Alternative formulation of the variational method for systems with an infinite number of degrees of freedom: application to Ising models

1990 J. Phys. A: Math. Gen. 23 4977

(http://iopscience.iop.org/0305-4470/23/21/033)

J. Phys. A: Math. Gen. 23 (1990) 4977-4993. Printed in the UK

Alternative formulation of the variational method for systems with an infinite number of degrees of freedom: application to Ising models

M Villani

Dipartimento di Fisica, Universita' di Bari, Italy and Istituto Nazionale di Fisica Nucleare, Sezione di Bari, Italy

Received 30 April 1990

Abstract. For systems with an infinite number of degrees of freedom, the Rayleigh-Ritz quotient of the standard variational method is analysed from an alternative point of view. Its numerator and denominator are treated as statistical models having the same correlation length. The approach is developed in the framework of the transfer matrix formalism and applied to Ising models. It allows us to obtain exactly the correlation length along a row or column of the anisotropic square Ising model, for $T \geq T_c$. Approximate applications of the method to the two-layer square and 3D simple cubic Ising models are also given. In these cases quite reliable critical curves are obtained through an essentially analytical procedure.

1. Introduction

In quantum mechanics, when we consider a system with a finite number of degrees of freedom, the variational approach provides a powerful method for the calculation of the ground-state energy (E_0). The starting point is the Rayleigh-Ritz (RR) quotient

$$\frac{(\psi, H\psi)}{(\psi, \psi)} \tag{1.1}$$

where H is the Hamiltonian and ψ a trial function with adjustable parameters $\{\alpha_i\}$. The minimum of the RR quotient with respect to the $\{\alpha_i\}$, allows us to obtain the best approximation for E_0. At the same time, we get an approximation for the ground-state wavefunction. However, while we can obtain accurate values for E_0, the properties of the ground-state wavefunction can be poorly described by the trial function ψ, with the parameters $\{\alpha_i\}$ fixed by the minimum condition. This feature of the variational method becomes more relevant when we consider systems with an infinite number of degrees of freedom. In quantum field theory, for example, there is a particular interest in the ground-state wave functional, the actual value of E_0 being subtracted from the spectrum of the Hamiltonian. In this case, the variational principle can be used for the determination of the parameters introduced in a trial functional. Usually these parameters are related to some physical properties of the system. Such a procedure has been analysed by Feynman [1]: it can happen that the minimum of the RR quotient is attained for values of the parameters which are very different from physically reasonable expectations. The origin of this difficulty has been called by Feynman sensitivity of the variational method to high frequencies. In field theory these give the most contribution to the energy.

0305-4470/90/214977+17$03.50

We have a similar situation in statistical mechanics. In this case the quantity which is usually calculated through a variational procedure is the free energy. The starting point can be the Gibbs variational principle [2], from which the mean-field approximation is deduced. On the other hand, if the partition function is expressed through a transfer matrix L, the standard variational method can be applied by considering again the RR quotient [3-5]

$$\frac{(\psi, L\psi)}{(\psi, \psi)} \tag{1.2}$$

which gives a lower bound to the highest eigenvalue of L. Then an approximation to the free energy is obtained through the

$$\sup_{\{\alpha_i\}} \frac{(\psi, L\psi)}{(\psi, \psi)}. \tag{1.3}$$

However, even a good approximation for the free energy can be unreliable for its derivatives, in which we are usually interested (magnetization, susceptibility, heat capacity). This is particularly true in the critical region, where a kind of discrepancy appears. A critical temperature can be deduced from (1.3), but the parameters $\{\alpha_i\}$ determined by the maximum condition lead to a vector ψ devoid of long-range correlations.

With the aim of overcoming this type of difficulty, we propose in this paper a reformulation of the variational procedure. In order to fix the ideas, the analysis will be made in the framework of the Ising models.

In this alternative formulation we make use of the transfer matrix formalism, but we avoid the determination of the parameters $\{\alpha_i\}$ of a trial vector ψ through (1.3). The basic ingredients are again the numerator and the denominator of the RR quotient, but these quantities are treated in a different way. First of all, we consider them as statistical models having as probability distributions

$$P_D(x) = \frac{1}{Z_D} \psi^2(x) \qquad P_N(x, y) = \frac{1}{Z_N} \psi(x) L(x, y) \psi(y) \tag{1.4}$$

where x, y are sets of random variables (the Frobenius-Perron theorem requires that $\psi(x)$ is positive). Then we demand that some relevant statistical quantities are the same both for $P_D(x)$ and $P_N(x, y)$. This requirement is simply a consequence of the property that, when $\psi(x)$ is the eigenvector of L associated to the highest eigenvalue, $P_D(x)$ is the marginal distribution for the set x, obtained from the joint distribution $P_N(x, y)$.

A fundamental quantity of probability distributions in statistical mechanics is the correlation length, which has a crucial role in the scaling hypothesis [2]. Our first condition, which we will call CLE (correlation length equality), is that P_D and $P_N(x, y)$ have the same correlation length. In this paper we report the consequences of this condition.

The details of the above approach are presented in section 2, in the framework of the anisotropic square Ising model. We consider only the region $T \geqslant T_c$. By making use of an ansatz already utilized in the standard variational procedure, we show that our CLE leads to the exact correlation length given by the Onsager solution, for all $T \geqslant T_c$. In section 3 we consider a two-layer Ising film, for which an exact solution is not known. In this case the CLE does not determine completely the correlation length. However, we show that it acts as an existence condition for the phase transition, allowing us to calculate quite reliable lower bounds to the critical point. Furthermore we show that, for T near T_c ($T > T_c$) the formalism can be developed on the basis of the perturbation method and the analysis can be made in an analytical way. The same type of results are obtained in section 4, for a cubic anisotropic three-dimensional Ising model. The analysis of this model makes use of the developments of the previous section. In the last section we discuss several open problems and we make some comments on further conditions to be imposed on $P_D(x)$ and $P_N(x, y)$ and on the possible role of other probability distributions, which are extensions of the numerator of the RR quotient.

2. The anisotropic two-dimensional square Ising model

Let us consider a two-dimensional square lattice with m rows and n columns. The anisotropic Ising Hamiltonian is

$$\mathscr{H} = -J_1 \sum_{i=1}^{m} \sum_{j=1}^{n} s_{i,j}s_{i+1,j} - J_2 \sum_{i=1}^{m} \sum_{j=1}^{n} s_{i,j}s_{i,j+1} \tag{2.1}$$

where $J_i > 0$ and $s_{i,j} = \pm 1$. We take, in general, $J_1 \neq J_2$. In fact such a situation allows us to see better the meaning of our approach. We impose toroidal boundary conditions.

Let $\sigma = (s_1, s_2, \ldots, s_m)$ be a spin configuration of a column and L_2 the symmetrized transfer matrix which connects two adjacent columns. The partition function can be written in the form

$$Z_{m,n} = \mathrm{Tr}\, L_2^n = \sum_{(\sigma)} L_2^n(\sigma, \sigma) \tag{2.2}$$

with

$$L_2(\sigma, \sigma') = \exp\left(\frac{K_1}{2} \sum_{i=1}^{m} s_i s_{i+1}\right) \exp\left(K_2 \sum_{i=1}^{m} s_i s_i'\right) \exp\left(\frac{K_1}{2} \sum_{i=1}^{m} s_i' s_{i+1}'\right) \tag{2.3}$$

($s_i = \pm 1$, $s_i' = \pm 1$, $K_i = J_i/kT$).

We can also consider the transfer matrix L_1 which connects two adjacent rows

$$L_1(\tilde{\sigma}, \tilde{\sigma}') = \exp\left(\frac{K_2}{2} \sum_{i=1}^{n} \tilde{s}_i \tilde{s}_{i+1}\right) \exp\left(K_1 \sum_{i=1}^{n} \tilde{s}_i \tilde{s}_i'\right) \exp\left(\frac{K_2}{2} \sum_{i=1}^{n} \tilde{s}_i' \tilde{s}_{i+1}'\right) \tag{2.4}$$

where $\tilde{\sigma} = (\tilde{s}_1, \ldots, \tilde{s}_n)$ is a spin configuration along a row. We will be interested in the thermodynamic limit $m, n \to \infty$.

Now, let us fix our attention on a given column j and consider the probability $P(\sigma)$ of a spin configuration σ on this column, regardless of the configurations of all other columns.

We have

$$P(\sigma) = \frac{L_2^n(\sigma, \sigma)}{\mathrm{Tr}\, L_2^n}. \tag{2.5}$$

In the limit $n \to \infty$, by making use of the spectral representation of L_2, we obtain [6, 7]

$$P(\sigma) = \frac{\phi_1^2(\sigma)}{\|\phi_1\|^2} \tag{2.6}$$

where $\phi_1(\sigma)$ is the eigenvector of $L_2(\sigma, \sigma')$ associated to the highest eigenvalue ($\|\phi_1\|^2 = \Sigma_{(\sigma)} \phi_1^2(\sigma)$). Equation (2.6) gives a very useful statistical interpretation of $\phi_1(\sigma)$, which will be the starting point of our analysis. In the following we will always take $T > T_c$.

Let ξ_1 be the correlation length of our model, along a column; we denote by ξ_2 the analogous quantity along a row. It is useful to give an explicit expression of ξ_1 and ξ_2 in terms of the transfer matrices L_1, L_2. If $\Lambda_1^{(i)}$ is the maximal eigenvalue of L_i and $\Lambda_2^{(i)}$ the nearest eigenvalue to it, then [7]

$$\frac{1}{\xi_1} = \ln \frac{\Lambda_1^{(1)}}{\Lambda_2^{(1)}} \qquad \frac{1}{\xi_2} = \ln \frac{\Lambda_1^{(2)}}{\Lambda_2^{(2)}}. \tag{2.7}$$

The mathematical mechanism which is responsible for the critical behaviour is the eigenvalues degeneracy [6, 7] ($\Lambda_1^{(2)} \simeq \Lambda_2^{(2)}$, $\Lambda_1^{(1)} \simeq \Lambda_2^{(1)}$). We will fix our attention on these fundamental quantities.

As a trivial consequence of (2.6), we have that the correlation length associated to the distribution $\phi_1^2(\sigma)/\|\phi_1\|^2$ is equal to ξ_1. The first step of our approach is the introduction of a trial distribution $\tilde{P}(\sigma)$ (or, equivalently, of a trial vector $\tilde{\phi}_1(\sigma)$) having the property to reproduce ξ_1 through the ratio of two eigenvalues. We consider the ansatz

$$\tilde{\phi}_1(\sigma) = \exp\left(A(T) \sum_{i=1}^{m} s_i s_{i+1} \right). \tag{2.8}$$

Now, $\|\tilde{\phi}_1\|^2$ is the partition function of a one-dimensional Ising model, to which is associated a 2×2 transfer matrix with eigenvalues

$$\gamma_1 = 2 \cosh 2A(T) \qquad \gamma_2 = 2 \sinh 2A(T). \tag{2.9}$$

Then $\tilde{P}(\sigma) = \tilde{\phi}_1^2(\sigma)/\|\tilde{\phi}_1\|^2$ has the desired property, if $A(T)$ is such that

$$\ln \frac{\gamma_1}{\gamma_2} = \frac{1}{\xi_1} \tag{2.10}$$

that is

$$A(T) = \tfrac{1}{2} \tanh^{-1} \mathrm{e}^{-1/\xi_1}. \tag{2.11}$$

We have again the degeneracy mechanism at the critical point, with

$$A(T) \xrightarrow[T \to T_c]{} +\infty.$$

As the next step, we generalize (2.6). Let $P(\sigma, \sigma')$ be the probability of a spin configuration (σ, σ') on two adjacent columns, regardless of the configurations of all other columns.

We have

$$P(\sigma, \sigma') = \frac{L_2(\sigma, \sigma') L_2^{n-1}(\sigma', \sigma)}{\mathrm{Tr}\, L_2^n}. \tag{2.12}$$

By making use again of the spectral representation of L_2, we obtain ($n \to \infty$)

$$P(\sigma, \sigma') = N \phi_1(\sigma) L_2(\sigma, \sigma') \phi_1(\sigma') \tag{2.13}$$

with

$$N = \frac{1}{\Lambda_1^{(2)} \|\phi_1\|^2}.$$

Of course $P(\sigma)$ is the marginal distribution for the set σ, deduced from the joint distribution $P(\sigma, \sigma')$. So, all the statistical quantities along a column are the same for both of them.

Now, let us introduce our ansatz $\tilde{\phi}_1(\sigma)$ in (2.13). We get a distribution

$$\tilde{P}(\sigma, \sigma') = \tilde{N}\tilde{\phi}_1(\sigma)L_2(\sigma, \sigma')\tilde{\phi}_1(\sigma') \tag{2.14}$$

which, as a rule, does not reproduce exactly $\tilde{P}(\sigma)$ after the summation over the set σ'. However, our ansatz will be consistent if at least the large distances behaviour of the pair correlation function along a column is the same for both $\tilde{P}(\sigma)$ and $\tilde{P}(\sigma, \sigma')$. This requirement, which we have called CLE, leads to an equation for ξ_1.

To the distribution $\tilde{P}(\sigma, \sigma')$ is associated the partition function of a $2\times m$ Ising lattice, with a 4×4 transfer matrix

$$l(s_1, s_2|s_1', s_2') = \exp\left(\frac{K_2}{2} s_1 s_2\right) \exp\left[\left(A(T)+\frac{K_1}{2}\right)(s_1 s_1' + s_2 s_2')\right] \exp\left(\frac{K_2}{2} s_1' s_2'\right). \tag{2.15}$$

Let δ_i $(i = 1, \ldots, 4)$ be the eigenvalues of l, with $\delta_1 > \delta_2 > \ldots$. It is easy to see that

$$\begin{aligned}\delta_1 &= 2\{\cosh K_2 \cosh(K_1+2A(T)) + [1+\sinh^2 K_2 \cosh^2(K_1+2A(T))]^{1/2}\}\\ \delta_2 &= 2\, e^{K_2} \sinh(K_1+2A(T)).\end{aligned} \tag{2.16}$$

The CLE demands that

$$\frac{\gamma_2}{\gamma_1} = \frac{\delta_2}{\delta_1}. \tag{2.17}$$

Explicitly, (2.17) gives

$$\tanh 2A(T) = \frac{e^{K_2} \sinh(K_1+2A(T))}{\cosh K_2 \cosh(K_1+2A(T)) + (1+\sinh^2 K_2 \cosh^2(K_1+2A(T)))^{1/2}}. \tag{2.18}$$

Formally, (2.18) looks like a generalized version of the usual mean-field equation for the order parameter, with the difference that the effective coupling constant $A(T)$ is involved.

Let us introduce the parameter K_1^* [8], through

$$\tanh K_1 = e^{-2K_1^*}.$$

Then, it can be easily checked that (2.18) is solvable only for $K_1^* \geqslant K_2$. For $K_1^* > K_2$, we have only one solution given by

$$\tanh 2A(T) = e^{-2(K_1^* - K_2)}. \tag{2.19}$$

From (2.11) we obtain

$$\frac{1}{\xi_1} = 2(K_1^* - K_2) \qquad (\forall K_1^* > K_2). \tag{2.20}$$

Our expression for ξ_1, agrees with the exact result [8]. The equation of the critical curve is

$$K_1^* - K_2 = 0$$

which is equivalent to the well known exact result

$$\sinh 2K_1 \sinh 2K_2 = 1.$$

The region $K_1^* \geqslant K_2$ corresponds to $T \geqslant T_c$.

It is useful to make some comment about the above procedure. In the standard variational approach, the ansatz $\tilde{\phi}_1(\sigma)$ has been already utilized for the isotropic square Ising model [3, 4]. The parameter $A(T)$ is determined by

$$\sup_{A(T)} \frac{(\tilde{\phi}_1, L_2\tilde{\phi}_1)}{\|\tilde{\phi}_1\|^2} = \sup_{A(T)} \frac{\Sigma_{(\sigma,\sigma')} \tilde{\phi}_1(\sigma) L_2(\sigma, \sigma')\tilde{\phi}_1(\sigma')}{\|\tilde{\phi}_1\|^2}. \tag{2.21}$$

In this way an approximation to the free energy (which is related to $\Lambda_1^{(2)}$) and an approximate value $\bar{T}_c$ of T_c is obtained. The results are good, but do not agree with the exact values. In the critical region it happens that the sup is attained for finite values of $A(T)$ [4]. Essentially we have shown that in the numerator and denominator of the RR quotient are hidden some statistical properties, which have a relevant role. In the usual variational procedure, based on (2.21), only γ_1 and δ_1 are present in the limit $m \to \infty$.

Our function $\tilde{\phi}_1(\sigma)$, with $A(T)$ given by (2.19), is not the exact eigenfunction of $L_2(\sigma, \sigma')$ associated to $\Lambda_1^{(2)}$. The RR quotient with this function gives only an approximation to $\Lambda_1^{(2)}$, which is generally worse than the value given by (2.21). However, $\tilde{\phi}_1(\sigma)$ through $P(\sigma, \sigma')$, allows us to reproduce exactly a relevant quantity associated to the other transfer matrix L_1, that is the ratio $\Lambda_1^{(1)}/\Lambda_2^{(1)}$. Of course, by exchanging the role of L_2 and L_1, we can also obtain exactly ξ_2. In the isotropic case $K_1 = K_2 = K$, we have only one transfer matrix L (if rows or columns are considered) and our $\tilde{\phi}_1(\sigma)$, while it does not reproduce Λ_1, allows us to obtain exactly the ratio Λ_1/Λ_2 of the two highest eigenvalues of L, which, in principle, are associated to two different eigenfunctions of L.

3. A two-layer Ising film

In this section we analyse the role of the CLE for a two-layer Ising film, made of two interacting isotropic square lattices. The Hamiltonian of the model is

$$\mathscr{H}_f = -J_1 \sum_{i=1}^{m} \sum_{j=1}^{n} (s_{i,j}s_{i+1,j} + s_{i,j}s_{i,j+1} + u_{i,j}u_{i+1,j} + u_{i,j}u_{i,j+1})$$
$$- J_2 \sum_{i=1}^{m} \sum_{j=1}^{n} s_{i,j}u_{i,j} \qquad (J_i > 0) \tag{3.1}$$

where J_1 is the «horizontal» isotropic coupling, while J_2 is the vertical coupling. We impose periodic conditions in the two horizontal planes and denote by $(\sigma, \tau) = (s_1, \ldots, s_m; u_1, \ldots, u_m)$ a spin configuration on a vertical section j of the layer. The symmetrized transfer matrix L which connects two adjacent vertical sections is given by

$$L(\sigma, \tau|\sigma, \tau') = \bar{L}(\sigma, \tau) \exp\left[K_1\left(\sum_{i=1}^{m} s_i s_i' + \sum_{i=1}^{m} u_i u_i'\right)\right] \bar{L}(\sigma, \tau') \tag{3.2}$$

with

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

$(K_i = J_i/kT)$.

We proceed as in the previous section, by fixing the attention on a vertical section. The probability $P(\sigma, \tau)$ of a spin configuration on this section is related to the eigenfunction of L associated to the highest eigenvalue. We parametrize $P(\sigma, \tau)$ by introducing

$$\tilde{\chi}_1(\sigma, \tau) = \exp\left[B(T)\left(\sum_{i=1}^{m} s_i s_{i+1} + \sum_{i=1}^{m} u_i u_{i+1}\right) + C(T)\sum_{i=1}^{m} s_i u_i\right] \tag{3.3}$$

so that

$$\tilde{P}_{\mathrm{D}}(\sigma, \tau) = \frac{\tilde{\chi}_1^2(\sigma, \tau)}{(\tilde{\chi}_1, \tilde{\chi}_1)}. \tag{3.4}$$

The trial function $\tilde{\chi}_1(\sigma, \tau)$ is the simplest generalization of (2.8), which takes into account of the vertical interaction through an effective coupling $C(T)$.

To $\tilde{P}_{\mathrm{D}}(\sigma, \tau)$ is associated a transfer matrix which connects two adjacent segments of the vertical section. This matrix has the same structure as the l of (2.15). Its highest eigenvalues, which control the large distance correlations, are given by

$$\begin{aligned} \varepsilon_1 &= 2[\cosh 2C(T) \cosh 4B(T) + (1 + \sinh^2 2C(T) \cosh^2 4B(T))^{1/2}] \\ \varepsilon_2 &= 2\,\mathrm{e}^{2C(T)} \sinh 4B(T) \end{aligned} \tag{3.5}$$

with $\varepsilon_1 > \varepsilon_2$, for finite $B(T)$. We take always $T > T_c$.

According to our previous considerations, we consider also the distribution for two adjacent vertical sections of the layer

$$\tilde{P}_{\mathrm{N}}(\sigma, \tau; \sigma', \tau') = \tilde{N}\tilde{\chi}_1(\sigma, \tau) L(\sigma, \tau; \sigma', \tau') \tilde{\chi}_1(\sigma', \tau') \tag{3.6}$$

To this distribution is associated a 16×16 transfer matrix t, given by

$$\begin{aligned} &t(s_1, s_2, s_3, s_4 | s_1', s_2', s_3', s_4') \\ &\quad = \bar{t}(s_1, s_2, s_3, s_4) \exp\left((B(T) + \tfrac{1}{2}K_1) \sum_{i=1}^{4} s_i s_i'\right) \bar{t}(s_1', s_2', s_3', s_4') \end{aligned} \tag{3.7}$$

with

$$\bar{t}(s_1, s_2, s_3, s_4) = \exp[\tfrac{1}{2}(C(T) + \tfrac{1}{2}K_2)(s_1 s_2 + s_3 s_4) + \tfrac{1}{2}K_1(s_2 s_3 + s_4 s_1)]$$

$(s_i = \pm 1, s_i' = \pm 1)$.

Let us denote by η_k $(k = 1, 2, \ldots, 16)$ the eigenvalues of t, η_1 and η_2 being the two highest, with $\eta_1 > \eta_2$. Our CLE demands that

$$\frac{\varepsilon_2}{\varepsilon_1} = \frac{\eta_2}{\eta_1}. \tag{3.8}$$

In order to handle the eigenvalue problem of the t matrix, it is useful to make use of the Kaufmann formalism [2]. Let us consider the unit 2×2 matrix I and the 2×2 Pauli spin matrices

$$X = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad Z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

We introduce the 16×16 matrices X_i, Z_i $(i = 1, \ldots, 4)$, defined through direct products

$$\begin{aligned} &X_1 = X \otimes I \otimes I \otimes I \qquad X_2 = I \otimes X \otimes I \otimes I \qquad \ldots \\ &Z_1 = Z \otimes I \otimes I \otimes I \qquad Z_2 = I \otimes Z \otimes I \otimes I \qquad \ldots. \end{aligned}$$

Let $|s_1, s_2, s_3, s_4\rangle$ be the orthonormal eigenstates of the Z_i matrices

$$Z_i|s_1, s_2, s_3, s_4\rangle = s_i|s_1, s_2, s_3, s_4\rangle$$

$$\langle s_1', s_2', s_3', s_4'|s_1, s_2, s_3, s_4\rangle = \prod_{i=1}^{4} \delta_{s_i s_i'}.$$

Then, following the method described in [2], we can write

$$t(s_1, s_2, s_3, s_4|s_1', s_2', s_3', s_4') = [2\sinh(2B(T)+K_1)]^2\langle s_1, s_2, s_3, s_4|\tilde{t}|s_1', s_2', s_3', s_4'\rangle \quad (3.9)$$

where

$$\tilde{t} = \tilde{t}_0^{1/2}\, \mathrm{e}^{\theta(T)\Sigma}\, \tilde{t}_0^{1/2} \quad (3.10)$$

with

$$\tanh\theta(T) = \exp(-(2B(T)+K_1)] \quad (3.11)$$

$$\tilde{t}_0^{1/2} = \exp\left[\frac{K_1}{2}(Z_2Z_3+Z_4Z_1)+\tfrac{1}{2}(C(T)+\tfrac{1}{2}K_1)(Z_1Z_2+Z_3Z_4)\right] \quad (3.12)$$

and

$$\Sigma = X_1+X_2+X_3+X_4.$$

From (3.9) it follows that

$$\frac{\eta_2}{\eta_1} = \frac{\tilde{\eta}_2}{\tilde{\eta}_1} \quad (3.13)$$

where $\tilde{\eta}_1, \tilde{\eta}_2$ are the two highest eigenvalues of $\tilde{t}(\tilde{\eta}_1 > \tilde{\eta}_2)$. So, for our problem, we can limit ourselves to the eigenvalue problem of the $\tilde{t}$ matrix. This can be further simplified by making use of parity projection operators [2], which allow us to split $\tilde{t}$ in the sum of two commuting matrices of effective order eight.

We can write

$$\tilde{t} = \tilde{t}_1 + \tilde{t}_2 = P_+\tilde{t}P_+ + P_-\tilde{t}P_- \quad (3.14)$$

where

$$P_+ = \tfrac{1}{2}(I+V) \qquad P_- = \tfrac{1}{2}(I-V)$$

and

$$V = \prod_{i=1}^{4} X_i.$$

These operators have the following well known properties. Let

$$|\Psi(s_1, s_2, s_3, s_4)\rangle = \frac{1}{\sqrt{2}}(|s_1, s_2, s_3, s_4\rangle + |-s_1, -s_2, -s_3, -s_4\rangle)$$

$$|\Phi(s_1, s_2, s_3, s_4)\rangle = \frac{1}{\sqrt{2}}(|s_1, s_2, s_3, s_4\rangle - |-s_1, -s_2, -s_3, -s_4\rangle). \quad (3.15)$$

Then

$$P_+|\Psi(s_1, s_2, s_3, s_4)\rangle = |\Psi(s_1, s_2, s_3, s_4)\rangle$$

$$P_+|\Phi(s_1, s_2, s_3, s_4)\rangle = 0$$

and an analogous equation for P_-.

If

$$C(T)+\tfrac{1}{2}K_2 = K_1 \quad (3.16)$$

then $\tilde{\eta}_1$ and $\tilde{\eta}_2$ could be determined analytically for every T, through the Kaufmann method [2]. But we expect that, in general, the above equation will not be satisfied. For example, when $J_2 = 2J_1$ (which corresponds to periodic conditions in the vertical direction), (3.16) gives $C(T) = 0$. However, within our approach, (3.9) is very useful for T near T_c, a region in which we are particularly interested. In fact, for such values of T, we can calculate $\tilde{\eta}_1$ and $\tilde{\eta}_2$ analytically by making use of the perturbation method.

For $T \to T_c$, there are long-range correlations along our vertical section. Then, according to the meaning of the distribution $\tilde{P}_D(\sigma, \tau)$, we must have

$$\lim_{T \to T_c} \frac{\varepsilon_1}{\varepsilon_2} = +1. \tag{3.17}$$

From (3.5) it follows that the above limit is satisfied only if

$$\lim_{T \to T_c} B(T) = +\infty \tag{3.18}$$

as can be expected. From (3.11) we deduce that

$$\lim_{T \to T_c} \theta(T) = 0 \tag{3.19}$$

and

$$\lim_{T \to T_c} \tilde{t} = \tilde{t}_0. \tag{3.20}$$

Equation (3.12) gives immediately the eigenvectors and the eigenvalues of $\tilde{t}_0$. So, for small $\theta(T)$, that is for T near T_c, $\tilde{\eta}_1$ and $\tilde{\eta}_2$ can be determined through the first terms of an expansion in powers of $\theta(T)$. To this end the decomposition given by (3.14) is useful. The highest eigenvalue of $\tilde{t}_0$, which we call λ_1, is doubly degenerate (we take $J_2 \neq 0$ and $C(T) \neq 0$). When $\theta(T) \neq 0$, we have the splitting of λ_1 into $\tilde{\eta}_1$ and $\tilde{\eta}_2$. However, λ_1 is not degenerate for the two matrices $P_+ \tilde{t}_0 P_+$ and $P_- \tilde{t}_0 P_-$, whose eigenvectors are given by (3.15). So $\tilde{\eta}_1$ and $\tilde{\eta}_2$ can be obtained by applying non-degenerate perturbation theory to $\tilde{t}_1$ and $\tilde{t}_2$ respectively. The details of this calculation are given in the appendix. The splitting of λ_1 appears at the fourth order of the perturbative expansion. We need also the sixth-order terms in order to have an expression for the correlation length.

Let

$$\begin{aligned} \tilde{\eta}_1 &= \lambda_1(1 + c_2(T)\theta^2(T) + c_4(T)\theta^4(T) + c_6(T)\theta^6(T) + O(\theta^8)) \\ \tilde{\eta}_2 &= \lambda_1(1 + d_2(T)\theta^2(T) + d_4(T)\theta^4(T) + d_6(T)\theta^6(T) + O(\theta^8)). \end{aligned} \tag{3.21}$$

Since $c_2(T) = d_2(T)$ (A.13), we have

$$\begin{aligned} \frac{\tilde{\eta}_2}{\tilde{\eta}_1} &= 1 - (c_4(T) - d_4(T))\theta^4(T) \\ &\quad + [d_6(T) - c_6(T) + c_2(T)(c_4(T) - d_4(T))]\theta^6(T) + O(\theta^8) \end{aligned} \tag{3.22}$$

Furthermore (3.5) and (3.11) give

$$\begin{aligned} \frac{\varepsilon_2}{\varepsilon_1} &= 1 - 2\,e^{4K1} \coth 2C(T)\theta^4(T) + \tfrac{8}{3}\,e^{4K_1} \coth 2C(T)\theta^6(T) + O(\theta^8) \\ &= 1 - 2\coth 2C(T)\,e^{-8B(T)} + O(e^{-16B(T)}). \end{aligned} \tag{3.23}$$

So, our correlation length $\xi(T)$, along a vertical section, is given by

$$\frac{1}{\xi(T)}=2\coth 2C(T)\,\mathrm{e}^{-8B(T)} \tag{3.24}$$

at the order of our calculations. It is useful to give the expansion of (3.22) in powers of $\exp(-4B(T))$. Then our CLE ((3.8) and (3.13)) leads to

$$\left(\frac{\tanh 2C(T)}{2}\right)^{1/2}\mathrm{e}^{-2K_1}\Delta(T)\frac{1}{\xi^{1/2}(T)}=c_4(T)-d_4(T)-2\,\mathrm{e}^{4K_1}\coth 2C(T) \tag{3.25}$$

for T near T_c, with

$$\Delta(T)=d_6(T)-c_6(T)+(c_2(T)-\tfrac{4}{3})(c_4(T)-d_4(T)). \tag{3.26}$$

Equation (3.25) gives the correlation length $\xi(T)$ approximately in terms of the original coupling K_1, K_2 of the model and of an additional effective coupling $C(T)$, which describes the vertical interaction between the spin of a vertical section resulting from the sum over all the other configurations. So, if we use the trial function of (3.3), we need a further condition for the complete determination of $\xi(T)$. This problem will be investigated elsewhere. However, as we show here, the CLE has a relevant physical consequence and quite reliable quantitative results can be deduced from (3.25), without the knowledge of $C(T)$.

The explicit expression of $\Delta(T)$ is given by (A.14). It turns out that $\Delta(T)$ is always positive and finite for any finite value of $C(T)$, K_1 and K_2. From (3.25) it follows that we must have

$$c_4(T)-d_4(T)>2\,\mathrm{e}^{4K_1}\coth 2C(T) \tag{3.27}$$

for $T>T_c$, while

$$c_4(T_c)-d_4(T_c)=2\,\mathrm{e}^{4K_{1c}}\coth 2C(T_c) \tag{3.28}$$

gives the critical curve.

It is useful to consider the ratio

$$\rho(T)=\frac{2\,\mathrm{e}^{4K_1}\coth 2C(T)}{c_4(T)-d_4(T)}$$

which, through (A.13) and (A.15), can be written explicitly in the form

$$\rho(T)=\frac{\mathrm{e}^{4K_1}\tanh(K_1+C(T)+\frac{1}{2}K_2)\coth 2C(T)}{\coth(K_1+C(T)+\frac{1}{2}K_2)[\coth 2K_1+\coth(2K_1+2C(T)+K_2)+\coth(2C(T)+K_2)]-2}. \tag{3.29}$$

For fixed T (or K_1, K_2), we consider $\rho(T)$ as a function of the unknown parameter $C(T)$. This function, which is always positive, has the following behaviour

$$\lim_{C(T)\to 0}\rho(Y)=+\infty \qquad \lim_{C(T)\to+\infty}\rho(T)=\mathrm{e}^{4K_1}\tanh 2K_1.$$

Furthermore it has a single finite minimum point $\bar{C}(T)$. Now, it results that, for any T lower than a finite value, which we call T_c^L, $\rho(T)$ is greater than one for any value of $C(T)$. Therefore, as a consequence of (3.25), we obtain the prediction that the disordered phase can exist only for $T>T_c^L$. Hence our procedure suggests that T_c^L is a lower bound of the critical temperature T_c

$$T_c^L\leqslant T_c. \tag{3.30}$$

For $T=T_c^L$, we have that the minimum of $\rho(T)$, as a function of $C(T)$, is equal to one. In other words, T_c^L is such that the equation of the critical curve is satisfied, and, at the same time, $\rho(T)$ is stationary with respect to the variations of the parameter $C(T)$. For $T>T_c^L$, $\rho(T)$ is lower than one only for a range of values of $C(T)$. This range becomes more broad as T increases.

The determination of lower bounds to the critical point is closely related to the problem of existence of a phase transition [9]. Usually, the standard variational methods (among these the mean-field approach) lead to upper bounds to T_c. Therefore it would be relevant to give a rigorous proof of (3.30). We are at present attempting to reach this goal.

It is useful to give the actual values of T_c^L. We state the results in terms of K_{1c} and K_{2c} ($K_{ic}=J_i/kT$). The lower bound to T_c becomes an upper bound to one of these quantities, the other being considered as a free parameter. Let

$$K_{1c}=f(K_{2c})$$

be the equation of the critical curve, with $K_{2c}\in(0,+\infty)$. The previous analysis allows us to calculate a critical curve

$$K_{1c}^A=f^A(K_{2c})$$

with $f^A(K_2)\geqslant f(K_2)$, $\forall K_2$. The function $f^A(K_{2c})$ is reported in figure 1 (upper curve). It is a monotonous decreasing function which reproduces correctly both the limits $K_{2c}\to 0$ and $K_{2c}\to+\infty$. We have

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

where $K_c^{d=2}=0.44068$ is the crittical point of the two-dimensional isotropic Ising model.

For the two-layer Ising film, some numerical results have been reported in the literature [10], in the two cases $K_1=K_2=K$ and $K_2=2K_1=2K$. The numerical calculations, based on series analysis, give

$$K_c=0.312 \qquad \text{(within an error of } \tfrac{1}{2}\%)$$

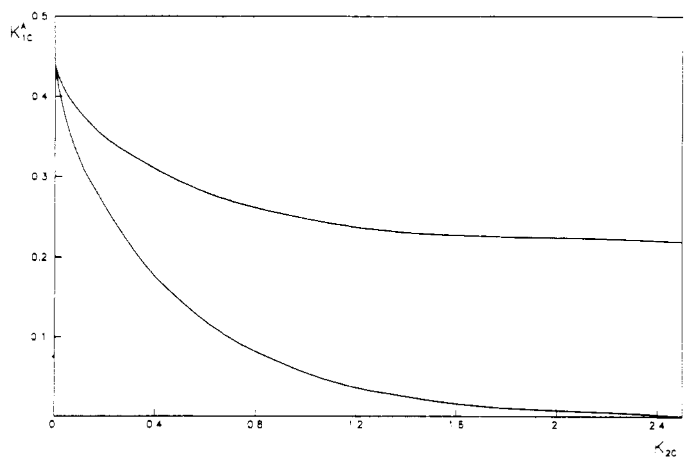

Figure 1. The critical curves $K_{1c}^A=f^A(K_{2c})$ for the two-layer Ising film (upper curve) and $K_{1c}^A=F^A(K_{2c})$ for the anisotropic cubic Ising model (lower curve).

when $K_1 = K_2$. In this case our result is $K_c^A = 0.3235$ ($T_c = 1.037 T_c^L$). For $K_2 = 2K_1$, the value

$$K_c = 0.276$$

is obtained, while we have $K_c^A = 0.2845$ ($T_c = 1.031\ T_c^L$).

The above approach can be extended to an n-layer Ising film. In this case we start from a probability distribution

$$\tilde{P}_D(\sigma_1, \sigma_2, \ldots, \sigma_n)$$

where σ_i denotes a spin configuration along a horizontal line of a vertical section of the layer. Besides $\tilde{P}_D(\sigma_1, \ldots, \sigma_n)$ we have to consider the distribution

$$P_D^{1/2}(\sigma_1, \ldots, \sigma_n) L(\sigma_1, \ldots, \sigma_n | \sigma_1', \ldots, \sigma_n') \tilde{P}_D^{1/2}(\sigma_1', \ldots, \sigma_n')$$

where L is the transfer matrix. If we introduce an effective horizontal coupling $B(T)$ for the spins of a vertical section, then we will have

$$\lim_{T \to T_c} B(T) = +\infty$$

for every finite n. So the analysis can be carried out through the perturbation method if T is near T_c. The Ising films have an important role in our approach, as will be seen in the next section.

4. A simple cubic Ising model

As the last example we consider a cubic Ising model, with an isotropic coupling J_1 in the horizontal planes and a coupling J_2 in the vertical direction. We denote by $[s_{i,j}]$ a spin configuration on a fixed horizontal plane. Following the same procedure of the previous sections we parametrize the probability of a configuration $[s_{i,j}]$, regardless of all the other configurations, by introducing

$$\tilde{\psi}_1([s_{i,j}]) = \exp\left[A(T)\left(\sum_{i=1}^{m}\sum_{i=1}^{n} s_{i,j}s_{i+1,j} + \sum_{i=1}^{m}\sum_{j=1}^{n} s_{i,j}s_{i,j+1}\right)\right] \tag{4.1}$$

so that

$$\tilde{P}_D([s_{i,j}]) = \frac{\tilde{\psi}_1^2([s_{i,j}])}{\|\tilde{\psi}_1\|^2} \tag{4.2}$$

$\|\tilde{\psi}_1\|^2$ is the partition function of an isotropic two-dimensional Ising model, with an effective coupling $2A(T)$, where T is the temperature of our three-dimensional model.

Let $\xi(T)$ be the correlation length associated to two spins which belong to the same plane of the cubic lattice and also to the same row or column of this plane. The parameter $A(T)$ can be related to $\xi(T)$. From (2.20) we have

$$\frac{1}{\xi(T)} = 4(A^*(T) - A(T)) \tag{4.3}$$

where $A^*(T)$ is given by

$$\tanh 2A(T) = e^{-4A^*(T)}.$$

For T near T_c, we can write

$$\frac{1}{\xi(T)} = 8(A(T_c) - A(T)) \tag{4.4}$$

with

$$A(T_c) = \frac{K_c^{d=2}}{2}. \tag{4.5}$$

The transfer matrix, which connects two adjacent vertical planes is

$$L([s_{i,j}]|[s'_{i,j}]) = \bar{L}([s_{i,j}]) \exp\left(K_2 \sum_{i,j} s_{i,j} s'_{i,j}\right) \bar{L}([s'_{i,j}]) \tag{4.6}$$

with

$$\bar{L}([s_{s,j}]) = \exp\left(\frac{K_1}{2} \sum_{i,j} (s_{i,j}s_{i+1,j} + s_{i,j}s_{i,j+1})\right).$$

Then, we consider the distribution

$$\tilde{P}_N([s_{i,j}]; [s'_{i,j}]) = N\tilde{\psi}_1([s_{i,j}])L([s_{i,j}]|[s'_{i,j}])\tilde{\psi}_1([s'_{i,j}]) \tag{4.7}$$

which, as we see, is related to a two-layer Ising film.

In order to calculate $\xi(T)$ through $\tilde{P}_N([s_{i,j}]; [s'_{i,j}])$ we make use of the results of the previous section. Due to equation (4.4) we are left only with the parameter $C(T)$. The critical curve is given by

$$\begin{aligned}\rho(T_c) = &\left[\exp(4A(T_c) + 2K_{1c}) \tanh\left(A(T_c) + \frac{K_{1c}}{2} + \frac{K_{2c}}{2} + \tilde{C}\right) \coth 2\tilde{C}\right] \\ &\times\left\{\coth\left(A(T_c) + \frac{K_{1c}}{2} + \frac{K_{2c}}{2} + \tilde{C}\right)\left[\coth 2\left(A(T_c) + \frac{K_{1c}}{2}\right)\right.\right. \\ &\left.\left. + \coth 2\left(A(T_c) + \frac{K_{1c}}{2} + \frac{K_{2c}}{2} + \tilde{C}\right) + \coth 2\left(\tilde{C} + \frac{K_{2c}}{2}\right)\right] - 2\right\}^{-1} \\ = &\, 1\end{aligned}$$

where $\tilde{C} = C(T_c)$ and $A(T_c)$ is given by (4.5).

As in the previous section we have the prediction that $T_c \geq T_c^L$, with T_c^L given by the condition $\rho(T_c) = 1$ and $\partial\rho(T_c^L)/\partial\tilde{C} = 0$. The critical curve

$$K_{1c}^A = F^A(K_{2c})$$

obtained in this way, is reported in figure 1 (lower curve). Again $F^A(K_{2c})$ gives correctly the limits $K_{2c} \to 0$ and $K_{2c} \to +\infty$. We have

$$\lim_{K_{2c}\to 0} F^A(K_{2c}) = K_c^{d=2} \qquad \lim_{K_{2c}\to\infty} F^A(K_{2c}) = 0.$$

For $K_1 = K_2 = K$, we obtain

$$K_c^A = 0.2419$$

while Monte Carlo calculations and series analysis give $K_c = 0.2217$ [11] ($T_c = 1.091\ T_c^L$). It is useful to compare this result with the lower bound T_{1c} obtained by Frohlich *et al* by making use of the infrared bounds [9]: $T_c = 1.14 T_{1c}$.

5. Conclusion

The previous approach can be extended to four dimensions. It can be seen that, in this case, we are led to consider a four-layer Ising film. For Ising films the CLE is not sufficient to determine completely all the involved effective couplings. In the two-layer case, as we have seen, we need a further condition in order to fix the effective vertical coupling $C(T)$. Of course, besides the correlation length, there are further relevant quantities associated to a probability distribution. Generally we can consider the moments. In statistical mechanics, an important property of a distribution is the susceptibility, which is related to the global behaviour of the pair correlation function through the 'fluctuation-dissipation' theorem [2]. In our approach, it will be useful to analyse the additional condition given by this relevant quantity. A further step can be the consideration of the heat capacity or of the four-points correlation function. Of course, if we have an external magnetic field H, there is a further effective coupling in the distribution $P_D(x)$ and a further condition involving the magnetization.

We can also fix the attention only on the CLE condition. In this case, besides the distribution

$$\psi(x)L(x,y)\psi(y)$$

we can consider the distributions

$$\psi(x)L(x,y)L(y,z)\psi(z),\ \psi(x)L(x,y)L(y,z)L(z,w)\psi(w),\ldots$$

which are generalizations of (2.13). They can act as further constraints on the undetermined parameters. We are led, in this way, to an alternative version of the moment problem in a Hilbert space [12], when the degrees of freedom are infinite.

These problems, as well as the region $T < T_c$, will be analysed in the next works.

Acknowledgments

I would like to thank G Cicuta, V Gorini, A Maritan and M Pellicoro for fruitful discussions.

Appendix

We call $\lambda_i (i = 1, \ldots, 8)$ the eigenvalues of $P_+\tilde{t}_0P_+$. These are also the eigenvalues of $P_-\tilde{t}_0P_-$. The eight eigenvectors of $P_+\tilde{t}_0P_+$ are ordered in the following way

$$|\Psi_1\rangle = |\Psi(1,1,1,1)\rangle \qquad |\Psi_2\rangle = |\Psi(1,1,1,-1)\rangle \qquad |\Psi_3\rangle = |\Psi(1,1,-1,1)\rangle$$
$$|\Psi_4\rangle = |\Psi(1,-1,1,1)\rangle \qquad |\Psi_5\rangle = |\Psi(1,-1,-1,-1)\rangle \qquad |\Psi_6\rangle = |\Psi(1,1,-1,-1)\rangle$$
$$|\Psi_7\rangle = |\Psi(1,-1,-1,1)\rangle \qquad |\Psi_8\rangle = |\Psi(1,-1,1,-1)\rangle \tag{A1}$$

where the $|\Psi(s_1, s_2, s_3, s_4)\rangle$ are given by (3.15). The same ordering defines the eight eigenvectors $|\Phi_i\rangle$ of $P_-\tilde{t}_0P_-$.

The λ_i are given by

$$\lambda_1 = \exp(2K_1 + 2C(T) + K_2) \qquad \lambda_2 = \lambda_3 = \lambda_4 = \lambda_5 = 1$$
$$\lambda_6 = \exp(-2K_1 + 2C(T) + K_2) \qquad \lambda_7 = \exp(2K_1 - 2C(T) - K_2) \tag{A2}$$
$$\lambda_8 = \exp(-2K_1 - 2C(T) - K_2).$$

We make a perturbative expansion of $\tilde{\eta}_1$ and $\tilde{\eta}_2$ to the sixth order of $\theta(T)$. To this end we write firstly

$$e^{\theta(T)\Sigma}=\sum_{k=0}^{6}\frac{1}{k!}\theta^k(T)\,\Sigma^k+O(\theta^7).$$

The action of the powers of Σ on the vectors $|\Psi_i\rangle$ and $|\Phi_i\rangle$, follows from the basic relations

$$X_1|s_1,s_2,s_3,s_4\rangle=|-s_1,s_2,s_3,s_4\rangle \qquad X_2|s_1,s_2,s_3,s_4\rangle=|s_1,-s_2,s_3,s_4\rangle.\ldots \tag{A3}$$

It turns out that

$$\Sigma^3=16\Sigma \tag{A4}$$

if the vectors $|\Psi_i\rangle$ are involved, while

$$\Sigma^2=4I \tag{A5}$$

when we consider the vectors $|\Phi_i\rangle$. As a consequence of these relations, we can write

$$\begin{aligned}\tilde{t}_1&=P_+\tilde{t}_0P_++P_+V_1P_+\\ \tilde{t}_2&=P_-\tilde{t}_0P_-+P_-V_2P_-\end{aligned} \tag{A6}$$

where the perturbations are given by

$$\begin{aligned}V_1&=\tilde{t}_0^{1/2}[(\theta+\tfrac{8}{3}\theta^3+\tfrac{32}{15}\theta^5)\Sigma+(\tfrac{1}{2}\theta^2+\tfrac{2}{3}\theta^4+\tfrac{16}{45}\theta^6)\Sigma^2]\tilde{t}_0^{1/2}\\ V_2&=\tilde{t}_0^{1/2}[(\theta+\tfrac{2}{3}\theta^3+\tfrac{2}{15}\theta^5)\Sigma+2\theta^2+\tfrac{2}{3}\theta^4+\tfrac{4}{45}\theta^6]\tilde{t}_0^{1/2}.\end{aligned} \tag{A7}$$

Let

$$G_1=\sum_{i=2}^{8}\frac{|\Psi_i\rangle\langle\Psi_i|}{\tilde{\eta}_1-\lambda_i} \qquad G_2=\sum_{i=2}^{8}\frac{|\Phi_i\rangle\langle\Phi_i|}{\tilde{\eta}_2-\lambda_i} \tag{A8}$$

Then

$$\begin{aligned}\tilde{\eta}_1&=\lambda_1+\sum_{k=1}^{6}\langle\Psi_1|(V_1G_1)^{k-1}V_1|\Psi_1\rangle+\ldots.\\ \tilde{\eta}_2&=\lambda_1+\sum_{k=1}^{6}\langle\Phi_1|(V_2G_2)^{k-1}V_2|\Phi_1\rangle+\ldots.\end{aligned} \tag{A9}$$

In the calculations of these matrix elements we keep only terms up to the sixth order of $\theta(T)$. The calculations are straightforward. We have

$$\begin{aligned}&\langle\Psi_1|V_1|\Psi_1\rangle=\lambda_1(2\theta^2+\tfrac{8}{3}\theta^4+\tfrac{64}{45}\theta^6)\\ &\langle\Psi_1|V_1G_1V_1|\Psi_1\rangle=4\lambda_1[\gamma_1\theta^2+(\tfrac{16}{3}\gamma_1+\alpha_1)\theta^4+(\tfrac{512}{45}\gamma_1+\tfrac{8}{3}\alpha_1)\theta^6]\\ &\langle\Psi_1|(V_1G_1)^2V_1|\Psi_1\rangle=32\lambda_1\{[2\gamma_1^2+\gamma_1\alpha_1]\tfrac{1}{2}\theta^4+[\tfrac{10}{3}(2\gamma_1^2+\gamma_1\alpha_1)+\tfrac{1}{4}\alpha_1^2]\theta^6\}\\ &\rangle\Psi_1|(V_1G_1)^3V_1|\Psi_1\rangle=16\lambda_1\gamma_1[\gamma_1\alpha_1\theta^4+(16\gamma_1^2+3\alpha_1^2+\tfrac{56}{3}\gamma_1\alpha_1)\theta^6]\\ &\langle\Psi_1|(V_1G_1)^4V_1|\Psi_1\rangle=32\lambda_1\gamma_1^2\alpha_1(3\alpha_1+8\gamma_1)\theta^6\\ &\langle\Psi_1|(V_1G_1)^5V_1|\Psi_1\rangle=64\lambda_1\gamma_1^3\alpha_1^2\theta^6\end{aligned} \tag{A10}$$

and

$$\begin{aligned}
\langle\Phi_1|V_2|\Phi_1\rangle &= \lambda_1(2\theta^2+\tfrac{2}{3}\theta^4+\tfrac{4}{45}\theta^6)\\
\langle\Phi_1|V_2G_2V_2|\Phi_1\rangle &= 4\lambda_1\gamma_2(\theta^2+\tfrac{4}{3}\theta^4+\tfrac{32}{45}\theta^6)\\
\langle\Phi_1|(V_2G_2)^2V_2|\Phi_1\rangle &= 8\lambda_1\gamma_2^2(\theta^4+\tfrac{5}{3}\theta^6)\\
\langle\Phi_1|(V_2G_2)^3V_2|\Phi_1\rangle &= 16\lambda_1\gamma_2^3\theta^6\\
\langle\Phi_1|(V_2G_2)^4V_2|\Phi_1\rangle &= O(\theta^8)\\
\langle\Phi_1|(V_2G_2)^5V_2|\Phi_1\rangle &= O(\theta^8)
\end{aligned}\qquad\text{(A11)}$$

where

$$\alpha_1=\sum_{i=6}^{8}\frac{\lambda_i}{\tilde\eta_1-\lambda_i}\qquad \gamma_\nu=\frac{1}{\tilde\eta_\nu-1}\qquad(\nu=1,2).$$

Now, if $\tilde\eta_1$ and $\tilde\eta_2$ are written as in (3.21), by expanding α_1 and γ_ν in powers of θ, we obtain from (A9) the coefficients c_i, d_i. Let

$$\alpha=\sum_{i=6}^{8}\frac{\lambda_i}{\lambda_1-\lambda_i}\qquad \beta=\lambda_1\sum_{i=6}^{8}\frac{\lambda_i}{(\lambda_1-\lambda_i)^2}.\qquad\text{(A12)}$$

We have

$$\begin{aligned}
c_2=d_2&=2\frac{\lambda_1+1}{\lambda_1-1}\\
c_4-d_4&=2\left[1+\frac{8}{\lambda_1-1}+\frac{12}{(\lambda_1-1)^2}+2\left(\frac{\lambda_1+1}{\lambda_1-1}\right)^2\alpha\right]
\end{aligned}\qquad\text{A13)}$$

$$\begin{aligned}
&d_6-c_6+(c_2-\tfrac{4}{3})(c_4-d_4)\\
&\quad\equiv\Delta=\frac{8}{(\lambda_1-1)^4}\{(\lambda_1^2-1)(\lambda_1+1)^2(\beta-\alpha^2)+(\lambda_1+1)^2\\
&\quad\times[6-(\lambda_1-1)^2]\alpha+2(\lambda_1^3+6\lambda_1^2+8\lambda_1+3)\}.
\end{aligned}\qquad\text{(A14)}$$

Our α and β can be written in the form

$$\begin{aligned}
\alpha&=\tfrac{1}{2}[\coth 2K_1+\coth(2K_1+2C(T)+K_2)+\coth(2C(T)+K_2)-3]\\
\beta&=\frac{1}{4}\left[\frac{1}{\sinh^2 2K_1}+\frac{1}{\sinh^2(2K_1+2C(T)+K_2)}+\frac{1}{\sinh^2(2C(T)+K_2)}\right].
\end{aligned}\qquad\text{(A15)}$$

References

[1] Feynman R P 1988 Difficulties in applying the variational principle to quantum field theory *Proc. Int. Workshop on Variational Calculations in Quantum Field Theory* ed L Polley del Pottinger (Singapore: World Scientific)

[2] Huang K 1987 *Statistical Mechanics* 2nd edn (New York: Wiley)

[3] Munster A 1974 *Statistical Thermodynamics* vol II (Berlin: Springer)

[4] Ruján 1979 *Physica* **96A** 379

[5] Tsang S K 1979 *J. Stat. Phys.* **20** 95

[6] Kac M 1968 *Mathematical Mechanism of Phase Transitions* (*Brandeis Lectures 1966*) New York: Gordon-Breach)

[7] Thompson C J 1972 *Mathematical Statistical Mechanics* (Princeton: Princeton University Press)

[8] Izyumov Yu A and Skryabin Yu N 1988 *Statistical Mechanics of Magnetically Ordered Systems* (New York: Consultant Bureau)

[9] Grosse H 1988 Model in statistical physics and quantum field theory *Trieste Notes in Physics* (Berlin: Springer)

[10] Fisher M E 1971 The theory of critical point singularities *Proc. Int. School of Physics E Fermi Course LI* ed M S Green (New York: Academic)
Capehart T W and Fisher M E 1976 *Phys. Rev.* B **13** 5021

[11] Wallace D J, Pawley A S, Swendson R and Wilson K G 1984 *Phys. Rev.* B **29** 4030
Adler J 1983 *J. Phys. A: Math. Gen.* **16** 3585

[12] Angelini L, Pellicoro M and Villani M 1988 *Phys. Rev.* D **38** 1183