

Algebraic Reduction of the Ising Model

R.J. Baxter

Received: 16 April 2008 / Accepted: 11 June 2008 / Published online: 16 July 2008
© Springer Science+Business Media, LLC 2008

Abstract We consider the Ising model on a cylindrical lattice of L columns, with fixed-spin boundary conditions on the top and bottom rows. The spontaneous magnetization can be written in terms of partition functions on this lattice. We show how we can use the Clifford algebra of Kaufman to write these partition functions in terms of L by L determinants, and then further reduce them to m by m determinants, where m is approximately $L/2$. In this form the results can be compared with those of the Ising case of the superintegrable chiral Potts model. They point to a way of calculating the spontaneous magnetization of that more general model algebraically.

Keywords Statistical mechanics · Lattice models · Transfer matrices

1 Introduction

Onsager calculated the partition function of the two-dimensional square-lattice Ising model in 1944 [1]. He did so by observing that the row-to-row transfer matrices generated a finite Lie algebra. This method was refined by Kaufman in 1949 [2], who showed that the matrices could be expressed in terms of a Clifford algebra. By taking the thermodynamic limit of a large lattice, one obtains the free energy per site.

The calculation of the spontaneous magnetization is much harder than that of the free energy. Onsager first announced his famous result for the spontaneous magnetization at a conference at Cornell University in August 1948 and later at a IUPAP conference in Florence [3, 4]. Onsager never published his derivation—the puzzle was finally solved by Yang in 1952 [5].

Here we wish to re-visit this problem, with a view to seeing if the techniques can be generalized to the N -state superintegrable case of the solvable chiral Potts model, which has properties very similar to the Ising model and reduces to it when $N = 2$. The author has previously used large-lattice functional relations and analytic methods to obtain the spontaneous magnetization of the general solvable chiral Potts model [6, 7], but it would still be

R.J. Baxter
Mathematical Sciences Institute, The Australian National University, Canberra, ACT 0200, Australia

interesting to have an algebraic method that may give more insight into the properties of the model on a finite lattice.

In Sect. 2 we define the Ising model on a cylindrical lattice of L columns, with fixed-spin boundary conditions on the top and bottom rows. We show how the magnetization can be expressed, firstly as the ratio of partition functions, and consequently as a matrix element between the two ground-state eigenvectors of the transfer matrix T . In Sect. 3 we introduce the hamiltonian \mathcal{H} which commutes with T and define ‘‘hamiltonian partition functions’’ \tilde{Z}, \tilde{W} . These are limiting cases of the usual partition functions and are slightly simpler to work with. This \tilde{Z} corresponds to the usual partition function, \tilde{W} to the partition function with an extra factor σ_1 , where σ_1 is some spin deep inside the lattice.

In Sect. 4 we set up the apparatus of the Clifford algebra and use it to write \tilde{Z}, \tilde{W} as square roots of L by L determinants. Up to this point our calculation parallels that of Yang [5]. Montroll et al. [4] also calculated the spontaneous magnetization using the more combinatorial pfaffian method and Szegő’s theorem on Toeplitz matrices [8]. In Sect. 5 we marry Yang and Montroll et al.’s techniques by using Szegő’s theorem to evaluate the appropriate large- L limit of \tilde{W} and hence derive Onsager’s famous formula for the spontaneous magnetization.

In Sect. 6 we show how the expressions for \tilde{Z}, \tilde{W} can be reduced to linear expressions in determinants of dimension $[(L - Q)/2]$, where $Q = 0$ or 1 . Then \tilde{Z} is precisely the hamiltonian limit of the corresponding partition function of the two-state superintegrable chiral Potts model. In Sect. 7 we write the result for \tilde{W} in terms of two orthogonal matrices B_+, B_- .

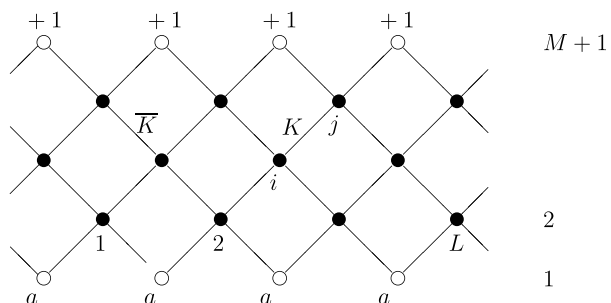
In Sect. 8 we comment on whether our result for \tilde{W} can be generalized to the N -state superintegrable chiral Potts model.

2 The Model

The model is defined on the square lattice \mathcal{L} , rotated through 45° , with $M + 1$ horizontal rows, each containing L spins, as in Fig. 1.

We impose cylindrical boundary conditions, so that the last column L is followed by the first column 1. At each site i there is a spin σ_i , taking the values $+1$ or -1 . The spins in the bottom row are fixed to have value $a = \pm 1$, those in the top row to have value $+1$. Adjacent spins σ_i, σ_j on southwest to northeast edges (with i below j) interact with Boltzmann weight $\exp(K\sigma_i\sigma_j)$; those on southeast to northwest edges with weight $\exp(\bar{K}\sigma_i\sigma_j)$.

Fig. 1 The square lattice \mathcal{L} turned through 45°



An important associated parameter is

$$k' = (\sinh 2K \sinh 2\bar{K})^{-1}. \tag{2.1}$$

The system is ferromagnetically ordered if K, \bar{K} are positive real and

$$0 < k' < 1. \tag{2.2}$$

To align with our notation for the chiral Potts model [6, 7] we are defining the RHS of (2.1) to be k' , rather than k as in Onsager [1, (2.1a)].

Partition Function

The partition function, which depends on a , is

$$Z_a = \sum_{\sigma} \prod_{(i,j)} \exp(K\sigma_i\sigma_j) \prod_{(i,j)} \exp(\bar{K}\sigma_i\sigma_j), \tag{2.3}$$

the products being over all edges of the two types. The sum is over all values of all the free spins. The partition function can be written as

$$Z_a = u_a^\dagger T^M u_+, \tag{2.4}$$

where T is the row-to-row transfer matrix, with elements

$$T_{\sigma,\sigma'} = \prod_{i=1}^L \exp(K\sigma_i\sigma'_{i+1}) \exp(\bar{K}\sigma_i\sigma'_i), \tag{2.5}$$

σ being the set of all spins $\sigma_1, \dots, \sigma_L$ in one row, and σ' being the set in the row above. Thus T is a 2^L by 2^L matrix. The vector u_a is of dimension 2^L , with entries

$$(u_a)_\sigma = \begin{cases} 1 & \text{if } \sigma_1 = \dots = \sigma_L = a, \\ 0 & \text{otherwise.} \end{cases} \tag{2.6}$$

Two transfer matrices T, T' , with different values of K and \bar{K} , commute provided they have the same value of k' [9, Sect. 7.5]

Here u^\dagger denoted the transposed conjugate of u . However, all our matrices are real, and T commutes with its transpose, so all the eigenvectors we shall discuss are also real: the complex conjugation is unnecessary.

We shall also need the matrices $S_1, \dots, S_L, C_1, \dots, C_L$, defined by

$$(S_j)_{\sigma,\sigma'} = \sigma_j \prod_{n=1}^L \delta(\sigma_n, \sigma'_n), \tag{2.7}$$

$$(C_j)_{\sigma,\sigma'} = \delta(\sigma_j, -\sigma'_j) \prod_{n=1}^L \delta(\sigma_n, \sigma'_n), \tag{2.8}$$

where the $*$ means that the term $n = j$ is excluded from the product, If I, S, C are the two-by-two matrices

$$I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad S = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad C = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \tag{2.9}$$

then we can write S_j, C_j as

$$S_j = I \otimes \cdots \otimes I \otimes S \otimes I \otimes \cdots \otimes I, \\ C_j = I \otimes \cdots \otimes I \otimes C \otimes I \otimes \cdots \otimes I,$$

the S, C on the RHS being in position j in the sequence of L factors.

Spontaneous Magnetization

Take $a = +1$, so all top and bottom boundary spins are $+$. Let 0 be a site deep within the lattice. The expectation value of its spin σ_0 is

$$\mathcal{M} = \langle \sigma_0 \rangle = Z_+^{-1} \sum_{\sigma} \sigma_0 \prod_{(i,j)} \exp(K \sigma_i \sigma_j) \prod_{(i,j)} \exp(\bar{K} \sigma_i \sigma_j). \tag{2.10}$$

We take the limit when the lattice is infinitely large, so $L, M \rightarrow \infty$, and site 0 is infinitely far from the boundaries.

The $\langle i, j \rangle$ products are unchanged by negating all spins σ_i , so if we imposed toroidal boundary conditions, then it would be true that

$$\langle \sigma_0 \rangle = -\langle \sigma_0 \rangle \tag{2.11}$$

and this would imply that $\mathcal{M} = 0$. At high temperatures ($k' \geq 1$), this is true also for our fixed-spin boundary conditions when we take the large-lattice limit. However, at lower temperatures ($0 < k' < 1$) the system has ferromagnetic long-range order and ‘‘remembers’’ the boundary conditions even in the limit of 0 deep inside a large lattice, so that

$$0 < \mathcal{M} < 1. \tag{2.12}$$

If 0 is in column i of row $j + 1$, then we can write (2.10) as

$$\mathcal{M} = u_+^\dagger T^j S_i T^{M-j} u_+ / u_+^\dagger T^M u_+. \tag{2.13}$$

Because of our cylindrical boundary conditions, we can cyclically permute the columns of \mathcal{L} , so the RHS must be independent of i .

The Sub-spaces

The operator that negates all the spins in a row is

$$R = C_1 C_2 \cdots C_L. \tag{2.14}$$

We can divide the full 2^L -dimensional space into two orthogonal subspaces $\mathcal{V}_+, \mathcal{V}_-$ such that

$$Rv = rv \quad \text{if } v \in \mathcal{V}_r, r = + \text{ or } -. \tag{2.15}$$

The operator R commutes with T and \mathcal{H} , so its eigenvectors lie either in \mathcal{V}_+ or \mathcal{V}_- . Let Λ_a be the largest eigenvalue of T in \mathcal{V}_a , ψ_a the corresponding eigenvector normalized so that $\psi_a^\dagger \psi_a = 1$, and define

$$v_+ = (u_+ + u_-)/\sqrt{2}, \quad v_- = (u_+ - u_-)/\sqrt{2}. \tag{2.16}$$

Then $T^j v_a \in \mathcal{V}_a$ for all j and

$$v_a^\dagger T^M v_b = 0 \quad \text{if } a \neq b, \tag{2.17}$$

$$v_a^\dagger T^j S_i T^{M-j} v_b = 0 \quad \text{if } a = b, \tag{2.18}$$

so (2.13) becomes

$$\mathcal{M} = \frac{v_+^\dagger T^j S_i T^{M-j} v_- + v_-^\dagger T^j S_i T^{M-j} v_+}{v_+^\dagger T^M v_+ + v_-^\dagger T^M v_-}. \tag{2.19}$$

Asymptotic Degeneracy

If $j, M - j$ are large we can take $T^j = \Lambda_+^j \psi_+ \psi_+^\dagger$ or $\Lambda_-^j \psi_- \psi_-^\dagger$ (depending on the sub-space in which T is acting). Similarly for T^{M-j} and T^M . Thus the four matrix elements in (2.19) will be proportional to $\Lambda_+^j \Lambda_-^{M-j}, \Lambda_-^j \Lambda_+^{M-j}, \Lambda_+^M, \Lambda_-^M$, respectively. However, for $0 < k < 1$, Λ_+ and Λ_- are *asymptotically degenerate*, in that for L large

$$\Lambda_+/\Lambda_- = 1 + O(e^{-L\zeta}), \quad (v_+^\dagger \psi_+)/ (v_-^\dagger \psi_-) = 1 + O(e^{-L\zeta}), \tag{2.20}$$

choosing the signs of the ψ_a appropriately. Here ζ is independent of L and is a measure of the interfacial tension between the two phases [9, Sect. 7.10]

That this should be so can be seen from a low-temperature series expansion of the eigenvalues Λ and eigenvectors ψ of T (or more conveniently the \mathcal{H} of the next sub-section) in increasing powers of k' .

One can start from a single zero-temperature configuration in which either all spins are $+1$, or all are -1 (u_+ or u_-). The expansion then proceeds identically for each choice (going from one choice to the other by negating all spins) until one reaches terms of order k'^L , when one has to consider the opposite state (all spins -1 or all $+1$) from the original. This has the same eigenvalue for $k' = 0$, so naive eigenvalue perturbation theory fails. Then and only then does one have to decide whether to symmetrize or anti-symmetrize the eigenvector with respect to R .

If ϕ_+, ϕ_- are the corresponding two near-eigenvectors, then up to this order the eigenvalues are the same and

$$u_-^\dagger \phi_+ = u_+^\dagger \phi_- = 0. \tag{2.21}$$

The actual eigenvectors of T are obtained by symmetrizing:

$$\psi_+ = (\phi_+ + \phi_-)/\sqrt{2}, \quad \psi_- = (\phi_+ - \phi_-)/\sqrt{2}. \tag{2.22}$$

The relations (2.20) then follow, using (2.16).

Hence in the limit of $j, M - j, L$ all large we can take $\Lambda_+ = \Lambda_-$, and (2.19) becomes

$$\mathcal{M} = \psi_+^\dagger S_i \psi_-. \tag{2.23}$$

3 The Associated Hamiltonian H

Consider the limit when K becomes small and \bar{K} becomes large while k' remains fixed. Then to first order in K we obtain from (2.5)

$$e^{-L\bar{K}}T = \mathcal{I} - K\mathcal{H} = e^{-K\mathcal{H}}, \tag{3.1}$$

where \mathcal{I} is the 2^L -dimensional identity matrix,

$$\mathcal{H} = \mathcal{H}_0 + k'\mathcal{H}_1, \tag{3.2}$$

where

$$\mathcal{H}_0 = -\sum_{i=1}^L S_i S_{i+1}, \quad \mathcal{H}_1 = -\sum_{i=1}^L C_i \tag{3.3}$$

and $S_{L+1} = S_1$. It follows that this hamiltonian \mathcal{H} commutes with the transfer matrix T and with R . Its ground state eigenvector (the one corresponding to the most negative eigenvalue of \mathcal{H} , and to the largest eigenvalue of T) in the subspace \mathcal{V}_a will therefore be ψ_a .

When $k' = 0$, the hamiltonian reduces to the diagonal matrix \mathcal{H}_0 and we see that the ground state eigenvectors are indeed u_+ and u_- , both with eigenvalue $-L$. It is convenient to define the closely related matrix

$$\mathcal{J} = \mathcal{H}_0 + L\mathcal{I} = \sum_{i=1}^L (\mathcal{I} - S_i S_{i+1}). \tag{3.4}$$

This has minimum eigenvalue 0, the corresponding eigenvectors being v_+ and v_- .

Hamiltonian Partition Functions

We could continue to look at the four matrix elements in (2.19), or attempt to evaluate directly the result (2.23). We prefer to follow an intermediate path and to consider the expressions

$$\tilde{Z}_+(\alpha) = v_+^\dagger e^{-\alpha\mathcal{H}} v_+, \quad \tilde{Z}_-(\beta) = v_-^\dagger e^{-\beta\mathcal{H}} v_-, \tag{3.5}$$

$$\tilde{W}(\alpha, \beta, x) = v_+^\dagger e^{-\alpha\mathcal{H}} e^{-\rho\mathcal{J}} S_1 e^{-\beta\mathcal{H}} v_-, \tag{3.6}$$

where

$$x = e^{-4\rho}. \tag{3.7}$$

This corresponds to taking the limits $K \rightarrow 0$, $j, M - j \rightarrow \infty$ in the matrix elements of (2.19), while keeping $\alpha = jK$ and $\beta = K(M - j)$ fixed.

The advantage of these expressions is that they are finite-lattice partition functions, but with T^j replaced by the rather simpler matrix $e^{-\alpha\mathcal{H}}$. For definiteness, we have taken the i in (2.19) to be 1.

We have introduced the factor $e^{-\rho\mathcal{J}}$ into \tilde{W} because the elements of the diagonal matrix \mathcal{J} take the values $0, 4, 8, \dots, 4[L/2]$, where $[x]$ is the integer part of x . Hence $\tilde{W}(\alpha, \beta, x)$ is a polynomial in x of degree $[L/2]$. This naturally manifests itself in the following working and provides a useful check against errors.

If $\rho \rightarrow \infty$, so $x \rightarrow 0$, then

$$e^{-\rho \mathcal{J}} S_1 \rightarrow u_+ u_+^\dagger - u_- u_-^\dagger = v_+ v_-^\dagger + v_- v_+^\dagger, \tag{3.8}$$

hence

$$\tilde{W}(\alpha, \beta, 0) = \tilde{Z}_+(\alpha) \tilde{Z}_-(\beta). \tag{3.9}$$

Also,

$$\begin{aligned} \tilde{W}(\alpha, 0, x) &= \tilde{Z}_+(\alpha), \\ \tilde{W}(0, \beta, x) &= \tilde{Z}_-(\beta). \end{aligned} \tag{3.10}$$

When α is large and positive, we can replace $e^{-\alpha \mathcal{H}}$ in the above definitions by $e^{-\alpha \Lambda_+ \psi_+ \psi_+^\dagger}$ or by $e^{-\alpha \Lambda_- \psi_- \psi_-^\dagger}$, according to the sub-space in which \mathcal{H} is acting. One can then verify that

$$\mathcal{M} = \lim_{\alpha, \beta, L \rightarrow \infty} \frac{\tilde{W}(\alpha, \beta, 1)}{[\tilde{Z}_+(2\alpha) \tilde{Z}_-(2\beta)]^{1/2}}. \tag{3.11}$$

We show in the following sections that we can use the techniques of Kaufman [2] and Yang [5] to express the \tilde{Z}, \tilde{W} expressions as square roots of L by L determinants, and hence evaluate \mathcal{M} by using Szegő’s theorem on Toeplitz forms [4, 8]. We further show that the determinants themselves can be expressed as squares of m by m determinants, where $L - 2 \leq 2m \leq L$.

4 The Clifford Algebra

General Remarks

Set

$$D_j = iC_j S_j = -iS_j C_j, \tag{4.1}$$

$$\Gamma_j = C_1 C_2 \cdots C_{j-1} D_j, \quad \Gamma_{L+j} = C_1 C_2 \cdots C_{j-1} S_j \tag{4.2}$$

for $j = 1, \dots, L$.

These $\Gamma_1, \dots, \Gamma_{2L}$ anti-commute:

$$\Gamma_i \Gamma_j + \Gamma_j \Gamma_i = 2\delta_{i,j} \mathcal{I} \tag{4.3}$$

for $1 \leq i, j \leq 2L$, so form a Clifford algebra. If P is any orthogonal (not necessarily real) $2L$ by $2L$ matrix and

$$\Gamma_i^* = \sum_{j=1}^{2L} P_{i,j} \Gamma_j, \tag{4.4}$$

then it is also true that

$$\Gamma_i^* \Gamma_j^* + \Gamma_j^* \Gamma_i^* = 2\delta_{i,j} \mathcal{I}. \tag{4.5}$$

If V is a 2^L -dimensional matrix such that

$$V \Gamma_i V^{-1} = \sum_{j=1}^{2L} v_{j,i} \Gamma_j, \tag{4.6}$$

for $i = 1, \dots, 2L$, then we say that the $2L$ by $2L$ matrix \widehat{V} with elements $v_{i,j}$ is the *representative* of V . If V is real and symmetric, then \widehat{V} is orthogonal and hermitian.

Such matrices form a group \mathcal{G} [10]. If V, V' have the property (4.6), with representatives $\widehat{V}, \widehat{V}'$, then so does VV' , and its representative is $\widehat{V}\widehat{V}'$. Also, if $\widehat{V} = \widehat{V}'$, then

$$V = cV',$$

where c is a scalar factor. This means that a representative matrix determines its parent to within a scalar factor.

If Γ_i, Γ_i^* are related by (4.4), then

$$V\Gamma_i^*V^{-1} = \sum_j v_{j,i}^* \Gamma_j^*, \tag{4.7}$$

where the $v_{j,i}^*$ are the elements of the matrix $P\widehat{V}P^T$. Thus mapping Γ_i to Γ_i^* takes \widehat{V} to $P\widehat{V}P^T$. Similarly, if a matrix 2^L -dimensional H satisfies

$$H\Gamma_i - \Gamma_i H = (H, \Gamma_i) = \sum_{j=1}^{2L} h_{j,i} \Gamma_j, \tag{4.8}$$

where the $h_{i,j}$ are elements of a matrix \widehat{H} , then the same mapping takes \widehat{H} to $P\widehat{H}P^T$. We call \widehat{H} the *H-representative* of H .

Any representative \widehat{V} is orthogonal and any H -representative \widehat{H} is anti-symmetric. If V (H) is hermitian, then (because the Γ_j are hermitian), then \widehat{V} (\widehat{H}) is also hermitian.

Exponentials of $\mathcal{H}, \widehat{\mathcal{H}}$

If \widehat{H} is hermitian, then it must be pure imaginary and anti-symmetric. It is diagonalizable by a unitary matrix. Further, its eigenvalues are real and occur in pairs λ and $-\lambda$, the corresponding eigenvectors being complex conjugates of one another. Adding and subtracting these eigenvectors, it follows that there is a real orthogonal transformation that takes $\widehat{\mathcal{H}}$ to the form

$$\mathcal{D} = \begin{pmatrix} 0 & iD \\ -iD & 0 \end{pmatrix}, \tag{4.9}$$

where D is a diagonal matrix.

By making such a transformation and then explicitly considering the two-by-two constituent blocks of \mathcal{D} , we find that $\exp(-\alpha\mathcal{H}) \in \mathcal{G}$, so

$$\text{representative of } \exp(-\alpha\mathcal{H}) = \exp(-\alpha\widehat{\mathcal{H}}). \tag{4.10}$$

Partition Functions

If V is hermitian, we can similarly find an orthogonal transformation (4.4) that takes \widehat{V} to block diagonal form, each block being hermitian and orthogonal, so of the form

$$\begin{pmatrix} \cosh 2\alpha_j & i \sinh 2\alpha_j \\ -i \sinh 2\alpha_j & \cosh 2\alpha_j \end{pmatrix}, \tag{4.11}$$

where $j = 1, \dots, L$. If we arrange the Γ_j^* so that this block is in rows and columns $j, L + j$, it follows that

$$V = c \exp \left(i \sum_{j=1}^L \alpha_j \Gamma_j^* \Gamma_{L+j}^* \right), \tag{4.12}$$

where c is some scalar factor.

Each term $\Gamma_j^* \Gamma_{L+j}^*$ commutes with all the other terms, so we have decomposed V into a direct product of two-by-two matrices, the j th such matrix having eigenvalues $e^{\alpha_j}, e^{-\alpha_j}$. Hence

$$\text{trace } V / (\det V)^{1/d} = \prod_{j=1}^L 2 \cosh \alpha_j, \tag{4.13}$$

where $d = 2^L$. On the other hand, from (4.11),

$$\det (I + \widehat{V}) = \prod_{j=1}^L 4 \cosh^2 \alpha_j, \tag{4.14}$$

so we have the identity

$$\text{trace } V / (\det V)^{1/d} = [\det (I + \widehat{V})]^{1/2}. \tag{4.15}$$

Note that the relations (4.10), (4.15) are quite general, being unchanged by the orthogonality transformation (4.4).

Representative of $\exp(-\alpha \mathcal{H})$

From (3.3), (4.2),

$$\begin{aligned} \mathcal{H}_0 &= iR\Gamma_L\Gamma_{L+1} - i \sum_{j=1}^{L-1} \Gamma_j\Gamma_{L+j+1}, \\ \mathcal{H}_1 &= i \sum_{j=1}^L \Gamma_j\Gamma_{L+j}, \end{aligned} \tag{4.16}$$

so we see that $\mathcal{H}_0, \mathcal{H}_1$ are both quadratic forms in the Γ_i , provided we introduce the spin-reversal operator R . We can do this, since in either sub-space \mathcal{V}_\pm , R commutes with T and \mathcal{H} , having the value $r = +1$ in \mathcal{V}_+ , and $r = -1$ in \mathcal{V}_- .

From now on we take \mathcal{H}_0 to be defined by (4.16) and \mathcal{J} by (3.4), with R fixed to be either $r = +1$ or $r = -1$. The minimum eigenvalue of \mathcal{J} is still zero, but is now unique, the corresponding eigenvector being either v_+ or v_- .

To calculate the matrix elements \tilde{Z}, \tilde{W} using (4.15), we shall need to write each as the trace of an operator. We can do this by introducing the operator $\exp(-\gamma \mathcal{J})$. In the limit $\gamma \rightarrow +\infty$ we have

$$e^{-\gamma \mathcal{J}} = v_r v_r^\dagger. \tag{4.17}$$

It follows that

$$\begin{aligned}
 [\mathcal{H}_0, \Gamma_i] &= 2i\Gamma_{L+i+1}, & [\mathcal{H}_0, \Gamma_{L+i}] &= -2i\Gamma_{i-1}, \\
 [\mathcal{H}_1, \Gamma_i] &= -2i\Gamma_{L+i}, & [\mathcal{H}_1, \Gamma_{L+i}] &= 2i\Gamma_i,
 \end{aligned}
 \tag{4.18}$$

for $i = 1, \dots, L$, provided that on the RHS we take $\Gamma_{2L+1} = -r\Gamma_{L+1}$ and $\Gamma_0 = -r\Gamma_L$.

We shall need some sparse L by L matrices. Let $0, I$ be the zero and identity L by L matrices, respectively, and let \mathcal{A}, \mathcal{B} be the one-off and two-off-diagonal L by L matrices

$$\mathcal{A} = \begin{pmatrix} 0 & 1 & 0 & \cdot & \cdot & \cdot & \cdot \\ \cdot & 0 & 1 & 0 & & & \\ \cdot & & & \cdot & & & \\ \cdot & & & & \cdot & & \\ \cdot & & & & & 0 & 1 & 0 \\ 0 & \cdot & & & & & 0 & 1 \\ -r & 0 & \cdot & \cdot & \cdot & \cdot & \cdot & 0 \end{pmatrix}, \quad \mathcal{B} = \begin{pmatrix} 0 & 0 & 1 & 0 & \cdot & \cdot & \cdot \\ \cdot & 0 & 0 & 1 & 0 & & \\ \cdot & & & & \cdot & & \\ \cdot & & & & & \cdot & \\ \cdot & & & & & & \cdot \\ 0 & & & & & 0 & 0 & 1 \\ -r & 0 & & & & & 0 & 0 \\ 0 & -r & 0 & \cdot & \cdot & \cdot & \cdot & 0 \end{pmatrix}.$$

Note that

$$\mathcal{A}\mathcal{A}^T = \mathcal{A}^T\mathcal{A} = I, \quad \mathcal{B} = \mathcal{A}^2.
 \tag{4.19}$$

Then from (4.18), the H -representatives are

$$\widehat{\mathcal{H}}_0 = 2i \begin{pmatrix} 0 & -\mathcal{A} \\ \mathcal{A}^T & 0 \end{pmatrix},
 \tag{4.20}$$

$$\widehat{\mathcal{H}}_1 = 2i \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix}.
 \tag{4.21}$$

Diagonalization of \mathcal{H}

From (3.2), the H -representative of \mathcal{H} is $\widehat{\mathcal{H}}_0 + k'\widehat{\mathcal{H}}_1$.

We calculate this matrix. First we define a $2L$ by $2L$ matrix M (not to be confused with the M of Sect. 2):

$$M = \begin{pmatrix} i & -i\mathcal{A} \\ \mathcal{A}^T & 1 \end{pmatrix}
 \tag{4.22}$$

(writing i for iI and 1 for I). Then

$$M^{-1}\widehat{\mathcal{H}}_0M = 2 \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix},
 \tag{4.23}$$

$$M^{-1}\widehat{\mathcal{H}}_1M = \begin{pmatrix} \mathcal{A} + \mathcal{A}^T & I - \mathcal{B} \\ I - \mathcal{B}^T & -\mathcal{A} - \mathcal{A}^T \end{pmatrix}.
 \tag{4.24}$$

Thus $M^{-1}\widehat{\mathcal{H}}_0M$ is diagonal.

For $j = 1, \dots, L$, define

$$\theta_j = \begin{cases} \pi(2j - 1)/L & \text{if } r = +, \\ 2\pi j/L & \text{if } r = -, \end{cases}
 \tag{4.25}$$

set

$$z_j = e^{i\theta_j},$$

and let \hat{P}, \hat{Q} be the L by L matrices with entries

$$\hat{P}_{i,j} = z_j^{i-1}/\sqrt{L}, \quad \hat{Q}_{i,j} = iz_j^{i-2}/\sqrt{L}. \tag{4.26}$$

These matrices are unitary, but not orthogonal. Set

$$P = \begin{pmatrix} \hat{P} & 0 \\ 0 & \hat{Q} \end{pmatrix}. \tag{4.27}$$

Then

$$\tilde{\mathcal{H}}_0 = P^{-1}M^{-1}\hat{\mathcal{H}}_0MP = 2 \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}, \tag{4.28}$$

$$\tilde{\mathcal{H}}_1 = P^{-1}M^{-1}\hat{\mathcal{H}}_1MP = 2 \begin{pmatrix} \tilde{C} & \tilde{S} \\ \tilde{S} & -\tilde{C} \end{pmatrix},$$

where \tilde{C}, \tilde{S} are diagonal L by L matrices with diagonal entries

$$\tilde{C}_{j,j} = \cos \theta_j, \quad \tilde{S}_{j,j} = \sin \theta_j. \tag{4.29}$$

Thus $\tilde{\mathcal{H}} = \tilde{\mathcal{H}}_0 + k'\tilde{\mathcal{H}}_1$ can be re-arranged as a matrix consisting of L two-by-two diagonal blocks. It is then straightforward to calculate its exponential, giving

$$\exp(-\alpha\tilde{\mathcal{H}}) = \begin{pmatrix} \tilde{U} & \tilde{V} \\ \tilde{V} & \tilde{T} \end{pmatrix}, \tag{4.30}$$

where $\tilde{U} = \tilde{U}_r(\alpha), \tilde{V} = \tilde{V}_r(\alpha), \tilde{T} = \tilde{T}_r(\alpha)$ are diagonal matrices with diagonal entries (for $j = 1, \dots, L$)

$$\begin{aligned} \tilde{U}_{jj} &= u_j(r, \alpha) = \cosh(2\alpha\lambda_j) + \frac{1 - k' \cos \theta_j}{\lambda_j} \sinh(2\alpha\lambda_j), \\ \tilde{V}_{jj} &= v_j(r, \alpha) = -\frac{k' \sin \theta_j \sinh(2\alpha\lambda_j)}{\lambda_j}, \\ \tilde{T}_{jj} &= t_j(r, \alpha) = \cosh(2\alpha\lambda_j) - \frac{1 - k' \cos \theta_j}{\lambda_j} \sinh(2\alpha\lambda_j) \end{aligned} \tag{4.31}$$

and

$$\lambda_j = (1 - 2k' \cos \theta_j + k^2)^{1/2}. \tag{4.32}$$

Calculation of \tilde{Z}

We want to calculate $\tilde{Z}_+(\alpha)$ and $\tilde{Z}_-(\beta)$ from (3.5). First consider $\tilde{Z}_+(\alpha)$, so take $r = +$ in the above equations. If we are working in the sub-space \mathcal{V}_+ , we can use (4.17) to write (3.5) as

$$\tilde{Z}_+(\alpha) = \text{trace } e^{-\gamma\mathcal{J}} e^{-\alpha\gamma\mathcal{H}} \tag{4.33}$$

in the limit $\gamma \rightarrow +\infty$. Since \mathcal{H}_0 and \mathcal{H}_1 are traceless,

$$\det e^{-\rho\mathcal{H}_0} = \det e^{-\alpha\mathcal{H}} = 1 \tag{4.34}$$

and from (3.4), $\det e^{-\gamma\mathcal{J}} = e^{-Ld\gamma}$, $\widehat{\mathcal{J}} = \widehat{\mathcal{H}}_0$. From (4.15) we therefore have

$$e^{2L\gamma} \widetilde{Z}_+(\alpha)^2 = \det(I + e^{-\gamma\widehat{\mathcal{H}}_0} e^{-\alpha\widehat{\mathcal{H}}}) = \det(e^{\gamma\widehat{\mathcal{H}}_0} + e^{-\alpha\widehat{\mathcal{H}}}). \tag{4.35}$$

Using the similarity transformation of (4.28), we can replace $\widehat{\mathcal{H}}_0, \widehat{\mathcal{H}}$ in (4.35) by $\widetilde{\mathcal{H}}_0, \widetilde{\mathcal{H}}$, giving

$$e^{2L\gamma} \widetilde{Z}_+(\alpha)^2 = \det \begin{pmatrix} e^{-2\gamma}I + \widetilde{U} & \widetilde{V} \\ \widetilde{V} & e^{2\gamma}I + \widetilde{T} \end{pmatrix}. \tag{4.36}$$

In the limit of γ large we expect the $e^{-2\gamma}I, \widetilde{V}, \widetilde{T}$ blocks in this determinant to become relatively negligible. The factors involving γ then cancel, leaving

$$\widetilde{Z}_+(\alpha)^2 = \det \widetilde{U} = \det \widetilde{U}_+(\alpha) = \prod_{j=1}^L u_j(+, \alpha). \tag{4.37}$$

Similarly,

$$\widetilde{Z}_-(\alpha)^2 = \det \widetilde{U}_-(\alpha) = \prod_{j=1}^L u_j(-, \alpha). \tag{4.38}$$

Calculation of \widetilde{W}

The function $\widetilde{W}(\alpha, \beta, x)$ is defined by (3.6). We note that the operator S_1 takes a vector in \mathcal{V}_+ to one in \mathcal{V}_- , and vice-versa. In particular,

$$S_1 v_r = v_{-r}$$

so

$$\lim_{\gamma \rightarrow +\infty} e^{-\gamma\mathcal{J}} S_1 = v_{-r} v_r^\dagger. \tag{4.39}$$

We can use this to write (3.6) as a trace:

$$\widetilde{W}(\alpha, \beta, x) = \text{trace } e^{-\gamma\mathcal{J}} S_1 e^{-\alpha\mathcal{H}} e^{-\rho\mathcal{J}} S_1 e^{-\beta\mathcal{H}}, \tag{4.40}$$

in the limit $\gamma \rightarrow +\infty$. We take the factors $e^{-\gamma\mathcal{J}}, e^{-\beta\mathcal{H}}$ to be acting in \mathcal{V}_- , the two other exponential factors to be acting in \mathcal{V}_+ .

The reason we are able to make further progress is that S_1 also belongs to the group \mathcal{G} . In fact $S_1 = \Gamma_{L+1}$. This is linear in the Γ_j , unlike \mathcal{H}_0 and \mathcal{H}_1 , which from (4.16) are quadratic. Even so,

$$S_1 \Gamma_j S_1^{-1} = \epsilon_j \Gamma_j, \tag{4.41}$$

where $\epsilon_{L+1} = +1$, else $\epsilon_j = -1$. The representative of S_1 is therefore

$$\widehat{S}_1 = \begin{pmatrix} -I & 0 \\ 0 & -E \end{pmatrix}, \tag{4.42}$$

where E is the diagonal matrix

$$E = \begin{pmatrix} -1 & 0 & \cdot & \cdot & \cdot & \cdot \\ 0 & 1 & 0 & & & \\ \cdot & & \cdot & & & \\ \cdot & & & 0 & 1 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & 1 \end{pmatrix}. \tag{4.43}$$

Remembering that $\widehat{\mathcal{J}} = \widehat{\mathcal{H}}_0$, the representative of the matrix whose trace is to be evaluated in (4.40) is therefore the $2L$ by $2L$ matrix

$$\widehat{\mathcal{W}}(\alpha, \beta, x) = e^{-\gamma \widehat{\mathcal{H}}_0} \widehat{\mathcal{S}}_1 e^{-\alpha \widehat{\mathcal{H}}} e^{-\rho \widehat{\mathcal{H}}_0} \widehat{\mathcal{S}}_1 e^{-\beta \widehat{\mathcal{H}}} \tag{4.44}$$

and from (4.15)

$$e^{2L(\gamma+\rho)} \widetilde{\mathcal{W}}(\alpha, \beta, x)^2 = \det [I + \widehat{\mathcal{W}}(\alpha, \beta, x)]. \tag{4.45}$$

We now use the similarity transformation of (4.28). We have to be careful because $A, M, \widehat{P}, \widehat{Q}, \mathcal{P}$ depend on r . We write them more explicitly as $A_r, M_r, \widehat{P}_r, \widehat{Q}_r, \mathcal{P}_r$. Similarly we write the θ_j, z_j of (4.25)–(4.27) as $\theta_{r,j}, z_{r,j}$. For the middle two exponential factors we take $r = +$, for the others we take $r = -$. The result is

$$\widetilde{\mathcal{W}} = \mathcal{P}_-^{-1} M_-^{-1} \widehat{\mathcal{W}} M_- \mathcal{P}_- = e^{-\gamma \widetilde{\mathcal{H}}_0} \widetilde{\mathcal{S}}_- e^{-\alpha \widetilde{\mathcal{H}}} e^{-\rho \widetilde{\mathcal{H}}_0} \widetilde{\mathcal{S}}_+ e^{-\beta \widetilde{\mathcal{H}}}, \tag{4.46}$$

where

$$\widetilde{\mathcal{S}}_r = \mathcal{P}_r^{-1} M_r^{-1} \widehat{\mathcal{S}}_1 M_r \mathcal{P}_r. \tag{4.47}$$

This threatens to become messy, but we find some remarkable simplifications. Firstly,

$$M_r \widehat{\mathcal{S}}_1 = \widehat{\mathcal{S}}_1 M_{-r}, \tag{4.48}$$

so $M_r^{-1} \widehat{\mathcal{S}}_1 M_{-r} = \widehat{\mathcal{S}}_1$ and (4.47) becomes

$$\widetilde{\mathcal{S}}_r = \mathcal{P}_r^{-1} \widehat{\mathcal{S}}_1 \mathcal{P}_{-r}. \tag{4.49}$$

The second surprise is that

$$\widehat{Q}_r^{-1} E \widehat{Q}_{-r} = \widehat{P}_r^{-1} \widehat{P}_{-r} = -F_r, \tag{4.50}$$

where F_r is the L by L matrix with entries

$$F_{i,j}^{(r)} = \frac{2z_{r,i}}{L(z_{-r,j} - z_{r,i})}, \tag{4.51}$$

satisfying

$$F_- F_+ = I. \tag{4.52}$$

Hence $\widetilde{\mathcal{S}}_r$ has the comparatively simple block-diagonal form

$$\widetilde{\mathcal{S}}_r = \begin{pmatrix} F_r & 0 \\ 0 & F_r \end{pmatrix}. \tag{4.53}$$

Let

$$\tilde{Y} = \tilde{S}_- e^{-\alpha\tilde{\mathcal{H}}} e^{-\rho\tilde{\mathcal{H}}_0} \tilde{S}_+ e^{-\beta\tilde{\mathcal{H}}}, \tag{4.54}$$

then from (4.45) and (4.46)

$$e^{2L(\gamma+\rho)} \tilde{W}(\alpha, \beta, x)^2 = \det \left(I + e^{-\gamma\tilde{\mathcal{H}}_0} \tilde{Y} \right) = \det \left(e^{\gamma\tilde{\mathcal{H}}_0} I + \tilde{Y} \right). \tag{4.55}$$

The RHS has the same structure as that of (4.36), so taking the limit $\gamma \rightarrow +\infty$, we obtain, similarly to (4.37),

$$e^{2L\rho} \tilde{W}(\alpha, \beta, x)^2 = \det \tilde{Y}_{11}, \tag{4.56}$$

where \tilde{Y}_{11} is the top-left L by L block of \tilde{Y} . Using (4.28) and (4.30), it follows that

$$\tilde{W}(\alpha, \beta, x)^2 = \det \left[F_- \tilde{U}_+(\alpha) F_+ \tilde{U}_-(\beta) + x F_- \tilde{V}_+(\alpha) F_+ \tilde{V}_-(\beta) \right]. \tag{4.57}$$

Set, for $r = +$ or $-$,

$$X_r(\alpha) = \left[\tilde{U}_r(\alpha) \right]^{-1} \tilde{V}_r(\alpha). \tag{4.58}$$

Then from (4.37) and (4.38), remembering that the \tilde{U}, \tilde{V} matrices are diagonal, so commute, we can write (4.57) as

$$\tilde{W}(\alpha, \beta, x)^2 = \tilde{Z}_+(\alpha)^2 \tilde{Z}_-(\beta)^2 \det \left[I + x F_+^{-1} X_+(\alpha) F_+ X_-(\beta) \right]. \tag{4.59}$$

To summarize thus far, (4.37), (4.38), (4.59) give $\tilde{Z}_+(\alpha), \tilde{Z}_-(\beta), \tilde{W}(\alpha, \beta, x)$ in terms of L by L determinants, and these formulae are exact for finite α, β, L . The spontaneous magnetization \mathcal{M} is then given by (3.11), taking the limit $\alpha, \beta, L \rightarrow +\infty$.

5 Calculation of \mathcal{M}

So far we have paralleled the method of Yang [5]. We now explicitly calculate \mathcal{M} from the above equations, but we use a method more like that of Montroll, Potts and Ward [4].

First set $x = 1$ and take the limit $\alpha, \beta \rightarrow +\infty$ in (4.37), (4.38), (4.59), using (4.31). We obtain

$$\mathcal{M}^2 = \xi_+ \xi_- \det \left[I + F_+^{-1} X_+(\infty) F_+ X_-(\infty) \right], \tag{5.1}$$

where

$$\xi_r = \prod_{j=1}^L \left(\frac{\lambda_{r,j} + 1 - k' \cos \theta_{r,j}}{2\lambda_{r,j}} \right)^{1/2}, \tag{5.2}$$

and $X_r(\infty)$ is the diagonal matrix with diagonal entries

$$[X_r(\infty)]_{j,j} = \frac{-k' \sin \theta_{r,j}}{\lambda_{r,j} + 1 - k' \cos \theta_{r,j}}. \tag{5.3}$$

From (4.50), $F_+ = -\hat{P}_+^{-1} \hat{P}_-$, so if we set

$$\tilde{X}_r = \hat{P}_r X_r(\infty) \hat{P}_r^{-1} \tag{5.4}$$

then (5.1) can be written

$$\mathcal{M}^2 = \xi_+ \xi_- \det \left[I + \tilde{X}_+ \tilde{X}_- \right]. \tag{5.5}$$

We could write this determinant as

$$\det \tilde{X}_+ \det \left[\tilde{X}_+^{-1} + \tilde{X}_- \right].$$

From the definition (4.26) of \hat{P} , the matrices \tilde{X}_+ , \tilde{X}_- are anti-cyclic and cyclic, respectively, i.e.

$$[\tilde{X}_r]_{i,j} = \eta_{i-j} = -r\eta_{L+i-j},$$

where η is defined by this equation. The matrix $\tilde{X}_+^{-1} + \tilde{X}_-$ is therefore Toeplitz, its elements i, j depending only on $i - j$, and one might hope to use Szegő's theorem [4, 8] to evaluate the determinant for large L .

Unfortunately this naive approach does not work because for large L the determinant is not dominated by its near-diagonal elements, in particular elements such as $(L, 1)$, $(2, L - 3)$ do not tend to zero. However, we can rescue this idea by making a particular bilinear transformation of \tilde{X}_+ , \tilde{X}_- .

We set

$$C_r = (I - i\tilde{X}_r)^{-1} (I + i\tilde{X}_r) = \hat{P}_r c_r \hat{P}_r^{-1}, \tag{5.6}$$

where c_r is a diagonal matrix with diagonal elements

$$(c_r)_{j,j} = \frac{\lambda_{r,j} + 1 - k' \exp(i\theta_{r,j})}{\lambda_{r,j} + 1 - k' \exp(-i\theta_{r,j})}. \tag{5.7}$$

Solving for X_r and substituting into (5.5), we obtain

$$\mathcal{M}^2 = \xi'_+ \xi'_- \det[(C_+ + C_-)/2], \tag{5.8}$$

where

$$\begin{aligned} \xi'_r &= \xi_r \det(I - i\tilde{X}_r) \\ &= \prod_{j=1}^L \left\{ \frac{[\lambda_{r,j} + 1 - k' \exp(-i\theta_{r,j})]^2}{2\lambda_{r,j}(\lambda_{r,j} + 1 - k' \cos \theta_{r,j})} \right\}^{1/2}. \end{aligned} \tag{5.9}$$

From (4.32),

$$\lambda_{r,j}^2 = [1 - k' \exp(i\theta_{r,j})][1 - k' \exp(-i\theta_{r,j})],$$

from which it follows that

$$\xi'_r = \prod_{j=1}^L \left[\frac{1 - k' \exp(-i\theta_{r,j})}{1 - k' \exp(i\theta_{r,j})} \right]^{1/4}. \tag{5.10}$$

The $\theta_{r,j}$ are either π or 2π , or occur in pairs $\theta_{r,j}, 2\pi - \theta_{r,j}$. It follows that

$$\xi'_+ = \xi'_- = 1. \tag{5.11}$$

Also, $(c_r)_{j,j}$ simplifies to

$$(c_r)_{j,j} = \left[\frac{1 - k' \exp(-i\theta_{r,j})}{1 - k' \exp(i\theta_{r,j})} \right]^{1/2}. \tag{5.12}$$

From (4.26) and (5.6), the elements of C_r are

$$(C_r)_{i,j} = \frac{1}{L} \sum_{m=1}^L (c_r)_{m,m} z_{r,m}^{i-j}, \tag{5.13}$$

where $z_{r,m} = e^{i\theta_{r,m}}$. The summand depends on m via $\theta_{r,m}$ and is an analytic function of $\theta_{r,m}$ on the real axis, periodic of period 2π . The $\theta_{r,m}$ are distributed uniformly throughout this period. If i, j are held finite and $L \rightarrow \infty$, it follows that

$$(C_r)_{i,j} = \frac{1}{2\pi} \int_0^{2\pi} f(\theta) e^{i(i-j)\theta} d\theta, \tag{5.14}$$

where

$$f(\theta) = \left[\frac{1 - k' \exp(-i\theta)}{1 - k' \exp(i\theta)} \right]^{1/2}. \tag{5.15}$$

This is actually the function $e^{i\delta^*}$ of Onsager [1, (89)] and Montroll et al. [4, (42)]. In their notation $z_1 = \tanh H'$, $z_2^* = \tanh H^*$ and our k' is $k' = z_2^*/z_1$. We have replaced ω by θ and taken the hamiltonian limit $z_1, z_2^* \rightarrow 0$.

Note that the limit (5.14) is the same for $r = +1$ and $r = -1$. Also, for large but finite L , the corrections to (5.14) vanish exponentially with L . The near-diagonal elements of C_+ and C_- therefore become equal when L is large.

Since $z_{r,m}^L = -r$, incrementing i or j in (5.13) by L negates $(C_r)_{i,j}$ for $r = +$, and leaves it unchanged for $r = -$. This means that the near-top-right and near-bottom-left elements of C_+ and C_- also tend to finite limits as L becomes large, but are equal and opposite. Their sum therefore approaches exponentially to zero.

Hence when L is large we expect (5.8) to become

$$\mathcal{M}^2 = \det C, \tag{5.16}$$

where C is the L by L matrix with elements (i, j) given by (5.14). These elements tend exponentially to zero as $|i - j|$ becomes large. This is a special case of the Toeplitz matrix discussed in [4] and we can use the general result (68) therein:

$$\lim_{L \rightarrow \infty} G^{-L} \det C = \exp \left(\sum_{n=1}^{\infty} n \kappa_n \kappa_{-n} \right), \tag{5.17}$$

where

$$G = \exp \left(\frac{1}{2\pi} \int_0^{2\pi} \log f(\theta) d\theta \right),$$

and

$$\log f(\theta) = \sum_{n=-\infty}^{\infty} \kappa_n e^{in\theta}.$$

Since $\log f(\theta)$ is an odd periodic function of θ , we have $G = 1$. Also, taking the logarithm of (5.15) and Laurent expanding, we readily find

$$2\kappa_n = k^m/n \quad \text{if } n > 0, \quad 2\kappa_n = -k'^{-n}/n \quad \text{if } n < 0. \tag{5.18}$$

It follows that $\det C = (1 - k^2)^{1/4}$, and hence

$$\mathcal{M} = (1 - k^2)^{1/8}. \tag{5.19}$$

This is of course Onsager’s famous result [3–5].

6 Connection with the Superintegrable Chiral Potts Model

The superintegrable case of the N -state chiral Potts model has some properties that closely resemble the Ising model. In particular, there is a spin-shift operator R that is the natural generalization of (2.14) and divides the vector space into N sub-spaces \mathcal{V}_Q , labelled by $Q = 0, 1, \dots, N_1$. Within each \mathcal{V}_Q , if one imposes the fixed-spin boundary conditions of Fig. 1, the transfer matrix T generates a yet smaller sub-space of dimension 2^m , where

$$m = m_Q = \left\lceil \frac{(N - 1)L - Q}{N} \right\rceil. \tag{6.1}$$

Here $[x]$ means the integer part of the real number x .

Within \mathcal{V}_Q , T is a direct product of m two-by-two matrices, and there are similarity transformations that reduce the associated hamiltonian to the direct sum [11, (2.20)]

$$\mathcal{H} = \mu_Q - N \sum_{j=1}^m [(1 - k' \cos \theta_j)S_j - k' \sin \theta_j C_j],$$

where

$$\mu = \mu_Q = 2k'Q + (1 + k')(mN - NL + L) \tag{6.2}$$

and $\theta_1, \dots, \theta_m$ are defined by

$$\cos \theta_j = (1 + w_j)/(1 - w_j), \quad 0 < \theta_j < \pi, \tag{6.3}$$

the w_1, \dots, w_m being the zeros of

$$P(z^N) = z^{-Q} \sum_{n=0}^{N-1} \omega^{(L+Q)n} [(z^N - 1)/(z - \omega^n)]^L, \tag{6.4}$$

which is a polynomial in $w = z^N$ of degree m .

6.1 The Partition Functions \tilde{Z}

One can explicitly calculate [12] the partition functions that generalize (4.37), (4.38):

$$\tilde{Z}_Q(\alpha) = e^{-\mu\alpha} u_1(\alpha) \cdots u_m(\alpha), \tag{6.5}$$

the function $u_j(\alpha)$ again being defined by (4.31), (4.32).

For $N = 2$ the model reduces to the Ising model with

$$r = 1 - 2Q, \quad m(r) = m_Q, \tag{6.6}$$

so $Q = 0$ corresponds to $R = +1$, while $Q = 1$ corresponds to $R = -1$. The result (6.5) of course agrees with (4.37), (4.38), but has a slightly different form. The partition function is no longer squared on the LHS and instead of there being L variables θ_j , there are only m , which lie between $(L - 2)/2$ and $L/2$. They are the same as the θ_j of the previous sections, but j takes only the values 1 to m and $0 < \theta_j < \pi$.

These differences are easily explained. For every θ_j with $j = 1, \dots, m$ there is a $\theta_{j'} = 2\pi - \theta_j$, where $j' = L + 1 - Q - j$ and $m < j' \leq L$. They have the same value of $u_j(\alpha)$ and each occurs in (4.37), (4.38), while only one occurs in (6.5). This accounts for the absence of the square in (6.5) and the presence of the factor $e^{-\mu\alpha}$ accounts for the exceptional cases $\theta_j = \pi$ or 2π , which are included in the product in (4.37), (4.38).

There are a total of four cases to consider.

- (1) $r = +1, Q = 0, L = \text{even}$: there are no exceptional cases and $m = L/2, \mu = 0$.
- (2) $r = +1, Q = 0, L = \text{odd}$: then $m = (L - 1)/2, \mu = -1 - k'$. There is an exceptional case at $j = (L + 1)/2 = m + 1$, where $\theta_j = \pi, u_j(\alpha) = \exp[2(1 + k')\alpha] = e^{-2\mu\alpha}$.
- (3) $r = -1, Q = 1, L = \text{even}$: then $m = (L - 2)/2, \mu = -2$. There are two exceptional cases at $j = L/2$ and L , with $\theta_j = \pi$ and $2\pi, u_j(\alpha) = \exp[2(1 + k')\alpha]$ and $\exp[2(1 - k')\alpha]$, so $u_{L/2}(\alpha)u_L(\alpha) = e^{-2\mu\alpha}$.
- (4) $r = -1, Q = 1, L = \text{odd}$: then $m = (L - 1)/2, \mu = k' - 1$. There is an exceptional case at $j = L$, with $\theta_j = 2\pi, u_j(\alpha) = \exp[2(1 - k')\alpha] = e^{-2\mu\alpha}$.

In each case the total contribution of the exceptional cases to the RHS of (4.37) or (4.38) is $e^{-2\mu\alpha}$, which accounts for the factor $e^{-\mu\alpha}$ in (6.5).

6.2 The Partition Function \tilde{W}

Equation (6.5) simplifies (4.37) and (4.38), expressing the $\tilde{Z}_Q(\alpha)$ as a product rather than the square root of a product, and reducing the number of factors from L to m or m' , where

$$m = m(+) = [L/2], \quad m' = m(-) = [(L - 1)/2]. \tag{6.7}$$

Can we similarly reduce the expression (4.57) for $\tilde{W}(\alpha, \beta, x)$?

The answer is yes, apart from simple factors that are independent of x and easily calculated. We show that the matrix

$$G = \tilde{U}_+(\alpha)F_+\tilde{U}_-(\beta) + x\tilde{V}_+(\alpha)F_+\tilde{V}_-(\beta) \tag{6.8}$$

can be reduced to block lower-triangular form and its determinant expressed as simple factors times the square of an m' by m' determinant.

For brevity, in the following two sub-sections, unless indicated otherwise we write

$$\theta_j = \theta_{+,j}, \quad \theta'_j = \theta_{-,j},$$

$$u_j = \tilde{U}_+(\alpha)_{j,j} = u_{+,j}(\alpha), \quad v_j = \tilde{V}_+(\alpha)_{j,j} = v_{+,j}(\alpha),$$

$$u'_j = \tilde{U}_-(\beta)_{j,j} = u_{-,j}(\beta), \quad v'_j = \tilde{V}_-(\beta)_{j,j} = v_{-,j}(\beta).$$

We use the definition (4.51) of F_{\pm} and the properties (4.52), (6.12). We also implicitly use

$$u_{L+1-j} = u_j, \quad v_{L+1-j} = -v_j, \quad u'_{L-j} = u'_j, \quad v'_{L-j} = -v'_j. \tag{6.9}$$

We do *not* in this section need the definitions (4.31).

The Case L Even

In this case note that the $\tilde{U}_+(\alpha)$, $\tilde{V}_+(\alpha)$ on the left of (6.8) depend on the θ_j , which satisfy $\theta_i + \theta_{L+1-i} = 2\pi$, so it is natural to combine rows i and $L + 1 - i$. The $\tilde{U}_-(\beta)$, $\tilde{V}_-(\beta)$ on the right involve θ'_j , satisfying $\theta'_j + \theta'_{L-j} = 2\pi$, so it is natural to combine columns j and $L - j$. Also, columns $L/2$ and L both correspond to exceptional values of θ'_j , where $\sin \theta'_j = 0$. The elements $v'_{L/2}, v'_L$ of $\tilde{V}_-(\beta)$ vanish for these two columns. From (6.7),

$$m = L/2, \quad m' = (L - 2)/2 = m - 1. \tag{6.10}$$

We perform the following equivalence transformations sequentially on G :

- 1) $G_{ij} \rightarrow G_{ij} + G_{L+1-i,j}, \quad 1 \leq i \leq m, 1 \leq j \leq L$
- 2) $G_{ij} \rightarrow G_{ij} + G_{i,L-j}, \quad m < j < L$
- 3) $G_{i,L} \rightarrow G_{i,L} - u'_L G_{i,m}/u'_m,$
- 4) $G_{i,j} \rightarrow G_{i,j} - u'_j G_{i,m}/u'_m, \quad 1 \leq j < m$
- 5) $G_{i,j} \rightarrow G_{i,j} - 2u'_j G_{i,m}/u'_m, \quad m < j < L$
- 6) $G_{i,j} \rightarrow G_{i,j} - (1 + c'_j)u'_j G_{i,L}/u'_L, \quad m < j < L$

where $c'_j = \cos(\theta'_j)$.

The first transformation corresponds to pre-multiplying G by some matrix, all the others to post-multiplying it. Steps (2) to (6) are for all values of i , i.e. $1 \leq i \leq L$. At every step we are merely incrementing rows or columns by linear combinations of other rows or columns, so the determinant of G is unchanged.

The elements $F_{i,j}$ of F_+ satisfy

$$F_{ij} + F_{L+1-i,L-j} = F_{i,L} + F_{L+1-i,L} = -2/L \tag{6.12}$$

for all i, j . It follows that the final matrix G is such that

$$G_{i,j} = 0 \quad \text{for } 1 \leq i \leq m, m < j \leq L. \tag{6.13}$$

Thus G is now block lower-triangular:

$$G = \begin{pmatrix} Y & 0 \\ \mathcal{T} & \mathcal{Z} \end{pmatrix}, \tag{6.14}$$

where the $Y, \mathcal{T}, \mathcal{Z}$ are m by m matrices. We find

$$Y_{i,j} = \frac{-2i(\sin \theta'_j u_i u'_j + x \sin \theta_i v_i v'_j)}{L(\cos \theta_i - \cos \theta'_j)}, \tag{6.15}$$

$$Y_{i,m} = -2u_i u'_m / L$$

for $1 \leq i \leq m, 1 \leq j < m$.

Note that the equivalence transformations are independent of x , so Y, T, Z are all linear in x , their coefficients being the transforms of the corresponding coefficients of G .

We define the m by m matrix Z by (6.14), but with the rows and columns re-arranged so that, for $1 \leq i, j \leq m$,

$$(Z)_{i,j} = (G)_{L+1-i,j'}, \tag{6.16}$$

where $j' = L$ if $j = m$, else $j' = L - j$. Let D, D' be the diagonal matrices with elements

$$D_{j,j} = \sin \theta_i, \quad D'_{j,j} = \sin \theta'_i \tag{6.17}$$

for $j = 1, \dots, m$, except that $D'_{m,m} = -iu'_L/u'_m$. Then we find that

$$Z = -D^{-1}YD'. \tag{6.18}$$

Hence

$$\det G = \det(-D')(\det Y)^2 / \det D \tag{6.19}$$

and we see that $\det G$ is basically the square of the m by m determinant $\det Y$.

We can do better yet and write it in terms of the square of an $m' = m - 1$ determinant. Expand the matrices G, Y in powers of x :

$$G = \mathbf{g}_0 + x\mathbf{g}_1, \quad Y = \mathbf{y}_0 + x\mathbf{y}_1, \tag{6.20}$$

so $\mathbf{g}_0, \mathbf{g}_1, \mathbf{y}_0, \mathbf{y}_1$ are x -independent matrices and $\mathbf{g}_0 = \tilde{U}_+(\alpha)F_+\tilde{U}_-(\beta)$. From (4.52), its inverse is

$$\mathbf{g}_0^{-1} = \tilde{U}_-(\beta)^{-1}F_-\tilde{U}_+(\alpha)^{-1}. \tag{6.21}$$

We can follow the above six steps (6.11) and evaluate \mathbf{g}_0^{-1} after the six equivalence transformations. It also becomes block lower-triangular, as in (6.14). In particular its top-left block is \mathbf{y}_0^{-1} , with elements

$$[\mathbf{y}_0^{-1}]_{i,j} = \frac{-2i \sin \theta'_i}{L(\cos \theta'_i - \cos \theta_j)u'_i u_j}, \tag{6.22}$$

$$[\mathbf{y}_0^{-1}]_{m,j} = -1/(u'_m u_j),$$

for $1 \leq i < m, 1 \leq j \leq m$. This matrix is necessarily the inverse of \mathbf{y}_0 and from (6.19)

$$\det \mathbf{g}_0^{-1}G = [\det \mathbf{y}_0^{-1}Y]^2. \tag{6.23}$$

We calculate $\mathbf{y}_0^{-1}Y$. From (6.15), the last column of \mathbf{y}_1 vanishes. Hence, using (6.20),

$$\mathbf{y}_0^{-1}Y = I + x\tilde{\mathcal{D}}, \tag{6.24}$$

where $\tilde{\mathcal{D}} = \mathbf{y}_0^{-1}\mathbf{y}_1$ and again the last column of \mathcal{D} vanishes. The determinant of the RHS is therefore the same as that of the RHS truncated to its first $m' = m - 1$ rows and columns.

We can therefore ignore the last column of Y and the last row of \mathbf{y}_0^{-1} when calculating the LHS, and contract Y to the non-square m by m' matrix with elements given by the first of (6.15), \mathbf{y}_0^{-1} to the m' by m matrix with elements given by the first of (6.22).

We define the m by m' matrix A_+ and the m' by m matrix A_- by

$$(A_+)_{i,j} = \frac{2 \sin \theta_i}{L(\cos \theta_i - \cos \theta'_j)}, \quad (A_-)_{i,j} = \frac{2 \sin \theta'_i}{L(\cos \theta'_i - \cos \theta_j)}. \tag{6.25}$$

Then from (6.15), (6.22),

$$\mathbf{y}_1 = -i\tilde{V}_+(\alpha)A_+\tilde{V}_-(\beta), \quad \mathbf{y}_0^{-1} = -i\tilde{U}_-(\beta)^{-1}A_-\tilde{U}_+(\alpha)^{-1}, \tag{6.26}$$

where we now take the $\tilde{U}_r, \tilde{V}_r, \tilde{X}_r$ to be the diagonal matrices defined by (4.31), (4.58), but truncated to the first $m(r)$ terms.

Hence

$$\tilde{\mathcal{D}} = -\tilde{U}_-(\beta)^{-1}\mathcal{D}\tilde{U}_-(\beta),$$

where

$$\mathcal{D} = -A_-X_+(\alpha)A_+X_-(\beta). \tag{6.27}$$

From (6.23), we obtain

$$\det G = \det \left(\tilde{U}_+(\alpha)F_+\tilde{U}_-(\beta) \right) [\det(I_{m'} + x\mathcal{D})]^2, \tag{6.28}$$

where $I_{m'}$ is the identity m' by m' matrix.

The relations (4.57), (4.59) therefore become

$$\tilde{W}(\alpha, \beta, x) = \tilde{Z}_+(\alpha)\tilde{Z}_-(\beta) \det[I_{m'} - xA_-X_+(\alpha)A_+X_-(\beta)]. \tag{6.29}$$

We have eliminated the square on the LHS and used only matrices of dimensions m or m' , involving only $\theta_1, \dots, \theta_m, \theta'_1, \dots, \theta'_{m'}$.

The Case $r = +, L$ Odd

Now (6.7) gives $m = m' = (L - 1)/2$, so the matrix G has one exceptional row at $i = m + 1$ and one exceptional column at $j = L$. The steps corresponding to (6.11) are

- 1) $G_{ij} \rightarrow G_{ij} + G_{L+1-i,j}, \quad 1 \leq i \leq m, 1 \leq j \leq L$
- 2) $G_{ij} \rightarrow G_{ij} - 2u_i G_{m+1,j}/u_{m+1}, \quad 1 \leq i \leq m, 1 \leq j \leq L$
- 3) $G_{ij} \rightarrow G_{ij} + G_{i,L-j}, \quad m < j < L$
- 4) $G_{i,j} \rightarrow G_{i,j} - 2u'_j G_{i,L}/u'_L, \quad m < j < L$
- 5) $G_{i,j} \rightarrow G_{i,j} - u'_j G_{i,L}/u'_L, \quad 1 \leq j \leq m.$

The first two transformations correspond to pre-multiplying G_r by some matrix, all the others to post-multiplying it. Steps (3) to (6) are for all values of $1 \leq i \leq L$.

Using (6.12), we find that the transformed G has the structure

$$G = \begin{pmatrix} Y & 0 & 0 \\ \cdot & 0 & -u_{m+1}u'_L \\ \mathcal{T} & \mathcal{Z} & \cdot \end{pmatrix}, \tag{6.31}$$

where $Y, \mathcal{T}, \mathcal{Z}$ are all m by $m' = m$ blocks and the middle row and last column have width one. Hence

$$\det G = u_{m+1}u'_L (\det Y)(\det \mathcal{Z}).$$

We also find

$$Y_{ij} = \frac{-4i \cos^2(\theta_i/2)[u_i u'_j \tan(\theta'_j/2) + x v_i v'_j \tan(\theta_i/2)]}{L(\cos \theta_i - \cos \theta'_j)}, \tag{6.32}$$

and again if we rearrange \mathcal{Z} so that $\mathcal{Z}_{ij} = G_{L+1-i, L-j}$ for $1 \leq i, j \leq m$, then

$$\mathcal{Z} = -D^{-1} Y D', \tag{6.33}$$

with D, D' given by (6.17) for all $1 \leq j \leq m$. Further,

$$(\mathbf{y}_0^{-1})_{ij} = \frac{-2i \sin \theta'_i}{L(\cos \theta'_i - \cos \theta_j) u'_i u_j}. \tag{6.34}$$

From (6.15) and (6.32), \mathbf{y}_0 is different for the two cases (L even and L odd). However, the wanted elements of \mathbf{y}_1 and of \mathbf{y}_0^{-1} are the same, so again $\tilde{W}(\alpha, \beta, x)$ is given by (6.29), with A_+, A_- given by (6.25).

From (4.31), (4.58), the diagonal matrices X_{\pm} have diagonal elements

$$\begin{aligned} [X_+(\alpha)]_{jj} &= -\frac{k' \sin \theta_j \sinh(2\alpha\lambda_j)}{\lambda_j \cosh(2\alpha\lambda_j) + (1 - k' \cos \theta_j) \sinh(2\alpha\lambda_j)}, \\ [X_-(\beta)]_{jj} &= -\frac{k' \sin \theta'_j \sinh(2\beta\lambda'_j)}{\lambda'_j \cosh(2\beta\lambda'_j) + (1 - k' \cos \theta'_j) \sinh(2\beta\lambda'_j)}, \end{aligned} \tag{6.35}$$

λ_j being defined by (4.32) and λ'_j by the same equation with θ_j replaced by θ'_j .

7 \tilde{W} in Terms of Orthogonal Matrices B_+, B_-

In this section we take the result (6.29) and write it in terms, not of the matrices A_+, A_- , but of related orthogonal matrices B_+, B_- .

In both the L even and odd cases, the matrices $\mathbf{y}_0^{-1}, \mathbf{y}_0^T$ above have a very similar structure, differing only by diagonal equivalence transformations. We can therefore construct orthogonal matrices B_+, B_- that differ from A_+, A_- only by such diagonal equivalence transformations.

Let E, E' be diagonal matrices with diagonal elements

$$\begin{aligned} E_{jj} &= \sin \theta_j, & E'_{jj} &= \sin \theta'_j & \text{if } L &= \text{even}, \\ E_{jj} &= \tan(\theta_j/2), & E'_{jj} &= \tan(\theta'_j/2) & \text{if } L &= \text{odd}. \end{aligned} \tag{7.1}$$

Then whether L is even or odd, we find from (6.15), (6.17), (6.22), (6.25), (6.32), (6.34) that

$$\mathbf{y}_0^{-1} = iD'A_+^T D^{-1}, \quad \mathbf{y}_0 = -iE^{-1}A_+E', \quad A_-^T = -D^{-1}A_+D', \tag{7.2}$$

\mathbf{y}_0^{-1} being the left-inverse of \mathbf{y}_0 .

Hence we can define matrices B_+, B_- :

$$B_+ = (DE)^{-1/2}A_+(D'E')^{1/2}, \quad B_- = (E'/D')^{1/2}A_-(D/E)^{1/2} \tag{7.3}$$

such that

$$B_- B_-^T = B_+^T B_+ = I_{m'}, \quad B_- = -B_+^T, \tag{7.4}$$

so, even though they are not necessarily square, they are orthogonal in this sense. Their elements are

$$\begin{aligned}
 (B_+)_{ij} &= -(B_-)_{ji} = \frac{2 \sin \theta'_j}{L(\cos \theta_i - \cos \theta'_j)} \quad \text{if } L = \text{even}, \\
 (B_+)_{ij} &= -(B_-)_{ji} = \frac{4 \cos(\theta_i/2) \sin(\theta'_j/2)}{L(\cos \theta_i - \cos \theta'_j)} \quad \text{if } L = \text{odd}.
 \end{aligned}
 \tag{7.5}$$

Set

$$E_+ = E, \quad E_- = E'^{-1}.
 \tag{7.6}$$

Then we can write the result (6.29) as

$$\tilde{W}(\alpha, \beta, x) = \tilde{Z}_+(\alpha) \tilde{Z}_-(\beta) \det[I_{m'} - x X_+(\alpha) E_+ B_+ X_-(\beta) E_- B_-].
 \tag{7.7}$$

We can slightly generalize this result. Write $\tilde{W}(\alpha, \beta, x)$ as $\tilde{W}_+(\alpha, \beta, x)$, and define a new function $\tilde{W}_-(\alpha, \beta, x)$ to be given by (3.6) with v_+, v_- interchanged. Thus (3.5), (3.6) become

$$\tilde{Z}_r(\alpha) = v_r^\dagger e^{\alpha \mathcal{H}} v_r, \quad \tilde{W}_r(\alpha, \beta, x) = v_r^\dagger e^{\alpha \mathcal{H}} e^{\rho \mathcal{J}} S_1 e^{\beta \mathcal{H}} v_{-r}.
 \tag{7.8}$$

Then we can generalize (7.7) to

$$\tilde{W}_r(\alpha, \beta, x) = \tilde{Z}_r(\alpha) \tilde{Z}_{-r}(\beta) \det[I - x X_r(\alpha) E_r B_r X_{-r}(\beta) E_{-r} B_{-r}],
 \tag{7.9}$$

for $r = \pm$, where I is the identity matrix of dimension $m(r)$. For $r = +$, (7.9) is (7.7). For $r = -$ it can be deduced from it by taking the hermitian conjugate of the definition (7.8) of \tilde{W}_- and interchanging α with β . We also need the general identity

$$\det(I + AB) = \det(I + BA),$$

which is true for all matrices A, B even when they are non-square.

8 Conclusions

We have considered the zero-field Ising model on a cylindrical lattice with fixed spin boundary conditions on the top and bottom rows, and have replaced the transfer matrix product by the exponential of the associated hamiltonian. This leaves the eigenvectors and spontaneous magnetization unchanged.

We have then used the Clifford algebra technique of Kaufman [2] to evaluate as L by L determinants the partition function \tilde{Z} , and the partition function \tilde{W} with a single-spin operator S_1 included. In Sect. 5 we show how to use Szegő's theorem on Toeplitz matrices to then obtain the spontaneous magnetization.

Much of this merely parallels Yang's derivation [5], but in Sect. 6 we go on to reduce \tilde{Z} and \tilde{W} to determinants of size m or m' where $L - 2 \leq 2m \leq L$.

In this form we can compare the Ising model results for \tilde{Z} with those of the N -state superintegrable chiral Potts model, taking the hamiltonian limit therein [11, 12]. The Ising model is the superintegrable chiral Potts model with $N = 2$, and the results of course agree.

All this leads up to the question whether we can generalize the Ising result (6.29) or (7.9) for \tilde{W} to the N -state superintegrable chiral Potts model, for general N . This should

open up the possibility of an algebraic derivation of the spontaneous magnetization of that model. We have numerical evidence that strongly suggests the answer to the question is yes, and that one can make a fairly immediate generalization of the form of the result given in (7.9) in terms of the orthogonal matrices B_+ , B_- . We shall present the conjecture in an accompanying paper [12].

References

1. Onsager, L.: Crystal statistics. I. A two-dimensional model with an order-disorder transition. *Phys. Rev.* **65**, 117–149 (1944)
2. Kaufman, B.: Crystal statistics. II. Partition function evaluated by spinor analysis. *Phys. Rev.* **76**, 1232–1243 (1949)
3. Onsager, L.: In: Proceedings of the IUPAP Conference on Statistical Mechanics, “Discussion e osservazioni”. *Nuovo Cimento (Suppl.)*, Ser. 9 **6**, 261 (1949)
4. Montroll, E.W., Potts, R.B., Ward, J.C.: Correlations and spontaneous magnetization of the two-dimensional Ising model. *J. Math. Phys.* **4**, 308–322 (1963)
5. Yang, C.N.: The spontaneous magnetization of a two-dimensional Ising model. *Phys. Rev.* **85**, 808–816 (1952)
6. Baxter, R.J.: Derivation of the order parameter of the chiral Potts model. *Phys. Rev. Lett.* **94**, 130602 (2005)
7. Baxter, R.J.: The order parameter of the chiral Potts model. *J. Stat. Phys.* **120**, 1–36 (2005)
8. Grenander, U., Szegő, G.: *Toeplitz Forms and Their Applications*. Univ. Calif. Press, Berkeley (1958)
9. Baxter, R.J.: *Exactly Solved Models in Statistical Mechanics*. Academic, London (1989). Dover Publications, Mineola (2007)
10. Baxter, R.J.: Corner transfer matrices of the eight-vertex model. II. The Ising model case. *J. Stat. Phys.* **17**, 1–14 (1977)
11. Baxter, R.J.: Superintegrable chiral Potts model: Thermodynamic properties, an “inverse” model, and a simple associated hamiltonian. *J. Stat. Phys.* **57**, 1–39 (1989)
12. Baxter, R.J.: A conjecture for the superintegrable chiral Potts model. *J. Stat. Phys.* **132** (2008). doi:[10.1007/s10955-008-9588-x](https://doi.org/10.1007/s10955-008-9588-x)