Saddle-point mean-field theory for long-range Ising spin models in terms of the eigenvalues and eigenvectors of the interaction matrix

1992 J. Phys. A: Math. Gen. 25 4723

(http://iopscience.iop.org/0305-4470/25/18/008)

J. Phys. A: Math. Gen. 25 (1992) 4723–4735. Printed in the UK

Saddle-point mean-field theory for long-range Ising spin models in terms of the eigenvalues and eigenvectors of the interaction matrix

Andrew Canning

Département de Physique Théorique, Université de Genève, CH-1211 Genève 4, Switzerland

Received 26 June 1991, in final form 27 March 1992

Abstract. The mean-field saddle-point theory for Ising spin models with long-range interactions is rewritten in terms of the eigenvalues and eigenvectors of the interaction matrix. This gives a natural division of long-range models into two classes: those where the rank of the interaction matrix is finite (in the limit $N \to \infty$); and the saddle-point integral for the partition function can be directly evaluated (subject to several other weaker conditions) and those where the interaction matrix has a divergent rank. In the latter case the saddle-point integral method cannot be directly applied to the partition function. In the former case of the simply solvable models there are natural order parameters associated with the eigenvectors of the non-zero eigenvalues which characterize the system. It will be shown that models in this class are also described by Curie–Weiss mean-field type equations $\langle S_i \rangle = \tanh \beta(\sum_j J_{ij}\langle S_j \rangle)$. This class of systems, where the interaction matrix has a finite rank, is very large and includes: (1) long-range ferromagnetic bond disorder model; (2) systems with Kac type interactions of the form $\gamma \exp(-\gamma|i-j|)$ (where $\gamma \to 0$ as $N \to \infty$); (3) separable random site spin glass models. These three types of models will be solved and interpreted in terms of the behaviour of the eigenvalues and eigenvectors of their interaction matrices. In the case of the site disorder spin glass models the van Hemmen model will be used to illustrate our approach and a simple general solution of the disordered phase of this model will also be presented.

1. Introduction

Ising spin models with long-range interactions are usually divided into two classes: separable and non-separable. These two classifications are usually used in relation to spin glass models (Benamira *et al* 1985) but can easily be extended to cover any long-range Ising spin model. In fact one of the purposes of this paper is to give a different, and very simple, mathematical description of long-range Ising spin models in terms of the rank of the interaction matrix. This very naturally divides them into two classes. These two classes can be associated with the terms separable and non-separable.

Separable models are systems that are considered to be exactly solvable (normally without recourse to a replica type calculation) and are described by a finite set of order parameters having a finite number of stable states. Solving these models typically involves rewriting the Hamitonian in the following way

$$H = -\sum_{ij} S_i J_{ij} Sj = H(\sigma_1, \sigma_2, \ldots, \sigma_p) = \sum_{k=1}^{p} \sigma_k^2 \tag{1}$$

0305-4470/92/184723+13$04.50

where p is finite and the new spin variables σ_k are linear combinations of the original spins, $\sigma_k = \sum_i a_k^i S_i$. Thus the Hamiltonian is an additively separable function in terms of the new variables σ hence the use of the name 'separable' to describe these systems. These models can then be solved by applying the Gaussian transformation to linearize the spin terms (allowing the spin trace to be performed) and then the saddle-point integral method can be used to evaluate the partition function in the thermodynamic limit. Some simple examples of separable site disorder spin glasses can be found in Grensing and Kühn (1987) and Choy and Sherrington (1984). In the next section of this paper we will show that a natural extension of the definition of models in this class are systems whose interaction matrix is of finite rank in the thermodynamic limit. The simplest example of a model of this type is the long-range ferromagnet studied by Kac (1968) where all the interactions take the same value $J_{ij} = J_0/N$, $\forall i, j \; i \neq j$, where N is the system size. The interaction matrix for this model has rank 1 and the model is characterized by only one order parameter, the magnetization, and only has two possible stable states at low temperature.

Non-separable models are models which cannot be described by a finite set of order parameters. Where solvable they are typically described in terms of order parameter functions or an infinite set of order parameter type equations. In the next section we will show that systems of this type have an interaction matrix with a diverging rank in the thermodynamic limit. The SK spin glass (Sherrington and Kirkpatrick 1975) is the best known model of this type (see Mezard *et al* (1986) for a review containing the most important papers on this model) where, in what is considered to be the true solution of this model given by Parisi (Mezrad *et al* 1986), the spin glass order parameter q is an order parameter function. For this model the TAP (Mezard *et al* 1986, Thouless *et al* 1977) equations, although different, are the analogy of the Curie–Weiss mean-field equations for separable models. The TAP equations cannot be reformulated in terms of a finite number of order parameter type equations like the separable models. It should be noted that there are other more easily solvable non-separable models of the site disorder type which were studied by van Hemmen *et al* (1986) and have order parameter functions as solutions.

In the next section of this paper we will present a generalization of the calculation used to solve the so called separable models. In doing this we will develop a set of conditions on the validity of the application of the saddle-point method to our resulting partition function which is related to the rank of the original interaction matrix. We will then show that all models having an interaction matrix of finite rank are described by the Curie–Weiss mean-field equations. In sections 3 and 4 we will study some examples of models with an interaction matrix of finite rank. More specifically in section 4 we will study and solve a long-range ferromagnetic bond disorder model and in section 4 we will study a one-dimensional ferromagnetic model with a Kac type exponential interaction of the form $J_{ij} = \gamma \exp(-\gamma|i-j|)$ in the limit $\gamma \to 0$ as $N \to \infty$. In section 5 we will look at the class of separable random site spin glass models in the light of our mean-field calculation using the van Hemmen model (van Hemmen *et al* 1986, Choy and Sherrington 1984) to illustrate the important points.

2. Generalized mean-field theory

Consider a system of N Ising spins S_i with no external field defined by the

Hamiltonian

$$H = -\tfrac{1}{2}\sum_{i,j} S_i J_{ij} S_j \tag{2}$$

the sum being over all values of i and j. As yet, we make no assumptions about the form of the interactions. Thus, with the appropriate choice of the interaction matrix **J**, this system can describe short-range, long-range or even disordered models. Orthogonally diagonalizing the interaction matrix **J** we can write the Hamiltonian as

$$H = -\tfrac{1}{2}\sum_{q=1}^{N} \lambda_q \left(\sum_i V_q^i S_i\right)^2 \tag{3}$$

where λ_q, $(q = 1, 2, \ldots, N)$ are the eigenvectors of **J**, with corresponding normalized eigenvectors $\boldsymbol{V}_q$. Writing the Hamiltonian like this we can see that all Hamiltonians are in a sense separable. The important property for the system to be simply solvable is that the number of separable terms must be finite. Using standard techniques the Gaussian transformation

$$\exp(a^2) = \frac{1}{\sqrt{2\pi}} \int \exp\left(-\tfrac{1}{2}y^2 + \sqrt{2}ay\right) \mathrm{d}y \tag{4}$$

can now be used to introduce variables x_q associated with each eigenvector giving

$$Z = \mathrm{Tr}\, S_i \frac{1}{(2\pi)^{N/2}} \int \prod_q \mathrm{d}x_q \exp\left[\frac{1}{2}\sum_q \left(-x_q^2 + x_q\sqrt{\beta\lambda_q}\sum_i V_q^i S_i\right)\right]. \tag{5}$$

The spins are now decoupled and the trace over the spins can be performed giving, after rescaling $(x \to \sqrt{N\beta}x)$,

$$Z = \left(\frac{\beta N}{2\pi}\right)^{N/2} \int \prod_i \mathrm{d}x_i \times \exp -N\left(\frac{1}{2}\beta\sum_i x_i^2 - \frac{1}{N}\ln\left[\prod_i 2\cosh\beta\sqrt{N}\sum_q x_q\sqrt{\lambda_q}V_q^i\right]\right). \tag{6}$$

The normal way to proceed in mean-field problems is now to try and apply the saddle-point method which for our purposes can be expressed as

$$\int \prod_\nu \mathrm{d}x_\nu \exp(Ng(\boldsymbol{x})) \simeq \sum_n C_n \exp(Ng(\boldsymbol{x}_n^*)) \tag{7}$$

in the limit $N \to \infty$ where $\boldsymbol{x}_n^*$ are the absolute maxima of the function $g(\boldsymbol{x})$ and C_n are finite constants. For the application of the saddle-point method to be valid, the function $g(\boldsymbol{x})$ must have a well defined functional form in the limit of large N. In particular, if $g(\boldsymbol{x})$ is a function of a fixed finite number of variables x_n, in the limit $N \to \infty$, and its functional form becomes independent of N in this limit then the saddle-point method can be applied to the integral. We can now develop a set of conditions on the interaction matrix **J** such that the saddle-point method can be applied to our integral equation 6. These are:

(i) the number of non-zero eigenvalues (the rank $R(\mathbf{J})$) must be finite and independent of N in the large N limit;

(ii) the values of the set of non-zero eigenvalues $\lambda_\nu, \nu = 1, \ldots, s$, s finite, must be themselves finite and independent of N in the large N limit;

(iii) the elements of the eigenvectors V_q, associated with the non-zero eigenvalues, must be such that the second term in the exponent (the $(1/N)\ln[\,]$ term) has a well defined limit independent of N in the large N limit.

In general the first condition will only be true when interactions depend explicitly on N. This condition allows all the $N - s$ variables in the integral corresponding to zero eigenvalues to be explicitly integrated out since they are now all simple Gaussian integrals. The remaining integral to be evaluated is now over a finite number of variables x_ν, $\nu = 1, \ldots, s$. Conditions (ii) and (iii) guarantee that the exponent in the integral has the desired functional form independent of N in the large N limit.

It is condition (i) which excludes this calculation from being valid for short-range models and also some random long-range models such as the SK spin glass (Sherrington and Kirkpatrick 1975, 1978) or the Hopfield neural network model (Amit *et al* 1985). In all of these systems the number of non-zero finite eigenvalues scales with the system size (Edwards and Jones (1976) and Opper (1989) for the eigenvalue spectra of the interaction matrix of the SK spin glass and the Hopfield neural network).

In the case of solvable long-range interaction models, which we will study in this paper, it is usually necessary to choose a non-zero value for the diagonal terms J_{ii} in order to satisfy condition (i). These diagonal elements have to be chosen to preserve the symmetry of the problem (see examples in the next sections) and can be absorbed into the Hamiltonian by the subtraction of a constant equal to their sum. This constant does not, in general, contribute to the free energy per site and does not affect the thermodynamics of the system.

Condition (ii) typically means that the interactions must be chosen to scale inversely with some power of N. We will see this explicitly with the examples studied in this paper. This condition is also important from a physical point of view so that we have a finite free energy per site in the thermodynamic limit. Condition (iii) is included for completeness and its physical interpretation is unclear. The $(1/N)\ln[\,]$ term in the integral is an infinite sum which must be convergent in the thermodynamic limit. We are not aware of any physically interesting choices of the interaction matrix which satisfy conditions (i) and (ii) and not condition (iii) but it may be possible to choose a pathological matrix satisfying conditions (i) and (ii) but not (iii).

Assuming conditions (i) to (iii) are met for a given interaction matrix we can apply the saddle-point method to the integral in equation (6) (after the $N - s$ other variables have been integrated out) and we can associate the function in the exponent with the free energy of the system. The free energy per site is thus given by

$$f = \frac{1}{2}\beta \sum_q \lambda_q m_q^2 - \frac{1}{N}\left[\prod_i 2\cosh \beta\sqrt{N} \sum_q m_q \lambda_q V_q^i\right] \tag{8}$$

the sum $\sum_q$ now being over the finite set of variables corresponding to the non-zero eigenvalues. We have defined a new set of variables $m_q = x_q/\sqrt{\lambda_q}$. The values of m_q in the free energy corresponding to thermal equilibrium are those which minimize it (this corresponds to evaluating the integral by the saddle-point approximation) and

these are given by

$$m_q = \frac{1}{\sqrt{N}} \sum_i V_q^i \tanh \left[\beta \sqrt{N} \sum_k m_k \lambda_k V_k^i \right]. \tag{9}$$

The free energy barriers between minima scale as N so in the thermodynamic limit all minima (global and local) of the free energy function will correspond to stable states (in the sense that their lifetimes diverge with N).

To obtain a physical interpretation of the parameters m_q we add external fields $h_q^i = h_q V_q^i$ at all sites which favour condensation into one of the minima in the free energy surface. It can then be shown, by applying the saddle-point approximation again, that the physical interpretation of m_q are order parameters given by

$$m_q = \lim_{h_q \to 0} \frac{\partial \ln Z}{\partial h_q} = \frac{1}{\sqrt{N}} \sum_i V_q^i \langle S_i \rangle \qquad (m_{q>s} = 0). \tag{10}$$

This calculation also shows that there is no ordering associated with the eigenvectors of zero or negative eigenvalues . Thus, we now have a finite set of order parameters with which to describe the different stable states of the system. Equations (9) and (10) are equivalent to the Curie–Weiss mean-field equations

$$\langle S_i \rangle = \tanh \beta \left(\sum_j J_{ij} \langle S_j \rangle \right) \tag{11}$$

and therefore we can think of conditions (i) to (iii) as sufficient conditions for the Curie–Weiss equations to be valid for a given model. It is also possible to arrive at this result by deriving TAP (Thouless *et al* 1977, Mezard *et al* 1967) type equations for a given model and showing that the Onsager reaction term is negligible. This approach would not directly bring out the role of the eigenvalue spectrum of the interaction matrix in the behaviour of the model but it is clear that the presence of the Onsager reaction term in the TAP equations can be associated with the divergence of the rank of the interaction matrix.

In order to interpret many of the results already published on models satisfying conditions (i) to (iii) it is important to notice that there are two other ways in which the interaction matrix **J** can be diagonalized. These two other diagonalized forms of the interaction matrix also allow the application of the Gaussian transformation and the saddle-point technique to the partition function but result in different order parameters. Writing **J** in the non-orthogonal diagonalized form

$$\mathbf{J} = \mathbf{B}\hat{\mathbf{A}}\mathbf{B}^{\mathrm{T}} \tag{12}$$

where $\hat{\mathbf{A}}$ is a diagonal matrix which is not constructed from the eigenvalues of **J**. This type of general diagonalization can be performed using a singular or non-singular matrix **B** which is not unique (even when the eigenvalues are all distinct). When **B** is non-singular the rank of **J** is equal to the rank of **A** and when **B** is singular the rank of **A** is greater than that of **J**. As in the previous calculation we will only be

interested in the case where the rank of **A** is finite. Using a diagonalization of this type we can rewrite the Hamiltonian as

$$H = -\frac{1}{2}\sum_l a_l \left(\sum_i b_l^i S_i\right)^2 \tag{13}$$

where a_l are the non-zero diagonal elements of $\hat{\mathbf{A}}$ with corresponding N component vectors b_l in **B**. We can now apply the Gaussian transformation to each of the a_l terms in the sum and go through the same type of calculation as we performed for the orthogonally diagonalized interaction matrix. This will give us order parameter type mean-field equations of the form

$$y_n = \frac{1}{\sqrt{N}}\sum_i b_n^i \tanh\left[\beta\sqrt{N}\sum_l y_l a_l b_l^i\right] = \frac{1}{\sqrt{N}}\sum_i b_n^i \langle S_i\rangle. \tag{14}$$

The variables y_n will be linearly independent when **B** is non-singular and will not be linearly independent when **B** is singular. This set of equations is totally equivalent to the orthogonal set of order parameter equations (9) and (10) and also the Curie–Weiss equations under the appropriate linear transformations. Thus there is no unique set of order parameters which describe the thermodynamics of the system although one might consider the orthogonal parameter equations the easiest to interpret and solve. Solving the orthogonal order parameter equations (9) to first order in m_q gives the well known result for Curie–Weiss mean-field theory that the critical temperature is equal to the largest eigenvalue. At this temperature the system condenses into a spin state, associated with the eigenvector of the largest eigenvalue, with a mean-field critical exponent of one-half.

We will see in section 5 that the solutions in the literature of the separable site disorder spin glass models can, in general, be interpreted as a diagonalization of the interaction matrix of one of the three types we have mentioned, i.e. orthogonal diagonalization or diagonalization with a singular or non-singular matrix. For the van Hemmen model in particular, we will show explicit forms of these three types of diagonalization which have been used to solve it in the literature.

3. Long-range ferromagnetic bond disorder models

In this simple example we will look at a ferromagnetic bond disorder model which satisfies conditions (i) to (iii). Bond disorder models are systems where each bond is chosen from some probability distribution and should be distinguished from site disorder problems where the interactions are calculated from random vectors sitting on each site (whose components are chosen from a probability distribution). Site disorder models will be studied in more detail in section 5. The model we are going to consider here has bonds given by the probability distribution

$$P(J_{ij}) = (1-c)\delta(J_{ij}) + c\delta\left(J_{ij} - \frac{1}{N}\right) \tag{15}$$

so that c is the concentration of bonds. We will study the case where c is finite so that we are not in the regime where we expect to see percolation phenomena since

the percolation threshold for infinite range models is $c_c = 0$ (see Stinchcombe (1983) for a review with references of percolation phenomena in low-dimensional models).

The semicircular eigenvalue spectrum for a large square matrix of elements determined from a probability distribution with fixed variance (all elements having the same variance) and mean zero was first determined by Wigner (1967) (for a derivation of this law see Mehta (1967)). Edwards and Jones (1976) then used a replica type calculation to rederive Wigner's result and also to extend it to the case where the mean is non-zero. They performed their calculation for a Gaussian distribution and found that the eigenvalue spectrum only depends on the first two moments of the distribution. Their results can easily be shown to be valid for any probability distribution with fixed variance and mean. They found that for a probability distribution with mean M/N and variance J^2/N in the limit $N \to \infty$ the eigenvalue spectrum of the corresponding $N \times N$ matrix is given by

$$\rho(\lambda) = \begin{cases} \rho_0(\lambda) + 1/N\,\delta[\lambda - (M + J^2/M)] & |M| > J \\ \rho_0(\lambda) & |M| < J \end{cases} \tag{16}$$

where

$$\rho_0(\lambda) = \begin{cases} (4J^2 - \lambda^2)^{1/2}/2\pi J^2 & |\lambda| < 2J \\ 0 & |\lambda| > 2J. \end{cases} \tag{17}$$

Thus the effect of the non-zero mean introduces a single eigenvalue which splits away from the semicircular continuum when $|M| > J$.

The mean of our probability distribution given by equation 15 is c/N and the variance is $c(1-c)/N^2$. This gives, in the notation used for the eigenvalue spectrum, $M = c$ and $J^2 = c(1-c)/N$. This means that all the eigenvalues lying in the semicircle are of order $1/\sqrt{N}$ and in the thermodynamic limit we have only one non-zero eigenvalue; the one that is not part of the semicircle and is given by $\lambda_m = c$. Thus this gives a quantitatively different (and very simple) system compared to models where the variance and mean are both of order $1/N$ such as the SK spin glass model (Sherrington and Kirkpatrick 1975, 1978). It is easy to show that to corrections of order $1/N$ the normalized eigenvector associated with λ_m is $V_m^i = 1/\sqrt{N}$. This means that the system has only one order parameter which is the magnetization whose value and physical interpretation are given by equations (9) and (10) i.e.

$$m = \tanh \beta c m \qquad \left(= \frac{1}{N} \sum_i \langle S_i \rangle \right). \tag{18}$$

Therefore the system undergoes a ferromagnetic phase transition at $kT_c = c$. This model behaves the same (at finite temperature) as a non-diluted ferromagnetic model with all the interactions given by $J_{ij} = c/N \ \forall i, j$. This non-diluted model has the same non-zero eigenvalue with associated eigenvector as our diluted model. It should be noted that at zero temperature the semicircle of $N - 1$ eigenvalues of order $1/\sqrt{N}$ can contribute to the saddle-point integral so that we may expect a large number of metastable states, possibly exponential in N (like the SK spin glass), at zero temperature. These metastable states will be separated by energy barriers of order $\sqrt{N}$ which are removed at finite temperature. This is in contrast to the SK model where the energy barriers are of order N and survive at finite temperature.

Finally it should be noted that any bond disorder system whose mean is of order $1/N$ and variance of order $1/N^2$ will behave in the same way as this model. Another example of a probability distribution having the first two moments of these orders is the folded Gaussian distribution where only the positive half of a Gaussian distribution of zero mean defines the interaction strengths.

4. One-dimensional model with interactions $J_{ij} = \gamma \exp(-\gamma|i-j|)$, $\gamma \to 0$

Kac (1968) and Baker (1961, 1962) studied a one-dimensional model of this type and showed that in the limit ($\gamma \to 0$) this model has a mean-field type phase transition. Kac showed this property by looking at the asymptotic degeneracy of a certain integral equation in this limit. We will now look again at this model in terms of the eigenvalue spectrum of the interaction matrix. Our calculation will give us no information about the case $\gamma \neq 0$ but will show that the model is described by Curie–Weiss mean-field theory in the limit $\gamma \to 0$

Considering a model with one-dimensional translational invariance we can write down the interactions as $J(r) = \gamma \exp(-\gamma r)$ where $r = |i-j|$ and $J(r) = J(N-r)$. Working with Fourier components this gives us, for the eigenvalues of the interaction matrix,

$$\lambda_q = \sum_r J(r) \exp 2\pi i r q \tag{19}$$

where q are the reciprocal lattice vectors

$$q = 0, \frac{1}{N}, \frac{2}{N}, \ldots, \frac{N-1}{N} \qquad \text{and} \qquad r = 0, 1, \ldots, N-1.$$

In Kac's calculation he first takes the limit $N \to \infty$ and then later the limit $\gamma \to 0$. These two limits result in a matrix of finite rank but to apply the saddle-point method to the integral we have to perform the calculation in a more restrictive way by forcing $\gamma \to 0$ at the same time as $N \to \infty$. Applying the limit $N \to \infty$ alone to the interaction matrix gives of order N non-zero eigenvalues so we cannot directly apply the saddle-point method to evaluate the partition function as condition (i) is violated. If we choose $\gamma \to 0$ such that, as $N \to \infty$, $\gamma N \to \infty$ then the interaction matrix only has one non-zero eigenvalue given by

$$\lambda_0 \neq 0 \ (\lambda_{q \neq 0} = 0) \qquad V_0^i = \frac{1}{\sqrt{N}} \qquad \forall i. \tag{20}$$

These limits are equivalent to Kac's as in first taking the limit $N \to \infty$ before the limit $\gamma \to 0$ he assumes terms of the form $N\gamma$ also tend to infinity. These limits also have a more physical motivation since if, for example, we had chosen γ to scale as $1/N^2$ (giving $N\gamma \to 0$) then each spin would have N interactions of order $1/N^2$ and so there would be no phase transition at a finite temperature and the spins would behave as if they were uncoupled. The special case $N\gamma \to a$ (when $N \to \infty$, $\gamma \to 0$) where a is finite gives an interaction matrix of finite rank, the rank now being greater than one. Thus this system is still described by Curie–Weiss mean-field theory and it behaves the same as the case $N\gamma \to \infty$ having a ferromagnetic phase transition associated with the largest eigenvalue λ_0.

Finally, we note that our model is described by only one order parameter which is, as in the previous example, the magnetization (given by $m_0 = \tanh \beta \lambda_0 m_0$). It thus behaves in the same way as a long-range ferromagnetic model with $J_{ij} = \lambda_0/N \ \forall i, j$.

5. Site disorder spin glass models

We now come to the largest class of models, which have been studied in the literature, whose interaction matrix (in the case of the separable models) satisfies conditions (i) to (iii). Long-range site disorder spin glass models (Benamira *et al* 1985, van Hemmen *et al* 1986, Grensing and Kühn 1987) typically have interactions given by

$$J_{ij} = \frac{1}{N} g(\xi_i, \xi_j) \tag{21}$$

where the ξ_i are stochastic variables of length p with each element chosen from the set $A_n = \{a_1, a_2, \ldots, a_n\}$ with some probability distribution and g is a bilinear function of the vectors ξ_i and ξ_j. Some papers (van Hemmen 1986, Grensing and Kühn 1986) also discuss the more general case where g is a symmetric function of the site randomness but we will leave the discussion of these models to the end of this section.

We will now proceed by directly diagonalizing $\mathbf{J}$ to give a simple proof of the upper bound of the rank of $\mathbf{J}$ and hence the maximum number of order parameters required to describe the system. If we construct an $N \times p$ matrix, which we shall call $\boldsymbol{\xi}$, where the ith row is given by ξ_i then we can write the interaction matrix as

$$\mathbf{J} = \boldsymbol{\xi} \mathcal{J} \boldsymbol{\xi}^{\mathrm{T}} \tag{22}$$

where $\mathcal{J}$ is the symmetric $p \times p$ matrix which defines the bilinear function g. Diagonalizing $\mathcal{J} = \mathbf{Q} \hat{\mathcal{J}} \mathbf{Q}^{\mathrm{T}}$ we can rewrite $\mathbf{J}$ as

$$\mathbf{J} = \boldsymbol{\xi} \mathbf{Q} \hat{\mathcal{J}} (\boldsymbol{\xi} \mathbf{Q})^{\mathrm{T}} \tag{23}$$

thus, since the maximum rank of $\mathcal{J}$ is p, the maximum rank of $\mathbf{J}$ is also p. Thus the maximum rank of $\mathbf{J}$ is independent of n, the number of variables from which each element is chosen. Depending on the choice of the set A_n and the choice of $\mathcal{J}$ equation (23) will be a diagonalization of the interaction matrix of one of the three forms discussed in section 2. For the most interesting models in the site disorder class $R(\mathbf{J}) = p$ (see for example the van Hemmen model in the next section). It should be noted that Grensing and Kühn (1987) showed that the maximum number of order parameters required to describe site disorder systems is p. Their calculation is more complicated than the one presented here and involves rewriting the Hamiltonian in terms of disjoint sublattices which reduces the problem to the diagonalization of an $n^p \times n^p$ matrix. Their calculation, as can most of the solutions of the separable site disorder spin glass models, can be regarded as a diagonalization of the interaction matrix. Details of this are given in the appendix. Thus, when p is finite (the separable models) the interaction matrix satisfies conditions (i) to (iii) and the system is described by the Curie–Weiss mean-field equations. The system is then described by a finite number of orthogonal order parameter equations (see equations (9) and (10)) and only has a finite number of stable states. When p is of order N then $R(\mathbf{J})$ is, in general, of order N and the system can exhibit 'true' spin glass behaviour in the sense that it can have of order N or more stable states which are not related by global rotation (Choy and Sherrington 1984). The Hopfield neural network storing of order N patterns (Amit *et al* 1985), $J_{ij} = 1/N \sum_{\mu=1}^{p} \xi_i^\mu \xi_i^\mu$ where ξ_i^μ is randomly chosen to be ± 1, is an example of a site disorder model of this type which exhibits true

spin glass behaviour. This crossover, in random site models, to random bond type spin glass behaviour when p is of order N had already been studied by Benamira *et al* (1985). They studied this crossover in the context of the extensivity of the logarithm of the characteristic function of the random couplings. The extensivity of the logarithm of the characteristic function is linked to the divergence of the rank of the interaction matrix as is 'true' spin glass behaviour. In fact we can think of the divergence of the rank of the interaction matrix in the limit $N \to \infty$ as a necessary condition for a system to have 'true' spin glass behaviour. We will now illustrate some of the points in this section by studying the van Hemmen model.

The van Hemmen model (van Hemmen *et al* 1983, Grensing and Kühn 1987) is usually defined in the following way,

$$J_{ij} = \frac{1}{N}(\xi_i \eta_j + \xi_j \eta_i) \tag{24}$$

where ξ_i and η_i denote the components of the two-dimensional vector ($p = 2$), ξ_i (in the notation of equation (21)). The model is normally defined with ξ_i and η_i being given by the same probability distribution and having zero average but, as we will see, these choices do not change the rank of the interaction matrix which is 2. It should be noted that often the van Hemmen model is defined as having a ferromagnetic interaction J_0/N added to the site disorder part. For simplicity we have omitted this term but it is easy to show (for all site disorder models (p finite) defined by equation (21)) that this term increases the rank of the interaction matrix by one.

There are many calculations solving the van Hemmen model (Grensing and Kühn 1987, Choy and Sherrington 1984) using the saddle-point technique to evaluate the free energy. The relation between these calculations is easy to understand when it is realized that there are many different ways to diagonalize the interaction matrix **J** and then apply the saddle-point technique. One example is the method used by Choy and Sherrington (1984) where they write the Hamiltonian in the separable form

$$\frac{1}{N}\sum_{i,j}(\xi_i\eta_j+\xi_j\eta_i)S_iS_j = \frac{1}{N}\left(\sum_i(\xi_i+\eta_i)S_i\right)^2 - \frac{1}{N}\left(\sum_i \xi_i S_i\right)^2 - \frac{1}{N}\left(\sum_i \eta_i S_i\right)^2. \tag{25}$$

Where we choose $J_{ii} = 2/N\xi_i\eta_i$ which, as discussed in section 2, does not affect the thermodynamics of the system. This is a diagonalization of the third type (as defined in section 2) by a singular matrix giving a number of separable terms higher than the rank of the interaction matrix. This means that the three order parameters introduced by a Gaussian transformation are not linearly independent. In fact, later in their paper they reduce the number of order parameters to two via a linear transformation. Grensing and Kühn (1987) formulated this model using their method of disjoint sublattices (see appendix). For this model their calculation corresponds to an orthogonal diagonalization or a diagonalization with a non-singular matrix depending on the probability distribution defining ξ_i and η_i.

The diagonalization procedure given in equation (23) can also be used to write the Hamiltonian in the separable form

$$\frac{1}{N}\sum_{i,j}(\xi_i\eta_j + \xi_j\eta_i)S_iS_j = \frac{1}{2N}\left(\sum_i(\xi_i+\eta_i)S_i\right)^2 - \frac{1}{2N}\left(\sum_i(\xi_i-\eta_i)S_i\right)^2. \tag{26}$$

This form is, in general, not orthogonal (see next paragraph). This type of diagonalization leads to mean-field equations of the same form as those given by Grensing and Kühn (1987) (see equation (38) in their paper).

A direct orthogonal diagonalization of the van Hemmen interaction matrix gives the following two non-zero eigenvalues and corresponding normalized eigenvectors

$$\lambda_{\pm} = \frac{1}{N}\left(\sum_i \xi_i\eta_i \pm \sqrt{\sum_i \xi_i^2 \sum_i \eta_i^2}\right) \qquad V_{\pm} = \frac{1}{\sqrt{|2N\lambda_{\pm} r|}}(\xi_i \pm r\eta_i) \tag{27}$$

where

$$r = \sqrt{\frac{\sum_i \xi^2}{\sum_i \eta_i^2}}. \tag{28}$$

The van Hemmen model is thus solvable for any choices of the ξ_i and η_i provided conditions (ii) and (iii) are met. Its orthogonal mean-field equations are given by equations (9) and (10) with the eigenvectors and eigenvalues defined by equation (27). These equations are, as stated in section 2, equivalent to the Curie–Weiss mean-field equations, and any of the other mean-field type equations stated in the literature (van Hemmen *et al* 1983, Choy and Sherrington 1984), under a linear transformation. Thus the model has a second-order phase transition at $kT_c = \lambda_+$ where it condenses with a critical exponent of one-half into the state associated with the eigenvector V_+. Detailed solutions of the order parameter equations for a few different probability distributions plus a ferromagnetic component in the interaction were presented in van Hemmen *et al* (1983). Our orthogonal solution of this model extends some of the results presented in this paper which are only valid for probability distributions having zero mean and fixed variance.

In the literature (van Hemmen *et al* 1983, Choy and Sherrington 1984) the order parameter equations of this model are usually written in terms of the two parameters $q_1 = (1/N)\sum_1 \xi_i\langle S_i\rangle$ and $q_2 = (1/N)\sum_1 \eta_i\langle S_i\rangle$ rather than their orthogonal forms. These are usually referred to as the natural order parameters of the system (Grensing and Kühn 1986). This brings up the question of what are the best set of order parameters to represent the system. It is true that, algebraically speaking, the order parameter equations are, in general, easier to derive in terms of q_1 and q_2 but these equations do not immediately yield the phase transition point and can be more difficult to solve than the orthogonal equations.

Returning now to the more general case where the function g (see equation (21)) is a symmetric function. For these models we cannot perform the direct diagonalization (see equation (23)) which worked for the bilinear models. In this case the technique of breaking up the system into disjoint sublattices, developed by Grensing and Kühn (1986, 1987 plus the appendix in this paper), can be used to diagonalize the interaction matrix. This calculation shows that $R_{\max}(\mathbf{J}) \leqslant n^p$ where n is the number of elements from which each component of ξ_i is chosen. Therefore the rank of the interaction matrix can now change with the probability distribution and it is only systems where n^p is finite that can be solved by orthogonal diagonalization of the interaction matrix. Van Hemmen *et al* (1986) have used other techniques to solve models where n^p is of order N and find order parameter functions as solutions.

6. Conclusion

The basic result of this paper is to show that the saddle-point mean-field type calculation can be reformulated in terms of the eigenvalues and eigenvectors of the interaction matrix. This calculation gives a systematic way to study the so called separable models such as the site disorder spin glasses. This calculation also shows that many of the solutions of these models in the literature can be regarded as a diagonalization of the interaction matrix and in this context the relation between the different solutions becomes more transparent. Our calculation also shows that a sufficient condition for the Curie–Weiss mean-field equations to be valid is that the rank of the interaction matrix is finite (in the limit $N \to \infty$). We also find that a necessary condition for 'true' spin glass behaviour is that the rank of the interaction matrix is divergent in the thermodynamic limit.

The formulation of the saddle-point integral in terms of the eigenvectors and eigenvalues also facilitates the study of a generalized class of long-range ferromagnetic models. Unlike the bond disorder ferromagnetic model these systems have some structure in their interactions which leads to stable states other than the pure ferromagnetic stable states. The study of these models will be the topic of a future paper (Canning 1992).

Acknowledgments

In the course of this work I have benefited from useful discussion with Stephen Cornell. I also wish to thank Carolynn Smith for a critical reading of this manuscript.

Appendix

In this appendix we show that the technique employed by Grensing and Kühn (1986, 1987) to solve separable site-disorder spin glass models can also be interpreted as a diagonalization of the interaction matrix. We will follow the same notation as that used in Grensing and Kühn (1986) (refer to this paper for a definition of the problem and the symbols we will use). If we define the disjoint sub-lattices via vectors given by,

$$\mathcal{M}_\gamma^i = 1 \qquad i \in \Omega_\gamma \qquad \mathcal{M}_\gamma^i = 0 \qquad i \notin \Omega_\gamma \tag{29}$$

where the sets Ω_γ are the disjoint sublattices then the interaction matrix can be written as

$$J_{ij} = \frac{J}{N} \sum_{\gamma\gamma'} V_{\gamma\gamma'} \mathcal{M}_\gamma^i \mathcal{M}_{\gamma'}^j. \tag{30}$$

Grensing and Kühn's sublattice magnetizations (see 1986, equation (5)) are now given by

$$M_\gamma = \sum_i \mathcal{M}_\gamma^i S_i. \tag{31}$$

Following their orthogonal diagonalization of $V_{\gamma\gamma'} = \sum_q Q_q^\gamma \lambda_q Q_q^{\gamma'}$ we can then rewrite the interaction matrix in the diagonalized form

$$J_{ij} = \frac{J}{N} \sum_q \lambda_q \sum_\gamma Q_q^\gamma M_\gamma^i \sum_{\gamma'} Q_q^{\gamma'} M_{\gamma'}^j \tag{32}$$

where, in the notation of equations (12) and (13) in this paper, $a_q = \lambda_q$ and $b_q^i = \sum_\gamma Q_q^\gamma M_\gamma^i$. Since the rank of V is n^p this shows that the maximum rank of J is also n^p. In general the diagonalized form given by equation (32) is not orthogonal but it is easy to show that it is orthogonal when all the disjoint sublattices are of the same size.

References

Amit D J, Gutfreund H and Sompolinsky H 1985 *Phys. Rev.* A **32** 1007; 1987 *Ann. Phys.* **173** 30

Baker G A 1961 *Phys. Rev.* **122** 1477; 1962 *Phys. Rev.* **126** 2071

Benamira F, Provost J P and Vallée G 1985 *J. Physique* **46** 1269

Canning A 1992 *J. Phys. A: Math. Gen.* submitted

Choy T C and Sherrington D 1984 *J. Phys. C: Solid State Phys.* **17** 739

Edwards S F and Jones R C 1976 *J. Phys. A: Math. Gen.* **9** 1595

Grensing D and Kühn R 1986 *J. Phys. A: Math. Gen.* **9** 1153

—— 1987 *J. Physique* **48** 713

Kac M 1968 *Statistical Physics, Phase Transitions and Superfluidity* vol 1, ed M Chretien, E P Gross and S Desser (New York: Gordon and Breach) p 241

Mehta M L 1967 *Random Matrices and the Statistical Theory of Energy Levels* (New York: Academic) p 240

Mezard M, Parisi G and Virasoro M A 1986 *Spin Glass Theory and Beyond (World Scientific Lecture Notes in Physics 9)* (Singapore: World Scientific)

Opper M 1989 *Europhys. Lett.* **8** 389

Sherrington D and Kirkpatrick S 1975 *Phys. Rev. Lett.* **32** 1792; 1978 *Phys. Rev.* B 17 4384

Stinchcombe R B 1983 *Phase Transitions and Critical Phenomena* vol 7, ed C Domb and J L Lebowitz (New York: Academic) p 152

Thouless D J, Anderson P W and Palmer R G 1977 *Phil. Mag.* **35** 593

van Hemmen J L, van Enter A C D and Canisius J 1983 *Z. Phys.* B **50** 311

van Hemmen J L, Grensing D, Huber A and Kühn R 1986 *Z. Phys.* B **65** 53

Wigner E P 1967 *SIAM Rev.* **9** 1