clue to the origin of the intrinsic anisotropy of the original model: the horizontal direction is associated with the original N-valued spins, but the vertical direction with "reduced" two-valued spins.

Finally, we have taken a limit when $\mathbf{T}_{\text{col}}$ becomes the identity operator. (This limit is unlike the Hamiltonian limit for, say, the eight-vertex model; §10.14 of ref. 12: then the transfer matrix becomes a shift operator.) By perturbing about this case, we have obtained the Hamiltonian $\mathscr{H}_{\text{col}}$ given by (8.20). Its coefficients are completely arbitrary, yet for all choices its eigenvalue spectrum is a direct sum. This property suggests that there may be an appropriate generalization of the corresponding Clifford algebra technique for the Ising model. Such a generalization may provide a more direct way of solving this (and possibly other) models.

APPENDIX A

In Eq. (7) of ref. 4, a function $F(a, b)$ is introduced, defined by

$$F(a, b) = \sum_{d=0}^{N-1} W(a-d)\, \bar{W}(b-d)\, W'(d)\, \bar{W}'(d) \tag{A1}$$

Here $W(n)$, $\bar{W}(n)$ are given by (2.1), so

$$W(n)\,\bar{W}(n) = \omega^n(1-x)/(1-x\omega^n) \tag{A2}$$

while $W'(n)$, $\bar{W}'(n)$ are given by the same equations, but with x, y replaced by $x' = y^{-1}$, $y' = \omega^s x^{-1}$. Thus

$$W'(n)\,\bar{W}'(n) = \omega^n(y-1)/(y-\omega^n) \tag{A3}$$

Substituting into (A1), we get

$$F(a,b) = \omega^b \mu^{a-b}(1-x)(1-y)\sum_{d=0}^{N-1}\phi(\omega^d) \tag{A4}$$

where the function ϕ is

$$\phi(z) = \frac{z}{(z-y)(z-y\omega^{a+1})}\prod_{j=a+2}^{b}\frac{z-x\omega^{j-1}}{z-y\omega^j} \tag{A5}$$

Provided $0 \leqslant a < b \leqslant N-1$, $\phi(z)$ is a rational function with only simple poles at $z = y$, $y\omega^{a+1}$,..., $y\omega^b$, and $\phi(\infty) = 0$. Hence $\exists$ x-independent coefficients A_k such that

$$\phi(z) = \sum_k A_k/(y-\omega^{-k}z) \tag{A6}$$

k taking the values 0, $a+1$, $a+2$,..., b. Hence

$$\sum_{d=0}^{N-1}\phi(\omega^d) = \sum_k A_k \sum_{n=0}^{N-1}(y-\omega^n)^{-1} \tag{A7}$$

(taking $n = d-k$ and using periodicity).

The rhs of (A7) is proportional to $A_0 + A_{a+1} + \cdots + A_b$, which, from (A6), is $y\phi(0)$. However, from (A5) it is obvious that $\phi(0) = 0$. Hence the rhs of (A7) vanishes and, from (A4),

$$F(a,b) = 0 \qquad \text{if} \quad 0 \leqslant a < b \leqslant N-1 \tag{A8}$$

This is the result used in ref. 4.

APPENDIX B. NUMERICAL CALCULATIONS

For various values of L, N, and Q (up to about $L = 7$), we numerically verified (2.19) and (2.20). Starting from v_Q, we successively built up a basis

for V_Q by repeated multiplications by $\mathscr{H}_0$ and $\mathscr{H}_1$, making a Gram–Schmidt orthogonalization of every new vector generated, so as to keep the basis linearly independent and orthogonal. In every case the basis turned out to be of dimension 2^m.

Our construction ensured that $\hat{\mathscr{H}}_0$ was diagonal. $\hat{\mathscr{H}}_1$ then has to be block-tridiagonal, with nonzero elements (i, j) only when $(\hat{\mathscr{H}}_0)_{ii} - (\hat{\mathscr{H}}_0)_{jj} = -1, 0,$ or 1. Since $\hat{\mathscr{H}}_0$ has repeated eigenvalues, we could apply further unitary rotations that left $\hat{\mathscr{H}}_0$ unchanged, while ensuring that any element (i, j) of $\hat{\mathscr{H}}_1$ vanished if $i \neq j$ and $(\hat{\mathscr{H}}_0)_{ii} = (\hat{\mathscr{H}}_0)_{jj}$. The quite startling effect of this was to make $\hat{\mathscr{H}}_1$ an even more sparse matrix (with only $m+1$ nonzero entries per row or column), with many of the nonzero entries equal. In fact, we found by inspection (to numerical accuracy) that $\mathscr{H} \leftrightarrows_0$ and $\hat{\mathscr{H}}_1$ were the k'-independent and k'-proportional parts of (2.20), respectively. Thus, with respect to this basis, $H = \hat{\mathscr{H}} = \hat{\mathscr{H}}_0 + k'\hat{\mathscr{H}}_1$, and $\mathbf{P} = I$.

Since the construction of this basis was independent of k', the calculations provided convincing evidence (but not of course a proof) that the unitary matrix $\mathbf{P}$ in (2.18) and (2.19) is independent of k'.

ACKNOWLEDGMENT

I thank Prof. M. N. Barber for numerous helpful comments and discussions.

REFERENCES

1. R. J. Baxter, J. H. H. Perk, and H. Au-Yang, *Phys. Lett. A* **128**:138 (1988).
2. R. J. Baxter, *J. Stat. Phys.* **52**:639 (1988).
3. G. Albertini, B. M. McCoy, J. H. H. Perk, and S. Tang, *Nucl. Phys. B* **314**:741 (1989).
4. R. J. Baxter, *Phys. Lett. A* **133**:185 (1988).
5. G. Albertini, B. M. McCoy, and J. H. H. Perk, *Phys. Lett. A* **135**:159 (1989); preprint ITP-SB-89-24, Institute for Theoretical Physics, State University of New York, Stony Brook, to appear in *Phys. Lett. A.*
6. R. J. Baxter, in *Advanced Studies in Pure Mathematics*, Vol. 19 (Kirokuniya-Academic, 1989).
7. K. Binder, in *Phase Transitions and Critical Phenomena*, Vol. 8, C. Domb and J. L. Lebowitz, eds. (Academic, London, 1983).
8. R. M. Hornreich, M. Luban, and S. Shtrikman, *Phys. Rev. Lett.* **35**:1678 (1975).
9. H. W. J. Blöte, J. L. Cardy, and M. P. Nightingale, *Phys. Rev. Lett.* **56**:742 (1986).
10. I. Affleck, *Phys. Rev. Lett.* **56**:746 (1986).
11. M. E. Fisher, *J. Stat. Phys.* **34**:667 (1984).
12. R. J. Baxter, *Exactly Solved Models in Statistical Mechanics* (Academic, London, 1982).
13. B. M. McCoy and T. T. Wu, *The Two-Dimensional Ising Model* (Harvard University Press, 1973).

14. E. W. Montroll, R. B. Potts, and J. C. Ward, *J. Math. Phys.* **4**:308 (1963).
15. P. J. Davis, *Circulant Matrices* (Wiley, New York, 1979).
16. F. Y. Wu, *Rev. Mod. Phys.* **54**:235 (1982).
17. L. Onsager, *Phys. Rev.* **65**:117 (1944).
18. B. Kaufman, *Phys. Rev.* **76**:1232 (1949).
19. R. J. Baxter, *Phil. Trans. Soc. Lond.* **289**:315 (1978).
20. T. D. Schultz, D. C. Mattis, and E. H. Lieb, *Rev. Mod. Phys.* **36**:856 (1964).