Parisi function for two spin glass models

1985 J. Phys. A: Math. Gen. 18 1255

(http://iopscience.iop.org/0305-4470/18/8/025)

J. Phys. A: Math. Gen. 18 (1985) 1255-1274. Printed in Great Britain

Parisi function for two spin glass models

P Sibani† and J A Hertz

Nordita, Blegdamsvej 17, 2100 Copenhagen Ø, Denmark

Received 2 November 1984

Abstract. The probability distribution function $P(q)$ for the overlap of pairs of metastable states and the associated Parisi order function $q(x)$ are calculated exactly at zero temperature for two simple models. The first is a chain in which each spin interacts randomly with the sum of all the spins between it and one end of the chain; the second is an infinite-range limit of a spin glass version of Dyson's hierarchical model. Both have non-trivial overlap distributions. In the first case the problem reduces to a variable-step-length random-walk problem, leading to $q(x)=\sin(\pi x)$. In the second model, $P(q)$ can be calculated by a simple recursion relation which generates devil's staircase structure in $q(x)$. If the fraction p of antiferromagnetic bonds is less than $(2)^{-1/2}$, the staircase is complete and the fractal dimensionality of the complement of the domain where $q(x)$ is flat is $\ln 2/\ln(1/p^2)$. In both models the space of metastable states can be described in terms of Cayley trees, which have, however, a different physical interpretation than in the SK model.

1. Introduction

Recent work on spin glasses (Parisi 1983, Young 1983, Mezard *et al* 1984, Gross and Mezard 1984) has emphasised the importance of the probability distribution function for the overlap

$$q^{\alpha\beta}=\frac{1}{N}\sum_{i=1}^{N} m_i^{\alpha} m_i^{\beta} \tag{1.1}$$

between a pair of metastable states whose magnetisations at site i are m_i^{α} and m_i^{β}. The purpose of the present work is to study this distribution in a couple of simple (if artificial) models with special features which allow some exact calculations. We hope that the results may provide some helpful guidance in sorting out the physics of more complicated systems which also exhibit condensation with many metastable states.

We use the term 'complex condensation' to refer to this kind of phase. We start from the supposition that the task of theory in this class of problem is to construct an appropriate statistical description of the metastable states and the relations among them. It is not yet known just what quantities it will be useful to know for understanding these qualities and their thermodynamic consequences, but it is clear from the start that one question we would like to have answered is 'how much do these various states resemble one another?'

The probability distribution $P(q)$ for the overlap (1.1) is designed, in part, to answer this question. Consider first two extreme situations: (i) a conventional ferromagnet,

† Alternative address: Det Fysiske Institut, Odense University, Odense, Denmark.

0305-4470/85/081255+20$02.25

which has just two metastable states; and (ii) a set of N independent, two-level systems, with 2^N states. In the first case $P(q)$ is just a pair of δ functions at $\pm m$, where m is the spontaneous magnetisation. In the second, $P(q)$ is a binomial distribution. In the large N limit, its width goes to zero (because we have normalised q in (1.1) to unity for the case of complete overlap), i.e. the different states have almost nothing in common with each other.

An intermediate case can be seen in a toy model studied by Palmer and Hertz (1983). It is like the independent, two-level-system case except that the different two-level systems represent ferromagnetic clusters of spins of widely differing sizes. This is shown schematically in figure 1. There are N spins in all: $N/2$ of them are in one cluster, $N/4$ in the next, and so on. There are $\log_2 N$ rigid clusters in all. The metastable states can be specified by a list

$$\alpha = (m_1, m_2, \ldots) \qquad m_i = \pm 1. \tag{1.2}$$

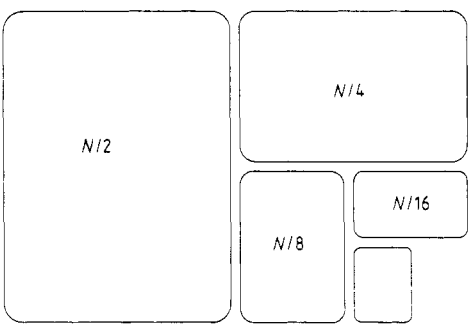

Figure 1. Visualisation of the toy model of Palmer and Hertz. The boxes represent independent clusters of spins of size $N/2$, $N/4$, etc, where N is the total size of the system.

The magnetisation of state α is then

$$M = N \sum_i m_i^\alpha 2^{-i} \tag{1.3}$$

and the overlap between two states α and β is

$$q^{\alpha\beta} = \sum_i m_i^\alpha 2^{-i}. \tag{1.4}$$

Now consider the quantity $r^{\alpha\beta} = \frac{1}{2}(q^{\alpha\beta} + 1)$, which lies between 0 and 1. From (1.4) we can write

$$r^{\alpha\beta} = \sum_i d_i^{\alpha\beta} 2^{-i} \tag{1.5}$$

where $d_i^{\alpha\beta}$ is randomly 0 or 1. But (1.5) is just the expansion of a number whose binary representation is 0. $d_1 d_2 d_3 d_4 \ldots$. Thus $r^{\alpha\beta}$ is uniformly distributed between 0 and 1, and consequently $P(q)$ is uniform on the interval from -1 to $+1$.

These examples illustrate what $P(q)$ tells us about the metastable states. In the kinds of situations we are used to dealing with in statistical mechanics, $P(q)$ turned out to have a trivial structure. It was necessary to have a whole spectrum of cluster sizes in order to get something other than one or two δ functions.

What else should we want to know about the metastable states besides the information contained in $P(q)$? Mezard and coworkers suggested that further interesting and useful information about the relation between the states could be found by examining higher-order, joint-probability distributions, such as that of the set of three overlaps $q^{\alpha\beta}$, $q^{\beta\gamma}$ and $q^{\gamma\alpha}$ between three states α, β and γ. They found in the SK model (Sherrington and Kirkpartick 1975), using the Parisi replica ansatz (Parisi 1979, 1980) that the form of this joint distribution suggested an ultrametric topology (i.e. a tree structure) in the space of states.

In this work we study two models which are much simpler than the SK model. They are so simple that, at least at zero temperature, the metastable states can be identified and classified quite directly for arbitrary bond configurations. This gives us the possibility of a more direct insight into the meaning of statistical characterisations of the states (such as $P(q)$ and the joint distributions mentioned above) than is possible in the SK model. Nevertheless, the models are not so simple as to be trivial in the sense that the first two examples we saw above were. We will find non-trivial structure in $P(q)$ in both cases. Our motivation in doing this is like that of Gross and Mezard, who have studied Derrida's random-energy model (1980, 1981) in a similar spirit.

Our first model is a one-dimensional one. In order to have non-trivial statistics, we need long-range interactions, which we introduce in the following fashion. We start constructing the chain from, say, the left end. Each spin is then taken to interact, via a bond of random sign, with the sum of all the spins to its left. In this way, interactions with longer and longer ranges are built up, up to the length of the chain. We will see that in a sense the range of this model is barely long enough to produce a non-trivial spin-glass state. The calculation of $P(q)$ turns out to reduce to a classic problem in random walk theory, with a simple solution.

Our second model is a special case of a class of models introduced for the ferromagnetic case by Dyson (1969, 1971). It is pictured in figure 2. We take a set of $N = 2^L$ spins S_i arranged as shown in a line. We introduce random interactions $\pm J_0$ between S_1 and S_2, between S_3 and S_4, and so on, as shown in the figure. (There are 2^{L-1} of these bonds.) We then add interactions between the sums of adjacent pairs, for example

$$\pm J_0 2^{-\lambda}(S_1 + S_2)(S_3 + S_4) \tag{1.6}$$

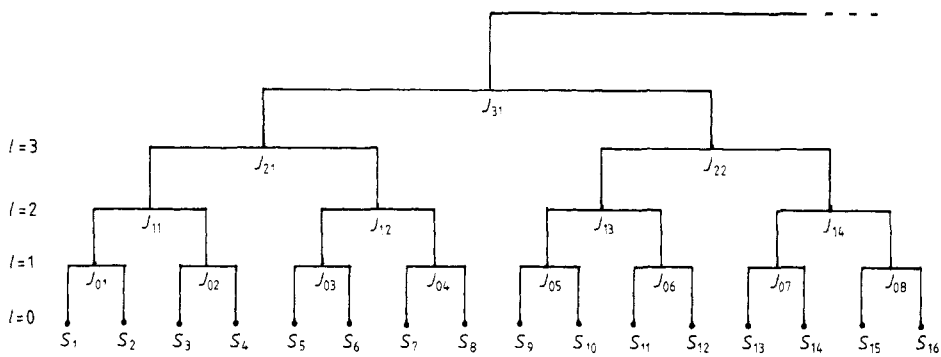

Figure 2. The Dyson model. The dots at zero level are actual spins, and the variables S_{lp} at the higher levels are block spin variables defined in a recursive manner by $S_{lp} = S_{l-1,p} + S_{l-1,2p-1}$. In our version of the model the coupling constants J_{lp} are random variables.

and so forth (2^{L-2} such bonds). The parameter regulates how fast the interactions fall off with distance along the chain. The next level interactions have the form

$$\pm J_0 2^{-2\lambda}(S_1+S_2+S_3+S_4)(S_5+S_6+S_7+S_8) \tag{1.7}$$

and so on. At the lth level we add interactions of magnitude $J_0 2^{-\lambda l}$ coupling pair of blocks of 2^l spins, until at level $l=L-1$ there is just one bond between the left- and right-halves of the system. Dyson introduced this kind of model as an approximation to one-dimensional ferromagnets with $r^{-\lambda}$ power law interaction, on the argument that the long-range part of the forces could be approximated by interactions between large blocks of spins. For $T=0$ and small λ the metastable states are readily identified and we are able to calculate $P(q)$ by iterating a simple functional relation.

The remainder of this paper comprises three sections, one for each model, followed by general comments and conclusions. A brief account of part of this work has been published in the proceedings of the Nobel symposium held at Graftavallen, Sweden, June 1984 (Hertz and Sibani 1985).

2. The tapeworm model

As mentioned in the introduction the tapeworm model is a one-dimensional chain of spins, each interacting in a random fashion with the sum of the previous ones. The form of the interaction is chosen in order to obtain maximal mathematical simplicity, admittedly at the price of introducing some unphysical features. Specifically we take the energy of the nth spin to be

$$E_n = \left[J_n \operatorname{sgn}\left(\sum_{i=1}^{n-1} \sigma_i \right) + \sigma_n \right]^2 \left[1 - \delta\left(\sum_{i=1}^{n-1} \sigma_i, 0 \right) \right] \tag{2.1}$$

where δ is the Kronecker delta, sgn is the sign function and J_n are independent stochastic variables taking the values $+1$ (antiferromagnetic bond) or -1 (ferromagnetic) with probabilities p and $(1-p)$ respectively. Each spin interacts with an 'effective field', stemming from the rest of the system, which is crudely represented by the sign function. If this field happens to vanish at the nth site, the last term in (2.1) ensures that the nth spin has zero energy, irrespective of its orientation.

Some insight in the consequences of (2.1) can be gained by considering the two extreme cases in which the bonds are (i) all ferromagnetic and (ii) all antiferromagnetic. In the former case the system becomes completely ordered at $T=0$. The order survives also at $T\neq 0$, since small fluctuations do not change the local field felt by a given spin at all. In the latter case an n-spin system splits into $n/2$ clusters, each containing two spins of opposite signs and each free to rotate with respect to the other with no change in energy. Exactly the same behaviour is present in the random Dyson model.

Returning to (2.1) we note that the ground state of the system is found by satisfying each bond. Then all E_n vanish separately and the ground-state energy is zero. In spite of the lack of frustration there are (lots of) degeneracies in the system. If $\Sigma_i^{n-1}\sigma_i$ vanishes for the first time at $n=n_0-1$, the n_0th spin has two degenerate orientations, each corresponding to a different macroscopic arrangement of the rest of the chain. Furthermore the n_0th and the first spin are equivalent and therefore the first bifurcation at n_0 is followed by an infinite series of bifurcations at $n_1, n_2, \ldots, n_k$. This situation is described graphically in figure 3 by a Cayley tree. The possible states of a system

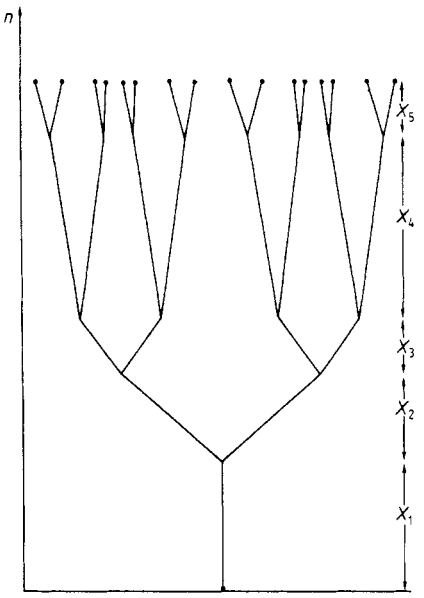

Figure 3. The Cayley tree of the tapeworm model. The height of the tree is the number of particles in the system, and each path in the tree corresponds to a clustering state of the system. The X_k is the number of spins in the kth cluster of the chain.

with n spins are the end points of a tree of height n and

$$X_k = n_k - n_{k-1} \tag{2.2}$$

is the height of the kth section of the tree.

The overlap between any two different states can be calculated by multiplying two paths in the tree section by section and by scaling by the total number of spins. Since the product of two sections is plus or minus the length of the section we are led to consider the distribution of the stochastic variable

$$Q_k = \frac{\lambda_1 X_1 + \lambda_2 X_2 + \ldots + \lambda_k X_k}{X_1 + X_2 + \ldots + X_k} \tag{2.3}$$

in the limit $k \to \infty$ of an infinitely high tree. In (2.3) the λ_i's are symmetrically distributed, independent random variables, each with range ± 1. Taking a given realisation of the λ corresponds to choosing two definite pure states of the system. X_k are functions of the couplings $J_1, J_2, \ldots, J_k$. Since we do not know what the couplings are, we consider X_k as stochastic variables which, as we shall see in short, can be identified with the times of return to the origin in a nearest-neighbour one-dimensional random walk. Finally, the Parisi function for the tapeworm at $T = 0$, $q(x)$, is the inverse of

$$x(q) = \int_0^q P_Q(z)\, \mathrm{d}z \tag{2.4}$$

where P_Q is the density of Q_k in the limit $k \to \infty$.

It is convenient to distinguish in the following treatment between two different cases: (i) $p = 1/2$, i.e. ferromagnetic and antiferromagnetic bonds are equally probable; and (ii) $p \neq 1/2$. Considering the first case, when $p = 1/2$ any spin has equal probability

of pointing in any direction. We can imagine that the spins are added to the system at equal time intervals and visualise the spin sum $\Sigma_{i=1}^{k}\sigma_i$ as the position at 'time' k of a random walker which started at the origin and moves by nearest-neighbour hops in a one-dimensional lattice, with equal probability for left and right hops. In this term X_k is simply the time (number of steps) elapsed between the $k-1$ and the kth returns to the origin. Referring back to (2.3), we can also give the two possible values of λ_i the following random walk interpretation. $\lambda_i=+1$ (-1) if the walker steps into the positive (negative) axis after the $(i-1)$th return. In these terms $+(-)X_i$ is the event; the walker spends time X_i on the positive (negative) axis after the $(i-1)$th return. Recalling that the denominator of (2.3) is the duration of the walk, it follows that

$$Q_k = P_k - N_k = 2P_k - 1 \tag{2.5}$$

where P_k and N_k are the fractions of time which are spent on the positive and negative axis in a k-step random walk. It is well known from the theory of probability (e.g. Feller 1970, vol 1) that the distribution of Q_k in the limit $k\to\infty$ is given by the so-called $\sin^{-1}$ law, which simply means that

$$P_Q(q) = (1/\pi)(1-q^2)^{-1/2} \tag{2.6}$$

and

$$x(q) = (1/\pi)\sin^{-1}(q). \tag{2.7}$$

Hence the Parisi function has the remarkably simple form

$$q(x) = \sin(\pi x) \qquad |x| < \tfrac{1}{2}. \tag{2.8}$$

Motivated by a recent publication of Mezard *et al* (1984) we now proceed to calculate the density $P(q_1, q_2, q_3)$ of the overlaps between the three possible pairs formed out of three randomly chosen states α, β and γ. The above authors investigated the metric properties of the state space of the SK model and found that at least two of the overlaps $q^{\alpha\beta}$, $q^{\alpha\gamma}$ and $q^{\beta\gamma}$ are equal while the third is smaller or equal to the two others, i.e.

$$P(q_1, q_2, q_3) = \tfrac{1}{2}\theta(q_1-q_2)\delta(q_3-q_2)P(q_1)P(q_2) + \text{cyclic permutations}$$
$$+ P(q_1)\left(\int_{-1}^{q_1} P(z)\,\mathrm{d}z\right)\delta(q_1-q_2)\delta(q_2-q_3). \tag{2.9}$$

Mezard *et al* also show that as a consequence of (2.9) the states of the SK model can be grouped in disjoint clusters, such that all the states in a cluster have overlap greater than a given q_J. Choosing a set $\{q_1, q_2, \ldots\}$ with $q_i > q_{i+1}$ gives a hierarchical structure of clusters such that clusters can be represented by a Cayley tree. This seems somehow analogous to the Cayley tree structure of the tapeworm; however, it must be stressed that the two concepts are quite distinct.

A cluster of the tapeworm can be conveniently defined as a string of spins which can be flipped without changing the energy. Defining the distance between two states as $d = m - n$, where m is the total number of bifurcations and n is the number of sections the two states have in common, we find as in the SK model a hierarchy of clusters within clusters, where all the spins belonging to a cluster have distances less than or equal to d. In contrast with the SK model this hierarchy is finite, since the sections of the tapeworm, $X_1, X_2, \ldots$ cannot be further subdivided. The main difference between the SK model and ours comes, however, from our definition of distance, which is related but by no means identical to that of overlap. In particular given any three

states α, β and γ at least two of the distances $d^{\alpha\beta}$, $d^{\alpha\gamma}$ and $d^{\beta\gamma}$ are equal, but the statement does not hold for the corresponding overlaps, and (2.9) does not apply.

In order to clarify this point it is useful to consider the overlap

$$Q_m^d = \frac{X_1 + \ldots + X_{m-d} + \ldots + \lambda_{m-d+1} X_{m-d+1} + \ldots + \lambda_m X_m}{X_1 + X_2 + \ldots + X_m} \tag{2.10}$$

between pairs of known separation d, in a tapeworm with m sections. It consists of a sure part $X_1 \ldots X_{m-d}$, which is the same for all the pairs separated by d, and a part which fluctuates from pair to pair, according to the realisation of the λ's. If d is small the overlaps of different pairs with the same d will be very similar, while for large d they will be uncorrelated. The key to understanding the properties of $P(q_1, q_2, q_3)$ lies therefore in the answer to the following question. Let (d_1, d_2) be the two distances associated in the usual way with three states α, β and γ, and assume $d_1 \geqslant d_2$. What is the joint density $P_{d_1,d_2}(n_1, n_2)$? As explained in detail in appendix 1, we find

$$P_{d_1,d_2}(n_1, n_2) = \begin{cases} (\frac{3}{4})(\frac{1}{2})^{2m-n_1-n_2} & n_1 > 0,\ 0 < n_2 < n_1 \\ (\frac{3}{2})(\frac{1}{2})^{2m-n_1} & n_1 > 0,\ n_2 = 0 \\ \delta_{n_2,0}(\frac{1}{4})^m & n_1 = 0. \end{cases} \tag{2.11}$$

Clearly, in the thermodynamic limit $m \to \infty$, a fraction of the total mass of P which is arbitrarily close to one is distributed in a finite region of the (n_1, n_2) plane close to m, i.e. the sure part of the overlap (2.10) only contains a finite number of terms which asymptotically do not contribute to the sum. This means that the correlations in the overlaps $q^{\alpha\beta}$, $q^{\alpha\gamma}$, $q^{\beta\gamma}$ induced by the fact that two of the distances $d^{\alpha\beta}$, $d^{\alpha\gamma}$ and $d^{\beta\gamma}$ are equal, vanish in the thermodynamic limit, i.e.

$$P(q_1, q_2, q_3) = P(q_1)P(q_2)P(q_3) \tag{2.12}$$

for an infinite system.

We have hitherto considered the case in which ferromagnetic and antiferromagnetic bonds are equally probable, i.e. $p = \frac{1}{2}$. If $p \neq \frac{1}{2}$ the treatment must be slightly modified and the result is also rather different. As a preliminary consideration we note that the random-walk (or coin-tossing) analogy must be changed since the hopping probabilities now depend on whether the walker is on the right-half or left-half line (if $p > \frac{1}{2}$, i.e. antiferromagnetic bonds are most probable, the walker will be biased towards the origin). The difficulty is, however, only a matter of notation. Since we are interested in return times, we can limit our discussion to a random walk on the positive axis which terminates when the walker returns to the origin. In terms of games this is the classical ruin problem for a player which has an infinitely rich adversary.

Let us first consider the case $p < \frac{1}{2}$. This biases the walk towards infinity. As is well known in such a biased random walk there is a non-zero probability of never returning to the origin, which is therefore only visited a finite number of times with probability one. Hence the Cayley tree has only a finite number of nodes, say k nodes, and its kth branch X_k grows without bound in the thermodynamic limit. A glance at (2.3) shows that the overlap will eventually be dominated by this one term. Hence $Q = \pm 1$ with equal probability and

$$P_Q(q) = \tfrac{1}{2}(\delta(q+1) + \delta(q-1)). \tag{2.13}$$

If $p > \frac{1}{2}$ the probability of return to the origin is enhanced. The distribution of return times is well known (Feller 1970), but we do not need its explicit form. The

important feature is that, in contrast with the $p = \frac{1}{2}$ case, this distribution has a finite mean

$$E = (2p-1)^{-1}. \tag{2.14}$$

By definition, E is the expectation of X_i for all i. It follows from a symmetry consideration that the variables $\lambda_i X_i$ all have zero expectation. Rewriting Q_k as

$$Q_k = \frac{(\lambda_1 X_1 + \ldots + \lambda_k X_k)/k}{(X_1 + \ldots X_k)/k} \tag{2.15}$$

we note that (Feller 1970, vol 2, p 234) by the law of large numbers the numerator converges to zero and the denominator to E. It follows that, in the thermodynamic limit

$$P_Q(q) = \delta(q) \tag{2.16}$$

and thus the point $p = \frac{1}{2}$ is a kind of critical point separating phases with the trivial behaviour of $P(q)$ which characterise high and low temperature phases of conventional broken symmetry systems.

3. The random Dyson model

Our second model is somewhat more complex than the first one and has a correspondingly more complex structure in its $P(q)$. The block spin variables illustrated in figure 2 are formally defined in a recursive manner by

$$S_{l,p} = S_{l-1,2p} + S_{l-1,2p-1} \tag{3.1}$$

with $S_{0,p} = S_p$, the actual spins. Then the Hamiltonian for a system of 2^N spins is

$$H = -\sum_{l=0}^{N-1} \sum_{p=1}^{2^{N-l-1}} J_{lp} S_{l,2p} S_{l,2p-1} \tag{3.2}$$

where J_{lp} are independent random variables taking on the values

$$J_{lp} = \pm J_0 2^{-\lambda l} \tag{3.3}$$

with probabilities $1-p$ (+case, ferromagnetic bond) and p (−case, antiferromagnetic bond) respectively. As we mentioned, Dyson introduced the ferromagnetic version of this model to study one-dimensional ferromagnets with power-law forces (Dyson 1969). However, as noted by Baker (1972), it can also be used to model higher-dimensional systems. In general a model with $J_l \propto 2^{-\lambda l}$ simulates $r^{-d\lambda}$ forces in d dimensions. The class of Dyson models should not be confused with another set of systems which are also often called hierarchical: the so-called Berker lattices. Spin glass behaviour has been studied in this kind of system (McKay *et al* 1982, 1984) from a somewhat different point of view. The present work complements that of McKay *et al* in that we are looking at long-range models and a mean-field limit, while they study short-range models and do not use such concepts as a mean-field order parameter.

As a model of a real spin glass, ours is obviously artificial, although perhaps hardly more so than the SK model. However, blocks of metal with magnetic impurities are not the only systems where spin glass theoretical concepts are relevant. There is a growing interest in the statistical mechanics of pseudospin models of information processing structures, in both computer science applications (Hubermann and Hogg

1984) and neurobiology (Anderson 1972, Hopfield 1982). While the Dyson hierarchical model is clearly a bad description of RKKY coupled moments in a metal, it is not such a bad starting point for describing realisable connections of logical elements or neurons. We also note that this model is not as fragile as the tapeworm model of the preceding section, in that there is nothing special about the case $p=\frac{1}{2}$. We will find non-trivial behaviour (at least in the mean-field limit) except for the purely antiferromagnetic and purely ferromagnetic limits. We note further that there is no symmetry between ferromagnetic and antiferromagnetic interactions, i.e. the model is not invariant under $p \leftrightarrow (1-p)$, but quite the opposite; the purely antiferromagnetic limit turns out to have 2^{N-1} independent two-spin clusters, while the purely ferromagnetic limit has only the two states of a conventional ferromagnet. This difference comes about because the block variables (3.1) are always *total* spins and never a staggered-spin-like combination. A truly symmetric model would have differences equally often as sums on the right-hand side of (3.1). Therefore we should not necessarily expect the kind of close correspondence between our model and a one-dimensional long-range spin glass (Kotliar *et al* 1983) that is found in the ferromagnetic case between Dyson's model and the $r^{-\lambda}$ chain.

We are interested in identifying the metastable states in the limit $T \to 0$, and by analogy with the tapeworm (or any system with a non-trivial Parisi function) we expect a large degenerate set of these which we refer to as the ground-state manifold. We can get a clue about the relevant energetics by examining the smallest non-trivial example of our model, one with only four spins and three bonds, J_{01}, J_{02} and J_{11}. (See the lower left corner of figure 2.) There are eight possible combinations for the signs of the three bonds. However, the sign of J_{11} is irrelevant to the energy; what matters, as we shall see in a moment, is whether this top bond should be satisfied, possibly at the expense of breaking one of the lower bonds. So we have four combinations left, of which only three are distinct: J_{01} and J_{02} both positive, both negative or of different signs. Let us find the lowest energy configurations for all the three cases.

(1) $J_{01}=J_{02}=+J_0$. Here there is no competition; all bonds are satisfied by taking $S_1=S_2$ and $S_3=S_4$ and letting the relative sign of the two pairs be fixed by that of J_{11}.

(2) $J_{01}=-J_{02}$. Here we have the possibility of frustration; if we satisfy the lower-level bond, the energy is $-2J_0$. (The upper bond does not contribute, since one of the blocks is zero.) If we break the antiferromagnetic lower-level bond and satisfy J_{11} instead, the energy is $-4J_0 2^{-\lambda}$. Which of the two two possibilities has the lowest energy clearly depends on λ; for $\lambda<1$ we satisfy the higher-level bond and for $\lambda>1$ this bond is irrelevant.

(3) $J_{01}=J_{02}=-J_0$. Here J_{11} is competing with both the J_{0i}. If we satisfy the J_{0i} we have $E=-2J_0$, while if we satisfy the upper-level bond, $E=2J_0-4J_0 2^{-\lambda}$. So, for any $\lambda>0$, the lower-level bonds win out.

It is then tempting to hope that this procedure may be repeated at the next level, where we combine two systems like those we have just examined and couple their total spins by a J_{21} and so on. Then we would say that for $0<\lambda<1$, the two blocks would be combined into one rigid cluster with probability $1-p^2$ and would remain independent of each other with probability p^2. This simple situation would then allow us to construct a recursion relation for $P(q)$, the probability density of overlap between two states. It is, however, not a simple matter to prove that this procedure actually yields the ground state of the system. We have postponed the discussion to appendix 2, where we show that for small λ the system is actually 'almost' completely polarised in its ground state, even in this very unfavourable and unprobable bond configuration. Small λ is a mean-field limit, but we cannot take λ exactly zero, since then, as we saw

in case (3) above, the states obtained by always satisfying the top bond are at least degenerate with those in which the lower bonds are satisfied.

In a brief account of part of this work published elsewhere (Hertz and Sibani 1985), it was stated that our recursion procedure gave the exact ground states for all λ less than a finite λ_c. This is incorrect. It is true that the states given by our algorithm are lower in energy than the unpolarised state (antiferromagnetic pairs) for $\lambda < \lambda_c$, but they are not quite the ground states. What we show in appendix 2 is that they are arbitrarily close to the true ground states for sufficiently small λ.

Having established a procedure for constructing the ground-state manifold, we can now easily find a simple recursion relation for $P_l(q)$, the overlap probability density in a l-level system (see figure 2). If we combine two blocks each of 2^l spins the only bonds that matter are the top ones within each of the blocks; if they are both negative (probabilty p^2) there is no correlation induced between spins by joining the blocks together; otherwise (probability $1-p^2$) they form a new rigid block with no internal free clusters. In the former case the stochastic variable representing the overlap at the $l+1$ level is just the arithmetic mean of the two independent overlaps at level l, and its density is just given by a scaled convolution of the level l densities. In the latter the overlap at level l is fixed to be ± 1 with equal probability. Weighting these two cases with the appropriate probabilities we obtain

$$P_{l+1}(q) = \frac{1-p^2}{2}(\delta(q+1)+\delta(q-1)) + 2p^2 \int_{-1}^{1} \mathrm{d}q'\, P_l(q') P_l(2q-q'). \tag{3.4}$$

This is our fundamental result. In terms of characteristic functions

$$f_l(k) = \int_{-1}^{1} \mathrm{d}q \exp(\mathrm{i}kq) P_l(q) \tag{3.5}$$

we have a simple functional recursion relation:

$$f_{l+1}(k) = (1-p^2)\cos k + p^2 f_l^2(k/2). \tag{3.6}$$

These equations must be iterated with the initial condition

$$P_0(q) = \tfrac{1}{2}(\delta(q+1)+\delta(q-1)). \tag{3.7}$$

It is not difficult to see that with this initial condition the limit of $P_l(q)$ for $l \to \infty$, $P_\infty(q)$, will also have a piece proportional to (3.7). Indeed, by writing the ansatz

$$P_\infty(q) = \tfrac{1}{2}a(\delta(q+1)+\delta(q-1)) + \tilde{P}(q) \tag{3.8}$$

where $\tilde{P}(\pm 1) = 0$, and substituting in (3.4) we obtain a simple quadratic equation for a with solution

$$a = \frac{1-[1-2p^2(1-p^2)]^{1/2}}{p^2}. \tag{3.9}$$

For $p = \frac{1}{2}$, for example, $a = 0.838$, i.e. most of the mass of $P(q)$ is in the spikes at $q = \pm 1$.

We have carried out the iteration of the recursion relation (3.4) on the computer. The results are shown for several values of p in figure 4. Instead of $P(q)$ we have plotted the Parisi function $q(x)$ (defined by $P(q) = \mathrm{d}x/\mathrm{d}q$). Thus the widths of the plateaux in $q(x)$ correspond to the heights of the spikes in $P(q)$. The broadest plateaux are observed for $q = \pm 1$ (these are not shown in fuil in figure 2), and for small p the widths of successively smaller plateaux falls off more rapidly for small than for large p.

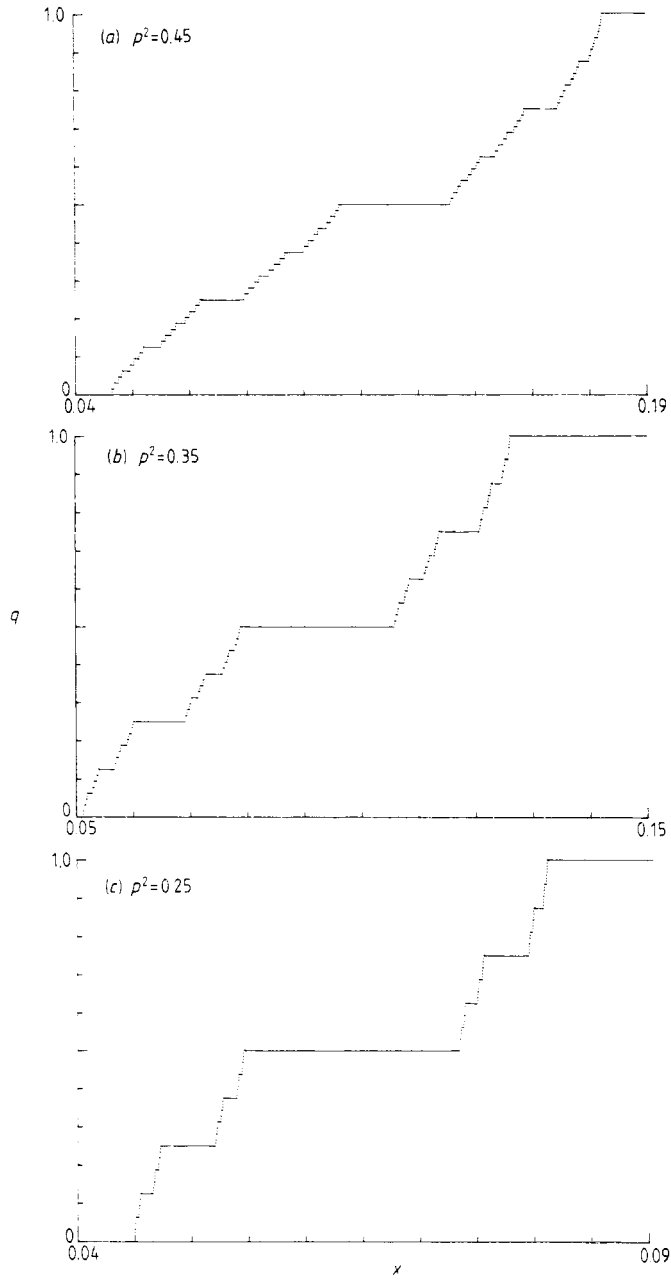

Figure 4. The Parisi function $q(x)$ of the random Dyson model is a devil's staircase. Here we show the staircase for values of p^2 of (a) 0.45, (b) 0.35 and (c) 0.25, where p is the probability that a given bond is antiferromagnetic. Note that the scales of the figures are at ratios 3 : 2 : 1.

The Parisi function clearly has a devil's staircase structure, since $P_\infty(q)$ has an infinite number of finite spikes. Is the staircase complete? This is the same as asking whether all the mass of $P_\infty(q)$ is in the spikes. We can answer the question in the

following way. Each iteration of the recursion relation produces a whole set of new spikes as well as adjusting the heights of the spikes already present. The number of spikes is the old number minus one, i.e. for large l the number of spikes is doubled at each iteration. If the total mass of the new spikes produced by the lth iteration goes to zero as $l \to \infty$, all the mass of $P_\infty(q)$ is in a finite number of spikes and the devil's staircase is complete.

In order to compute this quantity it is convenient to look at how the set of new spikes is generated by (3.4) without the rescaling which puts the new q in the interval $(-1, 1)$. We then have

$$P_{l+1}(q) = \frac{1-p^2}{2}(\delta(q, 2^{l+1}) + \delta(q, -2^{l+1})) + p^2 \sum_{q'} P_l(q') P_l(q - q') \tag{3.10}$$

where the δ are now Kronecker symbols. $P_l(q)$ is non-zero on the even integers between -2^l and 2^l. Now the key to the problem is in recognising that, if the spikes are numbered from the left (the left most spike at -2^l thus has number zero), the last set of spikes added in the iteration has an odd number, and $P_l(q)$ can naturally be divided into an even and an odd part. Thus, going back to the characteristic functions (3.5) we define $f_l^+(k)$ and $f_l^-(k)$ as the Fourier transform of the even and odd parts of $P_l(q)$ respectively. Since two even or two odd numbers add up to an even number, and an even and an odd add up to an odd number, we have

$$\begin{aligned} f_{l+1}^+(k) &= (1-p^2)\cos(k) + p^2[(f_l^+(k/2))^2 + (f_l^-(k/2))^2] \\ f_{l+1}^-(k) &= 2p^2 f_l^+(k/2) f_l^-(k/2) \end{aligned} \tag{3.11}$$

and since

$$f_l^+(0) + f_l^-(0) = 1 \tag{3.12}$$

(normalisation), we have a simple recursion relation for $f_l^-(0) = w_l$, the total weight in the odd (last generation) spikes:

$$w_{l+1} = 2p^2 w_l (1 - w_l). \tag{3.13}$$

We see immediately that for $p^2 \leqslant \frac{1}{2}$ the stable fix point is $w_\infty = 0$, i.e. asymptotically zero weight goes into the new spikes. In terms of $q(x)$, the plateaux cover all x in the interval $[-\frac{1}{2}, \frac{1}{2}]$, except for a set of zero measure; the devil's staircase is complete. For $p^2 > \frac{1}{2}$, w_∞ is non-zero, the staircase is incomplete and $q(x)$ must develop some continuous pieces. Since iterating this continuous piece does not give rise to new spikes, it is straightforward to see by integrating (3.10) that the total weight of the spikes of $P_\infty(q)$ is, for $p^2 > \frac{1}{2}$

$$m_\infty = (1-p^2)/p^2 \tag{3.14}$$

which is a smoothly decreasing function of p^2.

Returning to $p^2 < \frac{1}{2}$, where the staircase is complete, we can then use (3.13) to calculate the fractal dimensionality of the set of x for which $q(x)$ does not lie in a plateau. Following Jensen *et al* (1983) we define $S(r_l)$ as the fraction of the interval occupied by plateaux of width $\geqslant r_l$ and define the fractal dimensionality D of our set as

$$D = \lim_{r_l \to 0} \frac{\ln[(1 - S(r_l))/r_l]}{\ln(1/r_l)}. \tag{3.15}$$

From (3.13) we have

$$1 - S(r_l) = w_l \sim (2p^2)^l \tag{3.16}$$

for $l \to \infty$, while typically

$$r_l \simeq w_l / 2^{l-1} \tag{3.17}$$

since the weight is distributed among 2^{l-1} spikes. Thus

$$D = [\log_2(1/p^2)]^{-1}. \tag{3.18}$$

D is equal to $\frac{1}{2}$ for equal concentrations of ferromagnetic and antiferromagnetic bonds, increases as p increases and goes to unity as $p \to (2)^{-1/2}$, signalling that the devil's staircase is no longer complete for larger p.

By analogy to what we already have done for the tapeworm, we now want to establish the metric structure of the space of states of the random Dyson model. Consider therefore a system composed of 2^N spins. According to the preceding discussion the following possibilities arise: either the whole system is a rigid ferromagnetic cluster, with two possible states, or it is subdivided into two independent subclusters. The argument can be repeated for each of the subclusters and we end up with either one or two or three or ... 2^{N-1} independent rigid clusters. This can again be represented in terms of a Cayley tree: at each doubling of the number of spins the system can: (i) bifurcate or in general multifurcate according to how many states the neighbour block possesses; or (ii) return to or remain in a configuration with only two metastable states.

This corresponds to the two terms in the recursion relation (3.4). Should the system return from a configuration with many states to a two-state configuration, we simply replace the branches which meet again by a straight line, since this does not make any difference to the further evolution of the system as the size grows.

In conclusion we get again an ultrametric topology on the space of the pure states, with a distance defined in the same manner as in the tapeworm. There is however an important difference in the structure of $P(q_1, q_2, q_3)$. In the tapeworm this function factorises into the product $P(q_1)P(q_2)P(q_3)$ because the overwhelming majority of the states are very far apart in the thermodynamic limit. In the Dyson model on the contrary, there is a finite probability, for any size of the system, of having just two states related by an overall spin flip, which leads to strong correlations between the overlaps of any group of three states. Indeed the two possible ways of combining two blocks according to the signs of the bonds give

$$\begin{aligned} P_{l+1}(q_1, q_2, q_3) &= (1-p^2)P_0(q_1, q_2, q_3) \\ &\quad + 8p^2 \int_{-1}^{1}\int_{-1}^{1}\int_{-1}^{1} P_l(2q_1 - q_1', 2q_2 - q_2', 2q_3 - q_3')P_l(q_1', q_2', q_3') \end{aligned} \tag{3.19}$$

where

$$\begin{aligned} P_0(q_1, q_2, q_3) &= \tfrac{1}{4}(\delta(q_1-1)\delta(q_2+1)\delta(q_3+1) + \text{cyclic permutations} \\ &\quad + \delta(q_1-q_2)\delta(q_2-q_3)\delta(q_3-1)) \end{aligned} \tag{3.20}$$

is the joint overlap density for a system with only two states and also the initial condition of the recursion.

Rather than discussing in detail the properties of (3.19), we show by a simple argument that its solution cannot have the same form as the $P(q_1, q_2, q_3)$ given by Mezard *et al* for the SK model. Consider the case in which the system contains two independent sub-blocks, which are both rigid clusters. This happens with finite probability in a system of any size. Since the sub-blocks are rigid, each of them can be conveniently represented by just one spin, notwithstanding the fact that it has zero magnetisation. We thus have four states, which we label:

$$\alpha = \uparrow\uparrow \qquad \beta = \uparrow\downarrow \qquad \gamma = \downarrow\uparrow \qquad \delta = \downarrow\downarrow. \tag{3.21}$$

Clearly $q_{\alpha\beta} = q_{\alpha\gamma} = 0$ and $q_{\beta\gamma} = -1$, which is not compatible with (2.9), because the two overlaps which are equal are greater than (rather than smaller than) the third.

4. Conclusion

We have considered in this paper two simple models of spin systems with random interactions and have calculated exactly their zero-temperature Parisi functions. The models look very different: the first, which we have called the 'tapeworm', is simply a one-dimensional string of spins, each interacting with the sum of the previous ones. We find that the spins can be grouped into connected clusters; any cluster can be flipped independently without changing the energy of the system. In other words, each cluster corresponds to a local symmetry operation on the model. The tapeworm has an infinite number of clusters of finite size.

The second model is a random Dyson model, which also has clusters, defined as above. The distribution of cluster sizes differs, however, in the two cases: in the Dyson model any finite group of spins finally ends in one of the possible ferromagnetic configurations and one could say that, in the thermodynamic limit, only clusters of infinite size survive with non-zero probability. In any case, for a finite, arbitrarily large number of spins the state space of both models has an ultrametric topology. However, the distance between two states is defined in a different way than in the SK model.

There is one more similarity between our two models; they have exactly the same ground-state configurations if all the bonds are ferromagnetic or antiferromagnetic, as already mentioned in § 2. An important difference is that the Parisi function of the tapeworm is non-trivial only if the ferromagnetic and antiferromagnetic bonds are exactly equiprobable, while this is not the case for the Dyson model.

The non-triviality of the Parisi function requires two conditions to be satisfied. The first is the obvious one that the system should have many pure states. The existence of many states is, however not, sufficient to guarantee a non-trivial overlap density: one must also require that the cluster-size distribution is very broad. Specifically it must not have a finite mean. Examples of this fact are the set of N independent two-level systems considered in the introduction, and the tapeworm model in the case of prevalence of antiferromagnetic bonds. The clusters of the Dyson model and of the toy model of the introduction have clearly no finite mean size, and the same is true for the tapeworm when $p = \frac{1}{2}$, and all these models have a non-trivial Parisi function. Apart from these examples, the fact that a distribution of cluster sizes with a finite mean yields a trivial Parisi function follows directly from the law of large numbers, as is apparent from the arguments leading to (2.16).

We find it satisfying to have established a connection between old results from theory of random walks and new concepts like the Parisi function, and to find such features as fractal structure with variable Hausdorff dimensionality for the random Dyson model. It comes about because of both its hierarchical structure (which is what allows the recursive solution) and the discreteness of the random bonds. These are very special features of this particular model. However, we have seen that two seemingly different systems look rather similar if one considers the space of their pure states. This is also an *a posteriori* justification for considering simple examples as we have done here. It cannot be excluded that some of their features could be found in more complex models which are closer to the real world.

Acknowledgment

We thank R G Palmer for discussions which led us to this problem, T Witten for suggesting the relevance of Dyson models and the Institute of Theoretical Physics, University of California at Santa Barbara, and the Research Laboratory, Exxon Research and Engineering Company, for hospitality in the initial stages of this work.

Appendix 1

Consider a tapeworm with m sections and 2^m states. Take three states α, β and γ at random and ask for the smallest fraction of the system which contains them all. With probability $\frac{3}{4}$ this is the whole system, and with probability $\frac{1}{4}$ half of the system. In general if we subdivide the states into 2^k groups of 2^{m-k} contiguous element each, the probability that α, β and γ all fall into any of these groups is $(\frac{1}{4})^k$, which is also the probability that

$$d_1 = \max(d^{\alpha\beta}, d^{\alpha\gamma}, d^{\beta\alpha}) \leqslant m - k = n_1. \tag{A1.1}$$

It follows that

$$P(d_1 = n_1) = \begin{cases} -\frac{3}{4}(\frac{1}{4})^{m-n_1} & n_1 > 0 \\ (\frac{1}{4})^m & n_1 = 0. \end{cases} \tag{A1.2}$$

Now let $d_1 = n_1$, i.e. all the states are in a subgroup of 2^{n_1} elements and precisely two of them are in a subgroup of size 2^{n_1-1}. Repeating the previous argument they are both in a subgroup of size $2^{n_1-1-n_2}$ with probability

$$P(d_2 = n_2 | d_1 = n_1) = \begin{cases} \frac{1}{2}(\frac{1}{2})^{n_1-1-n_2} & n_1 > 0,\ 0 < n_2 < n_1 \\ (\frac{1}{2})^{n_1-1} & n_1 = 0. \end{cases} \tag{A1.3}$$

Equation (2.11) follows now from (A1.2) and (A1.3).

Appendix 2

In this appendix we justify by a detailed analysis the recursive method of constructing the ground state which is the basis of equation (3.4). The object of the present discussion is again the extreme case in which all the bonds in a 2^N spin block are negative, except for two of the three top bonds. We find that, for small λ, the block containing the +

bond is completely polarised and the other 'almost' completely polarised in the ground state of the system. Complete polarisation is only attained in the limit $\lambda \to 0$. However, the recursion (3.4) can be used for small non-zero λ, since the error in the overlap turns out to be negligible.

Unless otherwise stated, we shall use the word 'system' for a system with the above mentioned spin configuration and, for the sake of concreteness, place one of the two positive bonds in the right sub-block.

As a first step towards the construction of the ground state, we ask for the spin configuration yielding the lowest possible energy in a system in which all the bonds are antiferromagnetic and in which the total polarisation $S_{N,1}$ is kept fixed. The answer is given in the next lemma.

Lemma 1. The lowest energy configuration for the above system is obtained by recursively applying the following prescription for adding two units of polarisation, i.e. for the change $S_{N,1} \to S_{N,1}+2$. Divide the whole system into two sub-blocks. If these are equally polarised, choose one of them at random, otherwise take the less polarised sub-block, and within this sub-block repeat the process described above (subdivision and choice) again and again, until the chosen block for the first time has zero polarisation. Choose then a spin pair at random within the block and polarise it. As a corollary we note that any two blocks of the same size have polarisations differing by at most two units.

For the sake of continuity of the exposition we postpone the proof to the end of the appendix and proceed to apply the lemma. We also note that the prescription of the lemma also yields the lowest energy when the top bond of the block is positive rather than negative.

The main purpose of our discussion is to find the ground state of the system with the 'test' bond configuration. In this configuration we get a negative contribution to the energy from the two ferromagnetic top bonds, and a positive contribution from all the other broken bonds. A moment of reflection shows that in the ground state $S_{N-1,1} \leqslant S_{N-1,2}$. Furthermore, within each sub-block, the spin configuration must be according to lemma 1, since otherwise the positive part of the energy could be lowered by reshuffling the spins. We therefore search for the ground state among the metastable states obeying the two restrictions above, which we shall call 'low energy' metastable states. Fortunately there are only few of these, as shown by the next lemma.

Lemma 2. Let

$$C(\lambda) = \frac{2^{\lambda(N-1)}(1+2^{-\lambda}) - 2^{\lambda+1}}{2^{\lambda}-1} \qquad \lambda < \tfrac{1}{2}. \tag{A2.1}$$

There are no low energy metastable states for

$$S_{N-1,1} > \max\left(2^{\lambda}C(\lambda), \frac{(2^{1-\lambda}-1)C(\lambda)}{(2^{1-\lambda}-1)^2-(2^{\lambda}-1)}\right) \tag{A2.2}$$

unless

$$S_{N-1,2} = 2^{N-1} \qquad \text{and} \qquad S_{N-1,1} \geqslant (2^{1-\lambda}-1)(2^{N-1}-C(\lambda)). \tag{A2.3}$$

In other words, for any low energy metastable state, the relative polarisation $2^{-N}S_{N,1}$ is for small λ, either close to zero or to one.

Before proving the statement we need a little technical lemma, which is useful for getting sharp bounds.

Lemma 3. Consider a spin configuration generated by the rule of lemma 1, (all antiferromagnetic bonds). The energy change associated with the polarisation of one more spin pair

$$\Delta E = 2 \sum_{l=1}^{N-1} S_l 2^{-l\lambda} + 2 \tag{A2.4}$$

obeys the bound

$$\Delta E \leqslant 2\left(2 \frac{2^{-\lambda} - 2^{-N\lambda}}{1 - 2^{-\lambda}} + \frac{2^{-\lambda N} - 2^{-N-\lambda+1}}{2^{1-\lambda} - 1} S_{N,1} + 1 \right) \tag{A2.5}$$

where, for each l, S_l is the polarisation of the block which is a neighbour to the block containing the spin pair which is to be polarised.

Proof. It follows from lemma 1 that

$$S_l \leqslant \tfrac{1}{2} S_{l+1} + 1 \leqslant \frac{S_{l+n}}{2^n} + \sum_{i=0}^{n-1} 2^{-i}. \tag{A2.6}$$

Equation (A2.5) now follows by taking $l + n = N$ and by summing a geometric series.

We now proceed with the proof of lemma 2.

Proof. The energy change associated with the polarisation of a spin pair in the rightmost quarter of the system is

$$\Delta E_R = 2\left(1 + \sum_{l=1}^{N-3} S_l 2^{-\lambda l} - 2^{-\lambda(N-2)} S_{N-2,3} - 2^{-\lambda(N-1)} S_{N-1,1} \right). \tag{A2.7}$$

Using lemma 3 and the inequalities $S_{N-2,3} \geqslant \frac{1}{2} S_{N-1,2}$ and $S_{N-2,4} \leqslant \frac{1}{2} S_{N-1,2}$ we get the bound

$$\Delta E_R \leqslant 2\left[\frac{1 + 2^{-\lambda} - 2^{-(N-2)\lambda+1}}{1 - 2^{-\lambda}} + 2^{-\lambda(N-1)} \left(\frac{2^{\lambda} - 1}{2^{1-\lambda} - 1} S_{N-1,2} - S_{N-1,1} \right) \right]. \tag{A2.8}$$

The condition $\lambda < \frac{1}{2}$ ensures that $(2^{\lambda} - 1)/(2^{1-\lambda} - 1) < 1$, which is necessary to make the term in the inner parentheses negative, since as previously noted for any ground-state configuration $S_{N-1,1} \leqslant S_{N-1,2}$. By rearranging the terms we find that the energy change is negative for

$$S_{N-1,2} \leqslant S^M_{N-1,2} = S_{N-1,1} \frac{2^{1-\lambda} - 1}{2^{\lambda} - 1} - C(\lambda) \frac{2^{1-\lambda} - 1}{1 - 2^{-\lambda}}. \tag{A2.9}$$

Moreover $S^M_{N-1,2} > 0$ requires

$$S_{N-1,1} > C(\lambda) 2^{\lambda} \tag{A2.10}$$

which is the first condition on $S_{N-1,1}$.

Clearly, as long as (A2.9) and (A2.10) are satisfied, the energy can be lowered by increasing the polarisation of the right block. This process must stop when the critical

value $S^M_{N-1,2}$ is reached. At this point we break a bond in the left block, at an energy cost

$$\Delta E_l = 2\left(1 + \sum_{l=1}^{N-2} S_l 2^{-\lambda l} - 2^{-\lambda(N-1)} S^M_{N-1,2}\right) \tag{A2.11}$$

which by lemma 3 is bounded by

$$\Delta E_l \leq 2\left(\frac{1+2^{-\lambda}-2^{-(N-1)\lambda+1}}{1-2^{-\lambda}} + \frac{2^{-\lambda(N-1)}-2^{-N+2-\lambda}}{2^{1-\lambda}-1} S_{N-1,1} - 2^{-\lambda(N-1)} S^M_{N-1,2}\right). \tag{A2.12}$$

We now require that ΔE_l be negative, and by rearranging the terms we find:

$$S_{N-1,1} < (2^{1-\lambda}-1)(S^M_{N-1,2} - C(\lambda)). \tag{A2.13}$$

Inserting (A2.9) into (A2.13) finally yields

$$S_{N-1,1} > \left(\frac{(2^{1-\lambda}-1)2^{-\lambda}}{(2^{1-\lambda}-1)^2-(2^{\lambda}-1)}\right) C(\lambda) 2^{\lambda} \tag{A2.14}$$

which is, together with equation (A2.10) equivalent to (A1.5). The factor in parentheses is less than unity for $\lambda < 0.1$ and in general of order one. If (A1.5) is satisfied, the energy can be further lowered by polarising the left block further. This in turn allows $S^M_{N-1,2}$ to increase, which again means that $S_{N-1,1}$ can increase even more. The game comes to an end when $S^M_{N-1,2} = 2^{N-1}$ and $S_{N-1,1}$ obeys (A2.13). The lemma is thus proved. Next we check that the ground state has a high relative polarisation $2^{-N} S_{N,1}$.

Lemma 4. If $S_{N-1,1}$ is of the order of $2^{-\lambda(N-1)}$ then the system is not in its ground state.

Proof. The lemma follows from a comparison with the energy of the completely polarised state, which is $O(2^{N-1} 2^{(N-2)(-\lambda)})$, and negative for small λ, as shown by a simple calculation. If $S_{N-1,1}$ is $O(2^{-\lambda(N-1)})$ then $S_{N-1,2}$ is of the same order, according to (A2.9) and so the energy is of order $2^{-3\lambda(N-1)}$ which is much higher than in the completely polarised state.

In conclusion we have proved that the ground state of the system is very close to being completely polarised, i.e.

$$S_{N-1,2} = 2^{N-1} \qquad S_{N-1,1} \simeq (2^{1-\lambda}-1)(2^{N-1} - C(\lambda)).$$

For any $\lambda \neq 0$, $C(\lambda)$ is completely negligible compared to 2^{N-1} in the limit $N \to \infty$, and when λ is close to zero $2^{1-\lambda} - 1$ is close to one.

The last issue we shall discuss is the effect on the Parisi function of the approximation we have done by postulating a complete polarisation of the block. The overlap is not ± 1, but rather $\pm(1-\Delta)$ where $\Delta = 2(1 - S_{N,1} 2^{-N})$, i.e. we get a little shift in the position of the δ peaks. Since physical quantities are integrals over the overlap density, the error will be negligible provided that the quantity we are interested in changes smoothly in a little region close to one. This is clearly the case for all the moments of the Parisi function.

As the last part of this appendix, we present now the proof of lemma 1.

Proof. We note first that according to the prescription the overall polarisation cannot arise as the difference between two contributions of oppositely oriented blocks. Such

configurations have indeed higher energy than the ground state, as shown by the following simple argument. If $S_{N,1} = S_{N-1,1} - |S_{N-1,2}|$, the numerical value of both $S_{N-1,1}$ and $S_{N-1,2}$ can be decreased by two units without changing $S_{N,1}$. The ensuing energy change is

$$\Delta E = -2\left(\sum_{l=0}^{N-2} S_l^{\mathrm{R}} 2^{-\lambda l} + \sum_{l=0}^{N-2} S_l^{\mathrm{L}} 2^{-\lambda l} - [S_{N-1,1} + (S_{N-1,2} - 2)] 2^{-\lambda(N-1)} \right) \tag{A2.15}$$

where the notation S_l^{R} and S_l^{L} is analogous to the one used in lemma 3. The superscripts R and L refers to the right and left sub-blocks, and all the quantities can be taken as positive. Since

$$\sum_{l=0}^{N-2} S_l^{\mathrm{L}} 2^{-\lambda l} \geqslant 2^{-\lambda(N-2)} (S_{N-1,1} - 1) \tag{A2.16}$$

and a similar equation holds for the S_l^{R} we find

$$\Delta E \leqslant -2(S_{N,1} - 2)(2^{-\lambda(N-2)} - 2^{-\lambda(N-1)}) \leqslant 0 \tag{A2.17}$$

which rules out the possibility of such a configuration arising. As a consequence $S_{N,1}$ is a sum of positive contributions from broken antiferromagnetic bonds. Now let $S(p)$ be the statement that the lemma is true for $S_{N,1} \leqslant p$. $S(4)$ can be checked by direct computation. We now prove the implication $S(p) \rightarrow S(p+2)$, completing the proof by induction.

If $S(p)$ is assumed, a polarisation $S_{N,1} = p+2$ has been built according to the lemma. We now add the next two units of polarisation in such a way that the energy gain ΔE is minimised. This gives the minimum energy of the new configuration. There are two possible cases according to whether $(p+2)/2$ is (i) even or (ii) odd. In both cases the alternative is whether to place the next polarised pair within $S_{N-1,1}$ or $S_{N-1,2}$. All the subsequent choices, i.e. within the main sub-blocks, are by the induction hypothesis, determined by the rule.

(i) In this case $S_{N-1,1} = S_{N-1,2} = (p+2)/2$. The first choice cannot therefore influence ΔE, which proves $S(p+2)$.

(ii) Here we have $S_{N-1,1} = (p+2)/2 + 1$ and $S_{N-1,2} = (p+2)/2 - 1$ (or vice versa). We can either increase $S_{N-1,1}$ (violating the rule) or $S_{N-1,2}$ (complying to it). Let ΔE^{T} and ΔE^{R} be the energy change in the two cases. We now prove that $\Delta E^{\mathrm{T}} > \Delta E^{\mathrm{R}}$. Let $S_{N-1}^{\mathrm{R}} = S_{N-1,1}$, $S_{N-1}^{\mathrm{T}} = S_{N-1,2}$ and S_l^{R} (S_l^{T}) be the sub-block of $S_{N-1,2}$ ($S_{N-1,1}$) which, for each l, is closest to the block in which the last bond has been broken. Then

$$\Delta E^{\mathrm{T}} - \Delta E^{\mathrm{R}} = 2 \sum_{l=0}^{N-1} (S_l^{\mathrm{T}} - S_l^{\mathrm{R}}) 2^{-\lambda l}. \tag{A2.18}$$

In order to show that the right-hand side of (A2.18) is always positive we make the following remarks.

(1) $$\sum_{l=0}^{N-2} S_l^{\mathrm{T}} = S_{N-1,2} \qquad \sum_{l=0}^{N-2} S_l^{\mathrm{R}} = S_{N-1,1}$$

and

$$\sum_{l=0}^{N-2} S_l^{\mathrm{T}} - S_l^{\mathrm{R}} = 2. \tag{A2.19}$$

(2) All the terms in (A2.18) are either zero or two, i.e. they are all zero except one, say $S_q^T - S_q^R = 2$, for $q \leq N-2$. This can be shown by the following simple argument. Suppose for concreteness that $S_{N-2}^T = S_{N-2,1}$ and $S_{N-2}^R = S_{N-2,3}$. Then *either* $S_{N-2}^T = S_{N-2}^R + 2$ and $S_{N-2,2} = S_{N-2,4}$ *or* $S_{N-2}^T = S_{N-2}^R$ and $S_{N-2,2} = S_{N-2,4} + 2$. (This follows from the fact that no two blocks differ by more than two units, and that $S_{N-1,1} = S_{N-1,2} + 2$.) Since now S_{N-3}^T and S_{N-3}^R are sub-blocks of $S_{N-2,2}$ and $S_{N-2,4}$ respectively, in the first case we have proved our statement for $q = N-2$, while in the second we have postponed the problem to a lower level. Clearly, due to (A2.18), this second case cannot happen for all levels, which completes the argument.

Returning now to equation (A2.17) we can rewrite it as

$$\Delta E^T - \Delta E^R = 2[-2(2)^{-\lambda(N-1)} + 2(2)^{-\lambda q}] > 0 \tag{A2.20}$$

which completes the proof.

References

Anderson J A 1972 *Math. Biosci.* **14** 197
Baker G A 1972 *Phys. Rev.* B **5** 2622
Derrida B 1980 *Phys. Rev. Lett.* **45** 79
—— 1981 *Phys. Rev.* B **24** 2613
Dyson F J 1969 *Commun. Math. Phys.* **12** 91
—— 1971 *Commun. Math. Phys.* **21** 269
Feller W 1970 *An Introduction to Probability theory and its applications* vol I and II (New York: Wiley)
Gross D J and Mezard M 1984 *Preprint*
Hertz J and Sibani P 1985 *Phys. Scr.* **9** 199
Hopfield 1982 *Prod. Natl. Acad. Sci. USA* **79** 2554
Huberman B and Hogg T 1984 *Phys. Rev. Lett.* **52** 1048
Jensen M H, Bak P and Bohr T 1983 *Phys. Rev. Lett.* **50** 1637
Kotliar G, Anderson P W and Stein D L 1983 *Phys. Rev.* B **27** 602
McKay S and Berker A N 1984 *Phys. Rev.* B **29** 1315
McKay S, Berker A N and Kirkpatrick S 1982 *Phys. Rev. Lett.* **48** 767
Mezard M, Parisi G, Sourlas N, Toulouse G and Virasoro M 1984a *Phys. Rev. Lett.* **52** 1156
—— 1984b *J. Physique* **45** 843
Palmer R G and Hertz J A 1983 unpublished
Parisi G 1979 *Phys. Rev. Lett.* **43** 1754
—— 1980 *J. Phys. A: Math. Gen.* **13** L115, 1101, 1887
—— 1983 *Phys. Rev. Lett.* **50** 1946
Sherrington D and Kirkpatrick S 1975 *Phys. Rev. Lett.* **35** 1792
Young A P 1983 *Phys. Rev. Lett.* **51** 1206