Journal of Statistical Physics, Vol. 19, No. 5, 1978

Variational Approximations for Square Lattice Models in Statistical Mechanics

R. J. Baxter[1]

Received June 6, 1978

This paper concerns a square lattice, Ising-type model with interactions between the four spins at the corners of each face. These may include nearest and next-nearest-neighbor interactions, and interactions with a magnetic field. Provided the Hamiltonian is symmetric with respect to both row reversal and column reversal, a rapidly convergent sequence of variational approximations is obtained, giving the free energy and other thermodynamic properties. For the usual Ising model, the lowest such approximations are those of Bethe and of Kramers and Wannier. The method provides a new definition of corner transfer matrices.

KEY WORDS: Statistical mechanics; lattice statistics; Ising models; variational approximation; corner transfer matrices.

1. INTRODUCTION AND DEFINITION OF THE MODEL

In previous papers[1,2] variational approximations have been obtained for the free energy of two square lattice models in statistical mechanics; namely the monomer–dimer system and the zero-field Potts model. Here this method is applied to a fairly general square-lattice Ising model, including one-, two-, three-, and four-spin interactions. (Possible further generalizations are discussed in the summary.)

The method is interesting for two reasons. First, it provides a rapidly convergent sequence of approximations which can be thought of as improvements on the Bethe approximation. These can be used to locate phase transitions and critical points, and give very good numerical estimates of the non-critical thermodynamic properties.[1–3] Second, the method provides an alternative definition of the previously defined[4,5] "corner transfer matrices" (CTMs). (For the solvable models, i.e., the zero-field Ising and eight-

[1] Research School of Physical Sciences, The Australian National University, Canberra, Australia.

0022-4715/78/1100-0461$05.00/0

vertex models, these CTMs appear to have a very simple eigenvalue spectrum in the thermodynamic limit.)

The outline of the paper is as follows: the model is defined in this section, then the variational equations are obtained in Sections 2 and 3, the main result being (30). This equation defines certain matrices of dimension 2^n by 2^n, as well as the partition function per site κ. Taking $n = 0, 1, 2,...$ gives a hierarchy of variational approximations for κ. In Section 4 it is shown that the two lowest level approximations for the usual Ising model are the Bethe and Kramers–Wannier approximations.

In Section 5 the equations are given a graphical interpretation, which makes it obvious that for large n the matrices in (30) are the half-column, half-row, and corner transfer matrices of the square lattice.

The model considered here is defined as follows. To each site i of a square lattice assign a spin σ_i, with values $+1$ or -1 (+ or −). To each square face assign an energy $\epsilon(\sigma_i, \sigma_j, \sigma_k, \sigma_l)$, where i, j, k, l are the sites around the square, ordered as in Fig. 1, and

$$\begin{aligned} -\epsilon(a, b, c, d) = J_0 &+ \tfrac{1}{4}H(a + b + c + d) \\ &+ \tfrac{1}{2}J(ab + cd) + \tfrac{1}{2}J'(ac + bd) \\ &+ J''(ad + bc) + H_3(bcd + cda + dab + abc) + J_4 abcd \end{aligned} \quad (1)$$

Thus H, J, J', J'', H_3, and J_4 are the coefficients of one-spin, horizontal two-spin, vertical two-spin, diagonal two-spin, three-spin, and four-spin interactions, respectively. The total energy is

$$E = \sum \epsilon(\sigma_i, \sigma_j, \sigma_k, \sigma_l) \quad (2)$$

where the summation is over all square faces i, j, k, l of the lattice.

The partition function is

$$Z = \sum \prod w(\sigma_i, \sigma_j, \sigma_k, \sigma_l), \quad (3)$$

where the product is over all faces of the lattice, the sum is over all values (+ or −) of all the spins, and

$$w(a, b, c, d) = \exp[-\epsilon(a, b, c, d)/k_B T], \quad (4)$$

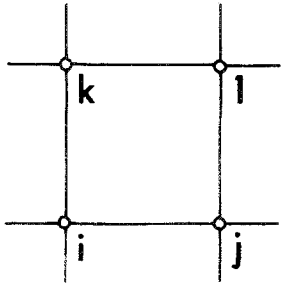

Fig. 1. A typical face of the square lattice.

k_B being Boltzmann's constant. The coefficients in (1) have been chosen so that

$$w(a, b, c, d) = w(c, d, a, b) \tag{5}$$

$$w(a, b, c, d) = w(b, a, d, c) \tag{6}$$

The property (5) ensures that the model is unchanged by reversing the order of the rows, while (6) ensures that it is unchanged by reversing the columns.

2. THE TRANSFER MATRIX

Take the lattice to have m columns and m' rows, and impose cyclic (toroidal) boundary conditions. Let $\sigma = \{\sigma_1, ..., \sigma_m\}$ be the spins on one row, and $\sigma' = \{\sigma_1', ..., \sigma_m'\}$ the spins on the row above. The contribution to the partition function of the interactions between these rows is

$$V_{\sigma,\sigma'} = \prod_{j=1}^{m} w(\sigma_j, \sigma_{j+1}, \sigma_j', \sigma_{j+1}') \tag{7}$$

taking $\sigma_{m+1} = \sigma_1$, $\sigma_{m+1}' = \sigma_1'$.

Let V be the 2^m by 2^m "transfer matrix" with elements $V_{\sigma,\sigma'}$. Then in the usual way one can show that, for m' large,

$$Z = \mathrm{Tr}\, V^{m'} \sim \Lambda^{m'} \tag{8}$$

where Λ is the maximum eigenvalue of V.

The conditions (5) ensure that V is symmetric, so Λ can be determined from the variational principle

$$\Lambda = \psi^T V \psi / \psi^T \psi \tag{9}$$

where ψ is a 2^m-dimensional vector chosen to maximize the RHS of (9); ψ^T is its transpose.

The object of statistical mechanics is to calculate the "partition function per site"

$$\kappa = Z^{1/mm'} = \Lambda^{1/m} \tag{10}$$

in the limit when m, m' are infinitely large.

3. VARIATIONAL APPROXIMATION

Let $\psi(\sigma)$ be the element σ of the vector ψ. Remembering that σ denotes the m spins $\sigma_1, ..., \sigma_m$, this can be written more explicitly as $\psi(\sigma_1, ..., \sigma_m)$. For reasons that will be given in Section 5, I use the following "trial function" for ψ:

$$\psi(\sigma_1, ..., \sigma_m) = \mathrm{Tr}[F(\sigma_1, \sigma_2)F(\sigma_2, \sigma_3)F(\sigma_3, \sigma_4) \cdots F(\sigma_m, \sigma_1)] \tag{11}$$

where each of $F(+,+)$, $F(+,-)$, $F(-,+)$, $F(-,-)$ is a 2^n by 2^n matrix and n is some given number, e.g., 3. The matrices $F(+,+),\ldots, F(-,-)$ must satisfy the symmetry conditions

$$F^T(+,+) = F(+,+), \quad F^T(+,-) = F(-,+), \quad F^T(-,-) = F(-,-) \tag{12}$$

(superscript T denoting matrix transposition). Otherwise they are arbitrary.

The program now is to calculate the RHS of (9), using (11), in the limit of m large; then to choose the four matrices $F(+,+),\ldots, F(-,-)$ to maximize the RHS, subject to the constraint (12).

Calculation of $\psi^T\psi$ and $\psi^T V\psi$

Let $F_{\lambda\mu}(a, b)$ be the element of $F(a, b)$ in row λ and column μ. From (11) it is readily found that

$$\psi^T\psi = \operatorname{Tr} R^m \tag{13}$$

where R is a 2^{2n+1} by 2^{2n+1} matrix. The element of R in row (λ, a, λ') and column (μ, b, μ') is

$$R(\lambda, a, \lambda'|\mu, b, \mu') = F_{\lambda\mu}(a, b)F_{\lambda'\mu'}(a, b) \tag{14}$$

When m is large it follows from (13) that

$$\psi^T\psi \sim \xi^m \tag{15}$$

where ξ is the maximum eigenvalue of R. Let X be the corresponding eigenvector and write its element λ, a, λ' as $X_{\lambda\lambda'}(a)$. This can be regarded as the element λ, λ' of a 2^n by 2^n matrix $X(a)$. Using (14) and (12), the eigenvalue equation for ξ can then be written

$$\sum_b F(a, b)X(b)F(b, a) = \xi X(a) \tag{16}$$

with $a = +$ or $-$.

The relations (12) imply that R is a symmetric matrix, so ξ can itself be determined from a variational principle, namely

$$\xi = \sum_{a,b} \operatorname{Tr}[X^T(a)F(a, b)X(b)F(b, a)] \Big/ \sum_a \operatorname{Tr} X^T(a)X(a) \tag{17}$$

Equations (16) ensure that (17) is stationary with respect to variations in the matrices $X(+)$ and $X(-)$.

Now consider the term $\psi^T V\psi$ in (9). From (7) and (11),

$$\psi^T V\psi = \operatorname{Tr} S^m \tag{18}$$

where S is a 2^{2n+2} by 2^{2n+2} matrix. The element of S in row $(\lambda, a, a', \lambda')$ and column (μ, b, b', μ') is

$$S(\lambda, a, a', \lambda'|\mu, b, b', \mu') = F_{\lambda\mu}(a, b)w(a, b, a', b')F_{\lambda'\mu'}(a', b') \tag{19}$$

The argument now closely parallels the derivation of (15)–(17). Let η be the maximum eigenvalue of S, and Y the corresponding eigenvector. Then

$$\psi^T V\psi \sim \eta^m \tag{20}$$

when m is large. The element λ, a, a', λ' of Y can be written as $Y_{\lambda\lambda'}(a, a')$ and regarded as the element λ, λ' of a 2^n by 2^n matrix $Y(a, a')$. Using (19) and (12), the eigenvalue equation for η can then be written

$$\sum_{b,b'} w(a, b, a', b')F(a, b)Y(b, b')F(b', a') = \eta Y(a, a') \tag{21}$$

for $a, a' = +, -$.

From (6), (12), and (19), S is a symmetric matrix, so

$$\eta = \left\{ \sum_{a,b,a',b'} w(a, b, a', b') \operatorname{Tr}[Y^T(a, a')F(a, b) \times Y(b, b')F(b', a')] \right\} \left\{ \sum_{a,a'} \operatorname{Tr}\, Y^T(a, a')Y(a, a') \right\}^{-1} \tag{22}$$

the RHS being stationary with respect to variations in the matrices $Y(+, +)$, $Y(+, -)$, $Y(-, +)$, $Y(-, -)$.

From (12), the eigenvalue equations (16) and (21) permit solutions satisfying the symmetry relations

$$X^T(a) = X(a), \qquad Y^T(a, a') = Y(a', a) \tag{23}$$

for $a, a' = +, -$. It seems reasonable to suppose that the desired maximum eigenvalues ξ, η will correspond to solutions satisfying these symmetry relations. This is assumed hereafter and is further justified in Section 5.

From (10), (9), (15) and (20), when m is large

$$\kappa = \Lambda^{1/m} = \eta/\xi \tag{24}$$

Thus in the thermodynamic limit the partition function per site is simply η/ξ.

Maximization of Λ

It remains to choose the matrices $F(+, +)$,..., $F(-, -)$ to maximize Λ, or equivalently $\kappa = \eta/\xi$. This can be done by using Eqs. (17) and (22) for ξ, η. The resulting expression for κ involves the matrices X and Y, as well as F, but is by definition stationary with respect to variations in X and Y.

Thus when differentiating κ with respect to F, the terms coming from the induced variations in X and Y do not contribute. This means that X and Y can be treated as constants in the differentiation. Using the symmetry constraints (12) and the properties (23), this gives

$$\sum_{a',b'} w(a, b, a', b')Y(b, b')F(b', a')Y(a', a) = \eta' X(b)F(b, a)X(a) \tag{25}$$

for $a, b = +, -$; where

$$\eta' = \kappa \sum_{a,a'} \mathrm{Tr}\ Y(a', a)Y(a, a') \Big/ \sum_{a} \mathrm{Tr}\ X^2(a) \tag{26}$$

Equations (16), (21) and (25) define the matrices $F(a, b)$, $X(a)$, and $Y(a, b)$ (to within normalization factors) and hence $\kappa = \eta/\xi$.

A further equation can be deduced by premultiplying both sides of (25) by $F(a, b)$ and summing over b. The LHS can then be simplified by using (21), and the RHS by using (16), giving

$$\sum_{a'} Y(a, a')Y(a', a) = \xi' X^2(a) \tag{27}$$

where $\xi' = \xi\eta'/\eta$.

Row–Column Symmetry

One rather unsatisfactory feature of the transfer matrix method is that it treats the rows of the lattice different than the columns, and so destroys the symmetry between them. It is therefore very gratifying to find that the symmetry is restored in Eqs. (16), (21), and (25).

To see this, define six 2^n by 2^n matrices $A(+)$, $A(-)$, $G(+, +)$, $G(+, -)$, $G(-, +)$, and $G(-, -)$ such that

$$B(a) = A^T(a) \tag{28a}$$

$$X(a) = B(a)A(a) \tag{28b}$$

$$Y(a, a') = B(a)G(a, a')A(a') \tag{28c}$$

for $a, a' = +, -$. (These equations do not define A, B, G uniquely, but ultimately I shall specialize to a representation in which they are well-defined.)

It appears that $X(+)$ and $X(-)$ are always positive definite, so $A(+)$, $A(-)$, $B(+)$, and $B(-)$ are nonsingular and can be chosen to be real.

The symmetry conditions (23) are now equivalent to

$$G^T(a, a') = G(a', a) \tag{29}$$

for $a, a' = +, -$. Substituting the expressions (28) into (16), (27), (21), and (25), the equations become

$$\sum_b F(a, b)B(b)A(b)F(b, a) = \xi B(a)A(a) \tag{30a}$$

$$\sum_b G(a, b)A(b)B(b)G(b, a) = \xi' A(a)B(a) \tag{30b}$$

$$\sum_{b,b'} w(a, b, a', b')F(a, b)B(b)G(b, b')A(b')F(b', a') = \eta B(a)G(a, a')A(a') \tag{30c}$$

$$\sum_{b,b'} w(a', a, b', b)G(a, b)A(b)F(b, b')B(b')G(b', a') = \eta' A(a)F(a, a')B(a') \tag{30d}$$

Equations (30) are the main result of this paper. Clearly they are symmetric with respect to interchanging A, F, ξ, η with B, G, ξ', η', while simultaneously replacing $w(a, b, a', b')$ with $w(a', a, b', b)$. However, this last is equivalent to rotating the lattice through 90°, i.e., to interchanging the rows with the columns. Thus this row–column symmetry is present in the equations.

Variational Expression for κ

From (17), (22), (28), and (29), the partition function per site $\kappa = \eta/\xi$ is

$$\kappa = r_1 r_4 / r_2 r_3 \tag{31a}$$

where

$$r_1 = \sum_a \mathrm{Tr}\, A(a)B(a)A(a)B(a) \tag{31b}$$

$$r_2 = \sum_{a,b} \mathrm{Tr}[A(a)F(a, b)B(b)A(b)F(b, a)B(a)] \tag{31c}$$

$$r_3 = \sum_{a,b} \mathrm{Tr}[B(a)G(a, b)A(b)B(b)G(b, a)A(a)] \tag{31d}$$

$$r_4 = \sum_{a,b,a',b'} w(a, b, a', b')\, \mathrm{Tr}[B(a')G(a', a)A(a)F(a, b)B(b)G(b, b')A(b')F(b', a')] \tag{31e}$$

Equation (31a), together with (31b)–(31e), is a variational expression for κ, in that it is stationary with respect to variations in the matrices $A(a)$, $B(a)$, $F(a, b)$, $G(a, b)$. This can be useful in numerical calculations, since it implies that if the A, B, F, G are calculated to within a certain error ϵ, and κ is then calculated from (31) [*without* using (30)], then the error in κ should only be of order ϵ^2.

Magnetization and Nearest-Neighbor Correlations

The derivatives of κ with respect to H, J, etc. can be obtained by differentiating (31). Since this expression is stationary with respect to the A, B, F, G, the induced variations in these matrices do not contribute. This means that the differentiations can be performed as if the A, B, F, G were constant. The magnetization M and nearest-neighbor horizontal correlation $\langle\sigma_i\sigma_j\rangle$ are thus readily found to be, after using (30) to simplify the resulting expressions,

$$\begin{aligned} M = \langle\sigma_i\rangle &= k_B T\,\partial(\ln\kappa)/\partial H \\ &= r_1^{-1}\sum_a a\ \mathrm{Tr}\ A(a)B(a)A(a)B(a) \end{aligned} \tag{32}$$

$$\begin{aligned} \langle\sigma_i\sigma_j\rangle &= k_B T\,\partial(\ln\kappa)/\partial J \\ &= r_2^{-1}\sum_{a,b} ab\ \mathrm{Tr}\ A(a)F(a,b)B(b)A(b)F(b,a)B(a) \end{aligned} \tag{33}$$

Representation with $A(\pm)$ Diagonal

Equations (30)–(33), and the symmetry relations (12), (28a), and (29), are unchanged by the following transformations:

$$\begin{aligned} A(a) &\to Q(a)A(a)P^T(a), \qquad B(a) \to P(a)B(a)Q^T(a) \\ F(a,b) &\to P(a)F(a,b)P^T(b), \qquad G(a,b) \to Q(a)G(a,b)Q^T(b) \end{aligned} \tag{34}$$

where $P(+)$, $P(-)$, $Q(+)$, and $Q(-)$ are 2^n by 2^n orthogonal matrices. From (28a) these orthogonal matrices can be chosen to ensure that the $A(a)$, $B(a)$ are diagonal.

Without loss of generality, the matrices $A(a)$, $B(a)$ can therefore be chosen to be diagonal. Equation (28a) then implies that

$$B(a) = A(a), \qquad a = + \text{ or } - \tag{35}$$

Choosing an appropriate normalization for the matrices $F(a, b)$, and for the matrices $G(a, b)$, one can also ensure that

$$\xi = \xi' = 1 \tag{36}$$

from which it follows that

$$\eta = \eta' = \kappa \tag{37}$$

The matrices $F(a, b)$, $G(a, b)$ are now completely determined by (30). The $A(a)$ are determined to within a single normalization factor. A convenient way to fix this is to require that the maximum diagonal element of $A(+)$ be unity.

Iterative Method of Solution

Using the A-diagonal representation above, define new 2^n by 2^n matrices as follows:

$$H(a, b|c) = \sum_d w(a, d, c, b)F^T(d, a)A(d)G(d, b) \tag{38}$$

$$\begin{aligned} F_1(a, b) &= A(a)F(a, b)A^{-1}(b) \\ G_1(a, b) &= A(a)G(a, b)A^{-1}(b) \end{aligned} \tag{39}$$

Now define two 2^{n+1} by 2^{n+1} matrices $U(+)$ and $U(-)$, and four 2^{n+1} by 2^n matrices $F_2(+)$, $F_2(-)$, $G_2(+)$, $G_2(-)$:

$$U(a) = \begin{pmatrix} H(+, +|a) & H(+, -|a) \\ H(-, +|a) & H(-, -|a) \end{pmatrix} \tag{40}$$

$$F_2(a) = \begin{pmatrix} F_1(+, a) \\ F_1(-, a) \end{pmatrix}, \qquad G_2(a) = \begin{pmatrix} G_1(+, a) \\ G_1(-, a) \end{pmatrix} \tag{41}$$

for $a = +, -$.

Then Eqs. (30c) and (26d) can be written, using (5),

$$U(a)F_2(a) = \kappa G_2(a)A(a) \tag{42a}$$

$$U^T(a)G_2(a) = \kappa F_2(a)A(a) \tag{42b}$$

for $a = +, -$.

Remember that $A(+)$ and $A(-)$ are diagonal and consider a particular value of a and a particular column of the matrix equations (40). It is readily seen that each column of $F_2(a)$ $[G_2(a)]$ is an eigenvector of $U^T(a)U(a)$ $[U(a)U^T(a)]$, the eigenvalue being the corresponding element of $\kappa^2A^2(a)$. Thus the columns are eigenvectors of a real, symmetric matrix and can be chosen orthogonal. If they are further chosen to be orthonormal, then

$$F_2^T(a)F_2(a) = G_2^T(a)G_2(a) = I \tag{43}$$

where I is the unit 2^n by 2^n matrix.

However, remembering that $\xi = \xi' = 1$ and the definitions (41) and (39), we see that (43) is precisely Eqs. (30a) and (30b).

This suggests the following iterative procedure for solving the equations for a given value of n, given a reasonable initial guess at the solution:

(a) Calculate $U(+)$ and $U(-)$ from (38) and (40).
(b) Obtain the normalized eigenvectors and eigenvalues of $U^T(a)U(a)$ and $U(a)U^T(a)$, $a = +, -$.
(c) Set κ^2 equal to the largest eigenvalue of $U^T(+)U(+)$.

(d) For each value of a there are 2^{n+1} eigenvalues. Select the 2^n of these that are closest to the previous values of the diagonal elements of $\kappa^2 A^2(a)$. Now redefine $A(a)$ so that the elements of $\kappa^2 A^2(a)$ are these eigenvalues.
(e) Using this selection and ordering of the eigenvalues, let $F_2(a)$ and $G_2(a)$ be the corresponding matrices of eigenvectors satisfying (42).
(f) Calculate the $F(a, b)$ and $G(a, b)$ from (41) and (39). Return to (a).

This procedure appears to converge fairly well, so long as the system is not near a critical point. Near such points a full Newton–Raphson iterative procedure may be necessary.

One subtle point is that at any intermediate stage in the calculation the symmetry conditions (12) and (29) may be violated. Nevertheless, the procedure should converge to a symmetric solution, provided the definition (38) of $H(a, b|c)$ is used exactly as written.

Given the solution for a particular value m of n, this procedure can also be used to obtain a reasonable guess at the solution for $n = m + 1$. One merely keeps *all* the eigenvalues and eigenvectors at stages (d) and (e), so that $F_2(a)$ and $G_2(a)$ are 2^{m+1} by 2^{m+1} matrices. From (41), $F_1(a, b)$ and $G_1(a, b)$ are 2^m by 2^{m+1} matrices. The matrices labeled $A(d)$ and $A(a)$ in (38) and (39) are 2^m by 2^m and are those of the old $n = m$ solution, while those labeled $A(b)$ are the new, expanded 2^{m+1} by 2^{m+1} matrices. The matrices $H(a, b|c)$ can now be calculated from (38) and are of dimension 2^{m+1} by 2^{m+1}. One can now calculate the $U(a)$ from (40) and proceed normally with $n = m + 1$.

Initial Guesses

These comments still beg the question of how to obtain a "reasonable" initial guess at the solution for some small value of n, notably $n = 0$. For high temperatures the simplest sensible $n = 0$ guess is

$$F(a, b) = G(a, b) = 2^{-1/2}, \qquad A(a) = 1$$

for $a, b = +, -$.

At low temperatures a sensible guess is one which gives an eigenvector ψ in (11) corresponding to the ground state of the system.

For instance, for a ferromagnetic system with $H + 4H_3 \geqslant 0$, the ground state has all spins up and the corresponding $n = 0$ guess (which should work for all sufficiently low temperatures) is

$$F(a, b) = G(a, b) = \delta(a, +)\delta(b, +)$$
$$A(+) = 1, \qquad A(-) = 0 \tag{44}$$

where $\delta(x, y) = 1$ if $x = y$, and otherwise is zero.

For a Hamiltonian dominated by antiferromagnetic nearest-neighbor two-spin interactions (i.e., J and J' large and negative), the $n = 0$ solution corresponding to the ground state is, for $a, b = +, -$,

$$F(a, b) = G(a, b) = \delta(a, -b), \qquad A(a) = 1 \tag{45}$$

In fact this is an exact solution of Eqs. (30). It gives the trivial $T = 0$ approximation for κ, but nontrivial useful approximations can be obtained from it by using the above procedure to obtain solutions for $n = 1, 2, 3, \ldots$.

If J'' is large and negative, then the Hamiltonian is dominated by antiferromagnetic next-nearest-neighbor interactions. If it is also true that $J \geqslant J'$, then the ground-state $n = 0$ solution is

$$F(a, b) = \delta(a, b), \qquad G(a, b) = \delta(a, -b), \qquad A(a) = 1 \tag{46}$$

This is also a trivial exact solution of (30), from which nontrivial solutions for higher n can be obtained by the above procedure.

Isotropic Case

If $J = J'$, then the Hamiltonian is unaltered by interchanging the rows with the columns. Provided this symmetry is not spontaneously broken (this can happen, for instance, if J'' is large and negative), the solution of (30) has the symmetry properties

$$G(a, b) = F(a, b), \qquad B(a) = A(a), \qquad \xi' = \xi, \qquad \eta' = \eta \tag{47}$$

The four matrix equations therefore reduce to two. The iteration procedure also simplifies, since $U^T(a) = U(a)$ and $G_2(a) = F_2(a)$, so (42) becomes the single eigenvalue equation

$$U(a)F_2(a) = \kappa F_2(a)A(a) \tag{48}$$

4. LOW-LEVEL APPROXIMATIONS FOR THE ISING MODEL

Bethe and Kramers–Wannier Approximations

As an example, consider the $n = 0$ solution when $J_0 = J = J' = H_3 = J_4 = 0$ and $J'' > 0$, so the model has only ferromagnetic next-nearest-neighbor interactions. It therefore factors into two independent nearest-neighbor Ising models, one on each sublattice.

Set

$$z = \exp(-2J''/k_BT), \qquad \nu = \exp(-\tfrac{1}{2}H/k_BT) \tag{49}$$

and define x, s by

$$e^{-s} = x = \nu(z + \nu x^3)/(1 + z\nu x^3) \tag{50}$$

Then the solution of (30) is

$$F(a, b) = G(a, b) = (2 \cosh 2s)^{-1/2} \exp[\tfrac{1}{2}s(a + b)]$$
$$A(+) = 1, \qquad A(-) = x, \qquad \xi = \xi' = 1 \tag{51}$$
$$\kappa = \eta = \eta' = (1 + z\nu x^3)^2/[\nu^2 z(1 + x^4)]$$

From (42), the magnetization is

$$M = (1 - x^4)/(1 + x^4) = \tanh 2s \tag{52}$$

These results are precisely those of the Bethe approximation.[6] (The μ and μ_1 of pp. 251–4 of Ref. 6 are related to ν and x by $\mu = \nu^4$, $\mu_1 = \nu x^3$.)

It is illuminating to consider the zero-field case $H = 0$, $\nu = 1$. Then (50) always has a solution $x = 1$, $s = 0$. There is a critical value T_c of T, given by

$$J''/k_B T_c = \tfrac{1}{2} \ln 2. \tag{53}$$

For $T < T_c$ there are two other real solutions of (50), having equal and opposite values of s. If H is regarded as infinitesimally small and positive, then the correct solution is the one for which s is positive.

Thus s is zero for $T > T_c$ and $F(a, b)$, $A(a)$ are unchanged by negating a and b. For $T < T_c$, s is positive and this spin-reversal symmetry of F and A is broken.

Such spontaneous symmetry-breaking is typical of Eqs. (30). If the Hamiltonian has a symmetry, then (30) will have a solution with the corresponding symmetry. If this symmetry is spontaneously broken at low temperatures, then (30) will also have two or more asymmetric solutions, which will then be the correct ones to use. The critical temperature can be obtained by increasing T until the asymmetric solutions coincide with the symmetric one. For finite n the critical exponents thereby obtained will usually have the classical values $\delta = 3$, $\beta = \frac{1}{2}$. There must be a "crossover phenomenon" as $n \to \infty$, since the exponents must then take the true two-dimensional values.

Another interesting example is the case $J_0 = J'' = H_3 = J_4 = 0$. The model then becomes the usual nearest-neighbor Ising model and it turns out that the $n = 0$ solution of (30) is precisely the Kramers–Wannier approximation.[7]

5. GRAPHICAL INTERPRETATION OF THE MATRICES

For a given value of n, the 2^n by 2^n matrices $A(a)$, $B(a)$, $F(a, b)$, and $G(a, b)$ are the solutions of Eqs. (30) and the symmetry relations (12), (28), and (29). This defines them to within normalization factors and the orthog-

onal transformations (34). The purpose of this section is to justify the trial function (11) and thereby give a graphical interpretation to A, B, F, and G.

Convergence of the Variational Approximations to the Exact Result

Suppose the lattice still has m' rows and m columns, but is wound on a vertical cylinder. To each square face assign a weight $w(\sigma_i, \sigma_j, \sigma_k, \sigma_l)$ as in Section 1. In addition, to each horizontal edge, between spins σ_i and σ_j, on the top row assign a weight $f(\sigma_i, \sigma_j)$ such that

$$f(\sigma_i, \sigma_j) = f(\sigma_j, \sigma_i) \tag{54}$$

Choose these top boundary weights to match the expected state of the system. For instance, at high temperatures a simple sensible choice would be $f(\sigma_i, \sigma_j) = 1$. For a low-temperature ferromagnet in nonnegative fields one could take $f(\sigma_i, \sigma_j) = \delta(\sigma_i, \sigma_j) = \delta(\sigma_i, +)\,\delta(\sigma_j, +)$; and for a low-temperature, nearest-neighbor antiferromagnet $f(\sigma_i, \sigma_j) = \delta(\sigma_i, -\sigma_j)$.

The probability that the m spins on the bottom row have values $\sigma_1, \ldots, \sigma_m$, respectively, is then

$$\chi(\sigma_1, \ldots, \sigma_m) = Z^{-1} \sum{}' \prod \prod w(\sigma_i, \sigma_j, \sigma_k, \sigma_l) \prod \prod f(\sigma_k, \sigma_l) \tag{55}$$

where the first product is over all the $m(m' - 1)$ square faces of the lattice, the second is over all m horizontal edges in the top row, and the sum is over all spins on the lattice other than those in the bottom row.

The summation in (55) can be arranged either by rows or columns. Grouping the spins by rows, (55) can be written as

$$\chi(\sigma_1, \ldots, \sigma_m) = Z^{-1}(V^{m'-1}\chi_0)_{\sigma_1, \ldots, \sigma_m} \tag{56}$$

where V is the 2^m by 2^m transfer matrix (7) and χ_0 is a 2^m-dimensional vector, with entries

$$\chi_0(\sigma_1, \ldots, \sigma_m) = f(\sigma_1, \sigma_2) f(\sigma_2, \sigma_3) \cdots f(\sigma_m, \sigma_1) \tag{57}$$

On the other hand, set $n = m' - 1$ and let $a, \lambda_1, \ldots, \lambda_n$ be the spins on one column of the lattice, and $b, \mu_1, \ldots, \mu_n$ the spins on the next column, arranged as in Fig. 2a. Let $\lambda = \{\lambda_1, \ldots, \lambda_n\}$, $\mu = \{\mu_1, \ldots, \mu_n\}$. Then the Boltzmann weight of the interactions between these columns is

$$\bar{F}_{\lambda\mu}(a, b) = w(a, b, \lambda_1, \mu_1) f(\lambda_n, \mu_n) \prod_{j=1}^{n-1} \prod w(\lambda_j, \mu_j, \lambda_{j+1}, \mu_{j+1}) \tag{58}$$

Normalize this definition by dividing it by $Z^{1/m}$. Then (55) can be written

$$\chi(\sigma_1, \ldots, \sigma_m) = \sum_{\lambda, \mu, \nu, \ldots, \tau} \bar{F}_{\lambda\mu}(\sigma_1, \sigma_2) \bar{F}_{\mu\nu}(\sigma_2, \sigma_3) \cdots \bar{F}_{\tau\lambda}(\sigma_m, \sigma_1) \tag{59}$$

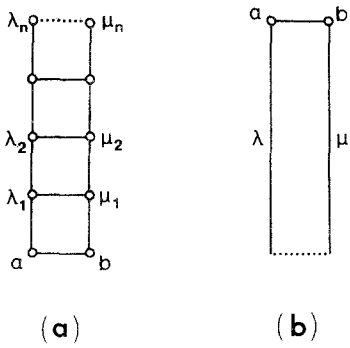

Fig. 2. (a) The half-column lattice segment with the weight $\bar{F}_{\lambda\mu}(a, b)$ given in (58); (b) the same segment with the row order reversed, internal edges omitted, and only the sites with spins a and b explicitly shown. These figures can also be used to represent any function of a, b, λ, μ; notably $F_{\lambda\mu}(a, b)$.

Regarding $\bar{F}_{\lambda,\mu}(a, b)$ as the element (λ, μ) of a 2^n by 2^n matrix $\bar{F}(a, b)$, (59) becomes

$$\chi(\sigma_1,..., \sigma_m) = \mathrm{Tr}\, \bar{F}(\sigma_1, \sigma_2)\bar{F}(\sigma_2, \sigma_3) \cdots \bar{F}(\sigma_m, \sigma_1) \tag{60}$$

From (54), (58), and (6), these matrices $\bar{F}(a, b)$ satisfy the symmetry relations

$$\bar{F}^T(a, b) = \bar{F}(b, a) \tag{61}$$

Now suppose that $n = m' - 1$ is large. From (56) it follows that in this limit

$$\chi(\sigma_1,..., \sigma_m) = \psi(\sigma_1,..., \sigma_m) \tag{62}$$

where ψ is the eigenvector of V corresponding to the maximum eigenvalue ψ. Combining (60) and (62), it follows that the trial function (11) is then *exactly* correct, and that

$$F(a, b) = \bar{F}(a, b), \qquad n \to \infty \tag{63}$$

This means that the sequence of variational approximations, obtained by taking $n = 0, 1, 2,...$ in Section 3, must converge to give the exact value of κ in the limit $n \to \infty$.

For finite n, (63) is no longer valid. [One could of course require that $F = \bar{F}$ in (11), but this would not be the optimal choice of the trial function.] Even so, (63) should be approximately true for finite n, particularly if the temperature is either high or low and a good choice of $f(a, b)$ is made. This approximate equality in (63) makes it possible to given an illuminating graphical interpretation of the matrices $F(a, b)$, $A(a)$, $B(a)$, and $G(a, b)$.

Half-Column, Half-Row, and Corner Transfer Matrices

Suppose that n is infinitely large, so F is $\bar{F}$. Then $F_{\lambda\mu}(a, b)$ is the Boltzmann weight of the lattice strip shown in Fig. 2a. Using (5), it is also the

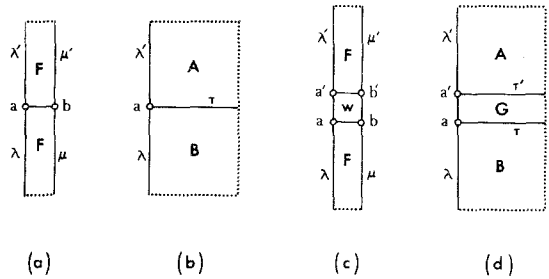

Fig. 3. The lattice segments corresponding to $R(\lambda, a, \lambda'|\mu, b, \mu')$, $X_{\lambda\lambda'}(a)$, $S(\lambda, a, a', \lambda'|\mu, b, b', \mu')$, and $Y_{\lambda\lambda'}(a, a')$, respectively.

weight of the strip shown in Fig. 2b. These are semiinfinite column strips of the lattice, so the $F(a, b)$ may meaningfully be called "half-column transfer matrices."

For simplicity, the internal edges have been omitted in the lattice segment in Fig. 2b, and the only sites explicitly shown are those of the spins a and b. These conventions are followed in subsequent figures.

From (14), $R(\lambda, a, \lambda'|\mu, b, \mu')$ is the weight of the fully infinite vertical strip shown in Fig. 3a. Thus R is a column-to-column transfer matrix. Operating by R a large number r of times on a given initial vector builds up the lattice of $2n + 1$ rows and $r + 1$ columns indicated in Fig. 3b. (For the moment, ignore the horizontal line labeled τ.) Thus the elements $X_{\lambda\lambda'}(a)$ of its maximal eigenvector are the probabilities of the $2n + 1$ spins on the left edge of this lattice having values λ, a, λ'.

Similarly, from (19), S is the column-to-column transfer matrix indicated in Fig. 3c; $Y_{\lambda\lambda'}(a, a')$ is the probability distribution for the spins on the left-hand edge of the lattice of $2n + 2$ rows and $r + 1$ columns shown in Fig. 3d.

The lattices in Figs. 3b and 3d can be partitioned as shown. Let $A_{\tau\lambda'}(a)$ be the probability distribution for the spins τ, a, λ', obtained by summing over all spins inside (or on the dotted boundaries) the upper half of the lattice in Fig. 3b. Similarly, let $B_{\lambda\tau}(a)$ be the distribution for the lower half; and let $G_{\tau\tau'}(a, a')$ be the Boltzmann weight of the interactions between the two adjacent rows shown in Fig. 3d.

Summing over τ [τ, τ'] in Fig. 3b [3d] gives the equation (28b) [(28c)]; the row-reversal symmetry (5) gives (28a) and (29); setting $r = n$ ensures that the matrices $A(a)$, $B(a)$ are square. Thus in the limit of n large the elements of the matrices $A(a)$, $B(a)$, $G(a, a')$ in Section 3 can be taken to be the probability distributions defined in the preceding paragraph.

From Fig. 3d it is therefore apparent that the $G(a, a')$ are "half-row transfer matrices" similar to the half-column matrices $F(a, b)$. The matrices

$A(a)$, $B(a)$ correspond to quadrants of a completely infinite lattice. Apart from the shape of the outer boundary, which is irrelevant in the limit $n \to \infty$, the 2^{n+1} by 2^{n+1} matrices

$$\begin{pmatrix} A(+) & 0 \\ 0 & A(-) \end{pmatrix}, \quad \begin{pmatrix} B(+) & 0 \\ 0 & B(-) \end{pmatrix} \tag{64}$$

are the "corner transfer matrices" previously defined for the eight-vertex model.[4,5]

Graphical Form of the Equations

Even if n is finite, the RHS of Eqs. (14), (28b), and (28c) can still be represented graphically as in Fig. 3, so long as the elements of F, A, B, and G are merely interpreted as undetermined functions of the outer spins on the corresponding segments.

One advantage of this graphical representation is that it makes the symmetries very clear. In particular, it can be seen that (23) is equivalent to requiring that the row-reversal symmetry be unbroken. It appears that this is always so, so (23) is valid. It is also immediately obvious that interchanging F and G, and A with B, is equivalent to interchanging the rows with the columns; i.e., to rotating the lattice through 90°.

Equations (30) can be represented graphically. Adopting the convention of summing over spins on unlabeled solid circles and edges inside a figure, Eqs. (30a) and (30c) can be represented as in Fig. 4. The diagrams representing Eqs. (30b) and (30d) can be obtained from these by rotating the figure anticlockwise through 90° and interchanging F, A, ξ, η with G, B, ξ', η', respectively.

The variational expression (31) is represented in Fig. 5.

A Simple Approximation for $\varkappa$

Since (31) is stationary with respect to variations in F, G, and A, reasonable approximations to F, G, A, and B therein should give good approximations to κ. For given n an obvious choice is to take F, G, A, and B to be

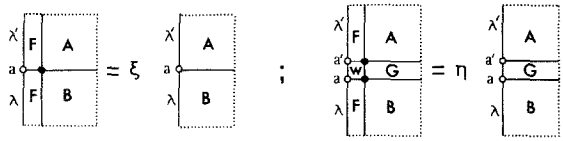

Fig. 4. Graphical representation of Eqs. (30a) and (30c).

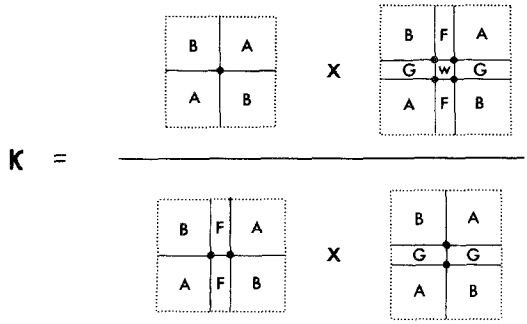

Fig. 5. Graphical representation of the variational expression (31) for κ.

exactly the half-column, half-row, and corner transfer matrices of a $2n + 1$ by $2n + 1$ square lattice. Setting $N = 2n + 1$, (31) then gives

$$\kappa = Z_{N,N}Z_{N+1,N+1}/Z_{N,N+1}Z_{N+1,N} \tag{65}$$

where $Z_{m,n}$ is the partition function of a lattice of m rows and n columns.

This is an approximate formula for κ. It is not the best result that can be obtained using 2^n by 2^n matrices, but it is quite good and has been discussed previously.[8] It is of course exactly true when n is infinitely large.

In calculating (65) one should introduce appropriate weights $f(\sigma_i, \sigma_j)$ for the top and bottom boundary edges, $g(\sigma_i, \sigma_j)$ for the left- and right-hand boundary edges, and $a(\sigma_i)$ for the four corner sites of the lattice. The best way to do this would be to choose them to maximize (65). Such a calculation would presumably be related to the work of Bolton and Gruen[9] on finite lattices with mean-field boundary conditions.

6. SUMMARY

Equations (30), together with the symmetry relations (12), (28a), and (29), define the 2^n by 2^n matrices $F(a, b)$, $G(a, b)$, $A(a)$, and $B(a)$ for $a, b = +, -$, to within normalization factors and the transformations (34). They also define the partition function per site $\kappa = \eta/\xi$. They simplify for a system in an isotropic state.

For large n the 2^{n+1} by 2^{n+1} matrices (64) are the corner transfer matrices of the system. For finite n, (64) can be taken to be the definition of the CTMs. In some ways this definition is more attractive than the previous one,[4] since no particular boundary conditions need be assumed. The equations, as it were, determine their own boundary conditions.

I did hope this would mean that for finite n the new CTMs of the zero-field eight-vertex model would turn out to have special properties similar

to those possessed by the Ising model ones. These might have pointed the way to an operator algebra solution of the eight-vertex model and a proof of the conjectured expressions for the order parameters.[10,11] Unfortunately I have so far made little progress in this direction, but have not completely abandoned hope.

As n increases, the approximations should rapidly converge to give the exact values for the thermodynamic functions. Tsang is investigating this convergence for the zero-field Ising case.[3]

Possible Generalizations

For definiteness I have supposed in deriving (30) that the spins take only the values $+1$ and -1, but this is not necessary. They can take any set of values so long as the symmetry relations (5) and (6) are satisfied. If they take q values, then it is natural to require the $F(a, b)$, $G(a, b)$, $A(a)$, and $B(a)$ to be q^n by q^n matrices.

Possibly even the symmetry requirements (5) and (6) can be weakened: certainly they do not seem to be essential in the graphical interpretations given in Figs. 4 and 5. However, the argument of Section 3 is much simpler and more satisfactory when V, R, and S are real, symmetric matrices.

REFERENCES

1. R. J. Baxter, *J. Math. Phys.* **9**:650 (1968).
2. S. B. Kelland, *Can. J. Phys.* **54**:1621 (1976).
3. S. K. Tsang, to be published.
4. R. J. Baxter, *J. Stat. Phys.* **15**:485 (1976).
5. R. J. Baxter, *J. Stat. Phys.* **17**:1 (1977).
6. C. Domb, *Adv. Phys.* **9**:149 (1977).
7. H. A. Kramers and G. H. Wannier, *Phys. Rev.* **60**:263 (1941).
8. I. G. Enting and R. J. Baxter, *J. Phys. A: Math. Gen.* **10**:L117 (1977).
9. H. C. Bolton and D. W. R. Gruen, *Phys. Rev. B* **15**:4544 (1977).
10. M. N. Barber and R. J. Baxter, *J. Phys. C: Solid State Phys.* **6**:2913 (1973).
11. R. J. Baxter and S. B. Kelland, *J. Phys. C: Solid State Phys.* **7**:L403 (1974).