

VARIATIONAL METHOD FOR LATTICE SYSTEMS: GENERAL FORMALISM AND APPLICATION TO THE TWO-DIMENSIONAL ISING MODEL IN AN EXTERNAL FIELD

P. RUJÁN

*Institute for Theoretical Physics, R. Eötvös University, H-1088 Budapest,
Puskin u. 5-7, Hungary*

Received 31 August 1978
Revised 10 November 1978

A simple variational approach to the eigenvalue problem of the transfer operator is proposed. After reducing the transfer operator according to the symmetries of the Hamiltonian, the leading eigenvalues of the irreducible blocks can be evaluated by elementary variational principles. Hence the thermodynamics and a large class of correlation functions of lattice systems can be calculated. Following a natural truncation scheme the results can be improved in a systematic way. The high accuracy and the convergence of the method is demonstrated by the two-dimensional Ising model. As a first application, the thermodynamics of the two-dimensional Ising ferro- and antiferromagnet in an external field is studied. We show how the same method can be used to obtain zero-temperature properties of interacting quantum lattice systems.

1. Introduction

The transfer operator (TO) theory is one of the most successful methods in dealing with the statistical mechanics of classical lattice systems with short range interactions[†]. Formal expressions involving only the eigenvalues and the eigenvectors of the TO are easily derived for any macroscopic quantity. In general the solution of the TO eigenvalue problem is an extremely difficult task, requiring a lot of ingenuity besides a deep knowledge of mathematical methods. Therefore, it is a natural attempt to look for general approximate methods. Except perhaps some early works⁴⁾ where the TO formalism is used to derive exact series expansions it seems that this powerful formalism has been used mainly in order to obtain rigorous results.

If one considers the transfer operator as a quantum operator, a fact suggested by the equivalence between different two-dimensional classical

[†] The few non-trivial models solved exactly were first treated using the transfer-matrix formalism^{1,2,3}).

systems and one-dimensional quantum systems⁵), one can directly apply the rich arsenal of quantum mechanics.

In this paper we propose to apply the Ritz-variational principle to the TO and to its corresponding quantum Hamiltonian. As usual we start in reducing the TO in irreducible blocks with respect to the global and the translational symmetry of the classical Hamiltonian. In the corresponding basis it is possible to construct trial functions acting on a block with given symmetry properties. As a non-trivial consequence one can follow the change of the leading eigenvalues as functions of temperature and then one can obtain informations about the mechanism of the phase transition and about the order parameter correlation length^{6,7}).

In the calculation of the needed matrix element we have used extensively the fact that the eigenvector corresponding to the largest eigenvalues of a positive, real, irreducible matrix is – like the ground state vector in quantum mechanics – nodeless⁸) and therefore can be exponentiated. The largest eigenvalue of the TO is rewritten in a Rayleigh ratio form which can be interpreted as the ratio of two partition functions for subsystems of a lower dimensionality but with an infinitely long range interaction.

Using a direct recursion method, recently we had shown⁹) that the interaction strength of these $(d - 1)$ -dimensional subsystems is steeply decreasing with the interaction range. On this basis we truncate the exact expansion of the “ground state vector” according to the range of interaction and consider the few short range coupling constants as variational parameters.

The calculation of the matrix element is therefore reduced to the calculation of two effective problems of lower dimensionality and short range interaction. The procedure can be repeated until one obtains exactly solvable models (e.g. one-dimensional models).

We show that our trial functions lead to exact results at high and low temperatures. The resulting free energy is everywhere an upper bound for the exact free energy.

The method was tested on the 2D Ising model. In a first approximation (2 variational parameters) the free energy per spin is within 0.3% of Onsager's¹) exact result. We compare our internal energy, specific heat and spontaneous magnetisation results with the exact results and find a fairly good agreement, except very near the critical point. The difficulty near the critical point is that our approach violates the isotropy of the correlation length. For the magnetisation exponent we obtain $\beta = 0.4$, ($\beta_{\text{exact}} = \frac{1}{8}$), quite close to the mean field result $\beta^{\text{MF}} = 0.5$. The critical point is $K_c = 0.413$ compared with the exact value 0.4407.

In order to have an idea about the convergence of the method we have also used a second approximation (4 variational parameters). In the high tem-

perature region the accuracy of the free energy is strongly increased, and the value of the critical point is moved to 0.422. Although every thermodynamic function is described better in the second approximation the critical exponent of the spontaneous magnetisation is not improved.

It seems that the systematic construction of trial functions presented below leads to better and better free energy and locations of the critical point but the correct asymptotic behaviour near the critical point can be obtained only taking into account infinitely long intralayer interactions.

As a first application we present numerical results concerning the thermodynamics of the two-dimensional Ising ferro- and antiferromagnet in an external field. The free energy, internal energy, magnetisation, specific heat and susceptibility are calculated at different external field values. The phase diagram of the isotropic antiferromagnet fits surprisingly well the probably exact result of Müller-Hartmann and Zittartz¹⁰). Sublattice magnetisation and susceptibility are also shown for different values of the applied field.

Finally we briefly show how one can apply the variational method presented here for the calculation of zero-temperature properties of quantum lattice systems (section 7). According to the trial function construction scheme presented in section 4 the ground state energy of the quantal system is expressed as a linear combination of short range correlation functions in a classical system with an appropriate Boltzmann-factor.

2. The transfer operator formalism

It has been known for a long time that the macroscopic properties of classical systems with short range interaction can be related to the eigenvalue problem of the so-called transfer operator¹¹). Here we summarize only some relations needed in the following. The reader is referred to one of the many comprehensive papers where the transfers matrix theory is treated in detail^{7,12,13}).

Let us consider a d -dimensional lattice with a classical variable (called spin) s_i at every lattice point. The variable s_i can have discrete (Ising-like spin) or continuous values as in the n -vector model¹⁵). The lattice is cut in N_2 ($d - 1$)-dimensional layers in such a way that only nearest neighbour layers interact:

$$H(\{s_i\}) = \sum_{i=1}^{N_2} h(\mu_i, \mu_{i+1}), \quad (1)$$

where μ denotes the configuration vector of a layer and the coupling constants of H include the $-\beta = -1/k_B T$ factor. In what follows, we suppose $h(\mu_i, \mu_{i+1})$ to be symmetric and independent of i .

The canonical partition function is:

$$Z = \int \left(\prod_{j=1}^{N_2} d\mu_j \right) \exp \sum_{i=1}^{N_2} h(\mu_i, \mu_{i+1}), \quad (2)$$

where $\int d\mu$ denotes the integral (sum) over the configuration space of a layer. The transfer operator eigenvalue problem is

$$\int d\mu \psi_\alpha(\mu) \exp h(\mu, \mu') = \lambda_\alpha \psi_\alpha(\mu'), \quad (3)$$

where

$$T: \int d\mu \dots \exp h(\mu, \mu'),$$

is the transfer operator and (3) can be put in a more abstract form

$$T|\psi_\alpha\rangle = \lambda_\alpha|\psi_\alpha\rangle. \quad (4a)$$

In the case when $h(\mu, \mu')$ is not symmetric one must solve also the left side eigenvalue problem

$$\langle \tilde{\psi}_\alpha | T = \lambda_\alpha \langle \tilde{\psi}_\alpha |. \quad (4b)$$

The right and the left side eigenvector of T are biorthogonal⁴²⁾:

$$\langle \tilde{\psi}_\alpha | \psi_\beta \rangle = \delta_{\alpha,\beta}$$

after a suitable normalization. The eqs. (8)–(12) remain valid after changing $\langle \psi_\alpha |$ to $\langle \tilde{\psi}_\alpha |$.

The transfer matrix elements are defined in the usual way:

$$T_{\alpha,\beta} = \langle \Phi_\alpha | T | \Phi_\beta \rangle = \int d\mu \int d\mu' \Phi_\alpha^*(\mu) e^{h(\mu,\mu')} \Phi_\beta(\mu'). \quad (5)$$

Generally one can choose a $\{\Phi_\alpha(\mu)\}$ basis in which the elements of the transfer matrix are non-negative. Therefore the conditions of the Perron-Frobenius theorem are satisfied⁸⁾. As a consequence the eigenvector corresponding to the largest eigenvalue of (5) has positive values for any layer configuration – it is nodeless.

Denoting by $\lambda_0 > \lambda_1 \dots$ and $|\psi_0\rangle, |\psi_1\rangle \dots$ the eigenvalues and the eigenstates of T , respectively, one can easily derive the following relations¹³⁾:

a) The partition function is

$$Z_{N_1 N_2} = \text{Sp } T^{N_2} = \sum \lambda_i^{N_2} \xrightarrow{N_2 \rightarrow \infty} \lambda_0^{N_2}, \quad (6)$$

where N_1 is the number of spins of a layer and N_2 the total number of layers.

In the thermodynamic limit $N_2 \rightarrow \infty$ the free energy per spin is:

$$f = \frac{1}{N_1} \ln \lambda_0 = -F/k_B T N_1 N_2. \quad (7)$$

b) A general correlation function in the thermodynamic limit has the following form:

$$\langle\langle A(\mu_i) B(\mu_k) \rangle\rangle \equiv \frac{\text{Sp } AB \exp H}{\text{Sp } \exp H} = \sum_i \left(\frac{\lambda_i}{\lambda_0} \right)^{|k-i|} \langle \psi_0 | A | \psi_i \rangle \langle \psi_i | B | \psi_0 \rangle. \quad (8)$$

Two kinds of correlation functions play a distinguished role, the short range or "two-layer" correlation functions and the long range correlation functions. The short range correlation functions can be expressed as:

$$\langle\langle A(\mu_i, \mu_{i+1}) \rangle\rangle = \frac{1}{\lambda_0} \langle \psi_0 | AT | \psi_0 \rangle. \quad (9)$$

The importance of such correlation functions arises from their connection with the internal energy – which for isotropic nearest neighbour interaction is:

$$U = \langle\langle H \rangle\rangle = \frac{q J N_1 N_2}{2} \langle\langle s_i s_{i+1} \rangle\rangle \equiv J N_1 N_2 u \quad (10)$$

(q is the coordination number); with the magnetisation

$$M = \mu \left\langle\left\langle \sum s_i \right\rangle\right\rangle \equiv \mu m N_1 N_2, \quad (11)$$

and more generally with the first derivatives of the free energy with respect to the coupling constants of $h(\mu, \mu')$.

The long range correlation functions are obtained from (8) in the $|k - i| \rightarrow \infty$ limit:

$$\langle\langle A(\mu_i) A(\mu_k) \rangle\rangle \sim \left(\frac{\lambda_\alpha}{\lambda_0} \right)^{|k-i|} \langle \psi_0 | A | \psi_\alpha \rangle \langle \psi_\alpha | A | \psi_0 \rangle, \quad (12)$$

where $|\psi_\alpha\rangle$ is the first eigenvector non-orthogonal to $A|\psi_0\rangle$. The corresponding correlation length can be defined as^{6,14)}

$$\xi_A^{-1} = \ln |\lambda_0 / \lambda_\alpha|. \quad (13)$$

Note that according to (9) the knowledge of $|\psi_0\rangle$ and λ_0 is sufficient for the calculation of thermodynamics and of two-layer correlation functions. The knowledge of λ_1 and $|\psi_1\rangle$ gives additional information on the long range correlation functions and on the correlation length. It may happen, however, that the phase transition has a more complicate structure than a simple crossing or degeneracy of λ_0 and λ_1 ¹⁴⁾. In such a situation more eigenvalues are needed.

3. The reduction of the transfer operator

We consider classical systems with global symmetries as e.g. the total spin-inversion in an 1/2 Ising model without odd interactions or a total rotation with an arbitrary α angle for the n -vector model¹⁵). In this paper only the Ising case will be discussed in detail, the generalization to the n -vector model is, however, straightforward¹⁶).

Our purpose is to construct an orthogonal basis in which the transfer matrix (5) decomposes into irreducible blocks with respect to the global symmetry and the translational invariance of the layer-Hamiltonian $h(\mu, \mu')$.

We start by expanding a function of the single s variable as:

$$g(s) = \sum_n a_n \phi_n(s), \quad (14)$$

in a basis irreducible with respect to an elementary transformation (one spin inversion or rotation). In the Ising case ($s = \pm 1$):

$$g(s) = a + bs, \quad (15a)$$

while in the n -vector model¹⁷) (s has n components and unit length):

$$g(s) = g(\cos \theta) = \sum_l z_l C_l^{n/2-1}(\cos \theta), \quad (15b)$$

where $C_l^{n/2-1}$ are the Gegenbauer hyperspherical functions and θ is the angle between s and the 1st axis.

A layer basis is constructed as the direct product of the (15) single spin basis:

$$\{\Phi(\mu)\} = \{\phi_{\alpha_1}(s_1)\} \otimes \{\phi_{\alpha_2}(s_2)\} \otimes \cdots \otimes \{\phi_{\alpha_{N_1}}(s_{N_1})\} \quad (16)$$

and is rearranged according to the irreducible representations of the *global* symmetry. For example in the Ising case a function G depending on a layer configuration vector is:

$$G(\mu) = \sum_{n=0}^{N_1} \sum_{\{\alpha_i\}} a_{\alpha_1 \alpha_2 \dots \alpha_n} s_{\alpha_1} s_{\alpha_2} \dots s_{\alpha_n}. \quad (17)$$

Then we have an expansion according to all possible 2^{N_1} spin products. The irreducible representations of the total spin inversion correspond to even and odd products, respectively.

If $G(\mu)$ is translational invariant

$$a_{\alpha_1 \alpha_2 \dots \alpha_n} = a_{\Delta_2 \Delta_3 \dots \Delta_n}^{(\alpha_1)},$$

where $\Delta_i = \alpha_i - \alpha_1$. After the Fourier transformation:

$$a_{\Delta_2 \Delta_3 \dots \Delta_n}^{(\alpha_1)} = \sum_{m=0}^{N_1-1} A_{\Delta_2 \Delta_3 \dots \Delta_n}^{(k)} \exp(2\pi i \alpha_1 k / N_1)$$

and

$$A_{\Delta_2 \Delta_3 \dots \Delta_n}^{(k)} = \frac{1}{N_1} \sum_{m=0}^{N_1} a_{\Delta_2 \Delta_3 \dots \Delta_n}^{(\alpha_1)} \exp(-2\pi i \alpha_1 k / N_1),$$

any block of the transfer matrix written in this basis will be characterized by its global symmetry quantum number(s) and by the value of the quasi-momentum.

Generally the leading eigenvalue λ_0 of the transfer matrix comes from the block with minimal quantum numbers, called the ground state block in the following. A phase transition occurs when λ_0 is degenerated (or crossed) by another λ^{OP} eigenvalue or eigenvalues coming from a different block, called the order parameter block. In the low temperature region ($\lambda_0 = \lambda^{\text{OP}}$) a quantity with matrix elements between the ground state and the order parameter blocks has a non-zero average value but vanishes when $\lambda_0 > \lambda^{\text{OP}}$ (high temperature region). Generally the simplest quantity with such properties is called the order parameter¹⁸).

In many systems the nature of the symmetry breaking is well known. In this case one can construct trial functions acting on the common subspace of the ground state and the order parameter blocks. The components of the trial function corresponding to the order parameter symmetry are zero if $\lambda_0 > \lambda^{\text{OP}}$ ($T > T_c$) but non-zero when $\lambda_0 = \lambda^{\text{OP}}$ ($T \leq T_c$).

4. A variational method

The variational method is one of the basic approaches of physics. It was applied with success to many problems also in statistical mechanics¹⁹). Actually the self-consistent and the mean-field like approaches are variational approaches.

There are two difficulties in any variational approximation: firstly one has to motivate the choice of a special form for the trial function by physical and numerical (or self-consistency) arguments and secondly one must calculate the matrix element needed. The differences between variational methods arise just from these two points.

After the reduction of the transfer operator we use the Perron–Frobenius theorem to rewrite the ground-state vector in an exponentiated form:

$$|\psi_0\rangle = \exp|\Phi_0\rangle. \quad (18)$$

Therefore

$$\begin{aligned}\lambda_0 &= \frac{\langle \psi_0 | T | \psi_0 \rangle}{\langle \psi_0 | \psi_0 \rangle} \\ &= \int d\mu \int d\mu' \exp[\Phi_0(\mu) + h(\mu, \mu') + \Phi_0(\mu')] / \int d\mu \exp 2\Phi_0(\mu) \\ &\equiv Z^{(n)} / Z^{(d)}.\end{aligned}\quad (19)$$

Eq. (19) can be interpreted as the ratio of two different partition functions: the numerator problem is a two-layer problem with the effective Hamiltonian:

$$H_{\text{eff}}^{(n)} = \Phi_0(\mu) + h(\mu, \mu') + \Phi_0(\mu'), \quad (20a)$$

while the denominator problem is a single layer problem with:

$$H_{\text{eff}}^{(d)} = 2\Phi_0(\mu). \quad (20b)$$

Using the expansion (17) it is easy to see that Φ_0 contains all possible even, translational invariant intralayer spin products, so that the Hamiltonians (20) have also infinite long range interaction terms. A simple calculation shows that any two-layer correlation function [eq. (9)] can be expressed as:

$$\langle\langle A(\mu_i, \mu_{i+1}) \rangle\rangle = \frac{1}{\lambda_0^{(n)}} \frac{\langle \phi_0^{(n)} | A T^{(n)} | \phi_0^{(n)} \rangle}{\langle \phi_0^{(n)} | \phi_0^{(n)} \rangle}, \quad (21)$$

where $\lambda_0^{(n)}$ and $|\phi_0^{(n)}\rangle$ denotes the largest eigenvalue and the corresponding eigenvector of the *numerator problem transfer matrix* $T^{(n)}$. Strictly speaking the transfer operator method cannot be applied to the systems (20a–b) because their long interaction range and the relation (21) is purely formal. However, it will be very useful within our approximation.

Now let us rearrange the sum over all possible products of intralayer spins in $\Phi_0(\mu)$ according to their interaction range. In the two-dimensional Ising model, for example, it looks like

$$\begin{aligned}\Phi_0(\mu) &= A_1 \sum s_i s_{i+1} + A_2 \sum s_i s_{i+2} + A_3 \sum s_i s_{i+3} \\ &\quad + A_4 \sum s_i s_{i+1} s_{i+2} s_{i+3} + \dots\end{aligned}\quad (22)$$

Our approximation consists in truncating the exact expansion (22) and considering the coupling constants of the few terms retained as variational parameters to maximize the Rayleigh ratio

$$\lambda_{\{A\}} = \frac{\langle \psi^{\text{Tr}} | T | \psi^{\text{Tr}} \rangle}{\langle \psi^{\text{Tr}} | \psi^{\text{Tr}} \rangle} \leq \lambda_0. \quad (23)$$

The necessary condition to get a maximum is

$$\frac{\partial}{\partial A_i} \lambda_{\{A\}} = 0, \quad (24)$$

for all A , or

$$\left\langle \left\langle \prod_{i \in A} s_i \right\rangle \right\rangle_n = \left\langle \left\langle \prod_{i \in A} s_i \right\rangle \right\rangle_d. \quad (25)$$

where $\prod_{i \in A} s_i$ is the product associated to a given A coupling constant and $\langle \rangle_n$, $\langle \rangle_d$ denotes an average according to the truncated form of (20a) and (20b) Hamiltonians, respectively.

Since we have taken into account only the short range part of the effective Hamiltonians (20) it is possible to repeat the whole procedure separately for the numerator and the denominator problem. Therefore the dimension of the problem can be reduced until one arrives at an one-dimensional or an exactly solvable model.

The simplest non-trivial models to which this method is easily applied are of course two-dimensional models because in this case the numerator-denominator problems have finite (or exactly solvable) transfer matrices. The calculation of three-dimensional models is more difficult because one obtains two-dimensional effective problems for which there are not exact results. Therefore one has to apply again the whole approximate procedure. In spite of this difficulty we hope that reliable results can be derived also in three dimensions.

The motivation for such an approach emerges from general arguments used in real-space renormalization theory²⁰), namely that the fixed point values of the coupling constants in the (22) or similar expansions are strongly decreasing functions of the interaction range and of the number of spins in the associated spin product⁹).

Noting that the leading terms of exact (high and low temperature) expansions are related to the effect of short range correlations on the "ground state" of the system it seems natural that the proposed construction of the trial functions leads to an exact fit of the first few terms of exact expansions at high and low temperature. We prove this conjecture for the 2D Ising model in the next section. The approximate free energy is an *upper bound* of the exact free energy.

Below the critical point T_c the largest eigenvalue of the transfer matrix is degenerate and the corresponding eigenvectors are in general a linear combination of a vector with ground state symmetry and a vector with order parameter symmetry. Therefore we add to the expansion (22) also terms with order parameter symmetry, for example, in the ferromagnetic case odd,

translational invariant terms:

$$\tilde{\Phi}_0(\mu) = A_0 \sum s_i + A_1 \sum s_i s_{i+1} + A_{12} \sum s_i s_{i+1} s_{i+2} + A_2 \sum s_i s_{i+2} + \dots \quad (26)$$

As long as the largest eigenvalue is non-degenerate the odd variational parameters are restricted to zero values but if $\lambda_0 = \lambda^{\text{OP}}$ they start working and are highly increasing the accuracy of the free energy in the low temperature region. Note that the form (26) of the trial function allows us to use (21), (11) for the direct calculation of the order parameter and of the critical point.

Now let us clarify the physical picture which can be attributed to our method. One has to recognize first that the truncated form of (26) can represent *exactly* the $|\psi_0\rangle$ of an (infinite long) strip of two, three, etc., spins with periodic boundary conditions. Secondly, let us remember the idea of the Bethe–Peierls approximation, namely that one solves a finite (generally small) cluster exactly and one puts it in the mean field of the remaining spins.

We can think of our method as solving exactly a strip of finite width and then putting it in the mean field of remaining spins. Note that our “cluster” contains an infinite number of spins. Moreover, the effective interaction with outer spins does not consist only of introducing a mean field like parameter A_0 , but also renormalises all intralayer interactions in the strip. Note that our method gives exact results in one-dimension, contrary to mean field theories.

For the above reasons we expect to obtain results much better than usual Bethe–Peierls approximations. The price we pay is that the symmetry of the original lattice is not preserved. In fact the most troublesome feature of the truncating approximation used for the construction of trial functions is that it is violating the symmetry property of the correlation functions, at least at large distances. This is especially clear near the critical point, where the correlation length ξ in the direction perpendicular to the layers diverges, but the correlation length in the direction parallel to the layers cannot diverge, because the corresponding numerator transfer matrix $T^{(n)}$ in (21) is finite.

This fact explains why our approach does not give accurate results for the critical indices, which are in fact mean field results.

5. Test of the method on the two-dimensional Ising model

In this section results are obtained for the Ising model with nearest neighbour interaction on a square lattice, where

$$h(\mu, \mu') = K \sum s_i s'_i + (K/2) \left(\sum s_i s_{i+1} + \sum s'_i s'_{i+1} \right).$$

They are compared with exact results^{1,21,31}). The simplest approximation is presented in detail to show how the general formalism works.

The simplest trial function is:

$$\psi^{Tr} = \exp\left(A_1 \sum s_i s_{i+1}\right) \tag{27}$$

and the corresponding Rayleigh ratio looks like

$$\lambda_{A_1} = \frac{\sum_{(\mu)} \sum_{(\mu')} \exp\{(K/2 + A_1)(\sum s_i s_{i+1} + \sum s'_i s'_{i+1}) + K \sum s_i s'_i\}}{\sum_{(\mu)} \exp\{2A_1 \sum s_i s_{i+1}\}} \tag{28}$$

The denominator problem is a simple one-dimensional problem with

$$\lambda_0^{(d)} = (2 \cosh 2A_1)^{N_1}, \tag{29a}$$

while the numerator partition function is obtained from the largest eigenvalue of a 4×4 matrix. However, taking into account the symmetries of the problem in the spirit of section 3 one knows that the largest eigenvalue belongs to the 2×2 block symmetric under total spin inversion *and* the change of spins $1 \rightleftharpoons 2, 1' \rightleftharpoons 2'$ (see fig. 1a). The result is

$$\lambda_0^{(n)} = \{2 \cosh K \cosh^2 L(1 + v_L + [(1 - v_L)^2 + 4v_L v_K]^{1/2})\}^{N_1}, \tag{29b}$$

where $L = K/2 + A_1, v_L = \tanh^2 L, v_K = \tanh^2 K$.

In general $\lambda_{A_1} = \kappa_0^{(n)}/\lambda_0^{(d)}$ is maximized by a computer. The numerator and denominator layer-Hamiltonians are (see also fig. 1):

$$h^{(n)} = K s_1 s_2 + L(s_1 s'_1 + s_2 s'_2) \tag{30a}$$

and

$$h^{(d)} = 2A_1 s_1 s'_1. \tag{30b}$$

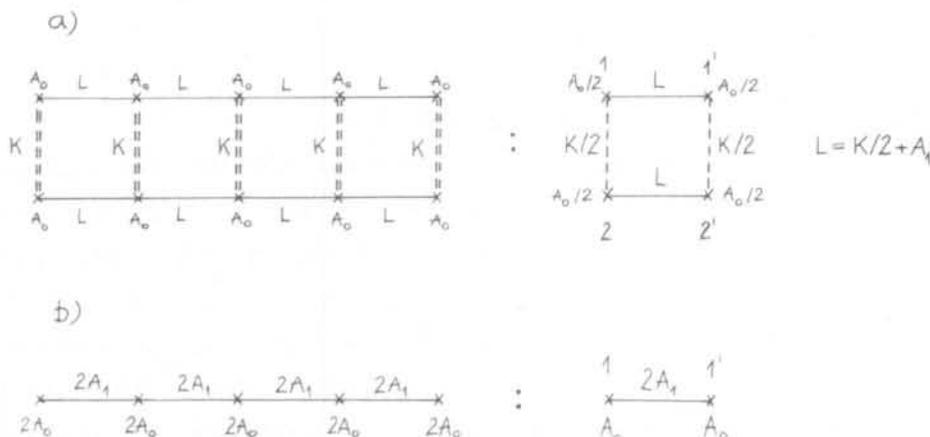


Fig. 1. a) The numerator two-row problem and its transfer operator kernel in first approximation; b) the denominator one-row problem and the denominator transfer matrix kernel.

Now let us show that the trial function (27) becomes exact at high temperatures. First let us suppose that A_1 is small if K is small. In first order the high temperature expansions for $\langle\langle S_1 S'_1 \rangle\rangle_n$ and $\langle\langle S_1 S'_1 \rangle\rangle_d$ are:

$$\langle\langle S_1 S'_1 \rangle\rangle_n \sim \tanh L \sim L = K/2 + A_1,$$

$$\langle\langle S_1 S'_1 \rangle\rangle_d \sim \tanh 2A_1 \sim 2A_1$$

and according to (25) the \bar{A}_1 value maximizing (28) is

$$\bar{A}_1 = K/2,$$

which is really small. The internal energy per spin can be expressed as [see eq. (10)]:

$$u(K) = 2\langle\langle S_1 S'_1 \rangle\rangle_n \sim 2K$$

and then the free energy per spin is in first order

$$f(K) = \int_0^K u(K) dK \sim K^2 + \ln 2,$$

as it should be. Since the free energy contains only even powers of K [$f(K) = f(-K)$] the next order correction to A_1 is at least of third order in K and the next coupling constant in the (22) expansion A_2 is also at least of third order. Similar arguments can be presented at low temperatures and can be generalised also for continuous spin models¹⁶).

Therefore we expect that the trial functions constructed as shown in the previous section are exact at high and low temperature even in more sophisticated models.

As mentioned earlier the low temperature results are highly improved if instead of (27) one uses the

$$\psi^{(1)} = \exp \left\{ A_0 \sum s_i + A_1 \sum s_i s_{i+1} \right\} \quad (31)$$

trial function, which we call first approximation.

We have constructed a computer program to calculate the Rayleigh ratio, its gradient vector and the first derivatives of the free energy per spin. At given K and given argument vector (input values) the largest eigenvalues and the corresponding (right-side and left-side) eigenvectors of the numerator and denominator problems are calculated. The gradient vector components are expressed through (9) and (25). A library routine is used to find the maximum of $\lambda_{(A)}$. Therefore we obtain the free energy per spin and its first derivatives for the given K coupling constant. All other thermodynamic functions are obtained by numerical derivation.

In the second approximation we used a trial function with the four free parameters A_0 , A_1 , A_{12} and A_2 :

$$\psi^{(2)} = \exp \left\{ A_0 \sum s_i + A_1 \sum s_i s_{i+1} + A_{12} \sum s_i s_{i+1} s_{i+2} + A_2 \sum s_i s_{i+2} \right\}. \quad (32)$$

The numerator and denominator layer-Hamiltonians are (see figs. 2a and 2b):

$$\begin{aligned} h^{(n)} = & K(s_1 s_3 + s_2 s_4) + L(s_1 s_2 + s_2 s'_1 + s_3 s_4 + s_4 s'_3) \\ & + A_2(s_1 s'_1 + s_2 s'_2 + s_3 s'_3 + s_4 s'_4) + A_0(s_1 + s_2 + s_3 + s_4) \\ & + A_{12}(s_1 s_2 s'_1 + s_2 s'_1 s'_2 + s_3 s_4 s'_3 + s_4 s'_3 s'_4), \end{aligned} \quad (33a)$$

and

$$\begin{aligned} h^{(d)} = & 2A_0(s_1 + s_2) + 2A_{12}(s_1 s_2 s'_1 + s_2 s'_1 s'_2) \\ & + 2A_1(s_1 s_2 + s_2 s'_1) + 2A_2(s_1 s'_1 + s_2 s'_2), \end{aligned} \quad (33b)$$

respectively.

Note that the (33a–b) layer-Hamiltonians cannot be put in a symmetric form and one has to use eqs. (8)–(12) according to (4b).

5.1. Discussion of numerical results

In fig. 3 the percentage errors of the first (dots) and second (crosses) approximations are presented. Even in the first approximation the free energy per spin is within 0.3% of the exact result (!), a result comparable with the lower bound renormalization group calculation of Kadanoff et al.²²). In second approximation the accuracy is further improved to 0.1% and at the same time the critical point is moved from $K_c^{(1)} = 0.413$ in first to $K_c^{(2)} = 0.422$ in second

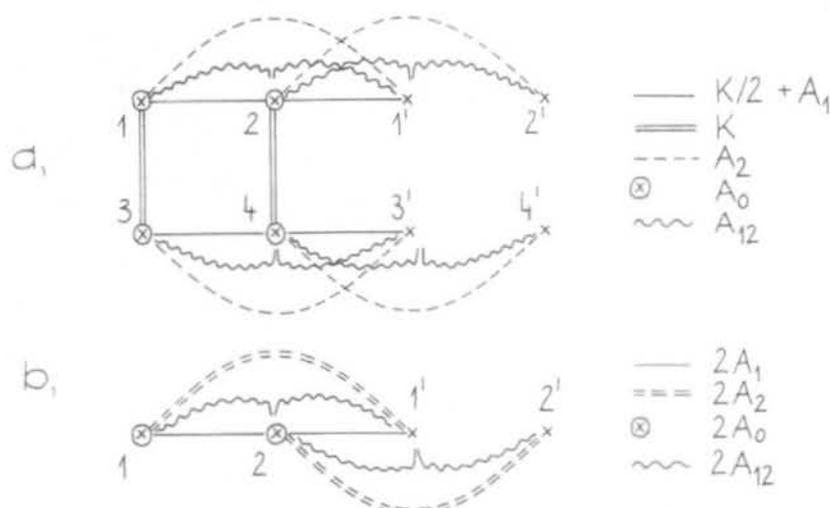


Fig. 2. The construction of the numerator (a) and denominator (b) transfer matrices in second approximation.

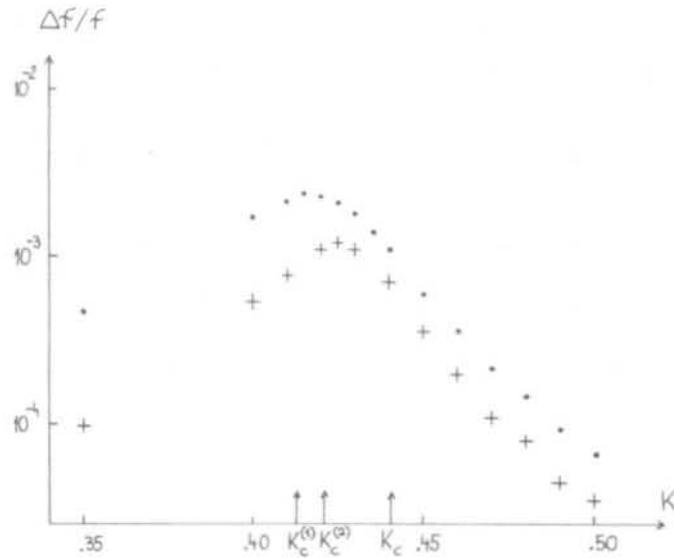


Fig. 3. Percentage errors of the free energy per spin in the first (dots) and second (crosses) approximations, $\Delta f/f = (f^{\text{exact}} - f^{\text{appr}})/f^{\text{exact}}$.

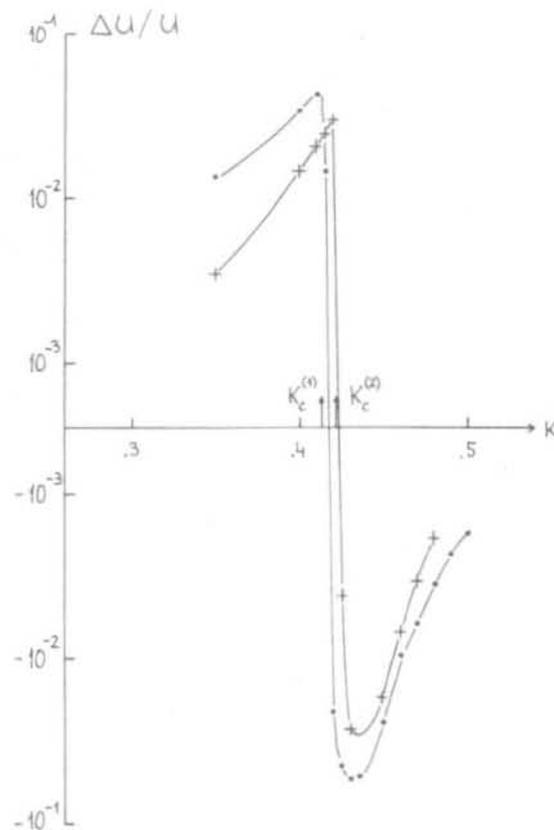


Fig. 4. Percentage errors and the internal energy per spin calculated in first (dots) and second (crosses) approximation. The solid lines are drawn only as aids to the eye.

approximation as compared with $K_c^{\text{exact}} = \frac{1}{2} \ln(\sqrt{2} + 1) = 0.4407$, $K_c^{\text{mean field}} = 0.25$, $K_c^{\text{Bethe}} = 0.346$. Note that the critical points are very accurate even as compared with RG calculations (see for example ref. 23)). The high temperature region is clearly highly improved in the second approximation. Probably a third order approximation (with 8 variational parameters) would lead to an extremely accurate free energy. However, such a calculation is expensive enough to surpass our computer facilities.

It is known from quantum mechanics that very good results can be obtained for the ground state energy even using a poor trial function. It is just the inaccuracy of the trial function (31) which causes the relatively large error (5–7%) of the internal energy† $\partial f/\partial K$ as presented in fig. 4. In fig. 5 first (dots) and second approximation (crosses) results are compared with the exact specific heat $K^2(\partial^2 f/\partial K^2)$. Since we expect that the free energy is more accurate in the low temperature region (the order parameter variational parameters are also working), it is not surprising to observe that the symmetry around T_c of the specific heat singularity is not preserved (the same effect is observed in the case of the susceptibility). The form of the specific heat is reminiscent of a higher Bethe approximation form²⁴).

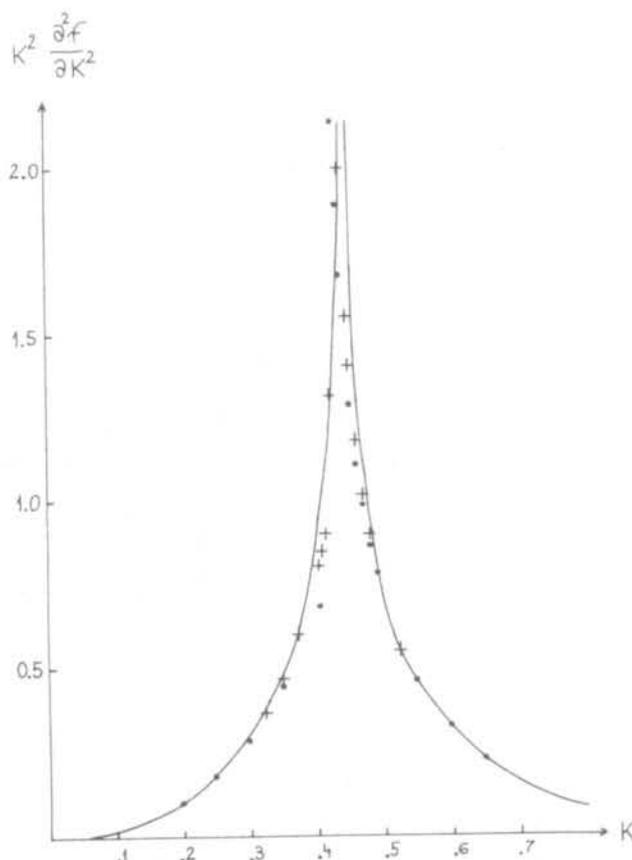


Fig. 5. The specific heat $K^2(\partial^2 f/\partial K^2)$ vs K in first (dots) and second (crosses) approximation compared with the exact result¹⁾ (solid line).

† Note that Kadanoff et al.²²) had obtained an error of 4% for the internal energy.

The spontaneous magnetisation is shown in fig. 6 where we have plotted m^8 as a function of K . Except for a relatively small region in the vicinity of T_c our results agree very well with the exact solution²¹⁾. However, the critical exponent β is about 0.4 and is closer to the mean field ($\beta = 0.5$) than to the exact value 0.125. Note that the second approximation is not improving the exponent at all.

The reduced susceptibility ($\partial^2 f / \partial h^2$) was evaluated in first approximation at $h = 10^{-5}$ and is shown in fig. 7. A linear fit to points above T_c (dots) on the log-log plot gets $\gamma = 1.24$ while below T_c one obtains $\gamma' = 1.8$. It is not clear why the latter result is so close to the exact $\gamma = \gamma' = 1.75$ result³¹⁾.

It is interesting to observe the behaviour of the maximizing value of the variational parameters. For example the parameter A_1 of the (27) approximation is represented by the solid line continued by the dashed line in fig. 8. The maximizing values of the first approximation parameters A_0 and A_1 are plotted by the dot-dashed and by the solid line, respectively. It is clear that the A_1 variational parameter is not *analytic* near T_c , its first derivative in K has a jump. This jump is reduced in the second approximation, but is

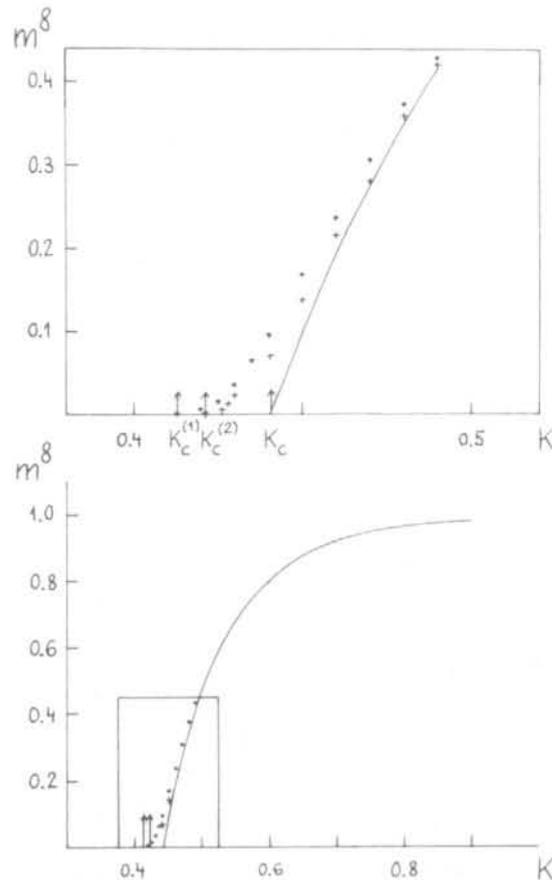


Fig. 6. m^8 vs K plot. The dots are first order, the crosses second approximation results compared with the exact Yang's²¹⁾ result (solid line).

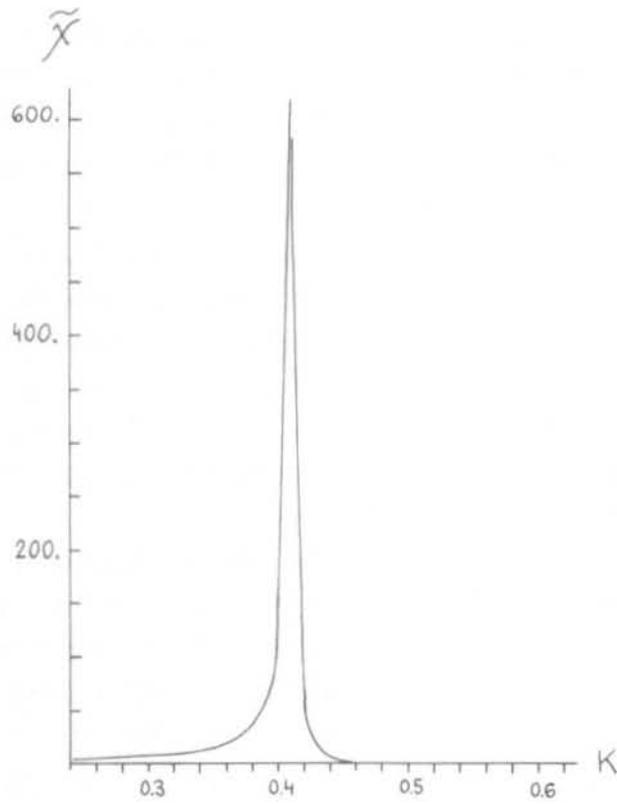


Fig. 7. The zero-field reduced susceptibility $\bar{\chi} = \partial^2 f / \partial h^2$ in first approximation.

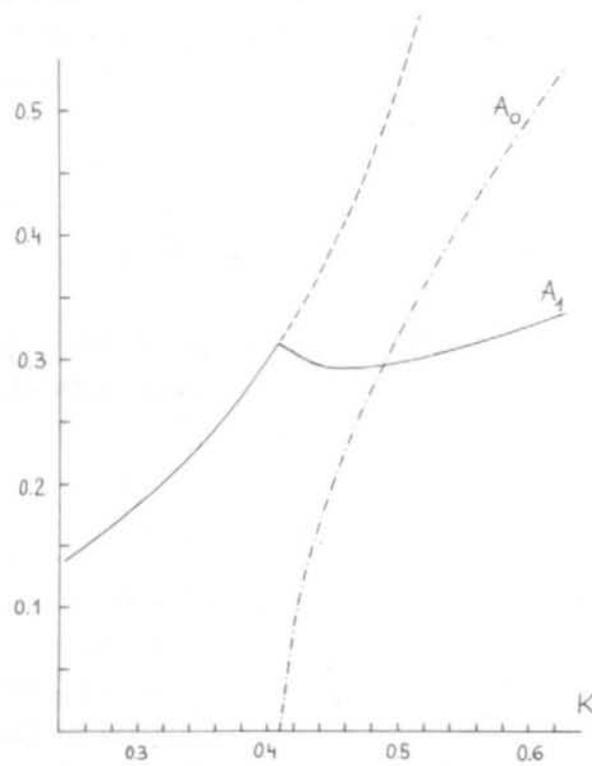


Fig. 8. Maximizing values of variational parameter A_1 in the (27) approximation (solid line continued by the dashed line) and of variational parameters A_0 and A_1 in first approximation (dot-dashed and solid line, respectively).

characteristic for any ground state variational parameters. Note also that the free energy goes abruptly to its low temperature form in the (27) approximation if A_1 increases above 0.5–0.7. It can be seen directly from the eqs. (29) that in that case λ_{A_1} equals $\exp N_1 K$ independently of A_1 . The fact that the choice (31) is much better is expressed also by the fact that in this case A_1 does not have to grow so much to ensure the correct low temperature behaviour. The study of the maximizing values of the variational parameters is very useful also to convince ourselves that the expansion (22) is convergent, that is $A_2 \ll A_1$, and $A_{12} \ll A_0$ in the second approximation. It turns out that A_0 and A_1 are greater at least by a 5 factor than A_{12} and A_2 , respectively, even at T_c , where one has the worst situation.

We conclude with a few remarks.

The violation of the isotropy of the order parameter correlation length introduces effectively two characteristic lengths, contrary to what happens in reality. Although the construction of the trial functions ensures the correct behaviour at high and low temperatures they cannot describe well situations when the true correlation length is much larger than the interaction range of the coupling constants taken into account by the trial functions. By a systematic improvement of the trial functions one can obtain excellent results for thermodynamics nearer and nearer to the critical point but close enough to T_c the asymptotic behaviour cross over to a mean field like behaviour.

In order to obtain accurate critical exponents the trial function must have a correct behaviour also near T_c . A first step in this direction has been recently done by Drell et al.²⁶⁾ and by Fernandez-Pacheco²⁷⁾.

6. Two-dimensional Ising model in an external field

In this section the two-dimensional Ising ferro- and antiferromagnet in an external field is studied applying our variational method. The model Hamiltonian is:

$$H = -J \sum_{\langle ij \rangle} s_i s_j - \mu B \sum s_i - \mu B_s \left[\sum_{\langle A \rangle} s_\alpha - \sum_{\langle B \rangle} s_\beta \right], \quad (34)$$

where $\mu = g\mu_B$ is the magnetic momentum of a spin, B is the external field and B_s the staggered field. $\langle ij \rangle$ means a sum over nearest neighbour spin pairs while $\langle A \rangle$ and $\langle B \rangle$ denote the sums over the two sublattices. When $J > 0$ the coupling is ferromagnetic while $J < 0$ corresponds to the antiferromagnetic case, well represented by materials as CoCs_3Br_5 , Rb_2CoF_4 , K_2CoF_4 ⁴³⁾.

From a theoretical point of view it is more natural to consider $K = J/k_B T$, $h = \mu B/k_B T$ and $h_s = \mu B_s/k_B T$ as independent variables instead of the usual

$1/K$; $\tau = h/K$; $\tau_s = h_s/K$ choice. Although the latter corresponds to an easily realisable experimental situation it is worth mentioning that the character of the thermodynamic functions in the two representations is almost the same. In our program we are able to calculate directly the free energy per spin (7), the internal energy per spin (10), the magnetisation (11) and the sublattice magnetisation per spin, respectively:

$$m_s \equiv \frac{M_s}{\mu N_1 N_2} = \frac{1}{2} \langle \langle s_A - s_B \rangle \rangle, \quad (35)$$

expressed through the desired independent variables. Therefore the specific heat is calculated as

$$C_B(K, \tau)/k_B N_1 N_2 = K^2 \left(\frac{\partial^2 f}{\partial K^2} \right)_\tau, \\ C_B(K, h)/k_B N_1 N_2 = K^2 \frac{\partial^2 f}{\partial K^2} + h^2 \frac{\partial^2 f}{\partial h^2} + 2Kh \frac{\partial^2 f}{\partial K \partial h}, \quad (36)$$

while the susceptibility and the sublattice susceptibility are simply

$$\tilde{\chi} \equiv \left(\frac{\partial^2 f}{\partial h^2} \right)_{K, h_s} = \frac{1}{K^2} \left(\frac{\partial^2 f}{\partial \tau^2} \right)_{K, \tau_s} = \frac{k_B T}{\mu^2 N_1 N_2} \chi, \quad (37a)$$

$$\tilde{\chi}_s \equiv \left(\frac{\partial^2 f}{\partial h_s^2} \right)_{K, h} = \frac{1}{K^2} \left(\frac{\partial^2 f}{\partial \tau_s^2} \right)_{K, \tau} = \frac{k_B T}{\mu^2 N_1 N_2} \chi_s. \quad (37b)$$

6.1. The ferromagnet

In the ferromagnetic case ($K > 0$) the staggered field is chosen to be zero. It is well known that the Ising ferromagnet in an external field does not show any phase transition. This fact is also obvious from our results. The numerical results were mainly obtained by using the first approximation (31). Generally, the calculation of a point in the (K, τ) plane is about 10-s in a CDC 3300 computer, but the same procedure takes about 3–4 min in second approximation, due to the increased number of variational parameters and to the increased dimension of numerator and denominator matrices. So, we have used the second approximation only for the calculation of a few points to check the first approximation results. For example the critical exponent defined as

$$m \sim \tau^{1/\delta} \quad \text{at } K = K_c, \quad (38)$$

was obtained from data calculated in second approximation. Our result is

$$\delta = 6.1 \quad (\delta^{\text{exact}} = 15, \delta^{\text{MF}} = 3).$$

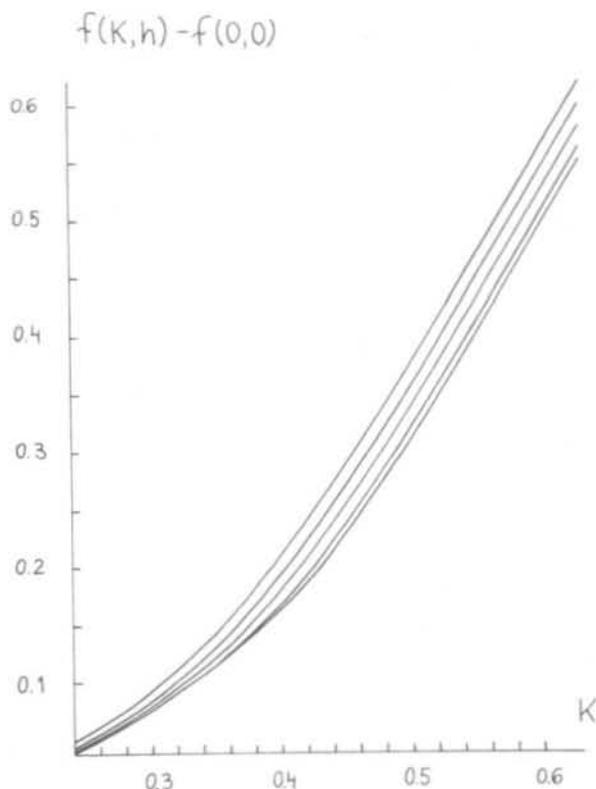


Fig. 9. The reduced free energy per spin $f(K, h) - f(0, 0)$ vs K for $h = 0; 0.01; 0.03; 0.05$ and 0.07 (in first approximation).

In fig. 9 we show the free energy for $h = 0, 1, 3, 5, 7 \times 10^{-2}$, respectively. The first derivatives of the free energy with respect to K and h (the latter one being just the magnetisation per spin) are drawn for the same h points on figs. 10 and 11. The results for the susceptibility are shown in fig. 12. Perhaps the most interesting quantity is the specific heat because a lot of exact results are known. Using the famous theorem of Lee and Yang on the zeros of the partition function Abe³⁵⁾ has shown that the specific heat depends logarithmically on the (small) field at the critical point K_c ($h = 0$), a result reconfirmed recently by Aronowitz³⁶⁾ using a perturbative approach:

$$C_B/k_B N = -\frac{8}{\pi} K_c^2(0) \ln \frac{h}{K_c(0)} \equiv a \ln \frac{1}{h} - b. \quad (39)$$

Our first approximation leads to

$$a = 0.339, \quad b = 0.0114, \quad (40a)$$

and the second approximation to:

$$a = 0.7079, \quad b = 0.3360. \quad (40b)$$

Although the constants a and b of (39) are not reproduced accurately, the fit to the logarithmic behaviour is remarkably good.

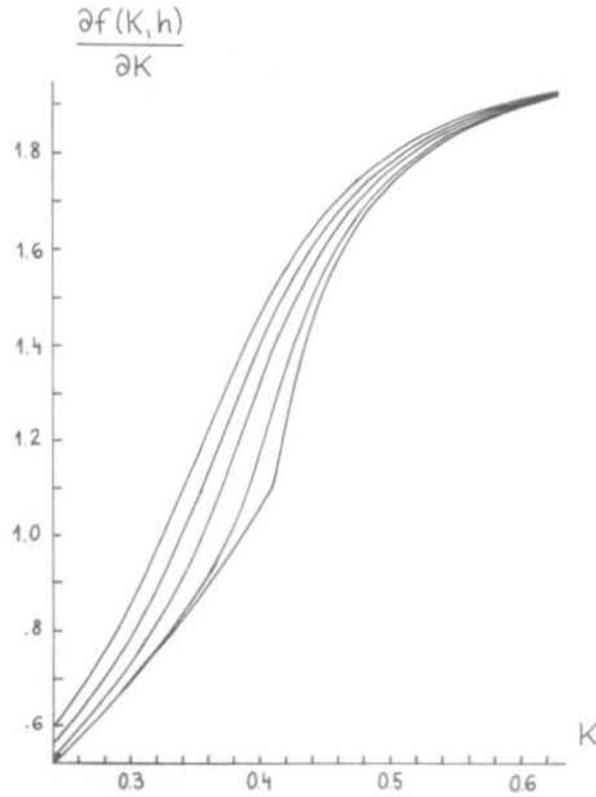


Fig. 10. $2\langle\langle s_i s_{i+1} \rangle\rangle_{K,h}$ vs K for the same h values as in fig. 9.

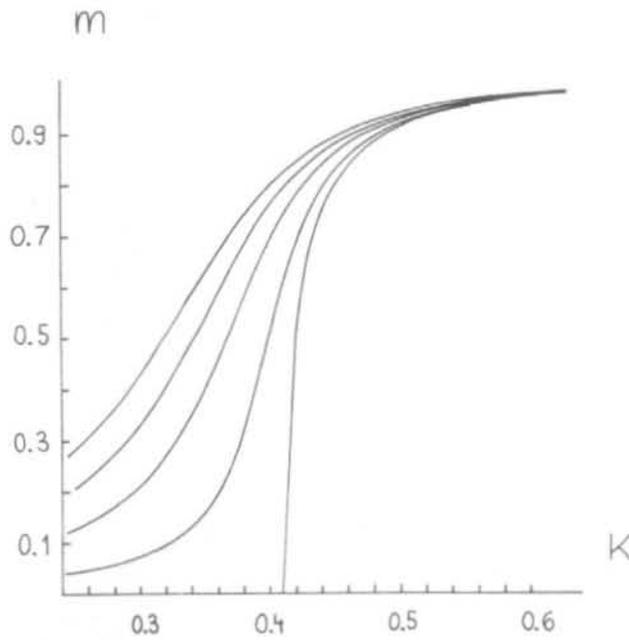


Fig. 11. The magnetisation $m = \partial f / \partial h$ vs K for fixed $h = 0; 10^{-2}; 3 \times 10^{-2}; 5 \times 10^{-2}; 7 \times 10^{-2}$ in first approximation.

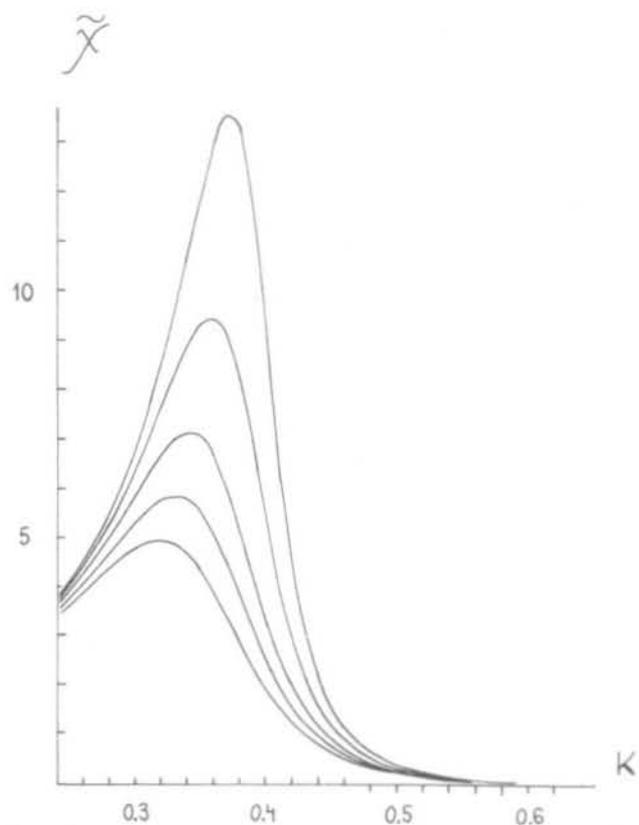


Fig. 12. The reduced susceptibility $\tilde{\chi} = \partial^2 f / \partial h^2$ vs K for $h = 2 \times 10^{-2}$; 3×10^{-2} ; 4×10^{-2} ; 5×10^{-2} ; 6×10^{-2} , in first approximation.

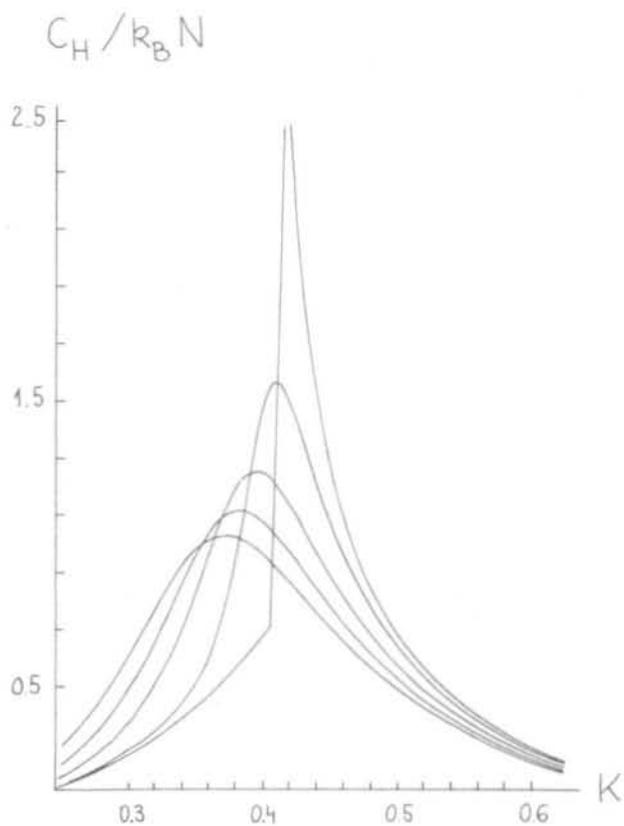


Fig. 13. The reduced specific heat $K^2(\partial^2 f / \partial K^2) + h^2(\partial^2 f / \partial h^2) + 2Kh(\partial^2 f / \partial K \partial h)$ vs K for $h = 0$, 10^{-2} ; 3×10^{-2} ; 5×10^{-2} ; 7×10^{-2} in first approximation.

6.2. The antiferromagnet

The situation is more interesting in the antiferromagnetic case. A phase transition occurs at a $T_N(h)$ temperature due to the competition between the energy of the external field (favouring a ferromagnetic arrangement) and the nearest neighbour coupling energy (favouring an antiferromagnetic phase). The global symmetry of the (34) Hamiltonian is in this case a total spin inversion *and* the interchange of the A and B sublattices ($h_s = 0$). In spite of this little difference in the nature of broken symmetry it is generally accepted that the Ising antiferromagnet belongs to the same universality class as the Ising ferromagnet. Now the order parameter is the sublattice magnetisation and the trial functions have to be constructed accordingly. In first approximation the trial function has three variational parameters:

$$\psi_{AF}^{(1)} = \exp \left\{ \alpha_0 \sum_{i \in A} s_i + \beta_0 \sum_{i \in B} s_i + A_1 \sum s_i s_{i+1} \right\}, \quad (41)$$

while in second approximation we have seven free parameters:

$$\begin{aligned} \psi_{AF}^{(2)} = \exp \left\{ \alpha_0 \sum_{i \in A} s_i + \beta_0 \sum_{i \in B} s_i + A_1 \sum s_i s_{i+1} + \alpha_{12} \sum_{i \in A} s_i s_{i+1} s_{i+2} \right. \\ \left. + \beta_{12} \sum_{i \in B} s_i s_{i+1} s_{i+2} + \alpha_2 \sum_{i \in A} s_i s_{i+2} + \beta_2 \sum_{i \in B} s_i s_{i+2} \right\}. \end{aligned} \quad (42)$$

The numerator and denominator Hamiltonians are trivially constructed as in the ferromagnetic case. Let us remark that in first approximation one has to define the numerator and denominator transfer matrices as the $T^{(n)} = T_1(\mu, \mu') T_1^+(\mu', \mu'')$ and $T^{(d)} = T_2(\mu, \mu') T_2^+(\mu', \mu'')$ produces, respectively, in order to ensure the translation symmetry of the sublattices. The correlation function matrices (9) must be also redefined accordingly. In second order approximation this problem is avoided by the fact that the numerator and denominator transfer-matrix layers are satisfying automatically the translation invariance of sublattices.

The only known exact solution for an antiferromagnet is the two-dimensional "super-exchange" antiferromagnet²⁸), a model in which non-magnetic spins are introducing an effective antiferromagnetic coupling between the basic Ising spins. This fact allows to use Onsager's solution¹) for the calculation of the free energy, internal energy, specific heat, magnetisation and susceptibility of the super-exchange ferromagnet. Where possible we will compare our results to Fisher's exact solution²⁸).

6.2.1. The phase diagram and the lattice-gas phase boundary

Recently Müller-Hartmann and Zittartz¹⁰) have used an ingenious method for the calculation of the interface energy of a two-dimensional Ising antifer-

romagnet. Taking into account only a special class of interface spin configurations they obtained a result which seems to be exact in spite of its unclarified foundation. According to Fisher's argument^{12b}), proved latter by Watson²⁹), the interface energy goes to zero at $T_N(B)$ as the inverse of ξ above T_N . The result¹⁰) confirms Griffiths' smoothness postulate³⁰) stating that only the location and the amplitude but not the exponent of the singularities of various thermodynamical functions depend on the value of the external field. From the interface energy function one can obtain the phase diagram $K_N(h)$ ¹⁰):

$$\cosh h_N = \sinh^2 \alpha K_N, \quad (43)$$

where

$$h_N = \mu B/k_B T_N(B); \quad K_N = J/k_B T_N(B); \quad \alpha = \ln(\sqrt{2} + 1)/K_N(0).$$

$K_N(0)$ and h_N are values obtained within a given approximation. In fig. 14 we plotted the critical curve of Müller-Hartmann and Zittartz. The points represent results obtained in the first approximation and the crosses results of the second approximation. The agreement is surprisingly good. Note that the result (43) is in very good agreement with the series results of Rapaport and Domb³²) who used longer series for the staggered susceptibility at zero and finite field values than in the original treatment of Bienenstock^{33a}) and Bienenstock and Lewis^{33b}). This form of phase diagram is also consistent with Fisher's result²⁸).

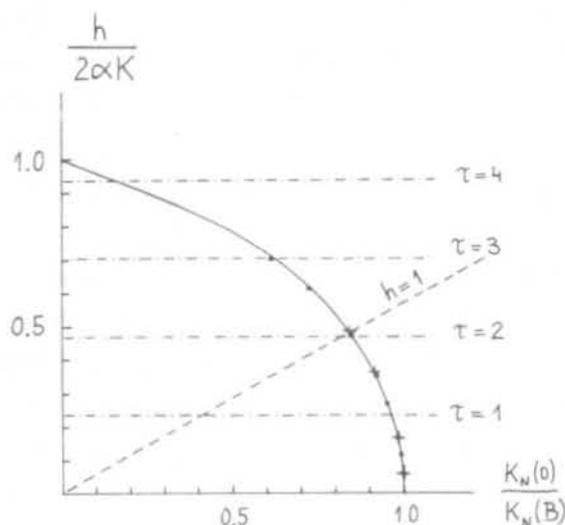


Fig. 14. The phase diagram obtained in first (dots) and second (crosses) approximation compared with the result (43) (solid line). The dashed line corresponds to the $h = 1$ plane, while the dot-dashed lines to the $\tau = 1, 2, 3, 4$, planes in first approximation [$\alpha = \ln(1 + \sqrt{2})/K_N(0) = 2.134$]. The curves representing the intersection of these planes with different thermodynamic functions are presented below.

The lattice-gas concentration $n(T_c)$ on the critical curve is obtained from

$$n_c(T_c) = \frac{1}{2} \left(1 + \frac{\partial f(K, h)}{\partial h} \right) \quad (44)$$

and is shown in fig. 15. Recently Subbaswamy and Mahan have calculated the lattice-gas phase boundary of the same model using a real space renormalization transformation²³). Our results, however, are closer to the high temperature series results^{33b}) and to the mean field results of Sato and Kikuchi³⁴).

6.2.2. The free energy, the $\langle\langle s_i s_{i+1} \rangle\rangle$ correlation function and the specific heat

The maximizing values of the first order approximation parameters (44) are shown in a typical case in fig. 16. We present typical free energy results in the usual (K, τ) variables in fig. 17.

The $-2\langle\langle s_i s_{i+1} \rangle\rangle$ is plotted at fixed τ in fig. 18. A comparison of these forms with the internal energy in the $B = 0$ case shows that the $\langle\langle s_i s_{i+1} \rangle\rangle$ correlation has a very similar form near the phase transition. At $B = 0$ its form agrees completely with the ferromagnetic form apart from the negative sign. Our results are in complete agreement with Fisher's results²⁸), apart from the mentioned crossover of our approximation to mean field in the immediate vicinity of T_N .

The specific heat plots are drawn in fig. 19. As expected, the amplitude of the specific heat singularity is decreasing with increasing $|K|$ and the specific heat has no more a singular behaviour if $B > B_c$. This picture is also in agreement with Fisher's results.

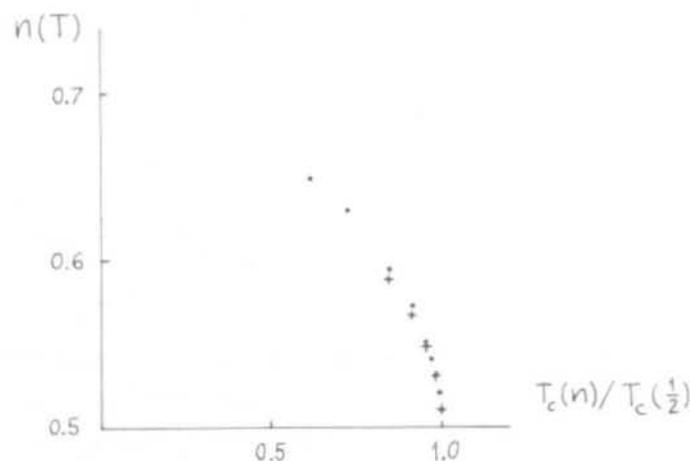


Fig. 15. The lattice-gas phase boundary obtained in first (dots) and in second (crosses) approximation.

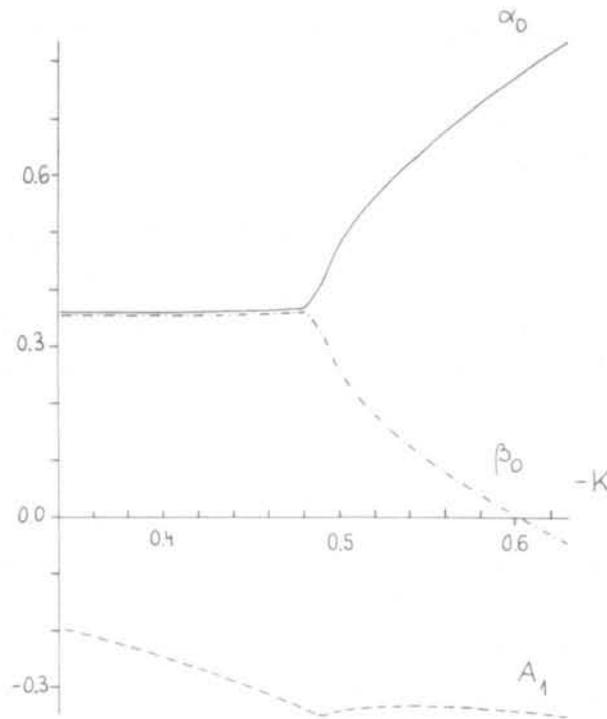


Fig. 16. Maximizing values of the variational parameters of the first approximation (41) vs $-K$ at $h = 1$.

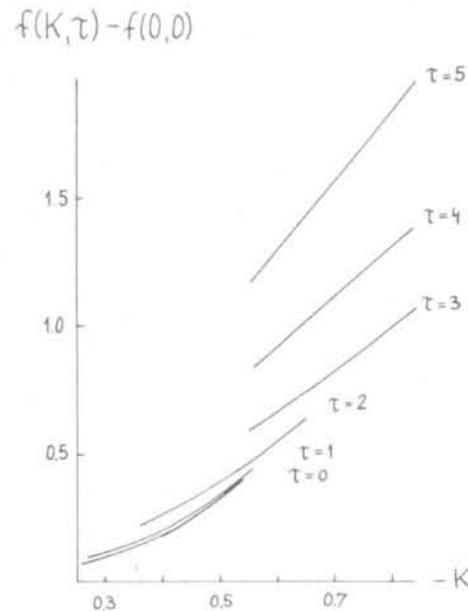


Fig. 17. The antiferromagnetic reduced free energy per spin $f(K, \tau) - \ln 2$ vs $-K$ for $\tau = 0, 1, 2, 3, 4, 5$.

6.2.3. The sublattice magnetisation and the staggered susceptibility

The spontaneous values of the sublattice magnetisation (fig. 20) at fixed values of h and τ are used for the location of the critical points. The similarity to the spontaneous magnetisation in the ferromagnetic case is evident. We have also calculated the zero staggered field sublattice reduced susceptibilities (fig. 21).

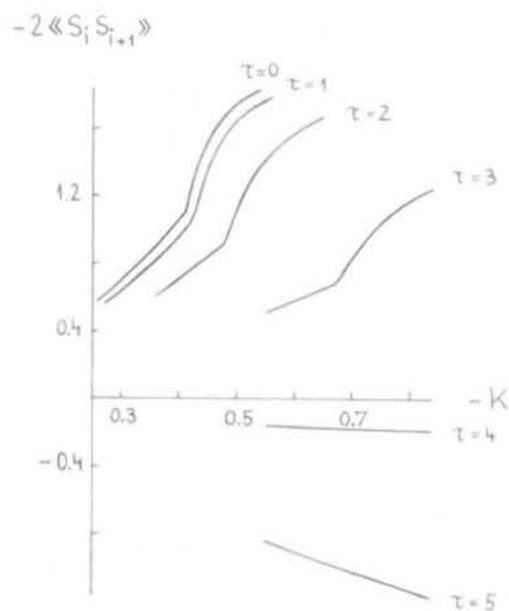


Fig. 18. The $-2\langle S_i S_{i+1} \rangle_{K,\tau} = -(\partial f / \partial K)_\tau + (\tau / K)(\partial f / \partial \tau)_K$ vs $-K$ for $\tau = 0, 1, 2, 3, 4, 5$.

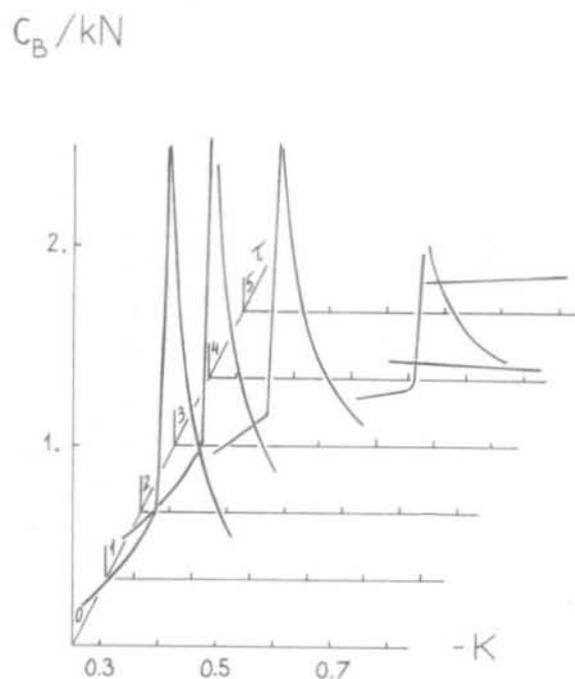


Fig. 19. The specific heat $C_B(K, \tau)$ vs $-K$ for $\tau = 0, 1, 2, 3, 4, 5$.

6.2.4. The magnetisation and the susceptibility

Perhaps one of the most interesting questions is the behaviour of the magnetisation and of the susceptibility. As long as the external field is smaller than the critical value the magnetisation is reduced to zero at low temperature by the growing antiferromagnetic order. This is clearly shown by our plots, fig. 22. At the transition point the susceptibility is singular (at finite external

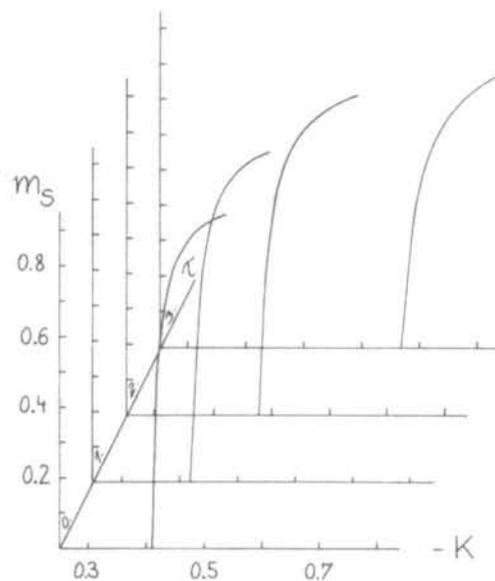


Fig. 20. The spontaneous sublattice magnetisation $(1/K)(\partial f/\partial \tau_s)_{\tau_s=0}$ vs $-K$ for $\tau = 0, 1, 2, 3$.

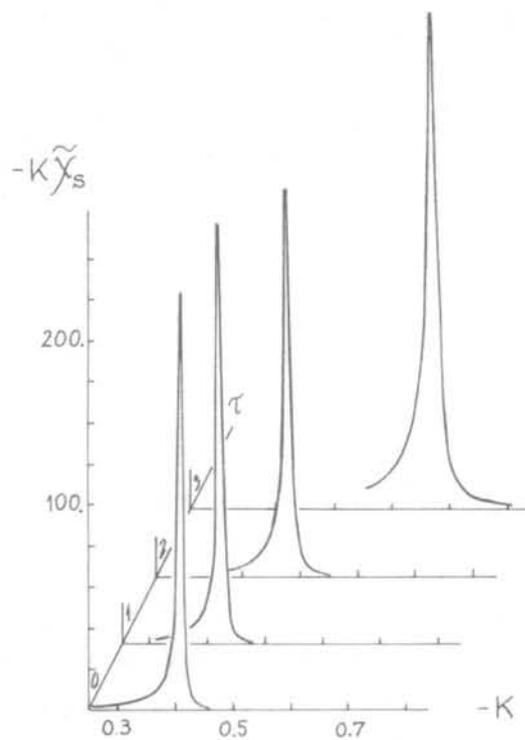


Fig. 21. The zero staggered field sublattice reduced susceptibility $K\tilde{\chi}_s(K, \tau, \tau_s = 0)$ vs $-K$ for $\tau = 0, 1, 2, 3$.

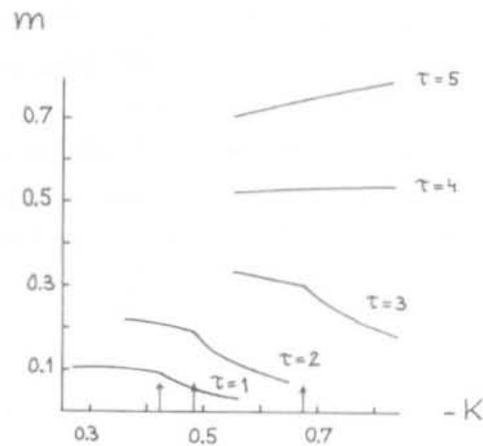


Fig. 22. The magnetisation $m(K, \tau)$ vs $-K$ for $\tau = 1, 2, 3, 4, 5$.

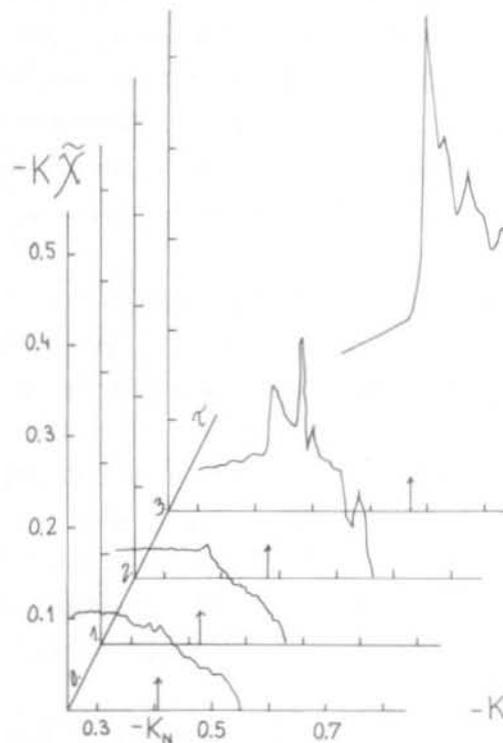


Fig. 23. The susceptibility $K\tilde{\chi}(K, \tau)$ vs $-K$ for $\tau = 0, 1, 2, 3$.

fields) as shown by fig. 23, but this seems to be a quite weak singularity, with a growing amplitude at higher fields. In fact in the super-exchange antiferromagnetic Fisher was able to show²⁸⁾ that at small fields the susceptibility behaves as

$$\tilde{\chi} = a - (b + c\tau^2) \ln|T - T_c(B = 0)| \quad (45)$$

(a, b, c are constants) and then at zero field it is proportional to the internal energy. Our result is too poor for such a comparison. In the susceptibility plots we have connected every calculated point by lines obtaining the curves of fig.

23. We cannot decide if the peaks appearing in the susceptibility below the critical temperature are due to the numerical inaccuracy or have a deeper physical significance. Although we cannot propose a physical picture for such a behaviour it will be interesting to check if this behaviour persists also in higher approximations. An other curious feature of the susceptibility plots is that the leading peak has its maximal values at temperature slightly lower than $T_N(\tau)$, see the plot of $K\bar{\chi}$ for $\tau = 3$ in fig. 23!.

7. Application to quantum systems

The main analogy used in the variational approach presented in this paper was the remarkable connection between (mostly exactly solvable) two-dimensional classical models and one-dimensional quantum Hamiltonians of interacting $\frac{1}{2}$ spins⁵). It is worth mentioning that the variational method can be applied not only to the analogue of a quantum problem but of course to the quantum system itself. The general formalism of section 4 can be used step by step in order to obtain the ground state properties (at $T = 0$) of quantum systems.

We restrict ourselves to the one-dimensional Ising chain in a transverse field at $T = 0$. It is generally accepted that this model belongs to the same universality class as the two-dimensional classical Ising model³⁷) in spite of the fact that the latter's transfer matrix does not commute with the one-dimensional Hamiltonian. Suzuki³⁸) has proved the equivalence between the two-dimensional Ising model and the ground state properties of a one-dimensional quantal anisotropic X - Y chain in a transverse field. A renormalisation group calculation³⁹) shows that the critical line of the anisotropic X - Y model goes always to the Ising fixed point, which explains the similar critical behaviour of these models.

Let us write the Hamiltonian of the one-dimensional Ising chain in a transverse field as⁴⁰):

$$H = \epsilon \sum_{i=1}^N \sigma_i^x - \Delta \sum_{i=1}^N \sigma_i^z \sigma_{i+1}^z, \quad (46)$$

where σ^x and σ^z are the Pauli matrices.

The trial function will be expressed on the orthogonal basis consisting of the direct products of the eigenvectors of σ^z :

$$\{|s_1 \dots s_N\rangle\} = \{|s_1\rangle |s_2\rangle \dots |s_N\rangle\}.$$

The Hamiltonian (46) has the same global symmetry as the two-dimensional Ising model. Using the fact that the ground state eigenvector is nodeless the

expansion (22) applies as:

$$|\psi_0\rangle = \exp\left\{\sum_{r=0}^N \sum_{\{\alpha_j\}} c_{\alpha_1\alpha_2\dots\alpha_n} \sigma_{\alpha_1}^z \sigma_{\alpha_2}^z \dots \sigma_{\alpha_n}^z\right\} |0\rangle = e^\Phi |0\rangle, \quad (47)$$

where $|0\rangle$ is defined as

$$|0\rangle = \sum_{\{s_i = \pm 1\}} |s_1\rangle |s_2\rangle \dots |s_N\rangle.$$

The ground state energy can be expressed in the following form:

$$\begin{aligned} E_0 &= \frac{\langle 0 | e^\Phi H e^\Phi | 0 \rangle}{\langle 0 | e^{2\Phi} | 0 \rangle} = \frac{\langle 0 | e^{2\Phi} H_1 | 0 \rangle}{\langle 0 | e^{2\Phi} | 0 \rangle} + \frac{\langle 0 | e^{2\Phi} B | 0 \rangle}{\langle 0 | e^{2\Phi} | 0 \rangle} \\ &= \langle\langle H_1(s_i) \rangle\rangle + \langle\langle B(s_i) \rangle\rangle, \end{aligned} \quad (48)$$

where H_1 is the part of H commuting with Φ and $B = e^{-\Phi}(H - H_1)e^\Phi$. $\langle\langle \rangle\rangle$ denotes a *statistical mechanical* average with the $e^{2\Phi}|0\rangle = e^{2\Phi(s_i)}$ "Boltzmann-factor" and the argument s_i is explicitly written to indicate a *function* of the $s_i = \pm 1$ *classical* variables ($i = 1, \dots, N$).

For example using the first approximation trial function (31)

$$|\psi^{(1)}\rangle = \exp\left\{A_0 \sum \sigma_i^z + A_1 \sum \sigma_i^z \sigma_{i+1}^z\right\} |0\rangle$$

one obtains

$$\begin{aligned} E_0 \leq E_{\{A\}} &= \frac{\sum_i \sum_{\{s\}} e^{A_0 \sum s_j + A_1 \sum s_j s_{j+1}} [\epsilon \sigma_i^x - \Delta s_i s_{i+1}] e^{A_0 \sum s_j + A_1 \sum s_j s_{j+1}}}{\sum_{\{s\}} e^{2A_0 \sum s_j + 2A_1 \sum s_j s_{j+1}}} \\ &= N\epsilon \langle\langle \exp - [2A_0 s_\alpha + 2A_1 s_{\alpha-1} s_\alpha + 2A_1 s_\alpha s_{\alpha+1}] \rangle\rangle - N\Delta \langle\langle s_\alpha s_{\alpha+1} \rangle\rangle, \end{aligned} \quad (49)$$

where the effect of σ_j^x is to flop s_j to $-s_j$ and the averages are independent of α (translational invariance). One easily recognizes that the calculation of $E_{\{A\}}$ is in fact reduced to the calculation of the $\langle\langle s_\alpha \rangle\rangle$, $\langle\langle s_\alpha s_{\alpha+1} \rangle\rangle$, $\langle\langle s_\alpha s_{\alpha+2} \rangle\rangle$ and $\langle\langle s_\alpha s_{\alpha+1} s_{\alpha+2} \rangle\rangle$ correlations functions in a one-dimensional *classical* Ising chain with nearest-neighbour interaction in an external field. The evaluation of these correlation functions can be done using the transfer matrix formalism of section 2.

In higher order approximations one will have longer even and odd interactions in the one-dimensional classical Hamiltonian and one must calculate all correlation functions until including spin products longer by one spin than the longest interaction term of the Hamiltonian. Note that the exact calculation of (48) E_0 requires in fact the calculation of all possible correlation

functions in a one-dimensional classical Ising chain with all possible interactions.

From a computational point of view the treatment of the quantum problem is more favourable than of the corresponding classical problem because the dimension of the matrix to be diagonalised is highly decreased.

Quite recently Fazekas and Kövér²⁴⁾ have independently worked out a quite similar approach. Although they have evaluated the Rayleigh-ratio (48) in a different way their results are in perfect agreement with our formula (50).

After completing the manuscript we have received a preprint by Pearson²⁵⁾ who emphasized the relation of this method with the Jastrow wave functions.

8. Conclusions

In this paper we have presented a simple variational method which can be applied to the transfer operator of classical systems or to equivalent quantum systems. The ground state trial functions are constructed in a way which takes into account all important symmetries of the problem. It will be useful to eventually incorporate the effect of more sophisticated symmetries, as the dual transformation^{12,27)}.

We used a truncating approximation in the exact exponentiated form of the ground state eigenvector taking into account only the short range part of the effective interaction between the spins lying on a transfer matrix layer. The systematic improvement of the trial function (depending mainly on computer facilities) leads to excellent results for thermodynamics (ground state properties at $T = 0$) except the immediate vicinity of the critical point, where the asymptotics changes to a mean field behaviour. However, *before* the crossover to the mean field region the thermodynamic functions (see figs. 5 and 6 for the specific heat and the magnetisation, respectively) represents quite well the behaviour of the *true* functions in the critical region. Therefore it seems possible to work out a consistent method for extracting better results for the critical indices.

An advantage of our formulation is the principal role played by the symmetry properties of the system. This gives us the possibility to follow the behaviour of the leading values of different symmetry blocks, at least to decide if they cross the ground state eigenvalue or not. The method can be applied also to models with continuous symmetry¹⁶⁾ and to three-dimensional models.

We hope to have proved that this old fashion variational approach is a competitive method even in the age of renormalization group calculations⁴¹⁾. Except a narrow region around T_c it leads to very accurate thermodynamic

functions in spite of its simplicity. Finally let me remark that it was amusing to recognize that *all* important information used in this paper was well known before my birth.

Acknowledgements

I am very grateful to Z. Rácz and A. Patkós for many valuable discussions and continued interest in this work. I am indebted to L. Larose and I. Kondor for critically reading the manuscript. I also benefitted an illuminating discussion with P. Fazekas on the relationship between our methods.

References

- 1) L. Onsager, Phys. Rev. **65** (1944) 117.
- 2) E.H. Lieb, Phys. Rev. Lett. **18** (1967) 692, 1046.
- 3) R.J. Baxter, Phys. Rev. Lett. **26** (1971) 832; Ann. Phys. (N.Y.) **70** (1972) 193.
- 4) J. Ashkin and W.E. Lamb Jr, Phys. Rev. **64** (1943) 159.
C. Domb, Proc. R. Soc. A **199** (1949) 199.
- 5) See for example P.W. Kasteleyn in: Fundamental Problems in Statistical Mechanics III, E.D.G. Cohen, ed. (North-Holland, Amsterdam, 1975).
M. Suzuki, Progr. Theor. Phys. **58** (1977) 755.
- 6) W.J. Camp and M.E. Fisher, Phys. Rev. B **6** (1972) 946.
- 7) M. Kac, in: Brandeis Lectures 1966, M. Chrétien, E.P. Gross and S. Desser, eds. (Gordon and Breach, New York).
- 8) S.B. Frobenius, Pressus. Akad. Wiss. (1908) 471.
O. Perron, Math. Ann. **64** (1907) 248.
R. Bellman, Introduction to Matrix Analysis (McGraw-Hill, New York, 1970).
See also the excellent book of P. Rózsa: Linear Algebra and its Applications (Műszaki Könyvkiadó, Budapest, 1974) in Hungarian.
- 9) P. Ruján, Physica **91A** (1978) 549.
- 10) E. Müller-Hartmann and J. Zittartz, Z. Phys. B **27** (1977) 261.
- 11) H.A. Kramers and G.H. Wannier, Phys. Rev. **60** (1941) 252, 263.
E.N. Lassetre and J.P. Howe, J. Chem. Phys. **9** (1941) 747.
E.W. Montroll, J. Chem. Phys. **9** (1941) 706.
- 12) C. Domb, Adv. in Phys. **IX** (1960) 149.
M.E. Fisher, J. Phys. Soc. Japan Suppl. **26** (1969) 87 and refs. therein. See also ref. 13.
- 13) M. Blume, N.A. Lurie and P. Heller, Phys. Rev. B **11** (1975) 4483.
- 14) J.D. Johnson, S. Krinsky and B.M. McCoy, Phys. Rev. A **8** (1973) 2526.
- 15) E.H. Stanley, Phys. Rev. **179** (1969) 570.
E.H. Stanley, Introduction to Phase Transitions and Critical Phenomena (Clarendon Press, Oxford, 1971).
- 16) A. Patkós and P. Ruján, to be published.
- 17) N. Vilenkin, Special Functions and the Representations of Groups, Ch. IX (Nauka, Moscow, 1965) in Russian.
- 18) S. Krinsky and D. Mukamel, Phys. Rev. B **16** (1977) 2313 and refs. therein.
- 19) R.P. Feynman, Statistical Mechanics (Benjamin, Reading, Massachusetts, 1972).

- 20) K.G. Wilson, *Rev. Mod. Phys.* **47** (1975) 773.
Th. Niemeijer and J.M.J. van Leeuwen, *Physica* **71** (1974) 17.
- 21) C.N. Yang, *Phys. Rev.* **85** (1952) 808.
- 22) L.P. Kadanoff, A. Houghton and M.C. Yalabik, *J. Statist. Phys.* **14** (1976) 171.
- 23) K.R. Subbaswamy and G.D. Mahan, *Phys. Rev. Lett.* **37** (1976) 642.
- 24) P. Fazekas and A. Kövér, KFKI-1978-48 preprint, submitted for publication to *Physica A*.
- 25) R.B. Pearson, FERMILAB-Pub-78/59-THY preprint (July 1978).
- 26) S.D. Drell, M. Weinstein and S. Yankielowicz, *Phys. Rev. D* **16** (1977) 1769.
- 27) A. Fernandez-Pacheco, SLAC-PUB-2099 preprint (1978).
- 28) M.E. Fisher, *Proc. R. Soc. A* **254** (1960) 66.
- 29) P.G. Watson, *J. Phys. C SER. 2*, **1** (1968) 575.
- 30) R.B. Griffiths, *Battelle Colloquium on Critical Phenomena*, R.I. Jaffee, ed. (McGraw-Hill, New York, 1972).
- 31) T.T. Wu, B.M. McCoy, C.A. Tracy and E. Barouch, *Phys. Rev. B* **13** (1976) 316.
- 32) D.C. Rapaport and C. Domb, *J. Phys. C* **4** (1971) 2684.
- 33) A. Bienenstock, *J. Appl. Phys.* **37** (1966) 1459.
A. Bienenstock and J. Lewis, *Phys. Rev.* **160** (1967) 393.
- 34) H. Sato and R. Kikuchi, *J. Chem. Phys.* **55** (1971) 677.
- 35) R. Abe, *Progr. Theor. Phys.* **38** (1967) 72.
- 36) S. Aronowitz, *Phys. Rev. B* **17** (1978) 2305.
- 37) P. Pfeuty and R. Elliott, *J. Phys. C* **4** (1971) 2370.
W. Dietrich, *Z. Phys.* **270** (1974) 239.
- 38) M. Suzuki, *Progr. Theor. Phys.* **46** (1971) 1337.
- 39) Z. Rácz, unpublished.
- 40) P. Pfeuty, *Ann. Phys. (N.Y.)* **57** (1970) 79.
- 41) S. Milosevič, in: *Phase Transitions in Condensed Matter*, Liblice Summer School Lecture Notes, M. Mátyas, ed. Prague (1978).
- 42) B.P. Demidovich and I.A. Maron, *Computational Mathematics* (MIR Publishers, Moscow, 1976).
- 43) For a review see L.J. de Jongh and A.R. Miedema, *Adv. in Phys.* **23** (1974) 1.