

Metastable states in spin glasses

To cite this article: A J Bray and M A Moore 1980 *J. Phys. C: Solid State Phys.* **13** L469

View the [article online](#) for updates and enhancements.

You may also like

- [A new tachometer](#)
- [A new design of motor headlight](#)
A Whitehead
- [Classified](#)

LETTER TO THE EDITOR

Metastable states in spin glasses

A J Bray and M A Moore

Department of Theoretical Physics, The University, Manchester M13 9PL, UK

Received 17 April 1980

Abstract. The number of solutions of the equations of Thouless, Anderson and Palmer is obtained as a function of temperature. The density of solutions with a given free energy is calculated for free energies greater than a (temperature-dependent) critical value.

Despite substantial effort during the past five years, a satisfactory mean field theory of spin glasses has not yet appeared. The replica method (Edwards and Anderson 1975, Sherrington and Kirkpatrick 1975 (SK)) requires, for an adequate description of the low-temperature phase, a breaking of the replica symmetry (de Almeida and Thouless 1978, Blandin 1978, Bray and Moore 1978, Parisi 1979). An alternative approach is provided by the mean field equations of Thouless, Anderson and Palmer (1977, referred to as TAP), which are exact for the long-range interactions of the SK model. In this Letter we show that there are a large number of solutions (of order $\exp(\alpha N)$, where N is the number of Ising spins in the system) of the TAP equations below T_c . One is interested in the distribution of solutions over free energy, and averages of observables over solutions or over solutions of given free energy. Since the solutions possess a range of free energies, all but those with the lowest free energy correspond to metastable states. We shall find that solutions with free energies exceeding a (temperature-dependent) critical value are *uncorrelated*, and their distribution over free energy is calculated exactly. Solutions of lower free energy are correlated, with the Edwards–Anderson order parameter (here called \hat{q}) as a measure of the correlation. Our approach is generalisable (at least for $T = 0$) to two- and three-dimensional systems with realistic interactions. The numerical work of Morgenstern and Binder (1980) suggests that for such systems the critical energy might coincide with the ground state energy.

We start from the set of N TAP equations for the magnetisation m_i of the i th spin:

$$\begin{aligned} G_i &\equiv \tanh^{-1} m_i + \beta^2 J^2 (1 - q) m_i - \beta \sum_j J_{ij} m_j = 0 \\ &\equiv g(m_i) - \beta \sum_j J_{ij} m_j \end{aligned} \quad (1)$$

with their associated free energy (divided by $Nk_B T$)

$$\begin{aligned} f &= -(\beta/N) \sum_{(ij)} J_{ij} m_i m_j - \frac{1}{4} \beta^2 J^2 (1 - q)^2 \\ &\quad + (1/2N) \sum_i \{ (1 + m_i) \ln[\frac{1}{2}(1 + m_i)] + (1 - m_i) \ln[\frac{1}{2}(1 - m_i)] \} \end{aligned} \quad (2)$$

where $\sum_{(ij)}$ means a sum over distinct pairs, $q = N^{-1} \sum_i m_i^2$ and J_{ij} is a random exchange interaction with probability distribution

$$P(J_{ij}) = (N/2\pi J^2)^{1/2} \exp(-NJ_{ij}^2/2J^2).$$

Substituting for $\sum_j J_{ij} m_j$ in equation (2) from equation (1), we express f as a sum of single-site terms:

$$f = N^{-1} \sum_i f(m_i) = N^{-1} \sum_i \left[-\ln 2 - \frac{1}{4} \beta^2 J^2 (1 - q^2) + \frac{1}{2} m_i \tanh^{-1} m_i + \frac{1}{2} \ln(1 - m_i^2) \right]. \quad (3)$$

The density of solutions associated with a particular free energy f is therefore

$$N_s(f) = N^2 \int_0^1 dq \int_{-1}^1 \prod_i (dm_i) \delta\left(Nq - \sum_i m_i^2\right) \delta\left(Nf - \sum_i f(m_i)\right) \times \prod_i \delta(G_i) |\det \mathbf{A}| \quad (4)$$

where $|\det \mathbf{A}|$ is the Jacobian normalising the delta functions:

$$A_{ij} = \partial G_i / \partial m_j = [(1 - m_i^2)^{-1} + \beta^2 J^2 (1 - q)] \delta_{ij} - \beta J_{ij} \equiv a_i \delta_{ij} - \beta J_{ij}. \quad (5)$$

(A term $-2\beta^2 J^2 m_i m_j / N$, which comes from differentiating q in equation (1), is negligible as $N \rightarrow \infty$ and has been dropped from equation (5)). Since we will find that the determinant is always positive, we will drop the modulus signs henceforth. Introducing integral representations for the delta functions gives

$$N_s(f) = N^2 \int_0^1 dq \int_{-\infty}^{\infty} \frac{d\lambda}{2\pi i} \int_{-\infty}^{\infty} \frac{du}{2\pi i} \int_{-\infty}^{\infty} \prod_i \left(\frac{dx_i}{2\pi i} \right) \int_{-1}^1 \prod_i (dm_i) \exp \left\{ -N(\lambda q + uf) + \lambda \sum_i m_i^2 + u \sum_i f(m_i) + \sum_i x_i g(m_i) - \beta \sum_{(ij)} J_{ij} (x_i m_j + x_j m_i) \right\} \times \det \mathbf{A} \{J_{ij}\}. \quad (6)$$

We wish now to average over the bond distribution. Since we anticipate that $N_s(f) \sim \exp(\alpha N)$, we should strictly average the extensive quantity $\ln N_s(f)$. This can be done by introducing replicas (see later). However, for the region (of free energies) in which the solutions are uncorrelated, the two types of averaging lead to the same final result. Therefore we construct the 'direct average'

$$\langle N_s(f) \rangle_J = \int \prod_{(ij)} (dJ_{ij} P(J_{ij})) N_s(f). \quad (7)$$

Terms involving J_{ij} are of the form

$$\begin{aligned} & \int_{-\infty}^{\infty} \prod_{(ij)} [dJ_{ij} (N/2\pi J^2)^{1/2}] \exp \left[-N \sum_{(ij)} J_{ij}^2 / 2J^2 - \beta \sum_{(ij)} J_{ij} (x_i m_j + x_j m_i) \right] \det \mathbf{A} \{J_{ij}\} \\ &= \exp \left[(\beta^2 J^2 / 2N) \sum_{(ij)} (x_i m_j + x_j m_i)^2 \right] \langle \det \mathbf{A} \{J_{ij} - (\beta J^2 / N) \\ & \quad \times (x_i m_j + x_j m_i)\} \rangle_J \end{aligned} \quad (8)$$

where the second line follows from the first by a simple translation of each integration variable. This translation is irrelevant as far as the final average in equation (8) is concerned, since it introduces a term of order N^{-1} into the matrix element A_{ij} (such terms are negligible and were already dropped from equation (5)). Thus equation (8) may be rewritten

$$\exp\left[\frac{1}{2}\beta^2 J^2 q \sum_i x_i^2 + (\beta^2 J^2 / 2N) \left(\sum_i x_i m_i\right)^2\right] \langle \det \mathbf{A} \{J_{ij}\} \rangle_J. \quad (9)$$

To compute the average of the determinant one can introduce replicas and use the representation

$$\det \mathbf{A} = \int_{-\infty}^{\infty} \prod_{i,\alpha} \left[\frac{d\xi_{i\alpha}}{(2\pi)^{1/2}} \right] \exp\left(-\frac{1}{2} \sum_{i,j,\alpha} \xi_{i\alpha} A_{ij} \xi_{j\alpha}\right) \quad (10)$$

where the replica labels α run from 1 to m , and analytic continuation to $m = -2$ is required at the end of the calculation. The J_{ij} integrals are gaussian and give

$$\begin{aligned} \langle \det \mathbf{A} \rangle_J = & \int_{-\infty}^{\infty} \prod_{i,\alpha} \left[\frac{d\xi_{i\alpha}}{(2\pi)^{1/2}} \right] \exp\left[-\frac{1}{2} \sum_{i,\alpha} a_i \xi_{i\alpha}^2 + \frac{\beta^2 J^2}{4N} \sum_{\alpha} \left(\sum_i \xi_{i\alpha}\right)^2\right. \\ & \left. + \frac{\beta^2 J^2}{2N} \sum_{\alpha < \beta} \left(\sum_i \xi_{i\alpha} \xi_{i\beta}\right)^2\right]. \end{aligned} \quad (11)$$

The squared terms are simplified by the Hubbard–Stratonovich identity

$$\begin{aligned} \exp(a^2/2) = & \int_{-\infty}^{\infty} (dx/\sqrt{2\pi}) \exp(-x^2/2 + ax). \\ \langle \det \mathbf{A} \rangle_J = & \int_{-\infty}^{\infty} \prod_{\alpha} \left[\left(\frac{N}{\pi}\right)^{1/2} dR_{\alpha} \right] \int_{-\infty}^{\infty} \prod_{\alpha < \beta} \left[\left(\frac{N}{2\pi}\right)^{1/2} dT_{\alpha\beta} \right] \int_{-\infty}^{\infty} \prod_{i,\alpha} \left(\frac{d\xi_{i\alpha}}{\sqrt{2\pi}} \right) \\ & \times \exp\left\{-\frac{1}{2} \sum_{i,\alpha} a_i \xi_{i\alpha}^2 - N \sum_{\alpha} R_{\alpha}^2 - \frac{N}{2} \sum_{\alpha < \beta} T_{\alpha\beta}^2 + \beta J \sum_{i,\alpha} R_{\alpha} \xi_{i\alpha}^2\right. \\ & \left. + \beta J \sum_{i,\alpha < \beta} T_{\alpha\beta} \xi_{i\alpha} \xi_{i\beta}\right\}. \end{aligned} \quad (12)$$

The integrals over $\{R_{\alpha}\}$ and $\{T_{\alpha\beta}\}$ are eventually performed by steepest descents. We adopt the solution $R_{\alpha} = R$ (for all α), $T_{\alpha\beta} = 0$ (for all (α, β)). One can show (details will be presented elsewhere) that this is the stable stationary point. With this choice the integrals over the $\xi_{i\alpha}$ are trivial and yield, after setting $m = -2$ and dropping multiplicative prefactors,

$$\langle \det \mathbf{A} \rangle_J = \prod_i (a_i - 2\beta J R) \exp(2NR^2) \quad (13)$$

with R to be determined variationally.

Assembling the various terms, using a further Hubbard–Stratonovich identity to simplify the term in $(\sum_i x_i m_i)^2$ in equation (9), and dropping multiplicative prefactors, yields

$$\begin{aligned} \langle N_s(f) \rangle_J = & \max \int_{-1}^1 \prod_i (dm_i) \int_{-\infty}^{\infty} \prod_i \left(\frac{dx_i}{2\pi i} \right) \exp\left\{N(-\lambda q - uf - \frac{1}{2}V^2 + 2R^2)\right. \\ & \left. + \frac{1}{2}\beta^2 J^2 q \sum_i x_i^2 + \sum_i x_i \tilde{g}(m_i) + \lambda \sum_i m_i^2 + u \sum_i f(m_i)\right\} \prod_i (a_i - 2\beta J R) \end{aligned} \quad (14)$$

where $\tilde{g}(m_i) = g(m_i) + \beta J V m_i$ and max indicates the maximum over the variables q, λ, u, V, R . Finally we set $V = -\beta J(1 - q) - \Delta/\beta J$ and $2R = \beta J(1 - q) - B/\beta J$, and integrate over the x_i to obtain the final result

$$\langle N_s(f) \rangle_J = \max \exp \{ N [-\lambda q - uf - (B + \Delta)(1 - q) + (B^2 - \Delta^2)/2\beta^2 J^2 + \ln I] \} \quad (15)$$

where

$$I = \int_{-1}^1 \frac{dm}{(2\pi)^{1/2} \beta J q^{1/2}} \left(\frac{1}{1 - m^2} + B \right) \exp \left[-\frac{(\tanh^{-1} m - \Delta m)^2}{2\beta^2 J^2 q} + \lambda m^2 + uf(m) \right] \quad (16)$$

and the maximum is taken over the five variables q, λ, u, Δ, B . The use of steepest descents is justified, in the limit $N \rightarrow \infty$, by the factor N inside the exponent in equation (15). The five stationarity equations become, after some manipulation,

$$\begin{aligned} q &= \langle m^2 \rangle, & f &= \langle f(m) \rangle \\ 0 &= B \{ 1 - \beta^2 J^2 \langle (1 - m^2)^2 / [1 + B(1 - m^2)] \rangle \} \\ \Delta &= -\frac{1}{2} \beta^2 J^2 (1 - q) + \langle m \tanh^{-1} m \rangle / 2q \\ \lambda &= B + \Delta - [1 - \langle (\tanh^{-1} m - \Delta m)^2 \rangle / \beta^2 J^2 q] / 2q \end{aligned} \quad (17)$$

where angular brackets mean an average over a probability distribution for m given by the integrand of equation (16) divided by I . The solutions of Sherrington and Kirkpatrick (1975) and of Sommers (1978) correspond to $B = \Delta = 0 = u = \lambda$ and to $B + \Delta = 0 = u = \lambda$ respectively. For both solutions, equation (14) gives $\langle N_s(f) \rangle_J = 1$. Both solutions are known to be unstable (de Almeida and Thouless 1978, Bray and Moore 1980, de Dominicis and Garel 1979).

Before proceeding further we observe that the third of equations (17) admits the

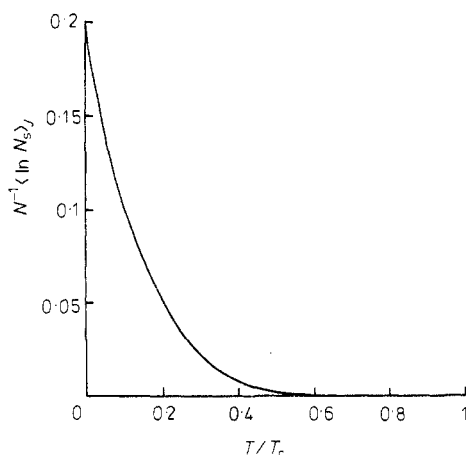


Figure 1. Logarithm of the total number of TAP solutions, divided by N , as a function of temperature. For T close to T_c , $N^{-1} \langle \ln N_s \rangle_J = \frac{8}{81} t^6 + O(t^7)$, where $t = 1 - T/T_c$, and the curve is therefore indistinguishable from the temperature axis for $T \gtrsim 0.6 T_c$.

solution $B = 0$. This is the solution we adopt since it may be shown (details to be presented elsewhere) that the other choice leads to an unstable stationary point. For $B = 0$, our expression for the determinant, equation (13), becomes

$$\langle \det \mathbf{A} \rangle_J = \prod_i (1 - m_i^2)^{-1} \exp(\frac{1}{2} N \beta^2 J^2 (1 - q)^2)$$

a positive quantity. This justifies *a posteriori* our earlier assumption that the modulus signs on the Jacobian can be dropped. The positivity of $\det \mathbf{A}$ suggests that nearly all solutions of the TAP equations are *minima* of the TAP free energy. This lends credence to our identification of TAP solutions as metastable states. At $T = 0$ the TAP equations reduce to $m_i = \text{sgn}(\sum_j J_{ij} m_j)$, and the identification is unambiguous in the sense that all solutions are stable against single spin-flips.

For the case $B = 0$, we have solved equations (17) numerically over the entire temperature range $0 \leq T \leq T_c$. The total number of solutions N_s is obtained by setting $u = 0$ (this removes the constraint imposed by the second delta function in equation (4)). The result is plotted in the form $N^{-1} \langle \ln N_s \rangle_J$ against T/T_c in figure 1 (the justification for placing the average outside the logarithm is given later). The variables q, λ, Δ, f thus obtained from equations (17) are appropriate to averages over all the TAP solutions. Close to T_c , analytic solutions may be obtained:

$$\begin{aligned} N^{-1} \langle \ln N_s \rangle_J &= (8/81)t^6 + O(t^7) \\ q &= t + t^2 - \frac{5}{9}t^3 + O(t^4) \\ \Delta &= \frac{2}{3}t^2 + \frac{10}{9}t^3 + O(t^4) \\ \lambda &= \frac{2}{3}t^2 + \frac{8}{9}t^3 + O(t^4) \\ f &= -\ln 2 - \frac{1}{4}\beta^2 J^2 + \frac{1}{6}t^3 + \frac{11}{24}t^4 + (79/90)t^5 + O(t^6) \end{aligned} \quad (18)$$

where $t = 1 - T/T_c$. These values are, in fact, characteristic of the overwhelming majority of all TAP solutions, since the integral for the total number

$$N_s = \int df N_s(f) = \int df \exp\{N(N^{-1} \ln N_s(f))\}$$

is dominated by a single value of f . (The equilibrium free energy, on the other hand, the lowest free energy for which TAP solutions exist, is determined from $\langle \ln N_s(f) \rangle_J = 0$ and cannot be calculated with the present methods.)

At $T = 0$, the number of metastable states may be obtained as the zero-temperature limit of the present theory, or by working directly with the equations $m_i = \text{sgn}(\sum_j J_{ij} m_j)$. Either method gives $N^{-1} \langle \ln N_s \rangle_J = 0.1992$ at $T = 0$, a result obtained independently by Tanaka and Edwards (1980) and by de Dominicis *et al* (1980), who have also obtained some of equations (18).

Consider now the effect of taking 'logarithmic' rather than 'direct' averages—in principle we should calculate $\langle \ln N_s(f) \rangle_J$ rather than $\ln \langle N_s(f) \rangle_J$, since $\ln N_s(f)$ is proportional to the size N of the system. The calculation utilises the standard replica trick, $\ln N_s = \lim_{n \rightarrow 0} (N_s^n - 1)/n$, and $\langle N_s^n(f) \rangle_J$ is calculated via simple generalisation of equation (4). Assuming no replica symmetry-breaking, one needs to introduce three new order parameters \hat{q}, η, ρ which couple to quantities which are off-diagonal in the

replica space. Details of the calculation will be presented elsewhere. The final result is (setting $B = 0$ as before)

$$N^{-1} \langle \ln N_s(f) \rangle_J = \max \left\{ -\Delta(1-q) - \Delta^2/2\beta^2 J^2 + \rho[\beta J(1-q) + \Delta/\beta J] \right. \\ \left. - \lambda q - uf - \frac{1}{2}\eta(q - \hat{q}) + \iint_{-\infty}^{\infty} \frac{dx dy}{2\pi} \exp[-\frac{1}{2}(x^2 + y^2)] \ln I \right\} \quad (19)$$

where

$$I = \int_{-1}^1 \frac{dm}{(2\pi)^{1/2} \beta J (q - \hat{q})^{1/2}} \frac{1}{1 - m^2} \exp \left\{ -\frac{(\tanh^{-1} m - \Delta m + \beta J \hat{q}^{1/2} x)^2}{2\beta^2 J^2 (q - \hat{q})} \right. \\ \left. + \lambda m^2 + uf(m) + \frac{m}{(\hat{q})^{1/2}} [\rho x + (\eta \hat{q} - \rho^2)^{1/2} y] \right\}. \quad (20)$$

This reduces to our previous result, equations (15) and (16), with $B = 0$, when η, ρ, \hat{q} are set zero (in that order). The order parameters q and \hat{q} have the physical significance

$$q = \langle m_i^2 \rangle_s, \quad \hat{q} = \langle m_i \rangle_s^2 \quad (21)$$

where $\langle \rangle_s$ means an *average over solutions* with the given free energy f . Thus it is \hat{q} , rather than q , which should be regarded as the Edwards–Anderson order parameter $\langle S_i^2 \rangle_J$ for this problem.

Setting to zero the derivatives of $N^{-1} \langle \ln N_s(f) \rangle_J$ with respect to $q, \Delta, \lambda, u, \hat{q}, \eta, \rho$ determines these parameters as a function of f . There are trivial solutions with $\hat{q} = q$ corresponding once more to the solutions of Sherrington and Kirkpatrick and of Sommers, although the interpretation is now different (for example, evaluation of $N^{-1} \langle \ln N_s \rangle_J$ for the Sommers solution now gives a *negative* result). Solution of the seven coupled equations for the non-trivial solutions is a formidable task, even numerically. We believe, however, that replica symmetry breaking will be required as soon as the off-diagonal order parameters \hat{q}, η, ρ become non-zero.

The free energy below which off-diagonal order develops can be determined by analysing the stability of equations (17) against off-diagonal fluctuations. Choosing u (instead of f) as independent variable for convenience we find (details elsewhere) that, provided u exceeds a critical value $u_c = -\frac{5}{6}t + O(t^2)$, or equivalently that f satisfies

$$f \geq f_c = -\ln 2 - \frac{1}{4}\beta^2 J^2 + \frac{1}{6}t^3 + \frac{11}{24}t^4 + (133/180)t^5 + O(t^6)$$

there is no off-diagonal ordering. This justifies our previous analysis of the case $u = 0$ (and that of Tanaka and Edwards 1980). The vanishing of the off-diagonal order parameters has the implication, via equation (20), that the TAP solutions for $f \geq f_c$ are *uncorrelated*. For $f < f_c$ the off-diagonal order parameters become non-zero and correlations develop between solutions. Close to T_c one finds, for $u \geq u_c$,

$$N^{-1} \langle \ln N_s(f) \rangle_J = \frac{8}{81}t^6 - \frac{1}{12}u^2 t^4$$

$$q = t + t^2 - \frac{5}{9}t^3 + \frac{1}{3}ut^2$$

$$f = -\ln 2 - \frac{1}{4}\beta^2 J^2 + \frac{1}{6}t^3 + \frac{11}{24}t^4 + (79/90)t^5 + \frac{1}{6}ut^4.$$

If one assumes that there is no breaking of the replica symmetry, one can solve equations (19) and (20) for $u < u_c$ and choose u such that $\langle \ln N_s(f) \rangle_J = 0$, corresponding to the lowest free energy consistent with the existence of TAP solutions. One then finds

$q = t + O(t^2)$ and $\hat{q} = 0.3471t + O(t^2)$, whereas intuitively one expects $\hat{q} = q$ for this lowest free energy. This is the origin of our belief that replica symmetry breaking is needed in this region, although a detailed stability analysis is needed to verify this.

For $T = 0$, we have calculated $\langle N_s(E) \rangle_J$, where $N_s(E)$ is the density of metastable states with energy NE , by averaging over solutions of the equations $m_i = \text{sgn}(\sum_j J_{ij} m_j)$.

A stability analysis shows that off-diagonal order parameters vanish (i.e. the metastable states are uncorrelated) for $E \geq E_c = -0.672J$, and $N^{-1} \langle \ln N_s(E_c) \rangle_J = 0.1254$. The energy for which the density of metastable states is maximal is $E_m = -0.506J$, and $N^{-1} \langle \ln N_s(E_m) \rangle_J = 0.1992$. The maximum energy of metastable states is $E_u = -0.286J$. The complete function $N^{-1} \langle \ln N_s(E) \rangle_J$ is given in figure 2. If the theory without off-diagonal order parameters is continued into the unstable region (broken curve in figure 2) one finds that $N^{-1} \ln \langle N_s(E) \rangle_J$ vanishes at $E_l = -0.791J$.

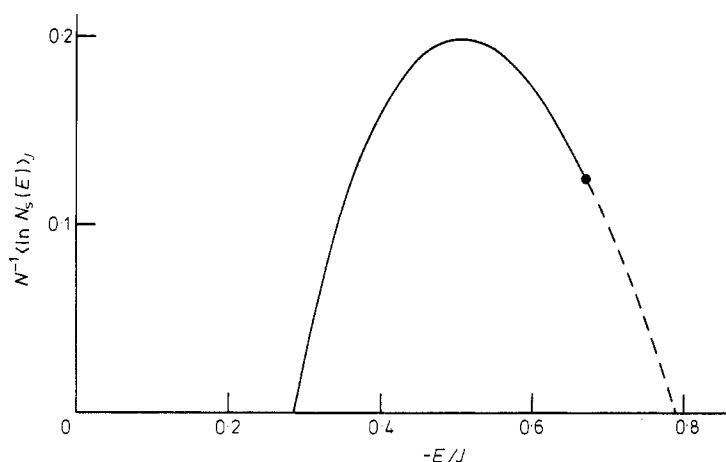


Figure 2. Logarithm of the density of metastable (i.e. one spin-flip stable) states, divided by N , for the SK model at $T = 0$. The broken curve corresponds to the range of energies for which the 'direct average' is unstable against off-diagonal fluctuations.

We conclude by noting a remarkable relationship between the results of the present work and those of a replica symmetry breaking scheme (the 'two-group model') introduced earlier by the authors (Bray and Moore 1978). If $\langle Z^n \rangle_J$ is calculated within the 'two-group model' (Z is the partition function of the SK model) and n is set to zero, one obtains not unity but a function identical to $\langle N_s \rangle_J$, the number of solutions of the TAP equations calculated above. We do not as yet understand the significance, if any, of this result.

During the final stages of this work we learned of the work of de Dominicis *et al* (1980). Their results agree with ours where they overlap, but the detailed results presented in figures 1 and 2, and the conclusion that the 'direct average' gives exact results for a range of free energies, are new here. We thank them for sending us a preprint of this work, and also T Garel for a copy of his thesis. We also wish to thank M G James and R Banach for useful discussions.

References

- de Almeida J R L and Thouless D J 1978 *J. Phys. A: Math. Gen.* **11** 983
Blandin A 1978 *J. Physique* **39** C6 1499
Bray A J and Moore M A 1978 *Phys. Rev. Lett.* **41** 1068
— 1980 *J. Phys. C: Solid St. Phys.* **13** 419
de Dominicis C, Gabay M, Garel T and Orland H 1980 *J. Physique* submitted
de Dominicis C and Garel T 1979 *J. Physique* **40** L575
Edwards S F and Anderson P W 1975 *J. Phys. F: Metal Phys.* **5** 965
Morgenstern I and Binder K 1980 preprint
Parisi G 1979 *Phys. Rev. Lett.* **43** 1754
Sherrington D and Kirkpatrick S 1975 *Phys. Rev. Lett.* **32** 1792
Sommers H-J 1978 *Z. Phys.* **B31** 301
Tanaka F and Edwards S F 1980 preprints
Thouless D J, Anderson P W and Palmer R G 1977 *Phil. Mag.* **35** 593