The replica method and solvable spin glass model

1979 J. Phys. A: Math. Gen. 12 563

(http://iopscience.iop.org/0305-4470/12/4/016)

J. Phys. A: Math. Gen., Vol. 12, No. 4, 1979. Printed in Great Britain

The replica method and a solvable spin glass model

J L van Hemmen† and R G Palmer
Duke University, Durham, N.C. 27706, USA

Received 14 July 1978, in final form 13 September 1978

Abstract. The replica method for random systems is critically examined, with particular emphasis on its application to the Sherrington–Kirkpatrick solution of a 'solvable' spin glass model. The procedure is improved and extended in several ways, including the avoidance of steepest descents and a reformulation which isolates the thermodynamic limit $N \to \infty$. Ideas of analyticity and convexity are employed to investigate the two most dubious steps in the replica method: the extension from an integer number (n) of replicas to real n in the limit $n \to 0$, and the reversal of the limits in n and N. The latter step is proved valid for the Sherrington–Kirkpatrick problem, while the non-uniqueness of the former is held responsible for the unphysical behaviour of the result.

1. Introduction

The central problem of equilibrium statistical mechanics is the calculation of the free energy, $F_N(\beta)$, of a system of N particles at inverse temperature $\beta = 1/kT$, and the subsequent use of the thermodynamic limit, $N \to \infty$, to derive the free energy per particle

$$f(\beta) = \lim_{N \to \infty} F_N(\beta)/N.$$

The existence of the thermodynamic limit is either proved or, more commonly, assumed on physical grounds. Exact solutions for $f(\beta)$ are known for a few non-trivial systems, but in most cases one has to resort to approximate methods or tricks. The thermodynamic limit is normally an integral part of such methods, being a requirement for their validity rather than a final hurdle. It is usually needed in the evaluation of the partition function, Z_N, *before* a logarithm is taken to derive $-\beta F_N(\beta) = \ln Z_N$, because it justifies the neglect of factors in Z_N of lower than exponential order in N.

This familiar scheme is severely upset by the introduction of randomness. In a random system the Hamiltonian, and hence $F_N(\beta)$, contains some random parameters representing the interactions or quenched configuration in a given sample. There are typically $O(N)$ or $O(N^2)$ random parameters, which we assume *independent*. In a few cases the final result, $f(\beta)$, can be shown to be the same for almost every such sample (van Hemmen, unpublished), but one usually has to average over all possible samples by averaging $F_N(\beta)$ over some probability distribution for the random parameters. This

† Present address: Institut für Angewandte Mathematik, Universität Heidelberg, Im Neuenheimer Feld 294, 6900 Heidelberg 1, Federal Republic of Germany

average must be computed *after* taking the logarithm, but *before* taking the thermodynamic limit, giving

$$-\beta f(\beta) = \lim_{N\to\infty} N^{-1}\langle \ln Z_N(\beta)\rangle. \tag{1}$$

We will always use the symbol $\langle\ldots\rangle$ to denote averaging over the randomness. The thermodynamic limit can no longer be used directly in the evaluation of $Z_N(\beta)$ and most conventional methods lose their validity. In addition, performing the average $\langle\ldots\rangle$ over many random parameters seems hopeless because the logarithm prevents any useful factorisation into a product of one-parameter averages.

The second of these problems is largely summounted by the 'replica trick' in which one replaces $\langle \ln Z\rangle$ by

$$\lim_{n\to 0} (\langle Z^n\rangle - 1)/n.$$

For positive integer n the computation of $\langle Z^n\rangle$ is frequently feasible, the average $\langle\ldots\rangle$ now being made possible by factorisation. Unfortunately, this introduces a new problem; the extension from positive integer n to real n in the neighborhood of $n=0$. Additionally, it has not proved possible to avoid the first problem, that of requiring the thermodynamic limit in the evaluation of Z or $\langle Z^n\rangle$ itself. One needs to employ the limit $N\to\infty$ *before* the limit $n\to 0$ in order to perform the trace implicit in Z, but the reverse order is clearly specified by the above framework. Previous authors have assumed that the limits may simply be interchanged, a question that we will examine in some detail.

The replica trick has a long history, dating back at least to Hardy *et al* (1934 § 6.8) as an identity for computing the average of a logarithm, but has become well known only recently since its application to the spin glass problem by Edwards and Anderson (1975); see also Kac 1968 and Lin 1970. In some applications, such as the spherical model spin glass treated by Kosterlitz *et al* (1976), the replica trick gives exactly the correct result obtainable by other methods, while in other cases the replica trick leads to a result that is definitely incorrect. The best example of this behaviour is in the solution of an Ising spin glass model by Sherrington and Kirkpatrick (1975), which we refer to as SK; their solution exhibits a negative entropy at low temperature. The same model has since been solved by Thouless *et al* (1977), giving physically reasonable low temperature predictions quite different from those of SK. The model therefore seems well defined, but SK's application of the replica trick fails entirely.

Several reasons may be proposed for the failure of the SK solution, including the reversal of the limits on N and n (SK and Palmer, unpublished), the passage from positive integer n to $n\approx 0$ (Klein 1977), the stability of SK's stationary point in their steepest descent method (de Almeida and Thouless 1978), and the steepest descent method itself (see comment before equation (26)). It is the purpose of this paper to address these potential problems, both for the replica trick in general and for the SK solution in particular. Although failing to find any useful *general* criteria for the effectiveness or otherwise of the replica trick, we are able to shed some light on its possible failures, and on the particular problems of the SK solution; these we blame on the extension from integer to real n. We succeed in avoiding the steepest descent method altogether, and in making the thermodynamic limit explicit. In passing, we also extend the work of SK to probability distributions for the random variables more general than Gaussian.

In § 2 we consider the replica trick in general and propose a three step process for its execution. The first step is essentially the calculation of $(\ln\langle Z_N^n\rangle)/N$ in the thermodynamic limit, for positive integer n; the second is the passage from integer n to n in the neighbourhood of zero; and the third is the proof of the equivalence of the result to that desired with the N and n limits reversed. The first step is performed for the standard SK problem in § 3, and for the SK problem with generalised probability distributions in § 4. The second and third steps are treated in § 5 from the viewpoint of analyticity, and in § 6 from the viewpoint of convexity. Our conclusions are summarised in § 7.

2. The replica method

Let us consider a spin system of N sites, with a Hamiltonian $\mathcal{H}_N$ that contains some random parameters. An example is the SK Hamiltonian, equation (10), in which the random parameters are the $\frac{1}{2}N(N-1)$ pair interaction strengths, J_{ij}, but in this section we consider no specific model. We are interested in calculating the partition function

$$Z_N(\beta)=\mathrm{Tr}\,\exp(-\beta\mathcal{H}_N) \tag{2}$$

and hence the *averaged* free energy per particle, $f(\beta)$, in the thermodynamic limit according to equation (1). For a classical Ising system, the trace in equation (2) is simply a sum over all spin configurations. We normally abbreviate $Z_N(\beta)$ to Z_N or Z.

The conventional replica trick employs the relation

$$\langle \ln Z\rangle=\lim_{n\to 0}\frac{\langle Z^n\rangle-1}{n} \tag{3}$$

but we find it more convenient to use the equivalent identity

$$\langle \ln Z\rangle=\frac{\mathrm{d}}{\mathrm{d}n}\ln\langle Z^n\rangle|_{n=0}. \tag{4}$$

There are several advantages of this formulation, as will soon become apparent. Defining

$$\phi_N(n)=N^{-1}\ln\langle Z_N^n\rangle \tag{5}$$

we must calculate

$$-\beta f(\beta)=\lim_{N\to\infty}N^{-1}\langle\ln Z\rangle=\lim_{N\to\infty}\left(\frac{\mathrm{d}}{\mathrm{d}n}\phi_N(n)|_{n=0}\right). \tag{6}$$

For positive integer n, apart from a factor $-\beta$, the quantity $\phi_N(n)$ is just the free energy per site of a replica system of N sites with n spins at each site. The replicas $1,2,\ldots,n$ all have the same random parameter set, but are uncoupled before averaging. The average $\langle\ldots\rangle$ in (5) is to be computed *before* taking the logarithm; this corresponds to an annealed system in which the random parameters may be regarded as thermodynamic variables with values weighted by their thermodynamic probability. We therefore expect $\phi_N(n)$ to be meaningful in the thermodynamic limit, at least for n a positive integer, and define

$$\phi(n)=\lim_{N\to\infty}\phi_N(n). \tag{7}$$

We also *assume*, for the present, that the limit in (7) exists for all *real* n. The function

$\phi_N(n)$ *is* defined, by (5), for real n and all finite N but it is not obvious that it tends to a limit at $N \to \infty$ unless n is a positive integer. The fact that it does indeed have a limit is central to the replica trick, and is essentially proved in § 6.

For positive integer n, we may expect to be able to evaluate $\phi(n)$ explicitly for a specific system, as we shall do in §§ 3 and 4 for the SK problem. The use of the thermodynamic limit will be essential in that evaluation, and in general our hopes for exact evaluation may be high for $\phi(n)$ but are vanishingly low for $\phi_N(n)$.

Unfortunately, the prescription (6) requires the differentiation at $n = 0$ *before* use of the thermodynamic limit. This naturally raises the question: Under what circumstances can we interchange the order of the limit $N \to \infty$ and the differentiation with respect to n at $n = 0$? Or, more formally,

$$\lim_{N\to\infty}\left(\frac{d}{dn}\phi_N(n)|_{n=0}\right) = \frac{\mathrm{d}}{\mathrm{d}n}(\lim_{N\to\infty}\phi_N(n))|_{n=0} \equiv \phi'(0)? \tag{8}$$

We note that both sides of (8) are well defined in the thermodynamic limit, in contrast to previous prescriptions (such as SK), which involve $\langle Z^n\rangle$ rather than $\langle Z^n\rangle^{N^{-1}}$.

The definition (5), and an examination of the computations required, gives us little confidence in the possibility of *explicitly* evaluating $\phi_N(n)$ for anything but positive integer n, and then only in the limit $N \to \infty$. We must therefore ask: Under what conditions can we extend $\phi(n)$ from positive integer n to real or complex n in the neighbourhood of $n = 0$? Symbolically,

$$\phi(n),\, n \in \mathbb{N} \to \phi(n), \qquad n \in \mathbb{R} \text{ or } \mathbb{C}? \tag{9}$$

There are many techniques that might appear useful in answering questions (8) and (9), but it seems in practice that these reduce to just two fundamental ideas. One is *analyticity*, to which we devote § 5, and the other is *convexity*, treated in § 6. Perhaps surprisingly, convexity turns out to be the more powerful tool.

Postponing the detailed discussion for subsequent sections, we see that we have effectively broken down the computation of $f(\beta)$ into a three step process:

(*a*) Calculate $\phi(n)$ for positive integer n;
(*b*) Find an extension of $\phi(n)$ to $n \approx 0$, and then compute $\phi'(0)$, showing that the value so obtained is unique;
(*c*) Prove that equation (8) is true, and hence that the result $\phi'(0)$ is indeed the required $-\beta f(\beta)$.

This procedure might justly be called the replica *method*, as opposed to the basic replica *trick* of equation (3) or (4).

We implement step (*a*) for the SK problem in §§ 3 and 4, and then return to a discussion of steps (*b*) and (*c*) in §§ 5 (analyticity) and 6 (convexity).

3. The Sherrington–Kirkpatrick problem

Sherrington and Kirkpatrick (1975; see also Kirkpatrick and Sherrington 1978) have proposed a model of a spin glass that apparently allows an exact solution for $f(\beta)$. Their solution uses the replica trick, and is subject to the general criticisms considered in § 2, as well as additional doubts about the steepest descent method employed. Further, the status of the thermodynamic limit is somewhat unclear because it is not made explicit.

We therefore rederive the essential results of SK in a simple but rigorous way, showing at the same time *why* the model is exactly soluble. We use the basic formulation of § 2 to make the thermodynamic limit explicit, and appeal to the molecular field techniques of van Hemmen (1978) and den Ouden *et al* (1976*a*, *b*) to avoid the steepest descent method.

The SK Hamiltonian may be written in the form

$$\mathscr{H}_N = -\frac{\tilde{J}}{N^{1/2}} \sum_{(ij)} J_{ij} S(i) S(j) - \frac{\tilde{J}_0}{N} \sum_{(ij)} S(i) S(j) - h \sum_i S(i). \tag{10}$$

It describes N Ising spins, $S(i) = \pm 1$, interacting with each other in pairs (ij) and with an external field h. Summations over pairs (ij) count each pair once only and exclude $i = j$ terms. The interactions J_{ij} are independent identically distributed random variables with mean zero and variance one. SK assumed a Gaussian probability distribution with density

$$p(J_{ij}) = (2\pi)^{-1/2} \exp(-\tfrac{1}{2} J_{ij}^2) \tag{11}$$

which we too adopt for now, before considering other distributions in the next section. We also simplify the model—without destroying the spin glass behaviour—by setting $\tilde{J}_0 = h = 0$; we return to the general case at the end of this section.

Let n be a positive integer. Then

$$Z^n = (\mathrm{Tr} \exp(-\beta \mathscr{H})^n = \mathrm{Tr} \exp \frac{\beta \tilde{J}}{N^{1/2}} \sum_{(ij)} J_{ij} \Big(\sum_{\alpha=1}^{n} S_\alpha(i) S_\alpha(j) \Big) \tag{12}$$

where α labels the n replicas and the trace, a finite sum, must now be taken over all nN spins $S_\alpha(i)$. We may immediately average over the randomness by integrating over the probability distribution (11), obtaining

$$\langle Z^n \rangle = \mathrm{Tr} \exp \frac{\gamma}{2N} \sum_{(ij)} \Big(\sum_{\alpha=1}^{n} S_\alpha(i) S_\alpha(j) \Big)^2 \tag{13}$$

where $\gamma = \beta^2 \tilde{J}^2$. A simple rearrangement gives

$$\langle Z^n \rangle = \exp(\tfrac{1}{4} \gamma n (N - n)) \, \mathrm{Tr} \exp \frac{\gamma}{2N} \sum_{(\alpha\beta)} \Big(\sum_{i=1}^{N} S_\alpha(i) S_\beta(i) \Big)^2 \tag{14}$$

and hence

$$\begin{aligned} \phi(n) &= \lim_{N \to \infty} \phi_N(n) = \lim_{N \to \infty} N^{-1} \ln \langle Z^n \rangle \\ &= \tfrac{1}{4} \gamma n - \gamma \lim_{N \to \infty} -(\gamma N)^{-1} \ln \mathrm{Tr} \exp \frac{\gamma}{2N} \sum_{(\alpha\beta)} \Big(\sum_{i=1}^{N} S_\alpha(i) S_\beta(i) \Big)^2. \end{aligned} \tag{15}$$

The limit in the second term is nothing but the free energy per site $f_n(\gamma)$ of a molecular field model at inverse temperature γ and Hamiltonian

$$\mathscr{H}_{\mathrm{MF}} = -\tfrac{1}{2} N \sum_{(\alpha\beta)} \Big(N^{-1} \sum_{i=1}^{N} S_\alpha(i) S_\beta(i) \Big)^2 \tag{16}$$

so that

$$-\gamma f_n(\gamma) = \lim_{N \to \infty} N^{-1} \ln \mathrm{Tr} \exp(-\gamma \mathscr{H}_{\mathrm{MF}}) \tag{17}$$

and

$$\phi(n)=\tfrac{1}{4}n\gamma-\gamma f_n(\gamma). \tag{18}$$

It is known (van Hemmen 1978, den Ouden *et al* 1976a, b) that a problem of the form (16)–(17) may be solved by the use of a simpler *effective* Hamiltonian, H, which is independent of N and gives *exactly* $f_n(\gamma)$, the quantity required *after* the thermodynamic limit, according to

$$-\gamma f_n(\gamma)=\ln \operatorname{Tr} \exp(-\gamma H). \tag{19}$$

It is precisely this 'removal' of the thermodynamic limit that makes the problem soluble.

The effective Hamiltonian appropriate to the present $\mathscr{H}_{\mathrm{MF}}$ may be written

$$H(\boldsymbol{q})=-\sum_{(\alpha\beta)}(q_{\alpha\beta}S_\alpha S_\beta-\tfrac{1}{2}q_{\alpha\beta}^2) \tag{20}$$

where $\boldsymbol{q}$ is a vector with $\frac{1}{2}n(n-1)$ components $q_{\alpha\beta}$, $1\leqslant\alpha<\beta\leqslant n$. $H(\boldsymbol{q})$ describes n interacting Ising spins S_α, with 'interactions' $q_{\alpha\beta}$ that must be chosen so that

(*a*) $q_{\alpha\beta}=\operatorname{Tr} S_\alpha S_\beta \exp(-\gamma H(\boldsymbol{q}))/\operatorname{Tr}\exp(-\gamma H(\boldsymbol{q}))$ (21)

(*b*) given (*a*), the $\boldsymbol{q}$-dependent free energy

$$f_n(\boldsymbol{q};\gamma)=-\gamma^{-1}\ln \operatorname{Tr}\exp(-\gamma H(\boldsymbol{q})) \tag{22}$$

be *minimal*.

The minimum is then the required $f_n(\gamma)$. Defining

$$\Xi_n(\boldsymbol{q})=\operatorname{Tr}\exp\Big(\gamma\sum_{(\alpha\beta)}q_{\alpha\beta}S_\alpha S_\beta\Big) \tag{23}$$

the condition (a) may be rewritten

$$\gamma q_{\alpha\beta}=\frac{\partial \ln \Xi_n(\boldsymbol{q})}{\partial q_{\alpha\beta}} \tag{24}$$

and equation (22) becomes

$$f_n(\boldsymbol{q};\gamma)=\tfrac{1}{2}\sum_{(\alpha\beta)}q_{\alpha\beta}^2-\gamma^{-1}\ln\Xi_n(\boldsymbol{q}). \tag{25}$$

We are now left with *one* problem, the minimisation of equation (25), because (24) is a necessary condition for an extremal value and may be dropped. We note that the condition (21) implies that minima are to be found only in the unit cube, $|q_{\alpha\beta}|\leqslant 1$ for all pairs $(\alpha\beta)$.

Surprisingly, the minimization of $f_n(\boldsymbol{q};\gamma)$ is highly non-trivial. Explicit differentiation shows that $\boldsymbol{q}=0$ is a *local* minimum when $\gamma<1$ and a local maximum when $\gamma>1$, but this is little help in locating the *absolute* minimum, even for $\gamma<1$. In Appendix A (much of which is based on the work of Elliott Lieb), we prove that the absolute minimum of $f_n(\boldsymbol{q};\gamma)$ is always realised on the positive diagonal, $q_{\alpha\beta}=q\geqslant 0$ for all pairs $(\alpha\beta)$. The minimisation is thus reduced to a one-dimensional problem along this diagonal. We also show in Appendix A that there are $2^{n-1}-1$ additional absolute minima elsewhere *equivalent* to any minimum on the positive diagonal (unless $\boldsymbol{q}=0$); however, since our prescription requires only one absolute minimum, these extra minima may safely be ignored.

Sherrington and Kirkpatrick *assumed* that all the analogous maxima in their steepest descent method were to be found on the positive diagonal, overlooking the

$2^{n-1}-1$ additional maxima. Logically however, the steepest descent approach requires the inclusion of *all* absolute maxima, and the result of the SK evaluation of $\langle Z^n\rangle$ should be multiplied by 2^{n-1}. Unfortunately this correction gives a divergence in the $n\to 0$ limit.

It is now easy to complete the calculation of $\phi(n)$. On the positive diagonal $q_{\alpha\beta}=q$,

$$\begin{aligned}\Xi_n(\boldsymbol{q}) &= \mathrm{Tr}\exp\Big(\gamma q\sum_{(\alpha\beta)}S_\alpha S_\beta\Big)\\ &= \exp(-\tfrac{1}{2}n\gamma q)\,\mathrm{Tr}\exp\bigg(\tfrac{1}{2}\gamma q\Big(\sum_{\alpha=1}^{n}S_\alpha\Big)^2\bigg)\\ &= \exp(-\tfrac{1}{2}n\gamma q)\int_{-\infty}^{\infty}\frac{\mathrm{d}z}{(2\pi)^{1/2}}\exp(-\tfrac{1}{2}z^2)(2\cosh(\gamma q)^{1/2}z)^n \end{aligned}\tag{26}$$

where the last equality follows from the identity

$$\exp(\lambda a^2)=\int_{-\infty}^{\infty}\frac{\mathrm{d}z}{(2\pi)^{1/2}}\exp(-\tfrac{1}{2}z^2)\exp[(2\lambda)^{1/2}az].\tag{27}$$

Equations (18), (25), and (27) now give, finally,

$$\phi(n)=\tfrac{1}{4}n\gamma\{1-2q-(n-1)q^2\}+\ln\int_{-\infty}^{\infty}\frac{\mathrm{d}z}{(2\pi)^{1/2}}\exp(-\tfrac{1}{2}z^2)(2\cosh(\gamma q)^{1/2}z)^n\tag{28}$$

where q may be determined by applying equation (24) along the positive diagonal and integrating by parts, giving the implicit equation

$$q=\frac{\displaystyle\int_{-\infty}^{\infty}\frac{\mathrm{d}z}{(2\pi)^{1/2}}\exp(-\tfrac{1}{2}z^2)z\tanh[(\gamma q)^{1/2}z]\cosh^n[(\gamma q)^{1/2}z]}{\displaystyle\int_{-\infty}^{\infty}\frac{\mathrm{d}z}{(2\pi)^{1/2}}\exp(-\tfrac{1}{2}z^2)\cosh^n((\gamma q)^{1/2}z)}.\tag{29}$$

When this equation has more than one solution, we must choose the one which gives the smallest value of $f_n(\boldsymbol{q};\gamma)$, and hence the largest value of $\phi(n)$. A numerical study (to be reported elsewhere) shows that this prescription gives $q=0$ when $\gamma\leqslant\gamma_c$, and $q>0$ when $\gamma>\gamma_c$, with γ_c a decreasing function of n. At $n=2$, $\gamma_c=1$ and q goes continuously to zero as $\gamma\to 1^-$, but for $n\geqslant 3$ there is a jump discontinuity in q at γ_c, and a corresponding discontinuity in $\partial f_n/\partial\gamma$, signalling a first order phase transition in the replica system.

If we simply treat n as a real variable in (28) and (29), which is certainly *one* extension from integer to real n, and assume the truth of equation (8), we arrive at precisely the SK result:

$$-\beta f(\beta)=(\tfrac{1}{2}\beta\tilde{J})^2(1-q)^2+\int_{-\infty}^{\infty}\frac{\mathrm{d}z}{(2\pi)^{1/2}}\exp(-\tfrac{1}{2}z^2)\ln[2\cosh(\beta\tilde{J}zq^{1/2})]\tag{30}$$

with

$$q=\int_{-\infty}^{\infty}\frac{\mathrm{d}z}{(2\pi)^{1/2}}\exp(-\tfrac{1}{2}z^2)\tanh^2(\beta\tilde{J}zq^{1/2}).\tag{31}$$

If $T\geqslant T_c\equiv\tilde{J}/k$ the only solution of equation (31) is $q=0$. When $T<T_c$ there is another solution, with $q>0$, that should presumably be selected, since it is a more natural continuation from the finite n results, and gives the larger value of $f(\beta)$.

The result for $\phi(n)$, equations (28) and (29), is rigorous, but the final SK result, equations (30) and (31), gives rise to a negative entropy at low temperatures (whichever solution is chosen for q), and is therefore certainly wrong. The fault *must* lie in the assumptions used between equations (29) and (30), which correspond to the questions (8) and (9) raised in § 2. We defer further discussion until §§ 5 and 6.

The generalisation to the full SK Hamiltonian, equation (10), presents no problem and will be sketched only briefly. The equivalent of equation (16) is

$$\mathscr{H}_{\mathrm{MF}} = -\tfrac{1}{2}\beta\tilde{J}^2 N \sum_{(\alpha\beta)} \left(N^{-1}\sum_i S_\alpha(i)S_\beta(i)\right)^2 - \tfrac{1}{2}\tilde{J}_0 N \sum_\alpha \left(N^{-1}\sum_i S_\alpha(i)\right)^2 - h\sum_\alpha\left(\sum_i S_\alpha(i)\right)^2 \tag{32}$$

and this time the inverse temperature is β itself. The corresponding effective Hamiltonian is

$$H(\boldsymbol{q}, \boldsymbol{m}) = -\beta\tilde{J}^2 \sum_{(\alpha\beta)} (q_{\alpha\beta}S_\alpha S_\beta - \tfrac{1}{2}q_{\alpha\beta}^2) - \tilde{J}_0 \sum_\alpha (m_\alpha S_\alpha - \tfrac{1}{2}m_\alpha^2) - h\sum_\alpha S_\alpha \tag{33}$$

where $q_{\alpha\beta}$ obeys the analogue of (21) and m_α satisfies

$$m_\alpha = \mathrm{Tr}\, S_\alpha \exp(-\beta H)/\mathrm{Tr}\exp(-\beta H). \tag{34}$$

Applying an extension of the theorem of Appendix A, we find that the appropriate minimum lies on the plane $q_{\alpha\beta} = q$, $m_\alpha = m$, where q and/or m may be zero in different regions of the phase diagram. The minimisation and evaluation then cause no difficulty and precisely reproduce the results of Kirkpatrick and Sherrington (1978).

4. Other probability distributions

Before continuing with our discussion of the replica method, we turn to other probability distributions, more general than the Gaussian used by SK. We show that almost any *even* probability distribution leads to the same results for the SK problem.

Let us retain the Hamiltonian (10), again taking $\tilde{J}_0 = h = 0$, but now let the J_{ij} be distributed according to an arbitrary *even* probability distribution with density $p(x)$. Equation (11) the Gaussian, represents one special case; another example is $p(x) = \tfrac{1}{2}\delta(x+1) + \tfrac{1}{2}\delta(x-1)$. As before, we assume of course that the J_{ij} are independent and identically distributed, and for convenience we take $\langle J_{ij}^2\rangle = 1$.

We require one further restriction on $p(x)$. We assume that

$$F_p(z) \equiv \langle\exp(xz)\rangle = \int_{-\infty}^{\infty} \exp(xz)p(x)\,\mathrm{d}x \tag{35}$$

is an *entire* function of z, implying the existence of all moments. When $z = \mathrm{i}t$, with t real, $F_p(z)$ is the characteristic function of p (Breiman 1968). For real z in the neighbourhood of the origin we have, by Taylor expansion of $\exp(xz)$,

$$F_p(z) = 1 + \frac{1}{2}z^2 + \frac{1}{4!}z^4\langle J^4\rangle + \mathrm{O}(z^4), \tag{36}$$

the odd moments vanishing because $p(x)$ is even.

The calculation of equations (12)–(13) may now be repeated, giving

$$\langle Z^n\rangle = \mathrm{Tr}\left[\exp \frac{\beta \tilde{J}}{N^{1/2}} \sum_{(ij)} J_{ij}\left(\sum_{\alpha=1}^{n} S_\alpha(i) S_\alpha(j)\right)\right] = \mathrm{Tr} \prod_{(ij)} F_p\left[\frac{\beta \tilde{J}}{N^{1/2}}\left(\sum_{\alpha=1}^{n} S_\alpha(i) S_\alpha(j)\right)\right]$$

$$= \mathrm{Tr}\exp\left(\sum_{(ij)} \ln[F_p(\beta \tilde{J}\lambda_{ij}/N^{1/2})]\right) \tag{37}$$

where

$$\lambda_{ij} = \sum_{\alpha=1}^{n} s_\alpha(i) s_\alpha(j). \tag{38}$$

It is important to note that $-n \leqslant \lambda_{ij} \leqslant n$ whatever N.

Putting $\gamma = (\beta \tilde{J})^2$, $\delta = \beta^4 \langle J^4\rangle$, and using (36), we obtain

$$F_p(\beta \tilde{J}\lambda_{ij}/N^{1/2}) = 1 + (\gamma/2N)\lambda_{ij}^2 + (\delta/24N^2)\lambda_{ij}^4 + o(1/N^2) \tag{39}$$

which enables us to evaluate the sum in (37):

$$\sum_{(ij)} \ln[F_p(\beta \tilde{J}\lambda_{ij}/N^{1/2}] = \sum_{(ij)} [(\gamma/2N)\lambda_{ij}^2 + \Lambda(\lambda_{ij})]. \tag{40}$$

The first term gives the desired contribution, while the correction term $\Lambda(\lambda_{ij})$ is of order N^{-2} and thus gives a bounded sum,

$$\left|\sum_{(ij)} \Lambda(\lambda_{ij})\right| \leqslant C \tag{41}$$

where the constant C may be chosen independently of N (and the λ_{ij}).

The correction is not obviously ignorable, but in fact becomes so in the thermodynamic limit. We therefore focus our attention on $\phi_N(n)$, the quantity that is relevant in this limit:

$$\phi_N(n) = N^{-1}\ln\langle Z^n\rangle = N^{-1}\ln \mathrm{Tr}\exp\left[\frac{\gamma}{2N}\sum_{(ij)}\left(\sum_{\alpha=1} S_\alpha(i) S_\alpha(j)\right) + \sum_{(ij)} \Lambda(\lambda_{ij})\right]. \tag{42}$$

Taking advantage of the Bogoliubov–Peierls inequality (Ruelle 1969)

$$N^{-1}|\ln \mathrm{Tr}(\exp(A)) - \ln \mathrm{Tr}(\exp(B))| \leqslant N^{-1}\|A - B\| \tag{43}$$

and the bound (41), we may now conclude that

$$|\phi_N(n;\Lambda) - \phi_N(n;0)| \leqslant C/N. \tag{44}$$

This proves that the details of the probability distribution, which are all contained in Λ, are quite irrelevant in the thermodynamic limit.

Three comments are appropriate. Firstly, the restriction that $F_p(z)$ be entire is needed for the model to make any sense at all, since $\beta\lambda_{ij}$ in equation (37) may take any value between $-\infty$ and $+\infty$. Secondly, the extension to the full Hamiltonian, without $\tilde{J}_0 = h = 0$, is straightforward and does not change the conclusion. Finally, we note that the *classical* structure of the model has been used in an essential way, particularly in equation (37).

5. Analyticity

Having completed step (a) of the replica method, the calculation of $\phi(n)$ for positive integer n, we turn to steps (b) and (c) as defined at the end of § 2. These require us to face questions (8) and (9), concerning reversal of limits and extension to real or complex n. In this section we consider these questions from the point of view of the analytic properties of $\phi_N(n)$ as a function of complex n.

Since Z_N is positive, whatever β and the J_{ij}, the quantity $\langle Z_N^n\rangle \equiv \langle \exp(n \ln Z_N)\rangle$ is well defined and *analytic* for all complex n. However, the function $\phi_N(n)$, defined by equation (5), involves the logarithm of $\langle Z_N^n\rangle$ and is only analytic in domains containing no zeros of $\langle Z_N^n\rangle$. There can be no zeros for real n, so we may choose a simply connected domain D_N that *includes* the whole real axis but *excludes* the zeroes of $\langle Z_N^n\rangle$; the bounds established in Appendix B (for SK) assure us that the zeroes lie a finite distance from the real axis. Unfortunately, D_N depends on N, and may cease to be a finite simply connected domain in the limit $N\to\infty$. For example, the zeroes of $\langle Z_N^n\rangle$ might approach the real axis at $n = 1$, say, cleaving D_N there in the thermodynamic limit. This picture is akin to the Yang and Lee (1952) description of a phase transition, with the 'transition' occurring as a function of n. If such a cleavage takes place, we are doomed to failure in any attempt to continue $\phi(n)$ through the 'transition' to $n=0$, and the calculation of $\phi(n)$ for positive integer n is essentially irrelevant for our purpose. However, the location of the zeroes of $\langle Z_N^n\rangle$ is still unknown, so let us tentatively assume that they are sufficiently well behaved to leave D_N intact in the thermodynamic limit.

We first consider step (c) of the replica method, the justification of the interchange of the thermodynamic limit and the differentiation in (8). To prove this, it is sufficient to show that $\phi_N(n)$ is *analytic* and *bounded* for n in a neighbourhood $\mathcal{N}$ of the origin, the bound and the domain being independent of N. For if these conditions are satisfied, the theorem of Vitali (Titchmarsh 1939 § 5.22) guarantees that we can (at least) find an infinite *sub*sequence $N_1, N_2, N_3\ldots$ such that $\phi_N(n)$ converges *uniformly* to an *analytic* function $\phi(n)$ along this subsequence, implying that the interchange of $N\to\infty$ (along the subsequence) and differentiation *is* allowed (Titchmarsh 1939 § 2.81). According to the previous discussion, $\phi_N(n)$ is analytic in such a domain $\mathcal{N}$ *if* no zero of $\langle Z_N^n\rangle$ approaches the origin as $N\to\infty$, but we have not succeeded in proving this. The boundedness of $|\phi_N(n)|$ is also hard to establish, even with the necessary assumption of no zeros of $\langle Z_N^n\rangle$, and we consider instead the function

$$\psi_N(n) = \langle Z_N^n\rangle^{N^{-1}} = \exp(\phi_N(n)) \tag{45}$$

which is analytic wherever $\phi_N(n)$ is. Bounds on $|\psi_N(n)|$ for the SK problem are derived in Appendix B, so the validity of the interchange of limits in (8), for $\psi_N(n)$ in place of $\phi_N(n)$, rests solely on the absence of zeroes of $\langle Z_N^n\rangle$ approaching the origin. It is easy to show that the truth of (8) for $\psi_N(n)$ implies that for $\phi_N(n) = \ln\psi_N(n)$, at least for real n, and we are therefore able to state the conclusion:

The interchange of the thermodynamic limit and differentiation at $n=0$, equation (8), is valid if:

(a) No zeroes of $\langle Z_N^n\rangle$ approach the origin $n=0$ as $N\to\infty$, and

(b) $|\langle Z_N^n\rangle^{N^{-1}}|$ is uniformly bounded (w.r.t.N) in some neighborhood of the origin.

Part (b) *is* true for the SK problem.

We now turn to step (b) of the replica method, the extension from positive integer n to a neighbourhood of $n=0$. Here too it is convenient to consider $\psi_N(n)$, and its limit $\psi(n) = \lim_{N\to\infty}\psi_N(n)$, rather than $\phi_N(n)$ and $\phi(n)$. By a calculation such as that performed in § 3 we know $\psi(n)$ at all positive integer n; there is clearly no difficulty in reversing the order of the thermodynamic limit and the exponential in equation (45). It is easy to find *an* extension to real n since, for example, equation (28) is itself well defined for real n, but we need to show that this extension is *unique*. Suppose, therefore, that we could construct two different extensions, $\psi_a(n)$ and $\psi_b(n)$, both agreeing with the calculated $\psi(n)$ at all the positive integers. We might now hope to demonstrate that $\psi_a(n)-\psi_b(n)$ is necessarily identically zero, giving a unique extension, by invoking Carleman's or Carlson's theorem (Titchmarsh 1939 §§ 3.71 and 5.81). But these theorems need analyticity in the whole right half-plane, or certainly in a sector thereof, and require a strong restriction on the growth rate of $|\psi(n)|$ as $|n|\to\infty$. We have already virtually abandoned hope of proving the necessary analyticity; we now show that the required growth condition is also unavailable for many models, including SK.

In Appendix B we show for the SK problem that the function $|\psi_N(n)|$ grows, as $|n|\to\infty$ in the right half-plane, according to

$$|\psi_N(n)|\sim\exp[C(\mathrm{Re}(n))^2] \tag{46}$$

where the constant C is independent of N. The same behaviour occurs in the Gaussian random field model considered by Schneider and Pytte (1977). This growth rate is too rapid for the application of Carlson's theorem, so we cannot expect a uniqueness proof. Indeed, we could add a function of the form $\exp(\alpha n^2)\sin\pi n$ $(0\leqslant\alpha<C)$ to a supposed extension to obtain another equally valid extension with a different derivative at the origin. A similar possibility has been noted by Klein (1977).

It might be thought that the function $\phi(n)=\ln\psi(n)$ would behave more cooperatively, but there is no obvious way to bound the growth of $\mathrm{Im}(\phi(n))=\arg\langle Z_N^n\rangle)/N$.

Carlson's theorem† *would* ensure a unique extension given analyticity, if the growth rate were only

$$|\psi_N(n)|\leqslant\exp(C\ \mathrm{Re}(n)). \tag{47}$$

Comparison with this growth condition may be a useful criterion for the success or failure of the replica method in other applications, but cannot be relied upon without further knowledge of analyticity, which depends in turn on the location of the zeros of $\langle Z_N^n\rangle$. In general, analyticity seems of little help in proving the uniqueness of the extension to real n. This is hardly surprising, because the replica method fails at precisely this stage, as we show in the next section.

6. Convexity

Considered as a function of a *real* variable n, the quantity $\phi_N(n)$ defined by equation (5) is *convex* (Hardy, Littlewood and Polya 1934, theorems 197 and 213). This valuable property allows us a much more satisfactory description of the failings of the replica method.

Our discussion hinges around a lemma due to Griffiths (1964) which reads, in terms of the present problem:

Lemma: Let $\phi_N(n)$ be a sequence of convex functions defined in an interval [a, b], and let $\phi'_N(n)$ be the first derivative of $\phi_N(n)$. If there exists a function $\phi(n)$ such

† In applying Carlson's theorem (Titchmarsh 1939 § 5.81), consider the function $\psi_N(n)\exp(-Cn)$; we do not need $C<\pi$ because (47) contains $\mathrm{Re}(n)$, not $|n|$.

that $\lim_{N\to\infty} \phi_N(n)=\phi(n)$ for every point in [a, b], then $\phi(n)$ is convex, and, furthermore,

$$\lim_{N\to\infty} \phi'_N(n)=\phi'(n) \tag{48}$$

at every point at which the derivative $\phi'(n)$ of $\phi(n)$ is continuous.

Equation (48) expresses, for $n=0$, exactly the equality that we wish to prove in step (c) of the replica method. Besides convexity, the Griffiths lemma requires two further conditions; the existence of the limit $\phi(n)$ (equation 7) for every real n in some finite interval containing the origin, and the continuity of $\phi'(n)$ at the origin. We consider these in turn.

Until now we have assumed that the limit (7) exists for all real n, although, as discussed in § 2, we only have good physical grounds for the assumption at positive integer n. A proof is now at hand however; if in any finite interval $[a, b]$, the sequence of convex functions $\phi_N(n)$ is bounded above and below independently of N, there necessarily exists a sub-sequence that converges uniformly to a function $\phi(n)$ in any interval contained in $[a, b]$ (Roberts and Varberg 1973). Upper and lower bounds on $\phi_N(n)$ are established in Appendix B for the SK problem, the bounds depending on n but not N, so for any finite interval $[a, b]$ we have $m\leqslant\phi_N(n)\leqslant M$ and uniform convergence of $\phi_N(n)$ to $\phi(n)$ (along a sub-sequence) is guaranteed. Although any finite interval containing the origin $n=0$ is sufficient for our purpose, we could extend the result to $[-\infty, +\infty]$ by a diagonal procedure (Reed and Simon 1972). The existence of the limit in (7) is thus proven (along a sub-sequence) for the SK problem; the approach of Appendix B should also be applicable to most other problems that possess a sensible thermodynamic limit.

The first part of the Griffiths lemma now assures us that $\phi(n)$ is itself convex, and therefore continuous. It is also, by virtue of convexity, differentiable *almost* everywhere, but we need to know, for the second part of Griffiths' lemma, whether in fact it has a continuous derivative at the origin. We have not succeeded in deciding this question for the general case, so we again specialise to the SK problem. In this case we have an explicit form for $\phi(n)$, equation (28) with the consistency condition (29), that is exact for all positive integer n. If we make the 'obvious' extension to real n, by *assuming* that these equations are also valid for real n, we can readily verify that the resulting $\phi'(n)$ *is* continuous at $n=0$. The Griffiths lemma then allows the interchange of limits and, with (6), gives us the final result obtained by Sherrington and Kirkpatrick, equations (30) and (31).

Stepping back for an overview, we see that we have rigorously justified all the dubious steps in the SK problem, including the interchange of limits, on the basis of a *single* assumption: that the correct extension of $\phi(n)$ from positive integer n to real n is simply given by interpreting n as a real variable in equations (28) and (29), the exact results for integer n. At the same time we know that the final result is *wrong* because it leads to a negative entropy at low temperature, and this is clearly impossible given the original Hamiltonian. We *must* therefore conclude that the above assumption is wrong and the 'obvious' extension from integer to real n is definitely incorrect. Our previous difficulties in proving the uniqueness of this extension are thus quite natural. We might have attempted to invoke convexity, as well as analyticity, to prove uniqueness, but we now perceive that this task is hopeless. There are at least two different extensions of equations (28) and (29); the unknown correct one, and the 'obvious', but incorrect, one used by SK.

Klein (1977) has also suggested that the extension to real n is non-unique and is to be blamed for the failure of SK. Our approach has gone further by eliminating all *other* loopholes in SK, thus *proving* that the faulty extension is *exclusively* responsible for this failure. In a different approach, de Almeida and Thouless (1978) have examined the stability of the minimum of $f_n(\boldsymbol{q}; \gamma)$, showing that an eigenvalue of the Hessian changes sign if extrapolated to $n = 0$. The existence of a negative eigenvalue for $n \approx 0$ suggests intuitively that we have the 'wrong' stationary point in this region. *At* any positive integer n, where the Hessian is well defined, there are no negative eigenvalues however—at $n = 0$ and $n = 1$ there are no eigenvalues at all—and so their result also points to a breakdown of the extension to real n.

It might be argued that $f(\beta)$ in equation (30) is right and that the negative entropy arises from an inpermissible reversal of $\partial/\partial\beta$ and the thermodynamic limit. But since $-\beta f(\beta)$ is a convex function of β for any finite system, and the thermodynamic limit exists, we can mimic the above arguments with n replaced by β, and conclude that this reversal is also allowable.

We warn the reader that our proof of the validity of the interchange of limits is not necessarily valid for versions of the replica method other than that formulated here. Prescriptions involving $\langle Z^n \rangle$ rather than $\langle Z^n \rangle^{N^{-1}}$ are particularly suspect. We also note that Klein (1977) has introduced a different representation for $\ln Z$,

$$\langle \ln Z \rangle = \lim_{\alpha \to 0} \int_0^\infty \mathrm{d}x \, \frac{\exp(-\alpha x)}{x} (\cos x - \cos(Zx)) \tag{49}$$

for which the interchange of the thermodynamic limit and the operations specified is definitely *in*valid; since $\cos x - \cos(Zx)$ is bounded, the quantity

$$\lim_{\alpha \to 0} \lim_{N \to \infty} N^{-1} \int_0^\infty \mathrm{d}x \, \frac{\exp(-\alpha x)}{x} (\cos x - \cos(Zx)) \tag{50}$$

is identically zero. The interchange *may* be valid when rewritten in terms of $Z^{N^{-1}}$, but this is not obvious and deserves further study.

Summarising this section:

The interchange of the thermodynamic limit and differentiation at $n = 0$, equation (8), is valid if

(*a*) The sequence of functions $\phi_N(n)$ has an upper and lower bound independent of N in some finite interval including the origin, and

(*b*) $\phi'(n)$ is continuous at $n = 0$.

Condition (*a*) *is* true for SK; condition (*b*) *is* true for the 'obvious' extension of SK to real n, which is therefore the *wrong* extension.

7. Conclusion

We have shown unambiguously that the trouble in the SK problem lies solely in the extension from integer to real n; the 'obvious' extension is definitely wrong. The rest of the SK calculation has been put on a firmer footing by making the thermodynamic limit explicit, by avoiding the steepest descent method (which gives nonsense if applied carefully), by proving the 'positive diagonal' theorem, and by proving the validity of the interchange of limits. We have also shown that a wide range of probability distributions gives the same results.

We have not actually proved the validity of the interchange of limits for the replica method *in general*, even within our own formulation, since the bounds of Appendix B, and the continuity of $\phi'(n)$ at $n=0$, are model-dependent, but it should be easy to follow similar arguments for models other than SK. In models possessing a slower growth rate, equation (47), we may reasonably *expect* the 'obvious' extension to be correct (and unique), but a proof of this depends on the location of the zeros of $\langle Z_N^n \rangle$.

We have shown, in the SK problem, that the extension to real n cannot be unique, since the 'obvious' extension is not the correct one. We *suspect* that zeros of $\langle Z_N^n \rangle$ do indeed approach the real axis as $N \to \infty$, giving $\phi(n)$ some sort of non-analyticity at small n, thus making the correct extension far from obvious and perhaps pathological. We tentatively *suggest* that the trouble occurs at or near $n=1$, noting that $\phi(n)$ is independent of q at this point. We also remark that $n=2$ appears to be a special point; for $n>2$ the transition in the replica system is first order and has a critical temperature (at which q jumps to zero) greater than T_c.

Acknowledgements

We thank Brian Buck, Joel Lebowitz, Elliott Lieb, and Barry Simon for helpful discussions. We are particularly grateful to Elliott Lieb for first proving the 'positive diagonal' theorem of Appendix A, and for allowing us to adapt his proof for presentation here. We also wish to express our gratitude to Michael Reed for his stimulating interest in this work.

Appendix A

We prove the following theorem:

Theorem: If $\gamma > 0$, and $\boldsymbol{q} = \{q_{\alpha\beta};\ 1 \leq \alpha \leq \beta \leq n\}$ is a vector in $\frac{1}{2}n(n-1)$ dimensional 'q-space' that maximises the function

$$Q_n(\boldsymbol{q}) = \operatorname{Tr} \exp\left[\gamma \sum_{(\alpha\beta)} (q_{\alpha\beta} S_\alpha S_\beta - \tfrac{1}{2} q_{\alpha\beta}^2)\right] \tag{A.1}$$

then $|q_{\alpha\beta}| = q \geq 0$ for all $(\alpha\beta)$. If $\boldsymbol{q} \neq 0$, there are 2^{n-1} equivalent $\boldsymbol{q}$'s, differing only in the signs of their components $q_{\alpha\beta}$, that all maximise $Q_n(\boldsymbol{q})$; one of these has all components positive. The absolute maximum of $Q_n(\boldsymbol{q})$ is thus always realised on the positive diagonal, $q_{\alpha\beta} = q \geq 0$ for all $(\alpha\beta)$.

Here, as elsewhere in this paper, $(\alpha\beta)$ stands for each distinct pair $\{1 \leq \alpha < \beta \leq n\}$, and the trace, a finite sum, is to be taken over all 2^n Ising spin configurations $\{S_\alpha = \pm 1;\ 1 \leq \alpha \leq n\}$. The theorem is trivial for $n=2$, so we assume $n \geq 3$ in what follows.

The conclusion of this theorem, for the *maximum* of $Q_n(\boldsymbol{q})$, is clearly equivalent to the result claimed in § 3 for the *minimum* of $f_n(\boldsymbol{q};\ \gamma) = -\{\ln Q_n(\boldsymbol{q})\}/\gamma$. It was also assumed, without proof, by Sherrington and Kirkpatrick in their original exposition, where, however, the existence of several equivalent maxima was overlooked. Although intuitively reasonable, the theorem cannot be regarded as obvious, and a proof is essential if the weaknesses of the replica method are to be systematically eliminated.

The theorem was first proved by Elliott Lieb of Princeton University. The first part of our proof, up to and including the lemma, is due to Lieb and is used by permission. The rest of our proof is simpler than Lieb's but was influenced by it.

We first examine the symmetry properties of $Q_n(\boldsymbol{q})$. There is an obvious *permutation* symmetry among the labels α, giving $n!$ points equivalent to a general point in q-space (and less for special points with some components equal). More importantly, there is a *sign* symmetry. If $\mu_\alpha = \pm 1$ for $\alpha = 1, 2, \ldots, n$, the *sign transformation*

$$T_\mu : \boldsymbol{q} = \{q_{\alpha\beta}\} \to \{\mu_\alpha \mu_\beta q_{\alpha\beta}\} \tag{A.2}$$

leaves Q_n unchanged; $Q_n(T_\mu \boldsymbol{q}) = Q_n(\boldsymbol{q})$. There are 2^{n-1} distinct sign transformations, including the identity, since $\{\mu_\alpha\}$ and $\{\mu'_\alpha = -\mu_\alpha\}$ are equivalent.

We remark that $Q_n(\boldsymbol{q})$ is a positive, continuous (in fact *smooth*) function of $\boldsymbol{q}$ that tends to zero uniformly as $\boldsymbol{q} \to \infty$. Therefore at least one absolute maximum exists, and, by the argument following equation (25), is to be found in the unit cube, $|q_{\alpha\beta}| < 1$ for all $(\alpha\beta)$.

It is convenient to break up the trace in (A.1) into an explicit sum over *two* spins, say S_1 and S_2, and a reduced trace, Tr$'$, over the remaining spins $S_3, \ldots S_n$. The components $q_{\alpha\beta}$ then fall naturally into four groups: q_{12}; $A = \{q_{1\alpha}; 3 \leq \alpha \leq n\}$; $B = \{q_{2\alpha}; 3 \leq \alpha \leq n\}$; $C = \{q_{\alpha\beta}; 3 \leq \alpha < \beta \leq n\}$. Writing $\boldsymbol{q} = (q_{12}, A, B, C)$ we now prove a lemma.

Lemma: If $\boldsymbol{q}'$ and $\boldsymbol{q}''$ are defined in terms of $\boldsymbol{q} = (q_{12}, A, B, C)$ by $\boldsymbol{q}' = (q_{12}, A, A, C)$ and $\boldsymbol{q}'' = (q_{12}, B, B, C)$ and $q_{12} \geq 0$, then

$$Q_n(\boldsymbol{q}) \leq \{Q_n(\boldsymbol{q}')Q_n(\boldsymbol{q}'')\}^{1/2}. \tag{A.3}$$

Further, if $q_{12} > 0$, equality holds in (A.3) if and only if $A = B$ (i.e. $\{q_{1\alpha} = q_{2\alpha}; 3 \leq \alpha \leq n\}$) and thus $\boldsymbol{q} = \boldsymbol{q}' = \boldsymbol{q}''$.

Proof: Consider first the quadratic form

$$K(g, h) \equiv \sum_{\{S_1, S_2 = \pm 1\}} g(S_1) \exp(\gamma q_{12} S_1 S_2) h(S_2) \tag{A.4}$$

and notice that the 2×2 matrix $\exp(\gamma q_{12} S_1 S_2)$ is positive semidefinite if $q_{12} \geq 0$ and strictly positive definite if $q_{12} > 0$. Therefore, by the Cauchy–Schwarz inequality,

$$|K(g, h)|^2 \leq K(g, g) K(h, h) \tag{A.5}$$

and when $q_{12} > 0$ this can only be an equality if g and h are linearly dependent, in which case $g(+1)h(-1) = g(-1)h(+1)$.

Letting $X = \{S_3, \ldots, S_n\}$, if we now set

$$g_1(S|X) = \exp\left[\gamma \sum_{\alpha=3}^{n} (q_{1\alpha} S_\alpha S - \tfrac{1}{2} q_{1\alpha}^2)\right] \tag{A.6a}$$

$$g_2(S|X) = \exp\left[\gamma \sum_{\alpha=3}^{n} (q_{2\alpha} S_\alpha S - \tfrac{1}{2} q_{2\alpha}^2)\right] \tag{A.6b}$$

$$V(X) = \exp\left\{\gamma \sum_{3 \leq \alpha < \beta \leq n} (q_{\alpha\beta} S_\alpha S_\beta - \tfrac{1}{2} q_{\alpha\beta}^2) - \tfrac{1}{2}\gamma q_{12}^2\right\} \tag{A.6c}$$

we see that

$$Q_n(\boldsymbol{q}) = \mathrm{Tr}' \, K(g_1, g_2 | X) V(X) \tag{A.7}$$

and the inequality (A.5) gives us (noting $V(X)>0$)

$$\begin{aligned} Q_n(\boldsymbol{q}) &\leqslant \mathrm{Tr}'\{K(g_1, g_1|X)K(g_2, g_2|X)\}^{1/2} V(X) \\ &\leqslant \{\mathrm{Tr}' K(g_1, g_1|X) V(X)\}^{1/2} \{\mathrm{Tr}' K(g_2, g_2|X) V(X)\}^{1/2} \\ &= \{Q_n(\boldsymbol{q}')Q_n(\boldsymbol{q}'')\}^{1/2} \end{aligned} \tag{A.8}$$

which proves (A.3). The second step in (A.8) is a straightforward application of the Cauchy–Schwarz inequality.

If $q_{12}>0$ the first inequality in (A.8) is strict *unless* $g_1(+1|X)g_2(-1|X)=g_1(-1|X)g_2(+1|X)$ for all X, which implies

$$\sum_3^n (q_{1\alpha}-q_{2\alpha})S_\alpha = 0$$

whatever the S_α. This sum cannot vanish both for $S_\alpha=+1$ and for $S_\alpha=-1$ unless $q_{1\alpha}=q_{2\alpha}$, for any given $\alpha \geqslant 3$. $A=B$ is thus necessary for equality in (A.3) if $q_{12}>0$; it is obviously sufficient.

Now suppose that $\boldsymbol{q}$ maximises $Q_n(\boldsymbol{q})$, and $q_{12} \geqslant 0$. With $\boldsymbol{q}'$ and $\boldsymbol{q}''$ defined in the sense of the lemma we have the inequality (A.3) and the additional inequalities $Q_n(\boldsymbol{q}) \geqslant Q_n(\boldsymbol{q}')$, $Q_n(\boldsymbol{q}) \geqslant Q_n(\boldsymbol{q}'')$, since $\boldsymbol{q}$ maximizes Q_n. These three inequalities can be satisfied simultaneously only if

$$Q_n(\boldsymbol{q}) = Q_n(\boldsymbol{q}') = Q_n(\boldsymbol{q}'') \text{ (if } q_{12} \geqslant 0 \text{ and } \boldsymbol{q} \text{ maximises } Q_n). \tag{A.9}$$

If $q_{12}>0$ this implies $q_{1\alpha}=q_{2\alpha}$ for all $\alpha \geqslant 3$.

The lemma and the deduction (A.9) single out spins one and two from the rest, but the permutation symmetry assures us that they could equally well be applied to any pair of spins (with obvious adaptions of the notation). Consider, therefore, a 'triangle' of any three 'points' λ, μ, ν $(\lambda<\mu<\nu)$ with their associated 'bonds' $q_{\lambda\mu}$, $q_{\lambda\nu}$, $q_{\mu\nu}$, and continue to assume that $\boldsymbol{q}$ maximises Q_n. If any one of the bonds is strictly positive the other two must be equal. If any one is strictly negative an appropriate sign transformation gives an equivalent maximising $\boldsymbol{q}$ in which it is strictly positive; the other two bonds must be equal in the transformed $\boldsymbol{q}$, and therefore equal in magnitude but opposite in sign (unless zero) in the original $\boldsymbol{q}$. These observations eliminate all but the following possibilities for the three quantities $q_{\lambda\mu}$, $q_{\lambda\nu}$, $q_{\mu\nu}$:

(*a*) All zero;
(*b*) All strictly positive and equal;
(*c*) All non-zero and equal in magnitude, one positive, two negative;
(*d*) Two zero, one non-zero.

In case (*d*), application of equation (A.9) to one of the zero bonds gives us a $\boldsymbol{q}'$ that has one zero and two non-zero bonds in the corresponding triangle. Since this belongs to none of the above cases, $\boldsymbol{q}'$ cannot maximize Q_n, and therefore, as $Q_n(\boldsymbol{q})=Q_n(\boldsymbol{q}')$, neither can $\boldsymbol{q}$. Thus case (*d*) must also be eliminated.

Since the above result applies to *any* triangle, a $\boldsymbol{q}$ that maximizes Q_n must either be identically zero or must have *all* components $q_{\alpha\beta}$ non-zero and equal in magnitude. If $\boldsymbol{q}$ is non-zero, each possible triangle must belong to case (*b*) or case (*c*) above, so, considering in particular the triangles 1–α–β $(2 \leqslant \alpha < \beta \leqslant n)$, we must have $\mathrm{sign}(q_{\alpha\beta}) = \mathrm{sign}(q_{1\alpha})\,\mathrm{sign}(q_{1\beta})$. Applying a sign transformation T_μ, with $\mu_1=1$, and $\mu_\alpha=\mathrm{sign}(q_{1\alpha})$ for $\alpha \geqslant 2$, generates an equivalent $\boldsymbol{q}$ with *all* components strictly positive and equal, *i.e.*

on the positive diagonal. Any $\boldsymbol{q}$ that maximises $Q_n(\boldsymbol{q})$ must therefore either be zero or belong to the set of $2^{n-1}\boldsymbol{q}$'s equivalent by a sign transformation to a $\boldsymbol{q}$ on the positive diagonal. This completes the proof of the theorem.

Appendix B

We establish upper and lower bounds, independent of N, for the function $\psi_N(n) = \langle Z^n \rangle^{N^{-1}}$ in the case of the SK model.

First consider positive *integer* $n \geq 1$. Equation (13) implies

$$\psi_N(n) \leq \{2^{Nn} \exp[\tfrac{1}{4}\gamma n^2(N-1)]\}^{N^{-1}} < \exp(Cn^2) \tag{B.1}$$

with $C = \ln 2 + \frac{1}{4}\gamma$, and this result is also true for $n = 0$, where $\psi_N(n) = 1$. The convexity of $\phi_N(n)$, and hence of $\psi_N(n)$, now provides a similar bound

$$\psi_N(n) \leq D \exp(Cn^2) \qquad (n \geq 0) \tag{B.2}$$

for all positive *real* n and some constant D.

For a lower bound, we realise that Z is defined by a sum (trace) of positive terms and must be larger than each, so in particular

$$Z > \exp\Big((\beta \tilde{J}/N^{1/2}) \sum_{(ij)} J_{ij}\Big). \tag{B.3}$$

Noting that the J_{ij} are *independent* random variables, we obtain, for non-negative real n,

$$\psi_N(n) = \langle Z^n \rangle^{N^{-1}} \geq \left(\int_{-\infty}^{\infty} \frac{dz}{(2\pi)^{-1}} \exp(-\tfrac{1}{2}z^2) \exp[n(\beta\tilde{J}/N^{1/2})z]\right)^{\frac{1}{2}(N-1)}$$

$$= \exp[\tfrac{1}{4}\gamma n^2(N-1)/N] \geq \exp(\gamma n^2/8) \qquad (n \geq 0) \tag{B.4}$$

(taking $N \geq 2$).

For negative n the first inequality in (B.4) must be reversed, leading to the upper bound

$$\psi_N(n) < \exp(\gamma n^2/4) \qquad (n < 0). \tag{B.5}$$

Using the Cauchy–Schwarz inequality,

$$\psi_N(n)\psi_N(-n) = \langle Z^n \rangle^{N^{-1}} \langle Z^{-n} \rangle^{N^{-1}} \geq \langle 1 \rangle^{2/N} = 1 \tag{B.6}$$

we may derive from (B.2) a lower bound for negative n:

$$\psi_N(n) \geq D^{-1} \exp(-Cn^2) \qquad (n < 0). \tag{B.7}$$

The inequalities (B.2), (B.4), (B.5), and (B.7) provide upper and lower bounds for $\psi_N(n)$ (and hence for $\phi_N(n)$) when n is real. For complex n we have the obvious inequality

$$|\langle Z^n \rangle| \leq \langle Z^{\mathrm{Re}(n)} \rangle \tag{B.8}$$

so that, in the right half-plane, the bound (B.2) gives

$$|\psi_N(n)| \leq D \exp[C(\mathrm{Re}(n))^2]. \tag{B.9}$$

Moreover, we note from (B.4) that $|\psi_N(n)|$ does actually realise a growth rate of this form, at least near the real axis.

References

de Almeida J R L and Thouless D J 1978 *J. Phys. A: Math. Gen.* **11** 983–90
Breiman L 1968 *Probability* (Reading, Mass: Addison-Wesley) ch 8
Edwards S F and Anderson P W 1975 *J. Phys. F: Metal Phys.* **5** 965–74
Griffiths R B 1964 *J. Math. Phys.* **5** 1215–22
Hardy G H, Littlewood J E and Pólya G 1934 *Inequalities* (Cambridge: Cambridge University Press)
van Hemmen J L 1978 *Fortschritte der Physik* **26** 397–439
Kac M 1968 *Arkiv for Det Fysiske Seminar i Trondheim* **11** 1–22
Kirkpatrick S and Sherrington D 1978 *Phys. Rev.* B**17** 4384–403
Klein M W 1977 *J. Phys. F: Metal Phys.* **7** L267–71
Kosterlitz J M, Thouless D J and Jones R C 1976 *Phys. Rev. Lett.* **36** 1217–20
Lin T F 1970 *J. Math. Phys.* **11** 1584–90
den Ouden L W J, Capel H W, Perk J H H and Tindemans P A J 1976a, *Physica* **85**A 51–70
den Ouden L W J, Capel H W and Perk J H H 1976b *Physica* **85**A 425–456
Reed M C and Simon B 1972 *Methods of Modern Mathematical Physics* vol 1 (New York: Academic Press) p 28
Roberts A W and Varberg D E 1973 *Convex functions* (New York: Academic Press) p 20
Ruelle D 1969 *Statistical Mechanics* (New York: Benjamin) p 17
Schneider T and Pytte E 1977 *Phys. Rev.* B**15** 1519–22
Sherrington D and Kirkpatrick S 1975 *Phys. Rev. Lett.* **35** 1792–6
Thouless D J, Anderson P W and Palmer R G 1977 *Phil. Mag.* **35** 593–601
Titchmarsh E C 1939 *The Theory of Functions* 2nd ed (London: Oxford University Press)
Yang C N and Lee T D 1952 *Phys. Rev.* **87** 404–9