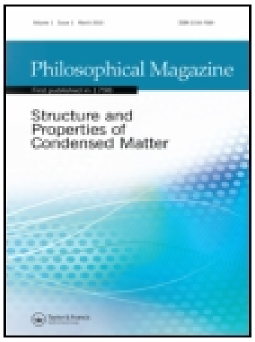


Taylor & Francis
Taylor & Francis Group

Philosophical Magazine

ISSN: 0031-8086 (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/tphm19

Solution of 'Solvable model of a spin glass'

D. J. Thouless , P. W. Anderson & R. G. Palmer

Philosophical Magazine, 1977, Vol. 35, No. 3, 593–601

Solution of 'Solvable model of a spin glass'

By D. J. Thouless
Department of Mathematical Physics, University of Birmingham, Birmingham, England

and P. W. Anderson†‡ and R. G. Palmer
Department of Physics, Princeton University, Princeton, New Jersey 08540, U.S.A.‡

[Received 12 October 1976]

Abstract

The Sherrington–Kirkpatrick model of a spin glass is solved by a mean field technique which is probably exact in the limit of infinite range interactions. At and above T_c the solution is identical to that obtained by Sherrington and Kirkpatrick (1975) using the $n \rightarrow 0$ replica method, but below T_c the new result exhibits several differences and remains physical down to $T=0$.

§ 1. Introduction

Sherrington and Kirkpatrick (1975) (SK) have proposed an idealized model of a spin glass which apparently allows an exact formal solution. Unfortunately, the solution is non-physical at low temperatures, leading in particular to a negative zero-point entropy. We present here a new solution of the SK model which behaves sensibly at low temperatures, while agreeing with the SK solution at and above the critical temperature T_c. Our analysis is based on a high temperature expansion, supplemented below T_c by a mean field theory which takes into account not only the average spin on each site, but also the mean square fluctuation from this average.

The Sherrington–Kirkpatrick Hamiltonian

$$\mathscr{H} = - \sum_{(ij)} J_{ij} S_i S_j \tag{1}$$

describes N Ising spins ($S_i = \pm 1$) interacting in pairs (ij) via infinite-range Gaussian-random exchange interactions :

$$\text{Prob}\,(J_{ij}) \propto \exp\left(\frac{-Z J_{ij}^2}{2\tilde{J}^2}\right) \tag{2}$$

with a variance $\tilde{J}^2/Z$ where Z is the number of neighbours of each spin, presumed effectively infinite ; we work in the limit $N \gtrsim Z \gg 1$. The $Z^{-1/2}$ dependence of the interactions is necessary to ensure a sensible thermodynamic limit. We consider only the case in which J_{ij} has zero mean, setting SK's J_0 parameter to zero. We also set $k_B = 1$ and $\beta = 1/T$ throughout.

---

† Also at Bell Laboratories, Murray Hill, New Jersey 07974, U.S.A.

‡ Work at Princeton University partially supported by National Science Foundation Grant No. DMR76–00886.

The SK solution follows the approach of Edwards and Anderson (1975) using the well-known trick of replacing $\ln Z$ by $\lim_{n\to 0}(Z^n-1)/n$ and regarding Z^n (for integer n) as the partition function of n replicas of the original system. This allows one to perform the average over the J distribution *before* taking the spin trace. It seems necessary, however, to use the thermodynamic limit $N\to\infty$ before taking $n\to 0$, and it is this improper reversal of limits that leads to SK's erroneous solution (R. G. Palmer, to be published). We therefore avoid the replica method and turn to a different approach.

§ 2. The high temperature region

For $T>T_c$ we make a high temperature series expansion for the free energy, using the standard identity

$$\exp(\beta J_{ij}S_iS_j)=\cosh\beta J_{ij}(1+S_iS_j\tanh\beta J_{ij}). \tag{3}$$

Thus

$$\begin{aligned}-\beta F&=\langle\ln \operatorname{Tr}\exp(-\beta\mathscr{H})\rangle_J\\ &=\langle\ln\prod_{(ij)}\cosh\beta J_{ij}\rangle_J+\langle\ln \operatorname{Tr}\prod_{(ij)}(1+T_{ij}S_iS_j)\rangle_J\\ &=N\beta^2\tilde{J}^2/4+0(N/Z)\\ &\quad+\langle\ln \operatorname{Tr}(1+\sum_{(ij)}T_{ij}S_iS_j+\tfrac{1}{2}\sum_{(ij)\neq(kl)}T_{ij}T_{kl}S_iS_jS_kS_l\ldots)\rangle_J,\end{aligned} \tag{4}$$

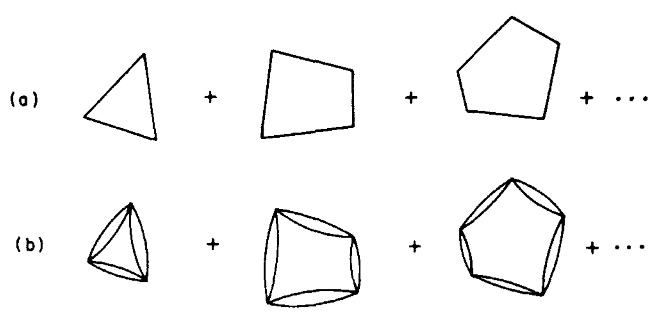
where $T_{ij}=\tanh\beta J_{ij}$. The expansion may be analysed diagramatically (each line representing a T_{ij}), noting the following conditions for a non-vanishing diagram :

(*a*) There must be an even number of lines at each vertex.
(*b*) No line may be double *before* taking the logarithm.
(*c*) Every line must be double *after* taking the logarithm (because $\langle J\rangle=0$).

We find no terms of order N (except the trivial $N\ln 2$), and a summable series in order N/Z, consisting of simple polygons (fig. 1 (*a*)) which become double (fig. 1 (*b*)) after taking the logarithm. We thus obtain

$$F=Nf_0+(N/Z)f_1+\text{lower order},$$

Fig. 1



The most important diagrams in order N/Z before (*a*) and after (*b*) taking the logarithm.

with

$$f_0 = -T \ln 2 - \tilde{J}^2/4T, \tag{5}$$

and

$$f_1 = -\tfrac{1}{4}T \ln (1-\beta^2\tilde{J}^2) + \text{non-singular part.}$$

The extensive part of the free energy is identical to the SK result for $T \geqslant T_c$. The contribution of order $N/Z \sim 1$ is negligible in the thermodynamic limit for $T > \tilde{J}$, but diverges as $T = T_c = \tilde{J}$ signaling the transition found by Edwards and Anderson (1975) and SK. We note that the divergent part f_1 is intrinsically positive, in contrast to the corresponding result of conventional mean field theory. This reflects the fact that the spin glass transition is a *blocking* effect of interference among the different interactions ; the free energy below T_c is *greater* than an analytic continuation of the high temperature result.

§ 3. Derivation of the mean field equation

Below T_c we must introduce a mean field in order to reconverge the series for F. We employ the usual identity

$$\operatorname{Tr} \exp(-\beta\mathscr{H}) = \operatorname{Tr} \exp(-\beta\mathscr{H}_0)\langle \exp(\beta\mathscr{H}_0 - \beta\mathscr{H})\rangle_{\mathscr{H}_0}, \tag{6}$$

where $\mathscr{H}_0$ is a soluble mean field Hamiltonian which is to be used in evaluating the diagrams generated by $\exp(\beta\mathscr{H}_0 - \beta\mathscr{H})$. An obvious ansatz is

$$(\mathscr{H}_0 - \mathscr{H})_{ij} = J_{ij}(S_i - m_i)(S_j - m_j) \tag{7}$$

so that

$$(\mathscr{H}_0)_{ij} = J_{ij}(m_i m_j - m_i S_j - m_j S_i)$$

where m_i is the mean spin on the ith site, to be determined self-consistently by the condition

$$\langle S_i \rangle_{\mathscr{H}_0} = m_i. \tag{8}$$

Ignoring the perturbation $\mathscr{H} - \mathscr{H}_0$ leads to the appealing (but incorrect) mean field equation

$$h_i = \sum_j J_{ij} m_j = T \tanh^{-1} m_i \tag{9}$$

which would imply a critical temperature of $2\tilde{J}$, since the largest eigenvalue of a Gaussian-random matrix (Mehta 1967) is $(J_\lambda)_{\max} = 2\sqrt{(NJ^2)} = 2\tilde{J}$. However, the series generated by $\exp(\beta\mathscr{H}_0 - \beta\mathscr{H})$ is still divergent with this choice for $\mathscr{H}_0$, showing that it is essential to consider correlations between spin fluctuations on different sites.

It is possible to proceed diagramatically (Thouless, Anderson, Lieb and Palmer, unpublished report), removing the most divergent diagrams by manipulating the $\mathscr{H}_0/\mathscr{H}$ separation. However, it is simpler, and perhaps more physical, to observe that the set of diagrams contributing in order N are just those which would remain on a ' Bethe lattice ' or Cayley tree ; all diagrams with no loops. Any diagram containing a closed ring is necessarily of order N/Z or less, since the internal connection reduces the number of site summations by one. The Bethe method (1935) is exact for the Ising model on a Cayley tree, and should therefore solve this problem wherever the order N/Z terms are convergent (and hence ignorable).

In the Bethe method, we consider a 'cluster' of a central site 0 and all its neighbours j. On the neighbours j we assume mean fields h_j which, for a Cayley tree, are the only effect *their* neighbours can have on them. Using the smallness of J_{0j} ($\propto Z^{-1/2}$), it is easy to arrive at the following expressions for m_0 and m_j:

$$\left.\begin{aligned} m_0 &= \tanh \beta \sum_j J_{0j} \tanh \beta h_j \\ m_j &= \tanh \beta h_j + m_0 \beta J_{0j}(1-\tanh^2 \beta h_j). \end{aligned}\right\} \tag{10}$$

We may now eliminate the h_js (again using the smallness of J_{0j}), obtaining the fundamental equation

$$\sum_j J_{0j} m_j - m_0 \beta \sum_j J_{0j}^2 (1-m_j^2) = T \tanh^{-1} m_0 \tag{11}$$

which supplants the incorrect eqn. (9), and must, of course, be valid for any choice of site 0. The correction term proportional to m_0 is more readily understood upon realizing that $\beta(1-m_j^2)$ is the single-site susceptibility, χ_j, as may easily be proved. Equation (11) may thus be written

$$m_0 = \tanh \beta \sum_j J_{0j}(m_j - m_0 J_{0j} \chi_j) \tag{12}$$

and the second term on the right-hand side is seen as the response of site j to the mean spin on site 0; this must be *removed* from m_j when computing m_0.

The corresponding free energy is not easily obtained from the Bethe method, and we therefore present it as a *fait accompli*, originally derived by diagram expansion:

$$F_{\mathrm{MF}} = -\sum_{(ij)} J_{ij} m_i m_j - \tfrac{1}{2}\beta \sum_{(ij)} J_{ij}^2 (1-m_i^2)(1-m_j^2) + \tfrac{1}{2} T \sum_i \left[(1+m_i) \ln \tfrac{1}{2}(1+m_i) + (1-m_i) \ln \tfrac{1}{2}(1-m_i)\right] \tag{13}$$

As it must, direct differentiation of eqn. (13) gives eqn. (11). Additionally, eqn. (13) is quite physically transparent: the first term is the internal energy of a frozen lattice; the second term is the correlation energy of the fluctuations, and is just the $N\tilde{J}^2/4T$ term of eqn. (5), modified for the effective 'freedom', $1-m_i^2$, of each spin; and the last term is the entropy of a set of Ising spins constrained to have means m_i.

We emphasize that the Bethe method, the use of the Cayley tree, and the resulting eqns. (11) and (13) are only meaningful if the terms of order N/Z (and lower) are convergent. Evaluation of these terms is rather awkward and will be discussed in detail elsewhere. The only simple region is near T_c, where we find

$$\tilde{J}(1-\overline{m^2}) \leqslant T \quad (T \sim T_c) \tag{14}$$

as the convergence criterion for the N/Z diagrams, and hence as a validity condition for our mean field theory.

Our problem is now reduced to finding solutions to the mean field eqn. (11), subject to the convergence condition (14), and then to using eqn. (13) to obtain the thermodynamics. This programme is not much easier than the original problem, since J_{ij} is a random matrix, and eqns. (11) and (13) hold for an individual realization, *not* for an ensemble average. We have been able to find solutions both near T_c and near $T=0$. In both cases the solutions involve some numerical conjectures checked by machine simulation, so that while we are reasonably certain of the general form of the solution in both regions, they are far from complete analyses. We also encounter some 'coincidences' which require further investigation. Details of our solutions will be given elsewhere, and we attempt here only a general description of the methods.

§ 4. The critical region

For T near T_c we expect m_i to be small and similar to the eigenvector M_i belonging to the largest eigenvalue $(J_\lambda)_{\max}=2\tilde{J}$ of the matrix J_{ij} :

$$\sum_j J_{ij}M_j=2\tilde{J}M_i \tag{15}$$

We first linearize eqn. (11), approximating $\sum_j J_{ij}{}^2\chi_j$ by $\tilde{J}^2\bar{\chi}$:

$$\sum_j J_{ij}m_j=\beta\tilde{J}^2(1-\overline{m^2})m_i+T(m_i+m_i{}^3/3+m_i{}^5/5+\dots). \tag{16}$$

We then expand m_i about M_i

$$m_i=M_i+\delta m_i, \tag{17}$$

chosing the R.M.S. amplitude

$$q=\overline{M_i^2} \tag{18}$$

of M_i such that M_i is orthogonal to δm_i. The components M_i have a Gaussian distribution, as may be proved from the invariance of a Gaussian-random matrix (with suitable diagonal elements) under orthogonal transformations. Using this fact to take a scalar product of eqn. (16) with M_i, we obtain

$$(2\tilde{J}-\beta\tilde{J}^2-T)q=(T-\beta\tilde{J}^2)q^2+3Tq^3+T\sum_i M_i{}^3\delta m_i+0(q^4). \tag{19}$$

The term in δm_i is essential—there is no solution without it—but is difficult to estimate. Analysing the projection of eqn. (16) orthogonal to M_i by a combination of eigenvector expansions and numerical estimates, we find finally

$$(2\tilde{J}-\tilde{J}^2/T-T)q-(T-\tilde{J}^2/T)q^2+(2T^2/\tilde{J}-3T)q^3=0. \tag{20}$$

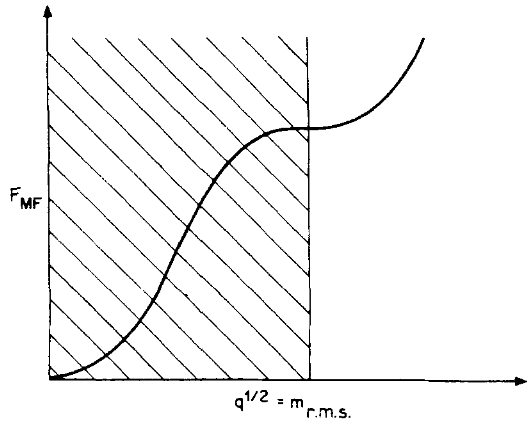
Near $T_c=\tilde{J}$ this equation has a double zero at

$$q=\overline{m^2}=1-T/T_c \tag{21}$$

which gives the mean field free energy, eqn. (13), a *saddle point* at T_c.

Near T_c the form of $F_{\rm MF}$ is very complicated and quite unlike that for typical phase transitions (fig. 2), and it is not at all surprising that the Edwards–Anderson and SK continuations come out on a wrong branch of the free energy function. It is important to note that $F_{\rm MF}$ is *not* a genuine

Fig. 2



The form of the mean field free energy for T slightly below T_c. The N/Z terms diverge in the shaded region.

free energy functional in the Ginzburg–Landau sense, in that the convergence condition (14) restricts the freedom of q, and in particular eliminates the spurious minimum at $q=0$ as soon as T falls below T_c. This behaviour, the q^3 term, and the saddle point in F_{MF}, are very reminiscent of the heuristic free energy functional of Harris, Lubensky and Chen (1976).

As far as we can see, our solution deviates only in higher order from SK near T_c. The cusp of the specific heat is the same, as is the T-dependence of $\overline{m^2}=q$.

§ 5. The low temperature region

In the low temperature regime, $T \ll T_c$, our analysis is based on the probability distributions of the fields $h_i = \sum_j J_{ij} m_j$ and the mean spins, m_i. *At* $T=0$ the mean field equation obviously selects a self-consistent lowest energy solution of

$$m_i = \text{sign}\,(h_i), \tag{22}$$

and we have generated a large number of such solutions numerically to investigate the distribution of h_i. We find that Prob $(|h_i|)$—hence written $p(h)$—becomes linear for small h as $N \to \infty$ (there is a finite offset $p(0)$ at finite N). As a by-product of this study, we find a ground state energy of $U_0 = -(0{\cdot}755 \pm 0{\cdot}010)\tilde{J}N$, which is certainly different from SK's value of $U_0 = -(2/\pi)^{1/2}\tilde{J}N = -0{\cdot}80\tilde{J}N$.

To derive the low temperature thermodynamics we assume

$$\lim_{h \to 0} p(h) = h/H^2 \tag{23}$$

and

$$q = \overline{m^2} = 1 - \alpha(T/\tilde{J})^2 \quad (T \ll T_c), \tag{24}$$

where H and α are parameters to be determined later. Equation (24) is easily justified *a posteriori*. Again approximating the $J_{ij}{}^2\chi_j$ term, m_i and h_i are related by

$$h_i = \alpha T m_i + T \tanh^{-1} m_i \tag{25}$$

and the definition

$$m^2 = \int_0^\infty m^2(h)p(h)\,dh \tag{26}$$

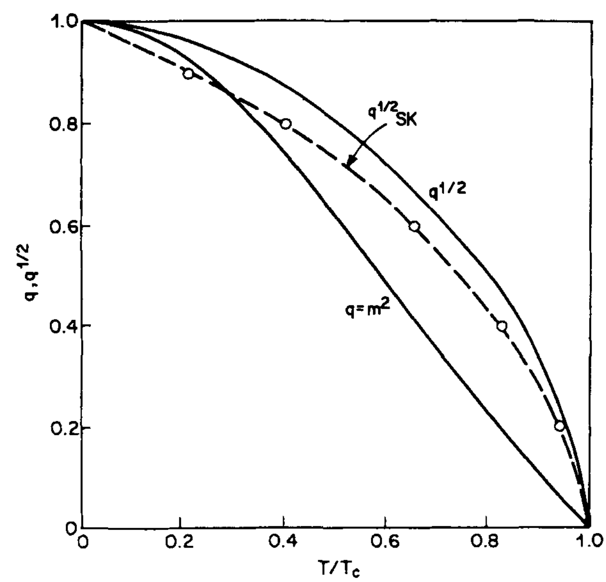
leads after some integration to

$$H^2/\tilde{J}^2 = \tfrac{1}{4}\alpha + (2 \ln 2 + 1)/3 + \ln 2/\alpha, \tag{27}$$

which leaves only one unknown parameter, α. The minimum acceptable value for H is $H = 1{\cdot}28\tilde{J}$ and we believe on the basis of our numerical work and general considerations that H is actually equal to this limiting value, giving

$$\alpha = 2\sqrt{\ln 2} = 1{\cdot}665. \tag{28}$$

Fig. 3



The order parameter $q = m^2$ as a function of temperature. The circles and broken line are the results of SK.

Figure 3 shows the resulting order parameter, q, fitted smoothly to the SK result near T_c. The convergence criterion, eqn. (14), is easily satisfied at low temperatures, but there are corrections to this criterion away from T_c. We suspect, but have not yet proved, that the solution coincides with the *true* convergence criterion, and the F_{MF} has the saddle point form sketched in fig. 2 at *all* temperatures below T_c, thus giving a line of critical points.

We may now calculate the entropy from eqn. (13):

$$S = -(1/2T^2)\sum J_{ij}^2(1-m_i^2)(1-m_j^2)$$
$$-\sum_i\left[\left(\frac{1+m_i}{2}\right)\ln\left(\frac{1+m_i}{2}\right)+\left(\frac{1-m_i}{2}\right)\ln\left(\frac{1-m_i}{2}\right)\right]$$
$$= -\frac{N\alpha^2}{4}(T/\tilde{J})^2 - N\int_0^\infty [\ldots]p(h)\,dh$$
$$= 0{\cdot}770N(T/\tilde{J})^2. \qquad (29)$$

The low temperature specific heat is thus quadratic:

$$C = 1{\cdot}54N(T/\tilde{J})^2. \qquad (30)$$

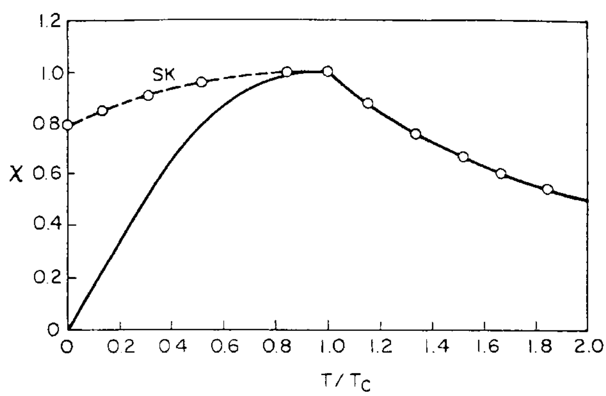
We have performed some finite temperature Monte Carlo simulations of this spin glass model and find a specific heat consistent with this result.

The low temperature susceptibility

$$\chi = \bar{\chi}_j = 1{\cdot}665T/\tilde{J} \qquad (31)$$

is linear in T, in contrast to SK's which has a finite zero-temperature intercept (fig. 4).

Fig. 4



The temperature dependence of the susceptibility, according to the present work (solid line) and SK (circles and broken line).

§ 6. Conclusion

In conclusion, we believe that we have shown that a *consistent* mean field theory of the Sherrington–Kirkpatrick 'solvable model' can be constructed. We believe that this mean field theory represents the actual thermodynamic behaviour of the model accurately to order $1/Z$. The infinite range interactions seem necessary at present to make the model tractable, but also make it somewhat unrealistic. We therefore caution others against any literal comparison of these model results with experiment, but emphasize that our solution strongly supports the essential conclusions of the Edwards–Anderson spin glass theory, that a sharp thermodynamic transition into a totally randomly ordered state can and does occur.

Acknowledgments

We wish to thank Elliot Lieb, David Sherrington, and Scott Kirkpatrick for many illuminating and helpful discussions, with particular gratitude to Kirkpatrick for keeping us abreast of his numerical work, which in several cases confirmed that referred to in the text.

We are also grateful to the Aspen Center for Physics for hospitality and assistance in the final preparation of this manuscript.

References

Bethe, H., 1935, *Proc. R. Soc.* A, **151,** 540 ; see Fowler, R. H., and Guggenheim, E. A., 1960, *Statistical Thermodynamics* (Cambridge University Press) for a good discussion.

Edwards, S. F., and Anderson, P. W., 1975, *J. Phys.* F, **5,** 965.

Harris, A. B., Lubensky, T. C., and Chen, J.-H., 1976, *Phys. Rev. Lett.*, **36,** 415.

Mehta, M. L., 1967, *Random Matrices and the Statistical Theory of Energy Levels* (London, New York : Academic Press).

Sherrington, D., and Kirkpatrick, S., 1975, *Phys. Rev. Lett.*, **35,** 1792.