J. Phys. A: Math. Gen. 25 (1992) 5423–5446. Printed in the UK

Constrained variational approach to second order phase transitions: application to a two-layer Ising film

L Angelini, D Caroppo, M Pellicoro and M Villani

Dipartimento di Fisica, Università di Bari, Italy and Istituto Nazionale di Fisica Nucleare, Sezione di Bari, Italy

Received 18 February 1992, in final form 18 June 1992

Abstract. Within the framework of the transfer matrix formalism, a variational method is elaborated with a constraint related to the fundamental role of the correlation length near a critical point. The method is applied to a two-layer ferromagnetic film, made of two horizontal Ising planes interacting through a vertical coupling. The critical curve of the model, as well as the correlation length above and near T_c are calculated. The critical exponent ν obtained agrees with the expected two-dimensional exact value. For the two particular values of the vertical coupling, for which there are numerical estimates of T_c through series analysis or the Monte Carlo method, an essential agreement with these estimates is found. Analytical expressions are given in the weak and strong vertical coupling regimes. Non-perturbative aspects in the weak regime are analysed and a singularity, of square root type, at zero vertical coupling is found. Also discussed is the transition from the weak to the strong vertical coupling regime and its physical manifestation through the rapidity with which the correlation length diverges at T_c.

1. Introduction

An alternative variational procedure for the analysis of systems with an infinite number of degrees of freedom, has been proposed in a previous paper [1]. This procedure allows us to obtain exactly, in a simple analytical way, the correlation length above T_c, along a row or column for the anisotropic two-dimensional square Ising model [1]. A further elaboration of the above approach and its application to a two-layer ferromagnetic Ising film is reported in this work. We obtain results concerning the critical curve and the correlation length above and near the critical temperature for this model.

The two-layer Ising film which we consider, is made of two interacting isotropic square lattices. The model is given by the Hamiltonian

$$H = -J_1 \sum_{i=1}^{m} \sum_{j=1}^{n} (s_{i,j} s_{i+1,j} + s_{i,j} s_{i,j+1} + u_{i,j} u_{i+1,j} + u_{i,j} u_{i,j+1}) - J_2 \sum_{i=1}^{m} \sum_{j=1}^{n} s_{i,j} u_{i,j} \qquad (J_i > 0,\ s_{i,j}, u_{i,j} = \pm 1). \tag{1.1}$$

J_1 is a horizontal isotropic coupling, while J_2 is the vertical coupling. Periodic conditions on the two horizontal planes are imposed.

0305-4470/92/215423+24$07.50

When $J_2 = 0$, the model is exactly soluble. In this case the Hamiltonian $H_0 = H[J_2 = 0]$ describes two uncoupled Ising planes or, equivalently, two free fermionic fields [2, 8, 19]. The partition function is then given by the Onsager solution. However, when $J_2 \neq 0$, the model becomes non-trivial and quite interesting both from a practical (with regard to the properties of the thin magnetic films) and a theoretical point of view, with relation to the problem of two interacting fermionic fields. In this case an exact solution is lacking.

In the current paper our attention is focused on the effects of the 'coupling' J_2, as this parameter is varied. The non-perturbative aspects for small J_2 are particularly analysed, as well as the strong coupling behaviour. This type of analysis is made near the critical point, within the framework of a proper variational procedure.

The general scheme of our approach is described in section 2. The variational method is constrained in such a way as to take into account the fundamental role of the correlation length when T is near T_c. In section 3 we introduce a trial effective Hamiltonian in order to describe the probability distribution for the spin configurations on a vertical section of the layer. Then, the probability distribution on two adjacent vertical sections induced by the transfer matrix, is analysed in section 4. These two distributions are related to the classical numerator and denominator of the Rayleigh–Ritz (RR) quotient. The basic equation of our approach is given in section 5. It is based on the requirement that the above two distributions have the same correlation length. A residual parameter of the effective Hamiltonian is fixed in section 6, through a variational procedure. Then we give, in this section, the critical curve and the correlation length for T near and above T_c. In section 7 we consider the weak and strong coupling regimes and we give, in these cases, the analytic expressions of our results. We obtain a square root singularity at $J_2 = 0$. Furthermore we discuss the physical manifestation of a transition from the weak to the strong coupling regime. Some comments are made in the last section.

We note furthermore that, within the general approach described in section 2, the Ising films are the starting point for studying model with $d \geqslant 3$ [1].

2. Constrained variational method near the critical point

In this section we give an outline of the general framework of our analysis of the model (1.1). The approach which we introduce is based on the transfer matrix formalism.

Let us denote by $(\sigma, \tau) = (s_1, \ldots, s_m; u_1, \ldots, u_m)$ a spin configuration on a vertical section of the layer along a column of the horizontal planes. We call Σ_v such a type of section. The symmetrized transfer matrix L associated with (1.1), which connects Σ_v with an adjacent Σ_v', is given by

$$L(\sigma, \tau | \sigma', \tau') = \overline{L}(\sigma, \tau) \exp\left[K_1 \sum_{i=1}^{m} (s_i s_i' + u_i u_i')\right] \overline{L}(\sigma', \tau') \quad (2.1)$$

with

$$\overline{L}(\sigma, \tau) = \exp\left[\frac{1}{2} K_1 \sum_{i=1}^{m} (s_i s_{i+1} + u_i u_{i+1}) + \frac{1}{2} K_2 \sum_{i=1}^{m} s_i u_i\right]$$

and

$$K_i = \frac{J_i}{kT} \qquad (i = 1, 2).$$

We will be interested in the thermodynamic limit $n, m \to \infty$. Let Λ_1 (which is related to the free energy) be the highest eigenvalue of L and $\Psi_1(\sigma, \tau)$ the corresponding eigenvector. If we consider the probability $P(\sigma, \tau)$ of a spin configuration (σ, τ) on a section Σ_v, regardless of the configurations of all other sections, we have [3,4]

$$P(\sigma, \tau) = \frac{\Psi_1^2(\sigma, \tau)}{(\Psi_1, \Psi_1)} \tag{2.2}$$

where $(\Psi_1, \Psi_1) = \sum_{\sigma,\tau} \Psi_1^2(\sigma, \tau)$. A consequence of (2.2) is that the pair correlation function of our model, involving two spins of a section Σ_v, is equal to the pair correlation function associated with the distribution $\Psi_1^2(\sigma, \tau)$. It follows that the singularity of Λ_1 at the critical temperature T_c is strictly connected to the appearance of long range order in the principal eigenvector Ψ_1 of L.

Let us write $P(\sigma, \tau)$ in a Boltzmann form through an effective Hamiltonian $h(\sigma, \tau)$ (in which we absorb a minus sign)

$$P(\sigma, \tau) = \frac{e^{h(\sigma,\tau)}}{Z(\Sigma_v)} \tag{2.3}$$

where $Z(\Sigma_v) = \sum_{\sigma,\tau} e^{h(\sigma,\tau)}$. We can write $h(\sigma, \tau)$ in the form [5]

$$h(\sigma, \tau) = A_1 \sum_{i=1}^{m} (s_i s_{i+1} + u_i u_{i+1}) + A_2 \sum_{i=1}^{m} s_i u_i + A_3 \sum_{i=1}^{m} (s_i u_{i+1} + u_i s_{i+1}) + A_4 \sum_{i=1}^{m} s_i u_i s_{i+1} u_{i+1} + A_5 \sum_{i=1}^{m} (s_i s_{i+2} + u_i u_{i+2}) + \dots \tag{2.4}$$

where the expansion is organized according to the range and the number of spins involved in a coupling. The effective coupling parametrs A_i are functions of the temperature T or of K_1 and K_2.

Now, according to the Frobenius–Perron theorem, $\Psi_1(\sigma, \tau)$ is unique (up to constant) and positive. Then, it follows from (2.2) that

$$\Psi_1(\sigma, \tau) = e^{h(\sigma,\tau)/2} \tag{2.5}$$

and

$$Z(\Sigma_\sigma) = (\Psi_1, \Psi_1). \tag{2.6}$$

It is useful to consider also the probability $P(\sigma, \tau; \sigma', \tau')$ of a spin configuration $(\sigma, \tau), (\sigma', \tau')$ on two adjacent sections Σ_v and Σ_v', regardless of the configurations of all other sections. A generalization of (2.2) gives [1]

$$P(\sigma, \tau; \sigma', \tau') = \frac{\Psi_1(\sigma, \tau) L(\sigma, \tau | \sigma', \tau') \Psi_1(\sigma', \tau')}{(\Psi_1, L\Psi_1)} \tag{2.7}$$

where

$$(\Psi_1, L\Psi_1) = \sum_{\sigma,\tau}\sum_{\sigma',\tau'} \Psi_1(\sigma,\tau) L(\sigma,\tau|\sigma',\tau')\Psi_1(\sigma',\tau')$$
$$\equiv \sum_{\sigma,\tau}\sum_{\sigma',\tau'} e^{h(\sigma,\tau;\sigma',\tau')} = Z(\Sigma_v, \Sigma'_v). \tag{2.8}$$

The effective Hamiltonian $h(\sigma,\tau;\sigma',\tau')$ of the lattice $\Sigma_v \cup \Sigma'_v$ can be easily deduced from (2.1) and (2.5). It follows from (2.7) that $h(\sigma,\tau)$, or the coupling parameters A_i, are determined by the marginality condition

$$\sum_{\sigma',\tau'} P(\sigma,\tau;\sigma',\tau') = P(\sigma,\tau). \tag{2.9}$$

After the determination of $h(\sigma,\tau)$, Λ_1 is given by

$$\Lambda_1 = \frac{Z(\Sigma_v, \Sigma'_v)}{Z(\Sigma_v)}. \tag{2.10}$$

Of course (2.9) and (2.10) do not reduce the complexity of our problem. However we will refer us to them in our discussion of appropriate approximate procedures.

In order to have a problem that can be dealt with, we will consider approximate descriptions $\tilde{h}(\sigma,\tau)$ of the effective interaction $h(\sigma,\tau)$, where only a finite number of couplings are present. To fix the ideas, we will take, as an example

$$\tilde{h}(\sigma,\tau) = \tilde{A}_1 \sum_{i=1}^{m}(s_i s_{i+1} + u_i u_{i+1}) + \tilde{A}_2 \sum_{i=1}^{m} s_i u_i$$
$$+ \tilde{A}_3 \sum_{i=1}^{m}(s_i u_{i+1} + u_i s_{i+1}) + \tilde{A}_4 \sum_{i=1}^{m} s_i u_i s_{i+1} u_{i+1}. \tag{2.11}$$

Formally, the structure of $\tilde{h}(\sigma,\tau)$ is obtained by truncating the expansion (2.4). However the couplings $\tilde{A}_i$ will be, in general, different from the parameters A_i ($i = 1,\ldots,4$). As a matter of fact, with the fixed structure (2.11), the actual values of the $\tilde{A}_i$ will depend on the particular procedure by which (2.9) will be approximately satisfied.

Let us introduce

$$\tilde{\Psi}_1(\sigma,\tau) = e^{\tilde{h}(\sigma,\tau)/2} \qquad \tilde{Z}_D = (\tilde{\Psi}_1, \tilde{\Psi}_1) \qquad \tilde{Z}_N = (\tilde{\Psi}_1, L\tilde{\Psi}_1) \tag{2.12}$$

and the probability distributions

$$\tilde{P}_D(\sigma,\tau) = \frac{\tilde{\Psi}_1^2(\sigma,\tau)}{\tilde{Z}_D}$$
$$\tilde{P}_N(\sigma,\tau;\sigma',\tau') = \frac{\tilde{\Psi}_1(\sigma,\tau) L(\sigma,\tau|\sigma',\tau') \tilde{\Psi}_1(\sigma',\tau')}{\tilde{Z}_N}.$$

A particular recipe, by which the parameters $\tilde{A}_i$ can be determined, is given by the requirement to obtain the best approximation to Λ_1. This is the standard variational method, which starts from the rigorous inequality

$$\frac{\tilde{Z}_N}{\tilde{Z}_D} \leqslant \Lambda_1 \tag{2.13}$$

valid for every choice of the $\tilde{A}_i$, and considers

$$\sup_{\{\tilde{A}_i\}} \frac{\tilde{Z}_N}{\tilde{Z}_D} = \tilde{\Lambda}_1 \leqslant \Lambda_1. \tag{2.14}$$

This approach gives preference to the actual numerical value of Λ_1. As a matter of fact, we are more interested to its analytic structure. So it is useful to see the meaning of the variational method from the point of view of our previous general discussion.

Let us denote by $\langle \ldots \rangle_D$ and $\langle \ldots \rangle_N$ the expectation values calculated through $\tilde{P}_D(\sigma, \tau)$ and $\tilde{P}_N(\sigma, \tau; \sigma', \tau')$, respectively. The necessary condition for a maximum of $\tilde{Z}_N/\tilde{Z}_D$

$$\frac{\partial}{\partial \tilde{A}_i} \frac{\tilde{Z}_N}{\tilde{Z}_D} = 0 \qquad (i = 1, \ldots, 4) \tag{2.15}$$

leads to [5]

$$\begin{aligned} &\langle s_i s_{i+1} \rangle_D = \langle s_i s_{i+1} \rangle_N, \ \langle s_i u_{i+1} \rangle_D = \langle s_i u_{i+1} \rangle_N \\ &\langle s_i u_i \rangle_D = \langle s_i u_i \rangle_N, \ \langle s_i u_i s_{i+1} u_{i+1} \rangle_D = \langle s_i u_i s_{i+1} u_{i+1} \rangle_N \end{aligned} \tag{2.16}$$

(by symmetry $\langle u_i u_{i+1} \rangle_D = \langle u_i u_{i+1} \rangle_N$, $\langle u_i s_{i+1} \rangle_D = \langle u_i s_{i+1} \rangle_N$).

So we see that the full condition (2.9), which leads to the equality, on Σ_v, of all the correlation functions associated with $P(\sigma, \tau)$ and $P(\sigma, \tau; \sigma', \tau')$, is replaced, in the standard variational method, by an approximate marginality requirement which involves only a finite number of short distances and low order correlation functions.

A consequence of the above remark is that (2.14) can be a reliable procedure, if we are far from the critical point, since then large distances and high order correlation functions have a negligible role, so that we have an effective reduction of the full set of conditions imposed by (2.9). If, on the other hand, we are near T_c, there is no trace in (2.14) or (2.16) of the physical and mathematical mechanism of a second-order phase transition. Large distance correlations on Σ_v and $\Sigma_v \cup \Sigma_v'$ having the same behaviour, are out of the content of (2.16).

However, when T is near T_c, an alternative procedure for the determination of the parameters $\tilde{A}_i$ can be formulated, which allows us to overcome the difficulties of the standard variational method when a large number of degrees of freedom are effectively involved.

First of all, in order to be sure that we are analysing the model for T near T_c (as a matter of fact we will consider $T > T_c$), we do not take the $\tilde{A}_i$ as completely free, but impose a constraint in such a way to be sure that $\tilde{h}(\sigma, \tau)$ leads actually

to large distances correlations on Σ_v and $\Sigma_v \cup \Sigma_v'$. After this, the crucial problem is to reduce the full condition (2.9) to some equations involving only the relevant quantities associated with a second order phase transition.

From the analytic point of view, the critical point is characterized by a singularity in Λ_1 at T_c. A consistent description of this situation, when we consider the ratio $\tilde{Z}_N/\tilde{Z}_D$ with an effective Hamiltonian $\tilde{h}$ having large distance correlations, would require that the singular points of $\tilde{Z}_N$ and $\tilde{Z}_D$ be coincident. Now the relevant quantity which is responsible of the singular behaviour is the correlation length. Then, on the basis of the above natural analytic requirement, we are led to the marginality condition that $\tilde{P}_D(\sigma,\tau)$ and $\tilde{P}_N(\sigma,\tau;\sigma',\tau')$ have the same correlation length. Due to the fundamental and unique role of this quantity when T is near T_c, we sum up, with the correlation length equality (CLE) [1], a relevant aspect of (2.9) for T near T_c.

In general, some residual parameters, which we call $\tilde{A}_m'$, will not be fixed by the CLE equation. Further independent marginality conditions are then needed. However, having satisfied the crucial large distance condition, we can come back to the variational inequality (2.13) and fix the residual $\tilde{A}_m'$ through

$$\sup_{\{\tilde{A}_m'\}} \frac{\tilde{Z}_N}{\tilde{Z}_D}. \tag{2.17}$$

In this way we arrive at a constrained maximum problem, with the CLE imposed on the numerator $\tilde{Z}_N$ and denominator $\tilde{Z}_D$ of the classical Rayleigh–Ritz variational method (RRVM). So, some kind of short range marginality conditions, compatible with the CLE equation, will be also satisfied.

This approach will be completely developed in the following sections. After the determination of the $\tilde{A}_i'$, we calculate the quantity to which we have essentially addressed the full condition (2.9), that is the correlation length along a vertical section Σ_v, which allows us to obtain the critical curve of our model.

As is well known, there are other variational approachs in statistical mechanics, which do not make use of the transfer matrix formalism. A classical procedure, which allows one to obtain closed form approximations, is the cluster variation method (CVM) (see, for example, [6, 7] and references therein). However also in this approach we have a situation like that of the unconstrained RRVM; in principle the CVM approximations are reliable in the absence of long range correlations [7]. As a matter of fact, the Kikuchi version of the CVM, in the case of the two-dimensional square Ising model, gives the same approximate critical point as the standard RRVM [6]. On the other hand, in this case, the CLE equation allows us to obtain the exact correlation length and then the exact critical point [1].

3. The trial effective Hamiltonian on a vertical section of the layer

Our explicit calculations are made by introducing, on a section Σ_v, the trial effective Hamiltonian

$$\tilde{h}_t(\sigma,\tau) = 2A\sum_{i=1}^{m}(s_is_{i+1} + u_iu_{i+1}) + 2B\sum_{i=1}^{m}(s_iu_{i+1} + u_is_{i+1}) + 2C\sum_{i=1}^{m} s_iu_i \tag{3.1}$$

which is modelled on the original Hamiltonian H, except for the effective coupling $2B$, which takes into account in the simplest way of the new types of couplings resulting from the summation over the spin configurations on all the other sections. A preliminary analysis of the case $B = 0$ was made in [1]. We will see that the effective coupling B will have an important role.

As in (2.12), we associate with $\tilde{h}_t(\sigma, \tau)$ the vector

$$\tilde{\Psi}_{1t}(\sigma, \tau) = e^{\tilde{h}_t(\sigma,\tau)/2} \tag{3.2}$$

and the probability distribution

$$\tilde{P}_{Dt}(\sigma, \tau) = \frac{\tilde{\Psi}^2_{1t}(\sigma, \tau)}{(\tilde{\Psi}_{1t}, \tilde{\Psi}_{1t})}. \tag{3.3}$$

By making use of the transfer matrix formalism, we give the correlation length ξ of $\tilde{P}_{Dt}$ as a function of the parameters A, B and C.

The function $\tilde{\Psi}^2_{1t} \times (\sigma, \tau)$ can be written in terms of a transfer matrix $t(s_1, s_2|s_1', s_2')$ which connects two adjacent segments of the vertical section Σ_v. We have the 4×4 matrix

$$\begin{aligned} t(s_1, s_2|s_1', s_2') &\equiv \langle s_1, s_2|t|s_1', s_2'\rangle \\ &= e^{Cs_1s_2} \exp(2A(s_1s_1' + s_2s_2') + 2B(s_1s_2' + s_2s_1'))\, e^{Cs_1's_2'}. \end{aligned} \tag{3.4}$$

Since t commutes with the parity operator, its eigenvectors have a definite parity. As a consequence t splits into two independent 2×2 blocks. The eigenvalues associated with the positive parity are given by

$$\begin{aligned} &e^{2C} \cosh 4(A + B) + e^{-2C} \cosh 4(A - B) \\ &\quad \pm \{[e^{2C} \cosh 4(A + B) - e^{-2C} \cosh 4(A - B)]^2 + 4\}^{1/2} \end{aligned} \tag{3.5}$$

while those of negative parity are

$$2e^{2C} \sinh 4(A + B) \tag{3.6}$$

and

$$2e^{-2C} \sinh 4(A - B). \tag{3.7}$$

We will call λ_1 the highest eigenvalue with positive parity (this is given by (3.5) with the plus sign), and λ_2 the highest eigenvalue with negative parity. In the region $T > T_c$, the correlation length ξ is given by [8]

$$\frac{1}{\xi} = \log \frac{\lambda_1}{\lambda_2} \tag{3.8}$$

For $T \leqslant T_c$, we must have $\lambda_1 = \lambda_2$.

Now, a ferromagnetic behaviour of our trial Hamiltonian demands that

$$\lambda_2 = 2e^{2C} \sinh 4(A + B). \tag{3.9}$$

Furthermore we require that A be positive and such that

$$A > |B|. \tag{3.10}$$

From (3.7) and (3.9) we deduce also

$$2C + 4B > 0. \tag{3.11}$$

For $T > T_c$ but near T_c, λ_1 and λ_2 must be nearly degenerate. From (3.5) and (3.9) it follows that this happens when A or B, or both are very large and positive. However (3.10) demands that, in any case, A must be large and positive, with

$$\lim_{T \to T_c^+} A = +\infty. \tag{3.12}$$

The effective coupling A is the relevant parameter of our trial Hamiltonian $\tilde{h}_t(\sigma, \tau)$.

It is useful to introduce the following notation

$$\begin{aligned} &z = e^{-2A} \qquad \omega = e^{-2B} \qquad \gamma = e^{-C} \\ &\lambda_1 = e^{2C+4(A+B)}\tilde{\lambda}_1 \qquad \lambda_2 = e^{2C+4(A+B)}\tilde{\lambda}_2 \end{aligned} \tag{3.13}$$

We have then

$$\begin{aligned} \frac{\lambda_1}{\lambda_2} &= \frac{\tilde{\lambda}_1}{\tilde{\lambda}_2} \\ &= \frac{1 + \omega^4 z^4 + \gamma^4(\omega^4 + z^4) + \{[1 + \omega^4 z^4 - \gamma^4(\omega^4 + z^4)]^2 + 16\gamma^4\omega^4 z^4\}^{1/2}}{2(1 - \omega^4 z^4)} \end{aligned} \tag{3.14}$$

For T near and above T_c, z is small and positive, with

$$\lim_{T \to T_c^+} z = 0 \tag{3.15}$$

and

$$\lim_{z \to 0^+} \frac{\tilde{\lambda}_1}{\tilde{\lambda}_2} = +1. \tag{3.16}$$

4. The induced probability distribution on two adjacent vertical sections

Now, according to the general discussion of section 2, we fix attention on the probability distribution

$$\tilde{P}_{Nt}(\sigma, \tau; \sigma', \tau') = \frac{\tilde{\Psi}_{1t}(\sigma, \tau) L(\sigma, \tau | \sigma', \tau') \tilde{\Psi}_{1t}(\sigma', \tau')}{(\tilde{\Psi}_{1t} L \tilde{\Psi}_{1t})} \tag{4.1}$$

which is induced on two adjacent sections Σ_v an Σ'_v, by the trial effective Hamiltonian $\tilde{h}_t(\sigma,\tau)$.

From (2.1) it follows that with this distribution is associated the 16×16 transfer matrix

$$\begin{aligned}\langle s_1, s_2, s_3, s_4|\ell|s'_1, s'_2, s'_3, s'_4\rangle &= \bar{t}(s_1, s_2, s_3, s_4) \\ &\times \exp\left((A + \tfrac{1}{2}K_1)\sum_{i=1}^{4} s_i s'_i + B(s_1 s'_2 + s_2 s'_1 + s_3 s'_4 + s_4 s'_3)\right) \\ &\times \bar{t}(s'_1, s'_2, s'_3, s'_4) \end{aligned} \tag{4.2}$$

with

$$\bar{t}(s_1, s_2, s_3, s_4) = \exp\left(\tfrac{1}{2}(C + \tfrac{1}{2}K_2)(s_1 s_2 + s_3 s_4) + \tfrac{1}{2}K_1(s_1 s_4 + s_2 s_3)\right).$$

We will analyse the properties of ℓ which are relevant for our approach.

First of all, by introducing a proper orthonormal basis, the operator ℓ splits in the direct sum of two 8×8 matrices, which we call ℓ^+ and ℓ^-. These give the eigenvectors of ℓ with parity $+1$ and -1, respectively.

Then, our problem is to calculate the highest eigenvalue of ℓ^+, which we call η_1, and the highest eigenvalue of ℓ^-, which we call η_2. From these we can deduce $\tilde{Z}_{Nt}$ and the correlation length on $\Sigma_u \cup \Sigma'_v$ associated with (4.1).

By exploiting the other symmetries of ℓ (which are related to the exchange of s_1 with s_2 and of s_3 with s_4, or (s_1, s_2) with (s_3, s_4)) we have that η_1 is the highest eingenvalue of a 5×5 matrix, while η_2 is the highest eingenvalue of a 2×2 matrix.

It is useful to write

$$\eta_{1,2} = \exp\left(4(A + B) + 2C + K_2 + 4K_1\right)\tilde{\eta}_{1,2}. \tag{4.3}$$

By making use of the notation (3.13) and of the following definition

$$\nu_1 = \mathrm{e}^{-2K_1} \qquad \nu_2 = \mathrm{e}^{-K_2/2} \tag{4.4}$$

we obtain $\tilde{\eta}_2$ in the form

$$\begin{aligned}\tilde{\eta}_2 = \tfrac{1}{2}(1 - \nu_1\omega^2 z^2)\{1 + \nu_1(\omega z)^2 + \nu_1(\gamma\nu_2)^2(\omega^2 + \nu_1 z^2) \\ + \left[(1 + \nu_1(\omega z)^2 - (\gamma\nu_2)^2\nu_1(\omega^2 + \nu_1 z^2))^2 + 16(\gamma\nu_1\nu_2 z\omega)^2\right]^{1/2}\}\end{aligned} \tag{4.5}$$

The most difficult part of the calculations is the determination of $\tilde{\eta}_1$ which is the highest eigenvalue of the 5×5 Hermitian matrix h_{mn} given by

$$\begin{aligned}&h_{11} = 1 + \nu_1^2\omega^4 z^4 \qquad h_{12} = 2(\nu_1\omega z)^2 \qquad h_{13} = 2(\nu_1\nu_2\gamma\omega z)^2 \\ &h_{14} = 2\nu_1(\nu_2\gamma\omega z)^2 \qquad h_{15} = 2\nu_1\nu_2\gamma\omega z(1 + \nu_1\omega^2 z^2) \qquad h_{22} = \nu_1^2 h_{11} \\ &h_{23} = \nu_1 h_{13} \qquad h_{24} = h_{13} \qquad h_{25} = \nu_1 h_{15} \qquad h_{33} = \nu_1^2\nu_2^4\gamma^4(\omega^4 + \nu_1^2 z^4) \\ &h_{34} = \nu_2^2\gamma^2 h_{13} \qquad h_{35} = 2\nu_1^2(\nu_2\gamma)^3\omega z[\omega^2 + \nu_1 z^2] \qquad h_{44} = h_{33}/\nu_1^2 \\ &h_{45} = h_{35}/\nu_1 \qquad h_{55} = 2h_{13} + \nu_1(\nu_2\gamma)^2(\omega^2 + \nu_1 z^2)(1 + \nu_1\omega^2 z^2)\end{aligned} \tag{4.6}$$

However, since we are interested in T above and near T_c, where z is small and positive (see (3.15)), we can calculate $\tilde{\eta}_1$, in an almost analytic way, by making use of the standard perturbation theory. This leads to an expansion of $\tilde{\eta}_1$ in powers of z^2

$$\tilde{\eta}_1 = 1 + a_1 z^2 + a_2 z^4 + a_3 z^6 + a_4 z^8 + \cdots \tag{4.7}$$

where the coefficients a_i are functions of $\omega, \gamma, \nu_1, \nu_2$. We will see that, in order to develop our approach, we need at least terms of order of z^8. Except a_1 and a_2, the other coefficients are given by cumbersome analytic expressions, which have been determined by making use of 'Mathematica' [9].

From (4.5) we deduce easily an analogous expansion of $\tilde{\eta}_2$

$$\tilde{\eta}_2 = 1 + b_1 z^2 + b_2 z^4 + b_3 z^6 + b_4 z^8 + \cdots . \tag{4.8}$$

It turns out that $b_1 = a_1$. We have, as in (3.16),

$$\lim_{z \to 0+} \frac{\tilde{\eta}_1}{\tilde{\eta}_2} = 1. \tag{4.9}$$

For $z^2 \neq 0$, we have a splitting of this degenerate eigenvalue which, however, appears at the order z^4.

5. The spectral gap equality for T near T_c

We call $\log \eta_1/\eta_2$ and $\log \lambda_1/\lambda_2$ the spectral gaps of the transfer matrices ℓ and t. The basic equation of our approach is the spectral gap equality, which we write in the form

$$\frac{\lambda_1}{\lambda_2} = \frac{\eta_1}{\eta_2}. \tag{5.1}$$

For $T > T_c$, (5.1) is equivalent to the correlation length equality for the distributions $\tilde{P}_{Dt}(\sigma, \tau)$ and $\tilde{P}_{Nt}(\sigma, \tau; \sigma', \tau')$. For $T \leqslant T_c$, we must have $\lambda_1/\lambda_2 = \eta_1/\eta_2 = 1$, according to the mechanism of the spontaneous symmetry breaking. It is useful to write (5.1) in the form

$$\frac{\tilde{\lambda}_1}{\tilde{\lambda}_2} = \frac{\tilde{\eta}_1}{\tilde{\eta}_2}. \tag{5.2}$$

We will study this equation for small values of z.

From (4.9), (4.10) and $a_1 = b_1$, we deduce

$$\frac{\tilde{\eta}_1}{\tilde{\eta}_2} = 1 + r_1 z^4 + r_2 z^6 + r_3 z^8 + \cdots . \tag{5.3}$$

We make an analogous expansion for $\tilde{\lambda}_1/\tilde{\lambda}_2$. From (3.14) we have

$$\frac{\tilde{\lambda}_1}{\tilde{\lambda}_2} = 1 + v_1 z^4 + v_3 z^8 + \cdots . \tag{5.4}$$

We note that the expansion of $\tilde{\lambda}_1/\tilde{\lambda}_2$ involves only powers of z^4. On the other hand, in (5.3) there generally appear odd powers of z^2, after the term $r_1 z^4$. Then, before imposing (5.2), we implement the trial Hamiltonian $\tilde{h}_t(\sigma,\tau)$ by requiring that, at least for T near T_c, $\tilde{\eta}_1/\tilde{\eta}_2$ and $\tilde{\lambda}_1/\tilde{\lambda}_2$ have the same analytic behaviour as functions of our relevant parameter A. So we are lead to the equation

$$r_2(\omega,\gamma,\nu_1,\nu_2)=0. \tag{5.5}$$

It is useful to see the meaning of (5.5), with regard to the discussion in section 2. Starting from the induced distribution $\tilde{P}_{Nt}(\sigma,\tau;\sigma',\tau')$ given by (4.1), we come back to the section Σ_v and consider the marginal Hamiltonian $\tilde{h}'_t(\sigma,\tau)$ given by

$$\frac{e^{\tilde{h}'_t(\sigma,\tau)}}{(\tilde{\Psi}'_{1t},\tilde{\Psi}'_{1t})}=\sum_{\sigma',\tau'}\tilde{P}_{Nt}(\sigma,\tau;\sigma',\tau'). \tag{5.6}$$

Now, as it was expected, $\tilde{h}'_t(\sigma,\tau)$ has a general structure different from that of $\tilde{h}_t(\sigma,\tau)$. There are in $\tilde{h}'_t(\sigma,\tau)$ many couplings that are absent in $\tilde{h}_t(\sigma,\tau)$. Therefore the equation $\tilde{h}'_t(\sigma,\tau)=\tilde{h}_t(\sigma,\tau)$, equivalent to (2.9), cannot be satisfied. However, as long as we fix the attention on the correlation length and we are near T_c, we can describe $\tilde{h}'_t(\sigma,\tau)$ through an equivalent effective Hamiltonian $\tilde{h}'_{t\ \mathrm{eff}}$ having the same spectral gap $(\log\tilde{\eta}_1/\tilde{\eta}_2)$ and a small number of effective couplings. Now, if we impose (5.5), the equation (5.3) can be reproduced up to the order z^8 through an $\tilde{h}'_{t\ \mathrm{eff}}$ having the same structure of $\tilde{h}_t(\sigma,\tau)$.

In order to analyse (5.5), which can be considered as a constraint between the parameters B and C, it is useful to introduce the quantity ε such that

$$\varepsilon=\gamma\omega \qquad \text{with} \qquad 0<\varepsilon<1 \tag{5.7}$$

due to (3.11). Then, (5.5) can be written in the form of a cubic equation in the variable ω^2. It turns out that there exists one and only one solution of this equation which is positive and continuous for $\varepsilon,\nu_1,\nu_2\in(0,1)$. The value of the parameter ω, so determined, will be denoted by ω^*. It is a function of ε,ν_1,ν_2, for which we have

$$0<\omega^*(\varepsilon,\nu_1,\nu_2)<1. \tag{5.8}$$

Some typical behaviour of ω^*, as a function of ε, is shown in figure 1.

We can see now the role of the coupling constant B. Without this parameter, (5.5) cannot be satisfied.

Having implemented our trial Hamiltonian $\tilde{h}_t(\sigma,\tau)$, we proceed to impose the condition (5.2). From (5.3), (5.4) and (5.5) we obtain

$$z^4[c_0(\varepsilon,\nu_1,\nu_2)+c_2(\varepsilon,\nu_1,\nu_2)z^4]=0 \tag{5.9}$$

where

$$c_0(\varepsilon,\nu_1,\nu_2)=r_1(\omega^*,\varepsilon,\nu_1,\nu_2)-v_1(\omega^*,\varepsilon)$$
$$c_2(\varepsilon,\nu_1,\nu_2)=r_3(\omega^*,\varepsilon,\nu_1,\nu_2)-v_3(\omega^*,\varepsilon).$$

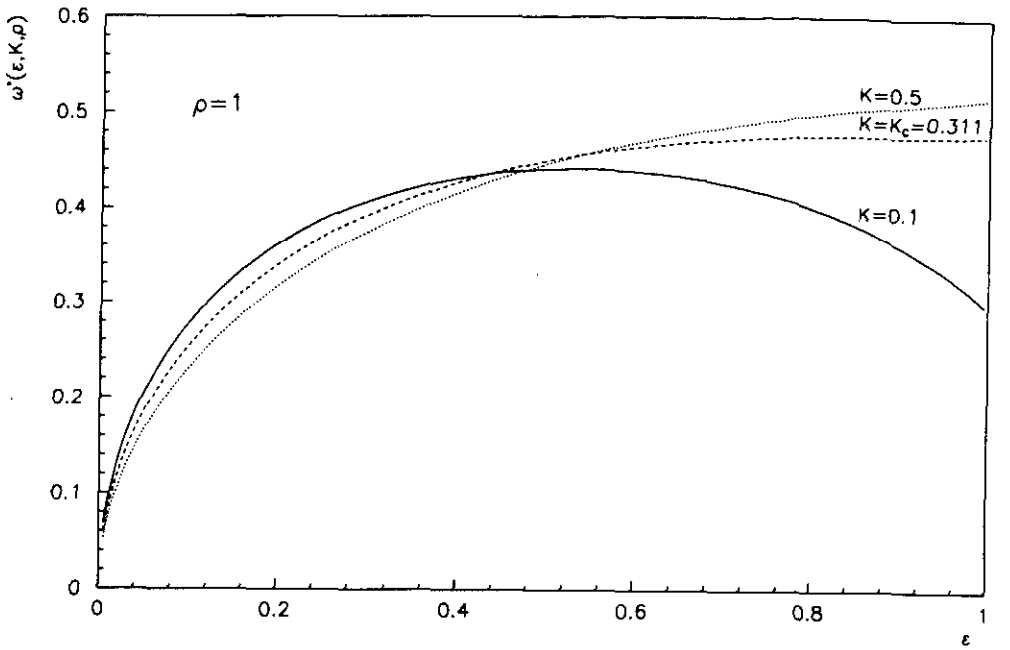

Figure 1. The solution ω^* of (5.5) as a function of ε, for $\rho = 1$.

It is useful to consider (5.9) as an equation for z. As we see, we always have the solution $z_{or} = 0$ (that is $A = +\infty$). However (5.9) can also have a solution, which we call z_{ds}, such that z_{ds} is positive and small.

In order to state the existence of z_{ds}, we make use of the parameters $K = K_1$ and $\rho = K_2/K_1$. Then, we consider the c_i as functions of ε, K, ρ. As we see from (5.9), the existence of z_{ds} is controlled by the quantity $c_0(\varepsilon, K, \rho)$.

Now, for fixed ρ, $c_0(\varepsilon, K, \rho)$ *has the following property: there exists a value* $\overline{K}(\rho)$ of K, which is unique, such that for $K > \overline{K}(\rho)$ $c_0(\varepsilon, K, \rho)$ is negative for every $\varepsilon \in [0,1]$; for $K < \overline{K}(\rho)$, but near $\overline{K}(\rho)$, there is an interval $I(K,\rho) = (\varepsilon_2(K,\rho), \varepsilon_1(K,\rho))$ of ε, with $0 < \varepsilon_2(K,\rho) < \varepsilon_1(K,\rho) < 1$, *such* that for $\varepsilon \in I(K,\rho)$, $c_0(\varepsilon, K, \rho)$ is positive, while for $\varepsilon \notin I(K,\rho)$ is negative. The $\varepsilon_i(K,\rho)$ $(i = 1,2)$ are the unique solutions of the equation

$$c_0(\varepsilon, K, \rho) = 0 \tag{5.10}$$

for $K < \overline{K}(\rho)$. The difference $\varepsilon_1(K,\rho) - \varepsilon_2(K,\rho)$ goes to zero as $K \to \overline{K}(\rho)$, so that

$$\lim_{K \to \overline{K}(\rho)} \varepsilon_2(K,\rho) = \lim_{K \to \overline{K}(\rho)} \varepsilon_1(K,\rho) = \overline{\varepsilon}(\overline{K}(\rho)) \equiv \overline{\varepsilon}. \tag{5.11}$$

Then, for K *near and below* $\overline{K}(\rho)$ and $\varepsilon \in I(K,\rho)$, $c_0(\varepsilon, K, \rho)$ is positive and small.

For fixed ρ, it can be concluded that there is a unique set S in the (K,ε) plane, where $c_0(\varepsilon, K, \rho)$ is as small as we want (positive or negative). We can write $S = I_{\overline{K}} \times I_{\overline{\varepsilon}}$, where $I_{\overline{K}}$ is an interval of values of K which is a neighbourhood of

$\overline{K}(\rho)$, while $I_{\bar{\varepsilon}}$ is an interval of values of ε , which is a neighbourhood of $\bar{\varepsilon}$. A typical behaviour of $c_0(\varepsilon, K, \rho)$ is shown in figure 2, while typical regions in the (K, ε) plane, where $c_0(\varepsilon, K, \rho)$ is positive, with the boundary determined by the $\varepsilon_i(K, \rho)$, are given in figure 3.

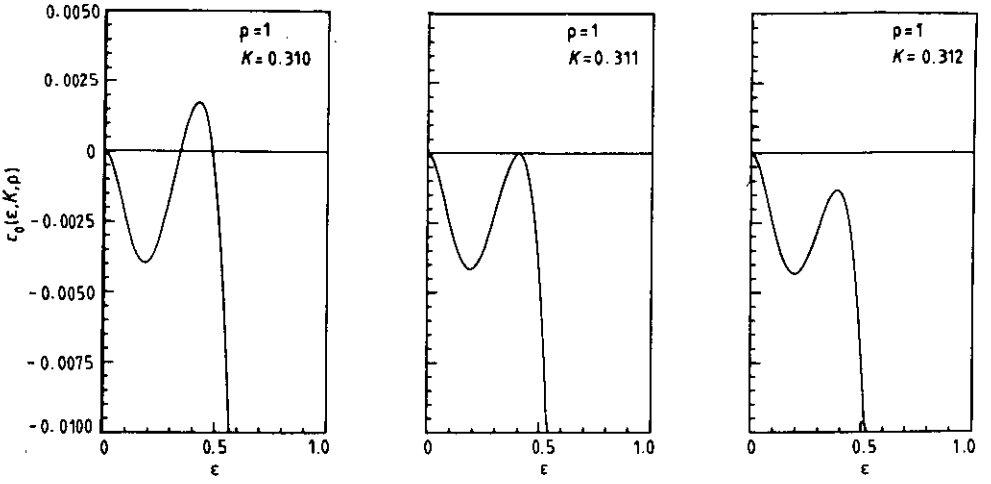

Figure 2. The behaviour of c_0 in (5.9), as a function of ε, for $\rho = 1$ and three values of K, which include $\overline{K}(1) = 0.311$.

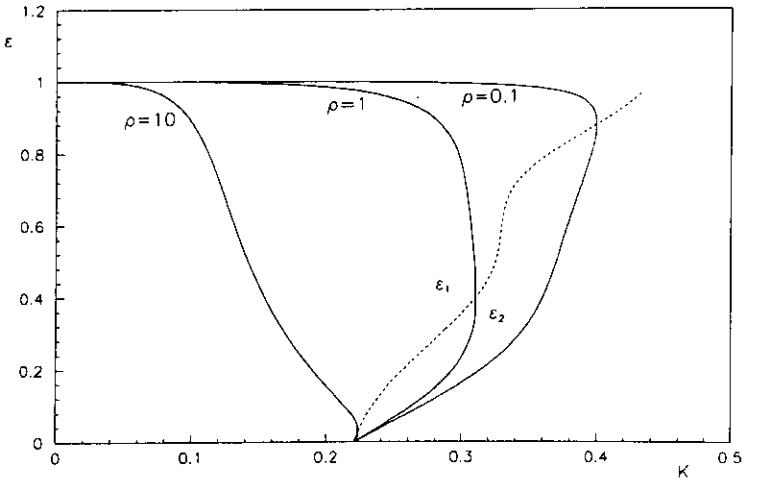

Figure 3. The domains in the (ε, K) plane, for three values of ρ, bounded by the ε and K axis and by the continuous lines, where c_0 is positive. The dashed line breaks the continuous curves into two branches, which give the solutions ε_1 and ε_2 of (5.10).

Now, (5.2) can be satisfied by small values of z, only if $(K, \varepsilon) \in S$. We can say that for $(K, \varepsilon) \in S$, we are in a neighbourhood of T_c (or K_c). Furthermore we have that $c_2(\varepsilon, K, \rho)$ is distinct from zero at the point $(\overline{K}(\rho), \bar{\varepsilon})$. In fact it turns out that $c_{2c} = c_2(\bar{\varepsilon}, \overline{K}(\rho), \rho)$ is finite and negative. This function of ρ is given in figure 4.

So we can conclude that there is a point $(\overline{K}(\rho), \bar{\varepsilon})$ in the (K, ε) plane, which is unique, such that there is a neighbourhood $\overline{S}$ of $(\overline{K}(\rho), \bar{\varepsilon})$ where (5.2) can be

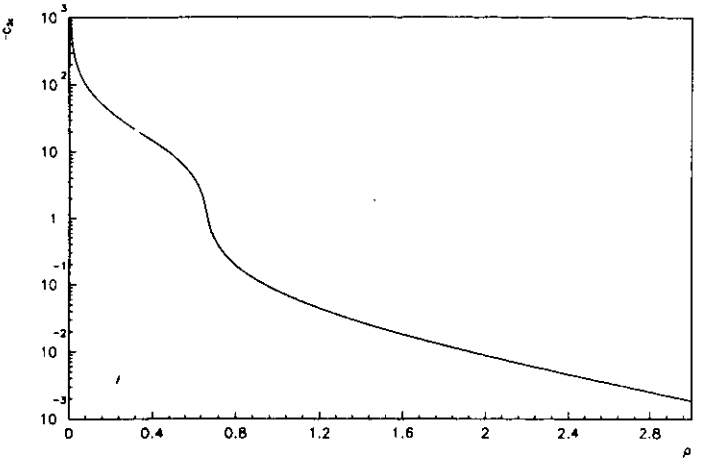

Figure 4. The coefficient c_{2c} of (5.9), as a function of ρ, when calculated at $\varepsilon = \bar{\varepsilon}$ and $K = \overline{K}(\rho)$.

satisfied by small values of z, which can be obtained through (5.9). Furthermore for $K > \overline{K}(\rho)$ and $(K, \varepsilon) \in \overline{S}$, (5.9) has only one acceptable (i.e. non-negative) solution given by $z_{or}(\varepsilon, K, \rho) = 0$. This type of solution exists also for $K < \overline{K}(\rho)$, if $\varepsilon \notin I(K, \rho)$. However for $K < \overline{K}(\rho)$, there is also a positive small solution

$$z_{ds}(\varepsilon, K, \rho) = \left(-\frac{c_0(\varepsilon, K, \rho)}{c_2(\varepsilon, K, \rho)}\right)^{1/4} \tag{5.12}$$

provided $\varepsilon \in I(K, \rho)$.

From the previous results we deduce that for $K > \overline{K}(\rho)$, we are in the ordered phase, so that $K_c \leqslant \overline{K}(\rho)$. For $K < \overline{K}(\rho)$, the kind of phase of our system is related to the value of the parameter ε. If $\varepsilon \notin I(K, \rho)$ we are again in the ordered phase, while if $\varepsilon \in I(K, \rho)$ we are in the disordered phase.

As we see the spectral gap equality gives a right framework for the analysis of a second order phase transition. However, in order to obtain an unambigous description for $K < \overline{K}(\rho)$, we have to fix ε. The variational method allows us to complete our analysis within the above framework.

6. The critical curve and the correlation length, above and near T_c, along a vertical section

We come back to the RR quotient of section 2 which, for our trial Hamiltonian, is given by

$$\frac{(\tilde{\Psi}_{1t}, L\tilde{\Psi}_{1t})}{(\tilde{\Psi}_{1t}, \tilde{\Psi}_{1t})} \tag{6.1}$$

In the limit $m \to \infty$, we have

$$\left\{\frac{(\tilde{\Psi}_{1t}, L\tilde{\Psi}_{1t})}{(\tilde{\Psi}_{1t}, \tilde{\Psi}_{1t})}\right\}^{1/m} = \frac{\eta_1}{\lambda_1} = \frac{1}{(\nu_1\nu_2)^2}\frac{\tilde{\eta}_1}{\tilde{\lambda}_1}. \tag{6.2}$$

For $K < \overline{K}(\rho)$, we consider the sup of this quantity, with respect to the residual parameter ε not fixed by the spectral gap equality. This is a constrained variational procedure with equality constraints given by (5.5) and (5.2). It is equivalent to some marginality condition at short distances, but not to the full set of short distance equations in the first line of (2.16).

Now, for $K < \overline{K}(\rho)$, but near $\overline{K}(\rho)$, we can write

$$\frac{\eta_1}{\lambda_1} = \frac{1}{(\nu_1\nu_2)^2}\left(1 + \frac{4(\varepsilon\nu_1\nu_2)^2}{1-\nu_1\nu_2^2\varepsilon^2} z^{*2}(\varepsilon, K, \rho)\right) \tag{6.3}$$

where $z^*(\varepsilon, K, \rho)$ is the non-negative solution of (5.9). This equation is obtained by expanding $\tilde{\eta}_1/\tilde{\lambda}_1$ in powers of z^2 and by taking the leading term.

So we see that the sup of η_1/λ_1 automatically selects the solution $z_{ds}(\varepsilon, K, \rho)$ given by (5.12), so that

$$\sup_{\varepsilon\in(0,1)} \frac{\eta_1}{\lambda_1} = \sup_{\varepsilon\in I(K,\rho)} \frac{1}{(\nu_1\nu_2)^2}\left[1 + \frac{4(\varepsilon\nu_1\nu_2)^2}{1-\nu_1\nu_2^2\varepsilon^2}\left(-\frac{c_0(\varepsilon, K, \rho)}{c_2(\varepsilon, K, \rho)}\right)^{\frac{1}{2}}\right] \tag{6.4}$$

The last sup gives for $K < \overline{K}(\rho)$ a strictly positive value to $z^*(\varepsilon, K, \rho)$ and it allows to fix in a unique way the value of ε, where the maximum is attained. We denote this values by $\varepsilon^*(K, \rho)$. For $K < \overline{K}(\rho)$, but near $\overline{K}(\rho)$, we can write

$$\varepsilon^*(K, \rho) \simeq \tfrac{1}{2}(\varepsilon_1(K, \rho) + \varepsilon_2(K, \rho)). \tag{6.5}$$

From (5.11) we have

$$\lim_{K\to\overline{K}^-(\rho)} \varepsilon^*(K, \rho) = \bar{\varepsilon}. \tag{6.6}$$

So, from the above procedure we deduce that, for $K < \overline{K}(\rho)$, we are in the disordered phase, with z having the positive value $z_{ds}(\varepsilon^*(K, \rho), K, \rho)$ and such that

$$\lim_{K\to\overline{K}^-(\rho)} z_{ds}(\varepsilon^*(K, \rho), K, \rho) = 0. \tag{6.7}$$

Then we conclude that the critical point is given by

$$K_c = \overline{K}(\rho). \tag{6.8}$$

Furthermore, for K below and near K_c, we obtain, from (3.8), (3.14) and (5.4), the correlation length $\xi(K, \rho)$ along a vertical section

$$\frac{1}{\xi(K, \rho)} = \frac{4\bar{\varepsilon}^4 + 2\bar{\omega}^4(1-\bar{\varepsilon}^4)}{1-\bar{\varepsilon}^4} z_{ds}^4(\varepsilon^*(K, \rho), K, \rho) \tag{6.9}$$

where only the leading term has been taken into account, and $\bar{\omega}$ is the value of ω^* at $\varepsilon = \bar{\varepsilon}$.

It can be verified that, if the parameters z or A are unconstrained and we consider the sup of η_1/λ_1 when they are freely varied, the maximum is attained always for values of z which are not small. The standard unconstrained variational

method, giving preference to regions where z is not small, is incompatible with the mathematical mechanism of a second order phase transition [3] and, as a consequence, with the CLE condition in the critical region. On the other hand, it can also be verified that, in the set $\overline{S}$, each of the short distance conditions in the first line of (2.16), cannot be satisfied separately.

We give in table 1 some values of $K_c(\rho)$ predicted by (6.8). In figure 5 we give also the critical curve, expressed as an equation

$$K_{1c} = f(K_{2c}) \tag{6.10}$$

where the critical value of the vertical coupling is considered as a free parameter in the range $(0, +\infty)$.

Table 1. Numerical values of the critical point $K_c(\rho)$, predicted by (6.8), for several values of ρ.

ρ	K_c
0.0	0.440687
0.2	0.380650
0.4	0.353656
0.6	0.334150
0.8	0.320901
1.0	0.311067
2.0	0.278453
4.0	0.246777

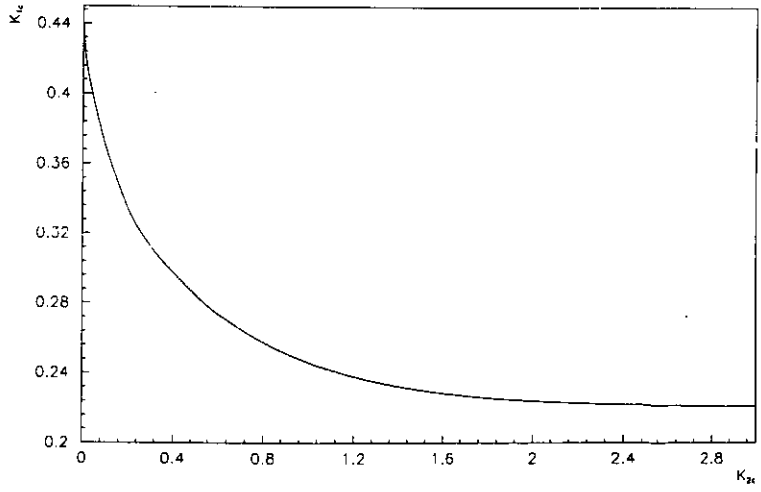

Figure 5. The critical curve $K_{1c} = f(K_{2c})$ of the two-layer Ising film, obtained from (6.8).

If we denote by

$$K_c^{d=2} = -\tfrac{1}{2}\log(\sqrt{2} - 1) = 0.440\,687$$

the critical point of the two-dimensional isotropic Ising model, we obtain from (6.8)

$$\lim_{K_{2c}\to 0} f(K_{2c}) = K_c^{d=2} \qquad \lim_{K_{2c}\to+\infty} f(K_{2c}) = \frac{K_c^{d=2}}{2} \tag{6.11}$$

in accord with the expected rigorous results.

Besides these limiting values, correctly given by our approach, there are no rigorous results about the critical curve of a two-layer Ising film. Numerical estimates of the critical temperature have been obtained in the two particular cases $\rho = 1$ and $\rho = 2$. The last case corresponds to periodic conditions in the vertical direction. It is useful to make a comparison between our predictions and these estimates.

In the case $\rho = 1$, that is $K_1 = K_2 = K$, by analysing the high-temperature series expansion of the susceptibility through the standard ratio or Padé approximant methods [10], the following value is obtained

$$(\tanh K_c)_{SE} = 0.302\,0 \pm 6. \tag{6.12}$$

There is also, for $\rho = 1$, a result based on the Monte Carlo calculation of the magnetization [11], which gives

$$(\tanh K_c)_{MC} = 0.298\,0. \tag{6.13}$$

We have a small discrepancy between the two estimates, whose origin has not been clarified. It has been argued that the uncertainty in the series calculations is bigger than that quoted in (6.12) [11]. For $\rho = 1$, our constrained variational approach gives

$$(\tanh K_c)_{CV} = 0.301\,41 \tag{6.14}$$

which agrees completely with the series estimate.

In the other case $\rho = 2$, that is $K_2 = 2K_2 = 2K$, there is only the estimate of T_c deduced from the expansion of the susceptibility [10]

$$(\tanh K_c)_{SE} = 0.269\,2 \pm 11 \tag{6.15}$$

while we obtain

$$(\tanh K_c)_{CV} = 0.271\,47. \tag{6.16}$$

If we were sure of the uncertainty reported in (6.15), then we have a discrepancy, but very small ($\sim 0.4\%$), between the two calculations. It would be interesting to see what kind of result the Monte Carlo approach gives in this case.

By making use of (5.12), the correlation length $\xi(K,\rho)$, given by (6.9), can be written in the form

$$\frac{1}{\xi(K,\rho)} = \frac{4\bar{\varepsilon}^4 + 2\bar{\omega}^4(1-\bar{\varepsilon}^4)}{1-\bar{\varepsilon}^4}\left(\frac{-c_0(\varepsilon^*(K,\rho),K,\rho)}{c_2(\bar{\varepsilon},K_c,\rho)}\right) \qquad (K<K_c) \tag{6.17}$$

with $\varepsilon^*(K,\rho)$ given by (6.5).

Now, for fixed ρ, as K goes to K_c, we have that $c_0(\varepsilon^*(K,\rho),K,\rho)$ tends to zero linearly, so that we obtain

$$\xi(K,\rho) \simeq \frac{\sigma_c}{K_c - K} \tag{6.18}$$

for K below and near K_c, where σ_c is a constant which depends on ρ. From (6.18) it follows that the critical exponent ν, which describes the singular behaviour of ξ near K_c, is given by

$$\nu = 1 \tag{6.19}$$

in accord with the expected two-dimensional value which is exactly known [8].

The above result raises a problem if there is some contact of our approach with the renormalization group (RG) method [12]. Eventually, a proper comparison could be made with the phenomenological renormalization [13] where, in the absence of a magnetic field, only a thermal field is considered. As a matter of fact, we note that in our analysis the parameter z has the role of a thermal field, the critical point being given by the equation $z = 0$. Moreover, if we refer to the effective Hamiltonians on Σ_v and $\Sigma_v \cup \Sigma_v'$, we can argue that $z = 0$ (that is $A(T_c) = \infty$) gives also the condition for a fixed point of the renormalization transformations, according to the RG treatment of the one-dimensional Ising model [14]. However, no scale factor related to RG transformations does appear in our approach. Then, on the one hand, we have an equation (i.e. (5.1)) for the thermal field, on the other hand the link between this quantity and the correlation length is obtained through a further equation (i.e. (6.9)).

Coming back to (6.18), we see that the divergence of $\xi(K, \rho)$ at K_c, is controlled also by the amplitude σ_c, which can be deduced from (6.17). This quantity, considered as a function of K_{2c}, is given in figure 6. As we see, we have a sharp maximum of σ_c for a particular value K_{2c}^* of K_c, with $K_{2c}^* \simeq 0.201$. This behaviour of σ_c, predicted by our approach, is related to a transition from the weak to the strong coupling regime, which will be discussed in the next section.

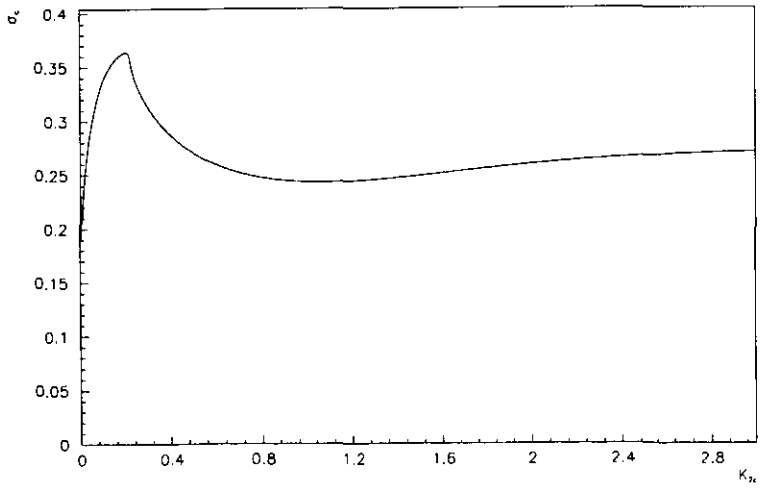

Figure 6. The amplitude σ_c of the correlation length, as a function of K_{2c}.

7. The weak and strong coupling regimes

A relevant aspect of our approach is that, in the two regions of weak or strong vertical coupling, the analysis can be made in a complete analytical way and the results can be

given by simple expressions. In the following we will call simply coupling the vertical coupling K_2 which is associated with the interaction between two Ising planes or two fermionic fields.

In the strong coupling regime we obtain for the critical curve

$$K_{1c} = f(K_{2c}) = \frac{K_c^{d=2}}{2} + \frac{\sqrt{2}-1}{2}e^{-2K_{2c}} + \frac{25\sqrt{2}-36}{4}e^{-4K_{2c}} + \cdots$$
$$= 0.220\,343 + 0.207\,107e^{-2K_{2c}} - 0.161\,165e^{-4K_{2c}} + \cdots \tag{7.1}$$

while, in the weak regime, we have

$$K_{1c} = f(K_{2c}) = K_c^{d=2} - \frac{\sqrt{6\sqrt{2}-10}}{8}\sqrt{K_{2c}} + \frac{251\sqrt{2}-172}{448}K_{2c} + \cdots$$
$$= 0.440\,687 - 0.153\,842\sqrt{K_{2c}} - 0.408\,41\,K_{2c} + \cdots. \tag{7.2}$$

From the last equation it follows that there is a singularity at zero coupling, which is predicted to be of square root type.

We note that, if we give the critical temperature as a function $T_c(J_2)$ of the vertical coupling and write, for small J_2

$$T_c(J_2) - T_c(0) \propto J_2^{1/\psi}$$

where ψ is the shift exponent [15, 16], we obtain that, in our approach, $\psi = 2$. On the other hand, on the basis of the generalized homogeneity postulate [16], the behaviour near $J_2 = 0$ can be described in terms of a crossover exponent ϕ, which is associated with a crossover temperature. It can be argued that $\phi = \psi$. Renormalization group arguments [17, 18] give support to this equality and allow us to obtain, as can also be deduced from dimensional considerations, $\phi = \gamma = 1.75$ (the susceptibility exponent of the two-dimensional square Ising model).

As we see in (7.1), the behaviour of the strong coupling expansion is very different, being given in terms of very different analytic functions. As a consequence, we can expect that there is a value of K_{2c} around which a transition occurs between the strong and weak coupling regimes. The problem is how to characterize this transition and what its physical implication is.

If we look at the critical curve, which is a smooth monotonous decreasing function of K_{2c}, there is no signal of such a transition. The situation is different if we analyse the behaviour of the effective coupling parameters of our trial Hamiltonian $\tilde{h}_t(\sigma, \tau)$. We fix the attention on the critical point, where $A \to +\infty$, and denote by B_c and C_c the values of B and C at this point. In the strong coupling regime we have

$$2B_c = K_{2c} - \frac{\log(5-3\sqrt{2})}{2} + \frac{71\sqrt{2}-122}{7}e^{-2K_{2c}} + \cdots$$
$$= K_{2c} + 0.138\,959 - 3.084\,405e^{-2K_{2c}} + \cdots \tag{7.3}$$
$$2C_c = K_{2c} - \frac{\log(2(\sqrt{2}-1))}{2} + (-7+5\sqrt{2})e^{-2K_{2c}} + \cdots$$
$$= K_{2c} + 0.094\,113\,2 + 0.071\,067\,8e^{-2K_{2c}} + \cdots. \tag{7.4}$$

We see that, at the critical point, for large values of the coupling K_{2c} between the two horizontal planes of our model, the parameters $2B_c$ and $2C_c$, which describe

the effective vertical coupling between the two horizontal lines of a section $\sum_v$, have nearly the same value of K_{2c}. There is only a small additive renormalization effect. In this respect, the strong coupling regime exhibits a simple behaviour. On the other hand, for small K_{2c}, with

$$V_0 = \log(1 - \tfrac{\sqrt{2}}{2})$$

we obtain

$$\begin{aligned} 2B_c &= -\frac{V_0}{2} + \frac{3\sqrt{7\,(675 + 251\sqrt{2})}}{392}\sqrt{K_{2c}} + \frac{231\,332 + 95\,829\sqrt{2}}{76\,832}K_{2c} + \cdots \\ &= 0.613\,974 + 0.649\,824\sqrt{K_{2c}} - 4.774\,764K_{2c} + \cdots \end{aligned} \tag{7.5}$$

and

$$\begin{aligned} 2C_c &= V_0 + \frac{5\sqrt{7(19 + 3\sqrt{2})}}{196}\sqrt{K_{2c}} + \frac{6\,483\sqrt{2} - 130\,784}{38\,416}K_{2c} + \ldots \\ &= -1.227\,95 + 0.325\,39\sqrt{K_{2c}} + 3.165\,755K_{2c} + \ldots \end{aligned} \tag{7.6}$$

Again there appears the square root singularity at $K_{2c} = 0$. But we have a further relevant phenomenon, which makes the weak regime more complex and interesting.

First of all we note that, for sufficiently small values of K_{2c}, the effective vertical coupling $2C_c$ involving pairs of nearest-neighbour spins, becomes antiferromagnetic, while the other vertical coupling $2B_c$, involving the next-nearest-neighbour spins, remains ferromagnetic. As a consequence some kind of frustration appears near $K_{2c} = 0$. However the ferromagnetic component is dominant, in the sense that, for $K_{2c} \neq 0$, we have always

$$2C_c + 4B_c > 0 \tag{7.7}$$

in agreement with (3.15), so that the global effective interaction between the two horizontal lines of a section $\sum_v$

$$2B_c \sum_{i=1}^{m}(s_i u_{i+1} + u_i s_{i+1}) + 2C_c \sum_{i=1}^{m} s_i u_i$$

is ferromagnetic.

The second aspect of (7.5) and (7.6), which is related in some way to the first one, is the finite value of B_c and C_c as $K_{2c} \to 0$

$$\begin{aligned} 2B_c(0) &= \lim_{K_{2c}\to 0+} 2B_c = -\tfrac{V_0}{2} \\ 2C_c(0) &= \lim_{K_{2c}\to 0+} 2C_c = V_0. \end{aligned} \tag{7.8}$$

This result is in contrast with the naive expectation that when the coupling K_{2c} tends to zero, the effective coupling B_c and C_c both also tend to zero. However, it happens that

$$2C_c(0) + 4B_c(0) = 0 \tag{7.9}$$

so that, when the two horizontal planes become uncoupled, the net effective interaction energy between the two horizontal lines of a section $\sum_v$, tends to zero. We can say that the effective vertical coupling of longer range, $2B_c$, is always ferromagnetic and remains finite as $K_{2c} \to 0^+$, due to some hysteresis mechanism. This fact, however, forces the other effective coupling $2C_c$ to become antiferromagnetic, in order to have a vanishing effective interaction energy when K_{2c} tends to zero. The mathematical origin of this phenomenon is the equation (5.5), which cannot be satisfied for $\omega = 1$ (that is $B = 0$).

From the analysis of the effective parameters, a physical picture emerges, which is quite different for the weak and strong coupling regime. It is very interesting, then, to determine where and how the transition between the two regimes occurs. In the figures 7(*a*) and 7(*b*) we give $2B_c$ and $2C_c$ as functions of K_{2c}. Both curves show clearly an upper and a lower branch associated with the strong and weak regimes, respectively. The transition from the upper to the lower branch happens in a very steep way, around the point $K'_{2c} \simeq 0.220$, where both the derivatives of $2C_c(K_{2c})$ and of $2B_c(K_{2c})$ get their maximum value. We see also that the transition, which looks like a smoothed jump, starts from the point $K^0_{2c} \simeq 0.270$ where C_c is zero.

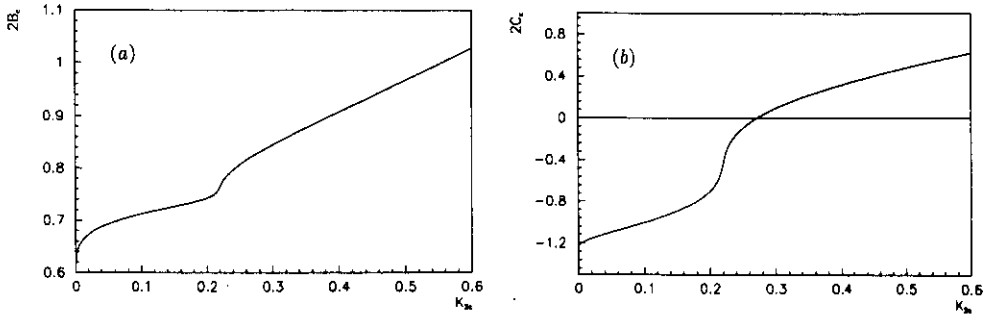

Figure 7. The effective coupling parameters $2B_c$ (*a*) and $2C_c$ (*b*)J at the critical point, as functions of K_{2c}.

Now we note that, in the interval where the behaviour of both $2B_c$ and $2C_c$ indicates the above transition from the weak to the strong coupling regime, we have also the point K^*_{2c} where σ_c gets its maximum. This maximum is an imitation of a mechanism analogous to the divergence of the correlation length at the critical point. It refers, however, to the rapidity σ_c with which the correlation length diverges at K_c. This further phenomenon of the appearence of the maximum, provides us with a physical means through which, in principle, the transition from the weak to the strong coupling regime can be detected and located.

For small values of K_{2c}, σ_c is given by

$$\sigma_c = \frac{3+\sqrt{2}}{32} + \frac{\sqrt{54-38\sqrt{2}}}{12\,544} \times \left(12\,701 + 9\,652\sqrt{2} + \frac{245+196\sqrt{2}}{\log(\sqrt{2}-1)}\right)\sqrt{K_{2c}} + \cdots = 0.137\,944 + 1.046\,83\sqrt{K_{2c}} + \cdots \tag{7.10}$$

while, in the strong coupling regime, we have

$$\sigma_c = \frac{85+16\sqrt{2}}{392} - \frac{69\sqrt{2}-53}{98\log(1+\sqrt{2})}K_{2c}e^{-2K_{2c}} + \frac{4701\sqrt{2}-6928}{2744}e^{-2K_{2c}} + \cdots$$
$$= 0.27465 - 0.516132K_{2c}e^{-2K_{2c}} + 0.101961e^{-2K_{2c}} + \cdots . \quad (7.11)$$

We note that, if we take initially $K_2 = 0$, we deduce from the equation [8]

$$\tfrac{1}{\xi^0} = 2(K^* - K) \quad (7.12)$$

that $\sigma_c^0 = 1/4$. Here $\xi^0 = \xi(K_2 = 0)$ and $e^{-2K^*} = \tanh K$.

On the other hand, from (7.10) we obtain

$$\lim_{K_{2c}\to 0+} \sigma_c \simeq 0.138 \neq \sigma_c^0. \quad (7.13)$$

This result is a physical manifestation of the finiteness of the effective parameters B_c and C_c when K_{2c} tends to zero, as expressed by the limits (7.8).

8. Conclusions

The approach proposed in this paper allows us to obtain quite reliable results which, on the one hand, are under control from the analytical point of view, and on the other hand are carriers of interesting qualitative physical pictures. This is a consequence of the particular role of the eigenvector $\Psi_1(\sigma,\tau)$ of the transfer matrix, on which we have focused our attention. From the point of view of quantum field theory $\Psi_1(\sigma,\tau)$ has the role of the vacuum state. So our results on the effective parameters B_c and C_c can be useful with regard to the problem of the non-perturbative aspects of the structure of the vacuum state.

A physical motivation of the reliability of our approach can be the analogous procedure followed in the low-energy quantum physics, where simple effective potentials (harmonic oscillators, double wells, ...) are introduced, in order to obtain a good description of the low part of the spectrum. As a matter of fact we consider in this paper, through $\tilde{h}_t(\sigma,\tau)$, a kind of effective potential of the simplest form, having the property to reproduce the energy gap between the ground state and the first excited state. However, it would be relevant to have, besides the heuristic arguments developed in this paper, a more rigorous basis with some control on the results. We will be faced with this problem in future work.

For $T > T_c$, besides the correlation length it would be interesting to calculate also the other important physical quantities, like the heat capacity and the susceptibility. With regard to the latter quantity, we have to introduce a small external magnetic field and, as a consequence, we need a modification of our trial Hamiltonian. The same procedure is required for $T < T_c$, if we want also to calculate the magnetization.

However, the heat capacity could be determined approximately through the second derivative, with respect to T, of the log of the RR quotient (6.1), which gives an approximation to the free energy. But, even if $\tilde{\eta}_1$ and $\tilde{\lambda}_1$ are singular at the same

point, we obtain a poor approximation to the free energy near the critical point. In fact, from (6.3) and (6.8) we see that we would obtain a free energy whose singular part does not satisfy the scaling relation $\sim (1/\xi)^d$ [12]. The Rayleigh–Ritz approximation to the free energy, while allowing us to get stability for $K < \overline{K}(\rho)$ with respect to the residual parameter ε, seems incompatible with the full content of the scaling hypothesis.

In order to obtain a reliable heat capacity we need an improvement upon the RR quotient which we have considered. In principle, if we consider the Hamiltonian $\tilde{h}'_t(\sigma,\tau)$ given by (5.6) and the vector

$$\tilde{\phi}_t(\sigma,\tau) = \exp\left(\tilde{h}'_t(\sigma,\tau) - \tfrac{1}{2}\tilde{h}_t(\sigma,\tau)\right) \propto (L\tilde{\Psi}_{1t})(\sigma,\tau) \tag{8.1}$$

the ratio

$$\frac{(\tilde{\phi}_{1t}, L\tilde{\phi}_{1t})}{(\tilde{\phi}_{1t}, \tilde{\phi}_{1t})} \tag{8.2}$$

would lead to an improved calculation of the free energy [19]. However, we have to check that the numerator and denominator in (8.2) have the same correlation length. This problem can be avoided if we consider an intermediate step. Let us write (5.6) in the form

$$\sum_{\sigma',\tau'} \tilde{\Psi}_{1t}(\sigma,\tau) L(\sigma,\tau|\sigma',\tau') \tilde{\Psi}_{1t}(\sigma',\tau') = \frac{(\tilde{\Psi}_{1t}, L\tilde{\Psi}_{1t})}{(\tilde{\Psi}'_{1t}, \tilde{\Psi}'_{1t})} e^{\tilde{h}'_t(\sigma,\tau)} \tag{8.3}$$

with

$$(\Psi'_{1t}, \Psi'_{1t}) = \sum_{\sigma,\tau} e^{\tilde{h}'_t(\sigma,\tau)}.$$

Now, to the extent that $\tilde{h}'_t(\sigma,\tau)$ is effectively described by $\tilde{h}_t(\sigma,\tau)$ as far as we are concerned with spin configurations having long range correlations, we are led to infer from (8.3) that

$$\frac{(\tilde{\Psi}_{1t}, L\tilde{\Psi}_{1t})}{(\tilde{\Psi}'_{1t}, \tilde{\Psi}'_{1t})} \tag{8.4}$$

could be a more consistent calculation of the free energy, when T is near T_c. Of course, the numerator and the denominator in (8.4) have the same correlation length.

This and the other problems mentioned above, as well as the case of higher dimensions, will be investigated elsewhere.

Acknowledgment

The authors are very grateful to F Cannillo for her special efforts in TEX typesetting.

References

[1] Villani M 1990 *J. Phys. A: Math. Gen.* **23** 4977

[2] Schultz T D, Mattis D C and Lieb E H 1964 *Rev. Mod. Phys.* **36** 856
[3] Kac M 1968 *Mathematical Mechanism of Phase Transitions (Brandeis Lectures 1966)* (New York: Gordon and Breach)
[4] Thompson C J 1972 *Mathematical Statistical Mechanics* (Princeton, NJ: Princeton University Press)
[5] Ruján P 1979 *Physica* **96A** 379
[6] Burley D M 1972 Closed form approximations for lattice systems *Phase Transitions and Critical Phenomena* vol 2, ed C Domb and M S Green (New York: Academic)
[7] An G 1988 *J. Stat. Phys.* **52** 727
[8] Izyumov Yu A and Skryabin Yu N 1988 *Statistical Mechanics of Magnetically Ordered Systems* (New York: Consultant Bureau)
[9] Wolfram Research Inc 1991 *Mathematica* (Champaign, IL)
[10] Capehart T W and Fisher M E 1976 *Phys. Rev.* B **13** 5021
[11] Binder K 1974 *Thin Solid Films* **20** 367
[12] Ma S K 1976 *Modern Theory of Critical Phenomena* (Reading, MA: Benjamin)
[13] Barber M N 1983 Finite-size scaling *Phase Transitions and Critical Phenomena* vol 8, ed C Domb and J L Lebowitz (New York: Academic)
[14] Grosse H 1988 *Models in Statistical Physics and Quantum Field Theory (Trieste Notes in Physics)* (Berlin: Springer)
[15] Aharony A 1976 Dependence of universal critical behaviour on symmetry and range of interaction *Phase Transitions and Critical Phenomena* vol 6, ed C Domb and M S Green (New York: Academic)
[16] Pfeuty P and Toulouse G 1977 *Introduction to the Renormalization Group and Critical Phenomena* (New York: Wiley)
[17] Grover M K 1973 *Phys. Lett.* **44A** 253
[18] Chang T S and Stanley H E 1973 *Phys. Rev.* B **8** 4435
[19] Mattis D C 1985 *The Theory of Magnetism 2* ch 3 (Berlin: Springer)