Coupled dynamics of fast spins and slow interactions in neural networks and spin systems

1993 J. Phys. A: Math. Gen. 26 3681

(http://iopscience.iop.org/0305-4470/26/15/018)

J. Phys. A: Math. Gen. 26 (1993) 3681–3695. Printed in the UK

Coupled dynamics of fast spins and slow interactions in neural networks and spin systems

R W Penney, A C C Coolen and D Sherrington

Department of Physics, Theoretical Physics, Oxford University, 1 Keble Road, Oxford OX1 3NP, UK

Received 17 March 1993

Abstract. We examine an Ising spin system in which both spins and the interactions between them may evolve in time, although on disparate timescales, such that the couplings change adiabatically. In thermal equilibrium we find a novel application of the replica method, but for finite replica number, representing the ratio of the temperatures of the spin and interaction systems. Regimes where the motion of the couplings has non-trivial effects are found in addition to those where solely the stochasticity of these interaction weights is significant, and this issue is closely related to the orders of the transitions between the various phases observed. Simulation results lend support to the analysis.

1. Introduction

In nature few properties of a system can be considered truly static; although observations may centre on relatively fast processes, there are likely to exist slower motions that may bear highly significantly upon the entire system if only given sufficient time. For example, if one examines neurophysiological tissue, even though the operation of this tissue may primarily relate to the behaviour of the neurons, it is clear that the relatively slow development of the synapses fundamentally influences the patterns of neuronal activity, with consequent implications for the functionality of that tissue. Moreover, in this case the evolution of the synapses strongly depends on the neuronal activity itself, producing a richly interacting system. Whilst one might focus on the dynamics of the neurons, it would be of considerable interest to examine how neurons and synapses evolve together. Such problems are, however, notoriously difficult to analyse. In this paper we offer a simple model which may offer insight into some of the general properties of such systems with two levels of dynamics.

For definiteness we will consider an idealized magnetic spin model, in which the interactions between the spins can evolve dynamically in response to their state, this itself being the product of some dynamical laws. The relation of this picture to that of McCulloch–Pitts neurons interacting via synaptic efficacies will allow some connection to be made between our model and neurophysiology, in addition to spin systems. Although studies of artificial neural networks have predominantly focused on either neuronal dynamics with fixed interactions (e.g. Amit *et al* 1987, Coolen and Ruijgrok 1988) or on evolution of synapses for fixed neuron states (e.g. Gardner 1988), there have been some attempts to unite these two approaches (e.g. Shinomoto 1987, Jonker and Coolen 1991). In view of the complexity of this goal, it is inevitable that the systems amenable to analysis will be somewhat contrived. Although our model is no exception, it promises some generality and also suggests new analytical techniques.

0305-4470/93/153681+15$07.50

The crucial feature of the systems with which we are concerned, is that the interactions between the fast variables (be they spins or neurons) change much more slowly than the fast variables themselves. For conventional magnets, although the states of the moments might predispose their mutual influences to change, slow atomic diffusion processes are usually necessary to alter the interaction strengths. Similarly, although with a less extreme separation of timescales, the dynamics of neurons are known to occur over far shorter times than does evolution of the synaptic efficacies. It is therefore reasonable to idealize this separation between spin and interaction-weight dynamics, and assume that changes in the interactions occur adiabatically, responding to the equilibrium character of the fast (spin) variables. Further, biological plausibility of the synaptic-weight dynamics, in terms of depending only on locally available information, leads to a simple form of dynamical law which can, under appropriate circumstances, be given as a gradient descent. This facilitates consideration of the equilibrium properties of the entire system, which will be the characteristics on which we will focus.

The form of dynamical laws which we propose will take us into the realm of the replica method (see, e.g. Mézard *et al* 1987), but with a finite number of replicas, n, without following the conventional limit $n \to 0$. The generalization of this method to finite n is an interesting matter in itself and our work will offer some investigation of the method for positive n.

2. Formulation of the model

We will be concerned with a system of N Ising spins (or formal neurons), $S_i \in \{\pm 1\}$ $i \in \{1..N\}$, connected by symmetric couplings, J_{ij}, and where both spins and interactions are endowed with dynamical laws. For the spins we will imagine a stochastic local field alignment, given as a single spin-flip Markov process

$$\frac{\mathrm{d}p(\{S_i\};t)}{\mathrm{d}t} = \sum_j \{W(-S_j \to S_j)p(S_1, \ldots -S_j, \ldots ; t) - W(S_j \to -S_j)p(\{S_i\};t)\} \tag{2.1}$$

in which $p(\{S_i\};t)$ is the microstate probability for a given spin configuration $\{S_i\}$, at time t, and the W's are spin transition rates. The choice

$$W(S_i \to -S_i) = \tfrac{1}{2}\big(1 - \tanh(\beta S_i h_i)\big) \qquad h_i = \sum_{j \neq i} J_{ij} S_j \tag{2.2}$$

corresponding to a Glauber dynamics, will allow us to immediately specify the equilibrium form of $p(\{S_i\};t)$ (i.e., $p(\{S_i\},\infty)$), dependent on the temperature $T = \beta^{-1}$ and the choice of J_{ij} (provided these are indeed symmetric, and remain essentially static over the timescale of the spin equilibriation). The interactions will be taken to evolve in response to the state of the spin system, in addition to externally imposed biases, according to some Langevin dynamics. Without careful choice of the dynamical laws for the weights, it is most improbable that such richly interdependent laws would be amenable to analysis except by simulation.

The separation of timescales between the two levels of dynamics suggested by nature, together with some biological constraints relevant to the neurological case, suggest a conveniently simple form of the weight dynamics. That the weights evolve only slowly suggests that they should depend on equilibrium correlation functions of the spin system, $\langle S_i S_j \ldots S_n \rangle$, which themselves depend on the current distribution of weights. However, neurophysiology imposes two extra significant requirements. Firstly, and unsurprisingly, we would like the weights to remain finite. Secondly, and more significantly, it is desirable that a synapse should evolve in response only to locally available information, hence, J_{ij} should only take account of $\langle S_i \rangle$, $\langle S_j \rangle$ and $\langle S_i S_j \rangle$. Requiring that the weights remain symmetric under the dynamics limits one to the combinations $\langle S_i \rangle \langle S_j \rangle$ and $\langle S_i S_j \rangle$. Motivated by the Hebbian trend in the synapses linking neurons in neurological tissue, we propose to consider interaction dynamics of the following prototypical form:

$$\tau \frac{\mathrm{d}J_{ij}}{\mathrm{d}t} = N^{-1}\big(\langle S_i S_j \rangle + K_{ij}\big) - \mu J_{ij} + N^{-1/2}\eta_{ij}(t) \qquad (i < j). \tag{2.3}$$

In addition to the (weight-dependent) equilibrium correlation between two spins, a weight-decay term along with biases, K_{ij}, and a Gaussian white noise, $\eta_{ij}(t)$ are involved. The noise is of mean zero and has covariance given by

$$\langle \eta_{ij}(t)\eta_{kl}(t') \rangle = 2\tilde{\beta}^{-1}\tau\delta_{ij,kl}\,\delta(t-t') \qquad (i < j,\ k < l) \tag{2.4}$$

whereby we define the temperature of the interaction system ($\tilde{T} = \tilde{\beta}^{-1}$). Factors of N, where present, are necessary to ensure non-trivial behaviour in the thermodynamic limit $N \to \infty$. (The weight dynamics considered by Shinomoto (1987) closely resemble (2.3), but without a noise term.) Henceforth, in our discussions of the dynamics of the weights, we will focus only on the representative upper triangle of the weight matrix, i.e. J_{ij}, K_{ij} and η_{ij} with $i < j$. The chosen dynamics imply that during the timescale over which the J_{ij}'s evolve, the spins are taken to come to complete thermal equilibrium, and the interactions then respond to the equilibrium character of the spin system. Of course, the randomness contained within the couplings is likely to lead to spin-glass character in the spin-system, which is likely to be associated with large timescales for the former equilibriation. This complication is a price of the simplicity of our chosen dynamical law, and also of the complete connectivity implicit in our model.

It is advantageous to simplify the analysis by one further step, by confining attention to the equilibrium properties of the entire system, i.e., to the equilibrium of the J_{ij}, which implies that of the spin-system itself. Our choice of a first-order Langevin equation will allow this equilibrium behaviour to be readily characterized, and will lead us into a novel application of the replica method. By defining a Hamiltonian for the spin system

$$H\big(\{S_i\},\{J_{ij}\}\big) = -\sum_{i<j} S_i J_{ij} S_j - \sum_{\mu i} h_\mu \xi_i^\mu S_i \qquad (\xi_i^\mu \in \{\pm 1\}) \tag{2.5}$$

and hence a partition function (dependent on the choice of the *symmetric* interactions)

$$Z_\beta\big(\{J_{ij}\}\big) = \operatorname*{Tr}_{\{S_i\}} \exp\big(-\beta H\big(\{S_i\},\{J_{ij}\}\big)\big) \tag{2.6}$$

we may identify the right-hand side of (2.3) as the gradient of a potential plus the noise term

$$\tau \frac{\mathrm{d}J_{ij}}{\mathrm{d}t} = -\frac{\partial}{\partial J_{ij}}\left\{-\frac{1}{\beta N}\ln Z_\beta - \frac{1}{N}\sum_{i<j} J_{ij}K_{ij} + \tfrac{1}{2}\mu\sum_{i<j}(J_{ij})^2\right\} + \frac{1}{\sqrt{N}}\eta_{ij}(t). \tag{2.7}$$

It is well known that the equilibrium probability distribution for such a diffusion in a potential is given by the Boltzmann form, and hence we may focus attention on the partition function of the interaction system, and employ it both as a generating functional of statistical averages of the interactions in equilibrium, and also as a source of order parameters for the spin system

$$\tilde{Z}_{\tilde{\beta}} = \int \prod_{i<j} \left\{ \frac{\tilde{\beta}\mu N}{2\pi} \right\}^{1/2} \mathrm{d}J_{ij} \; [Z_\beta]^{\tilde{\beta}/\beta} \exp\left\{ \tilde{\beta} \sum_{i<j} J_{ij} K_{ij} - \tfrac{1}{2}\tilde{\beta}\mu N \sum_{i<j} (J_{ij})^2 \right\}. \tag{2.8}$$

Thus, for a dynamical system with adiabatically evolving interactions between the fast variables, within the (broad) framework imposed by the dynamical laws (2.2) and (2.3), the equilibrium properties of the entire system may be investigated using equilibrium statistical mechanics.

It is worthwhile assimilating (2.8) before proceeding with its analysis. Defining the ratio of the temperatures $\tilde{\beta}/\beta$ to be a parameter n, we note that if n happened to be a positive integer, then (2.8) would represent an integer moment of the partition function of the spin system. Further, if n was taken towards zero (meaning that the interactions are affected more by the stochastic noise, $\eta_{ij}(t)$, than the spins), then we would be led towards an average of the free energy of the spin system, and (2.8) would show close similarities with the starting point of the Sherrington–Kirkpatrick (Sherrington and Kirkpatrick 1975, hereafter referred to as SK) model of a spin-glass, involving quenched random interactions. Both these cases have been analysed using the replica method. In the former case ($n \in \mathbb{Z}$, treated in Sherrington 1980) the absence of any need to make analytic continuations makes the use of the replica method seem wholy trustworthy, but in the second application ($n \to 0$) it is well known that the replica method develops subtleties absent from the former case. (Also discussed in Sherrington (1980) is the limit $n \to \infty$, where the weights are influenced more by the Hebbian stimulus $\langle S_i S_j \rangle$ than the η_{ij}'s.) Given that it would be unduly restrictive to confine attention to an integer ratio of temperatures, we will be forced to address both these regimes, although the analytical pathologies of the small n scenario will themselves provide some insight into our system. The SK model has itself also been considered for finite n (Kondor 1983), although under more restricted conditions than (2.8) permits.

Before analysis becomes finally tractable, we will need to address the biases K_{ij} which steer the weights towards some desired values. In the spirit of Hebb's rule, it would be reasonable to adopt a separable form for K_{ij}, i.e.

$$K_{ij} = K p^{-1/2} \sum_{\mu=1}^{p} \xi_i^\mu \xi_j^\mu \tag{2.9}$$

in which, in common with the chosen field term in (2.5), we imagine that a set of binary patterns, ξ_i^μ, are to be favoured states of the spins. As a first step we will confine attention to the situation of a single pattern ($p = 1$), which may be formally eliminated by the gauge transformations $S_i \to \xi_i S_i$, $J_{ij} \to \xi_i J_{ij} \xi_j$ and $K_{ij} \to \xi_i K_{ij} \xi_j$. The extension to greater numbers of patterns although tractable would be somewhat involved.

3. Analysis

Having simplified our problem to a level amenable to analysis, we may proceed to evaluate the interaction partition function $\tilde{Z}_{\tilde{\beta}}$. We will invoke replica mean field theory, in which we formally assume the ratio of temperatures, n, to be integral, but then attempt to analytically continue our expressions to real n (without any restriction in magnitude, although we will not permit negative n). Using standard manipulations, we may reduce $\tilde{Z}_{\tilde{\beta}}$ to an extremization problem over the space of some replica-dependent order parameters. By this stage our problem has yielded a single-site spin Hamiltonian by virtue of our imposition of complete connectivity.

$$\tilde{Z}_{\tilde{\beta}} = \exp\left\{\frac{\tilde{\beta}NK^2}{4\mu} + \frac{N\tilde{\beta}}{4\mu n} + Nn\ln 2\right\}\exp N\left\{\operatorname*{extr}_{m^\gamma\, q^{\gamma\delta}} F(\{m\},\{q\})\right\} \tag{3.1}$$

in which $\gamma, \delta \in \{1, \ldots, n\}$ are replica indices, and

$$\begin{aligned} F(\{m\},\{q\}) &= -\sum_\gamma \frac{m^{\gamma 2}\tilde{\beta}K}{2\mu n} - \sum_{\gamma<\delta}\frac{q^{\gamma\delta 2}\tilde{\beta}}{2\mu n^2} + G(\{m\},\{q\}) \\ G(\{m\},\{q\}) &= \ln\left[2^{-n}\operatorname*{Tr}_{\{S^\gamma\}}\exp\left\{\sum_\gamma\left(\frac{m^\gamma\tilde{\beta}K}{\mu n} + \frac{\tilde{\beta}h}{n}\right)S^\gamma + \sum_{\gamma<\delta}\frac{q^{\gamma\delta}\tilde{\beta}}{\mu n^2}S^\gamma S^\delta\right\}\right]. \end{aligned} \tag{3.2}$$

The interpretation of these order parameters is as follows:

$$m^\gamma = \overline{\langle S_i^\gamma\rangle} \qquad q^{\gamma\delta} = \overline{\langle S_i^\gamma\rangle\langle S_i^\delta\rangle} \tag{3.3}$$

where $\langle A\rangle$ denotes a Boltzmann average of A with energy given by the (replicated) spin Hamiltonian, $\sum_\gamma H(\{S_i^\gamma\},\{J_{ij}\})$, cf (2.5), for a given choice of weights. $\overline{B}$ represents an average of B over the dynamics of $\{J_{ij}\}$ after these have reached equilibrium. These order parameters closely parallel those of spin-glass and neural network problems, save that the disorder average is replaced by a time average over the dynamics of the interactions.

In order to make the correspondence of these expressions with those of the SK model (and of Sherrington 1980) more lucid, it is advantageous to define new parameters as follows:

$$J_0 = \frac{K}{\mu} \qquad \tilde{J} = \frac{1}{\tilde{\beta}\mu} = \frac{1}{\beta n\mu} \tag{3.4}$$

and also to promote β, in preference to $\tilde{\beta}$, as the fundamental (inverse) temperature scale. Thus, for the SK model, J_0/N would represent the mean-value coupling between spins, and $\tilde{J}/N$ would be their variance. However, these interpretations are closely tied with the limit $n \to 0$ of this model. For the present problem, the first two moments of the couplings to leading order in N are

$$N\tilde{\beta}^{-1}\left.\frac{\partial\ln\tilde{Z}_{\tilde{\beta}}}{\partial K_{ij}}\right|_{K_{ij}=K} = N\overline{J_{ij}} = J_0 + \beta\tilde{J}\sum_\gamma m^{\gamma 2} \qquad N\left(\overline{J_{ij}^2} - \overline{J_{ij}}^2\right) = \tilde{J}. \tag{3.5}$$

Furthermore, there is a finite correlation between weights, in contrast to the SK model

$$N^2\left(\overline{J_{ij}J_{ik}} - \overline{J_{ij}}\,\overline{J_{ik}}\right) = \beta^2\tilde{J}^2\left\{\sum_\gamma m^{\gamma 2} + 2\sum_{\gamma<\delta}m^\gamma q^{\gamma\delta}m^\delta - \left(\sum_\gamma m^{\gamma 2}\right)^2\right\} \qquad j \neq k \tag{3.6}$$

$$N^2\left(\overline{J_{ij}J_{kl}} - \overline{J_{ij}}\,\overline{J_{kl}}\right) = \mathrm{O}(N^{-1}) \qquad (ij) \neq (kl). \tag{3.7}$$

In common with spin-glass problems we are faced with an extremization in (3.1) which must be analytically continued to real n. For positive integer n the analysis of Lieb (discussed in van Hemmen and Palmer 1979) shows that the order parameters m^γ and $q^{\gamma\delta}$ are replica symmetric i.e., $m^\gamma = m \ \forall\gamma$ and $q^{\gamma\delta} = q \ \forall\gamma < \delta$ (to within possible sign changes due to the potential for antiparallel spin alignment in different replicas). The assumption of the same form of order parameters is known to be incorrect for $n \to 0$, but little is known for general n. For $n \to 0$ it is believed that Parisi's choice of $q^{\gamma\delta}$ (Parisi 1980a,b) produces the correct thermodynamic behaviour for the SK model, and phenomena in agreement with simulations of this system, although a demonstration of the uniqueness or proof of the validity of the Parisi scheme has not come to our attention. Given that the considerable successes of the Parisi hierarchical replica symmetry breaking formalism as applied to the SK model, we will adopt this ansatz for the present problem. This formalism also has the attractions of augmenting the replica symmetric ansatz with additional degrees of freedom, while remaining straightforwardly analytically continuable to arbitrary n, However, for simplicity we will consider only a one-step RSB for $q^{\gamma\delta}$, rather than the full infinite hierarchy. Computationally, in areas of our parameter space where $q^{\gamma\delta}$ is found to reproduce the replica symmetric results we will directly employ the replica symmetric theory with enhanced confidence (and a reduced CPU budget). Where replica symmetry is observed to break, in the conventional interpretation of this phenomenon, we expect diverging dynamical timescales for the spin system, and hence partial freezing of the system.

Regarding the order parameters m^γ, guided by the need to have a sensible limit $n \to 0$ for the SK problem, and also the observation in de Almeida and Thouless (1978) that the 'dangerous' fluctuations away from the replica symmetric saddle-point have no components transverse to $m^\gamma = m$, we will assume that replica symmetry in this order parameter is always preserved, i.e., all replicas have the same net magnetization, even though they may have inequivalent ways of achieving it (hence a non-replica symmetric $q^{\gamma\delta}$).

In view of the possibility of a first-order transition from a replica symmetric state to one having broken replica symmetry (as is observed by Krauth and Mézard 1989) in addition to that of a continuous transition as seen in the Sherrington–Kirkpatrick model, the extremization in (3.1) has been attempted in acknowledgement of both these pathologies. The first danger is related to the global stability of the replica symmetric saddle-point, and has been tackled by searching for a number of possible extremum values and choosing the best extremum found (we will have more to say about the label 'best' shortly). We thereby hope that should a first-order transition be exhibited by the dynamical system we would not overlook this in our numerical analysis of (3.1). Secondly there is the issue of local stability of the saddle-point chosen, with respect to fluctuations towards greater levels of replica symmetry breaking. Given that the full Parisi solution of the SK model is known to be only marginally stable (de Dominicis and Kondor 1983), there is little point in examining the local stability of a one-step replica symmetry broken solution, unless this has special features (such as found in Krauth and Mézard 1989). Therefore we will only examine the local stability of a replica symmetric solution of (3.1), following de Almeida and Thouless (de Almeida and Thouless 1978, hereafter referred to as AT). Thus where replica symmetry is found to be locally stable and globally so with respect to a one-step breaking, confidence in the replica symmetric solution seems justifiable.

Following Mézard *et al* (1987) in interpreting replica symmetry breaking in terms of the existence of disjoint ergodic components of phase space (forming 'pure' thermodynamic states), which together make up the full Gibbs state, the need to clarify the thermal average in (2.3) becomes apparent. A pure-state average, $\langle S_i S_j \rangle_\psi$ for state ψ, would be the most natural object dynamically, as this reflects the excursions of the spin system over finite times, within a single deep valley (labelled 'ψ') of its energy landscape. In contrast, a full Gibbs average would consider contributions to such an average over long enough timescales such that the spins could explore all deep valleys of the energy landscape. Although this second possibility represents a less realistic average in terms of the dynamics which we wish to address, it is the form produced by the replica method. It should be recognised that specifying over which pure state to perform the average $\langle S_i S_j \rangle$ would require more intimate knowledge of the dynamical variables than we have allowed for. The expression provided by the replica method thus represents the best guess for this quantity, in the absence of more information. Even under conditions where replica symmetry maintains, so that strictly only a single pure state contributes, the difficulty of separating broken ergodicity on the timescale of an experiment (or a simulation) from true thermodynamic broken ergodicity remains considerable.

A few comments about the search for the 'best' extremum in (3.1) are in order. Given that this extremization derives from a saddle-point integration associated with our system size, N, becoming large, is would be logical to assume that this extremum should produce the maximum of the exponent. For positive integral n there should be no disputing this, but for $n \to 0$, the analysis of Sherrington and Kirkpatrick (1975) indicates that the physically acceptable extremum is actually a minimum with respect to q. This can be accounted for by recognising that the q's are $\frac{1}{2}n(n-1)$ in number, a quantity that changes sign at $n = 1$, turning a positive Hessian into a negative one. However, no such metamorphosis occurs for the order parameter m, with respect to which our exponent should always be maximized. Therefore, for $n > 1$ we may maximize our exponent with respect to all variables, but for $n < 1$ the two classes of order parameter must be treated separately. Whilst this foible of the replica method seems to be devoid of physical implications, it is an additional complication of the numerical analysis of $\tilde{Z}_{\tilde{\beta}}$, particularly when replica symmetry breaking is entertained.

The object of central concern will be the precursor of the free energy, $F(\{m\}, \{q\})$, which we consider in two basic forms. Adopting a replica symmetric ansatz, $F(\{m\}, \{q\})$ becomes

$$
\begin{aligned}
F_{\mathrm{RS}} &= -\tfrac{1}{2}n \cdot \beta m^2 J_0 - \tfrac{1}{2}n(n-1) \cdot \tfrac{1}{2}\beta^2 q^2 \tilde{J} + G_{\mathrm{RS}} \\
G_{\mathrm{RS}} &= \ln\left[\int \mathrm{D}x \, \cosh^n\left(\beta\left[m J_0 + h + x\sqrt{q\tilde{J}}\right]\right)\right] - \tfrac{1}{2}n\beta^2 q \tilde{J}.
\end{aligned}
\tag{3.8}
$$

(We adopt the shorthand $\mathrm{D}x = \mathrm{e}^{-\frac{1}{2}x^2}\mathrm{d}x/\sqrt{2\pi}$.) For a one-step breaking in which

$$
q^{\gamma\delta} = \begin{cases} q_0 & I(\gamma/r) \neq I(\delta/r) \\ q_1 & I(\gamma/r) = I(\delta/r) \end{cases} \qquad I(x) = \inf\{y \in \mathbb{Z} : y \geqslant x\}
\tag{3.9}
$$

we obtain

$$
\begin{aligned}
F^1_{\mathrm{RSB}} &= -\tfrac{1}{2}n \cdot \beta m^2 J_0 - \tfrac{1}{2}n(r-1) \cdot \tfrac{1}{2}\beta^2 q_1^2 \tilde{J} - \tfrac{1}{2}n(n-r) \cdot \tfrac{1}{2}\beta^2 q_0^2 \tilde{J} + G^1_{\mathrm{RSB}} \\
G^1_{\mathrm{RSB}} &= \ln\left[\int \mathrm{D}x \left\{\int \mathrm{D}y \, \cosh^r\left(\beta\left[m J_0 + h + x\sqrt{q_0\tilde{J}} + y\sqrt{(q_1 - q_0)\tilde{J}}\right]\right)\right\}^{n/r}\right] \\
&\quad - \tfrac{1}{2}n\beta^2 q_1 \tilde{J}.
\end{aligned}
\tag{3.10}
$$

F^1_{RSB} is then to be extremized with respect to all order parameters, m, q_1, q_0 and r. For this one-step breaking, one may obtain the distribution of overlaps, $P(q)$ as

$$
P(q) = \frac{1-r}{1-n}\delta(q - q_1) + \frac{r-n}{1-n}\delta(q - q_0)
\tag{3.11}
$$

from which we deduce that the order parameter r must lie in the interval $[n, 1]$, if $P(q)$ is to be non-negative (Mézard *et al* 1987). We note in passing that $P(q)$, or its moments, should be preferred to the parameters $\{q_0, \ldots, q_s, r_1, \ldots, r_s\}$ (in an s-step breaking) as a means of characterising the symmetry breaking present. Different, although physically equivalent, choices of these order parameters will have identical $P(q)$'s; e.g. there are three ways in which $F^1_{\rm RSB}$ may reduce to $F_{\rm RS}$, namely $q_1 = q_0$, $r = n$ or $r = 1$.

Under conditions of replica symmetry, we may give concise expressions for the spin moments

$$\overline{\langle S_i \rangle^p} = \left[\int \mathrm{D}x \cosh^n(\Xi) \tanh^p(\Xi)\right]\left[\int \mathrm{D}x \cosh^n(\Xi)\right]^{-1}$$
$$\Xi = \beta\big(mJ_0 + h + x(q\tilde{J})^{1/2}\big) . \tag{3.12}$$

The AT-eigenvalue associated with replica symmetry breaking, is then given as

$$\lambda = \beta^2 \tilde{J}\big\{1 - \beta^2 \tilde{J}\big[1 - 2\overline{\langle S_i \rangle^2} + \overline{\langle S_i \rangle^4}\big]\big\} . \tag{3.13}$$

If $\lambda > 0$ the replica symmetric saddle-point is locally stable.

Having introduced the relevant order parameters, and the issues affecting them, we may now proceed to examine the thermodynamics of our coupled-dynamical system.

4. Theoretical results

Considering first the case $J_0 = h = 0$, we may ignore the order parameter m, and look for extrema of $F(0, \{q\})$, which would represent paramagnetic phases if $q^{\gamma\delta} = 0$, or spin-glass phases if $q^{\gamma\delta} \neq 0$, signalling an inhomogeneous average alignment of the spins. For $n \geqslant 2$ and integral, this problem reduces to that investigated by Sherrington (1980) (where the spin-glass phase corresponds to an inter-replica ordering). It was observed that both second- and first-order transitions from paramagnetic to spin-glass states could occur, although only for $n = 2$ was the former nature possible. For the present problem we find that $F_{\rm RS}$ leads to a natural interpolation between the points considered by Sherrington (1980), and that this replica symmetric solution appears to locally stable and equivalent to that produced by offering a one-step breaking (i.e., employing $F^1_{\rm RSB}$) at least for $n > 1$. If $n > 2$ we find exclusively first order transitions in $\tilde{Z}_{\tilde{\beta}}$ (as predicted by Sherrington 1980), but for $n < 2$ all such phase changes appear to be second order, and moreover to occur for $\beta\tilde{J}^{1/2} = 1$. Comparison with the analogous transition temperature for the SK model shows the two to be identical, i.e., for $n < 2$ ($T < 2\tilde{T}$) the dynamics of the couplings exert negligible influence on this transition, except insofar as these dynamics give the interaction weights a random distribution, characterized by the imposed values of J_0 and $\tilde{J}$. If $n > 2$ the mutual influence of the spins and weights produces a non-trivial change in this phase transition, altering both its order and its location (such that $\beta_c\tilde{J}^{1/2} < 1$). The transition temperature for a range of n is indicated in figure 1.

In order to explore the incidence of replica symmetry breaking, we have also considered varying the replica number n, for fixed β. We have found no evidence of first-order transitions from replica symmetric to replica symmetry broken states, only continuous transitions occurring at the point where the RS solution becomes locally unstable (where $\lambda \to 0^+$, equation (3.13)). Graphs illustrative of this effect are offered in figure 2.

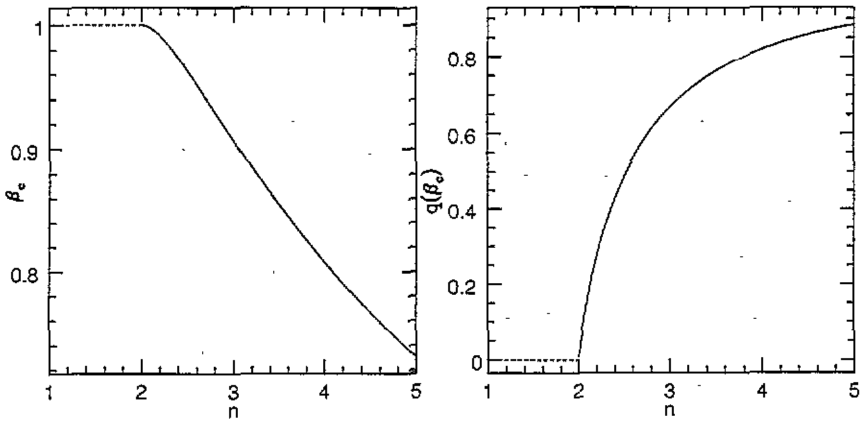

Figure 1. P → SG transition for various n, at $J_0 = 0$, $\tilde{J} = 1$. For $n > 2$ the transition is first order, but is second order elsewhere. The inverse spin temperature, β, at the transition is shown on the left; the corresponding value of q (in the SG phase) is indicated on the right.

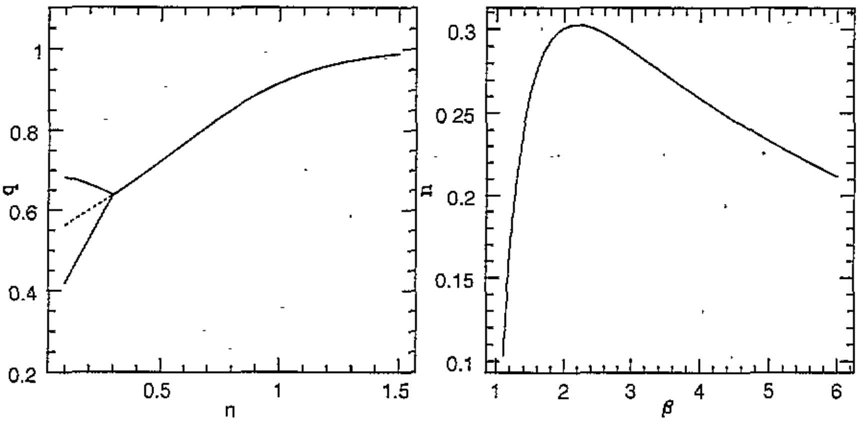

Figure 2. Indicated on the left are order parameters q_{RS} (dotted curve) and $\langle q^1_{\mathrm{RSB}}\rangle \pm \sigma(q^1_{\mathrm{RSB}})$ (full curves) at $\beta = 2$, $J_0 = 0$, $\tilde{J} = 1$. The AT line for these J's is shown on the right.

Near the transition temperature of a similar model, and for $n < 2$, Kondor (1983) has applied the full Parisi hierarchical symmetry-breaking (cf Parisi 1980b) and has demonstrated that the condition of local stability of the saddle-point assumed is more stringent than other requirements which the replica method should fulfil, such as monotonicity and convexity in n. The possibility of first-order transitions was not pertinent to this work.

It is interesting to note that replica symmetry breaking, at least as signalled by the change in local stability on the AT line, occurs only for $n < 1$ and that the value of n_{AT} appears to have a local extremum. We have not been able to make analytic statments about n_{AT} for general n, although for large β, $n_{\mathrm{AT}} \sim \beta^{-1}\tilde{J}^{-1/2}\big[2\ln(\beta\tilde{J}^{1/2})\big]^{1/2}$, as given by Kondor (1983). It would seem possible that the pathologies of the replica method, manifested as replica symmetry breaking, could show up as early as $n < 2$, where there are fewer than two replicas even though the order parameter q purports to be the mutual overlap of two replicas.

Seemingly this is not significant until rather smaller values of n. We have considered F^1_{RSB} for $n = 1 + \varepsilon$, under which circumstances we find that replica symmetry appears to hold at least to lowest order in ε, and moreover, the order parameter q may be defined in the limit $n \to 1$, and has a value which interpolates between those obtained on either side of this value. For such n, the saddle-point conditions for m and q decouple, with the condition on q being non-trivial, such that $q > m^2$

$$m = \tanh\big(\beta(mJ_0 + h)\big) \tag{4.1}$$

with q then given by minimising

$$-\tfrac{1}{4}\beta^2 q\tilde{J}(q+2) + \int \mathrm{D}x\ \ln\{\cosh(\Xi)\}\cosh(\Xi)\exp\left(-\tfrac{1}{2}\beta^2 q\tilde{J}\right) \times \left[\cosh\big(\beta(mJ_0 + h)\big)\right]^{-1}. \tag{4.2}$$

If extrema of $F(\{m\}, \{q\})$ are sought with non-zero magnetization, m, in the absence of a magnetic field, h, the parameter J_0 is always significant. With J_0 finite, there are expected to be three distinct phases: paramagnetic (P, $m = 0 = q$, indicating lack of spin ordering), spin-glass (SG, $m = 0$, $q \neq 0$, denoting a spatially-disordered local alignment of spins) and now ferromagnetic (F, $m \neq 0 \neq q$, signalling an overall net alignment of moments, with greater local order if $q > m^2$). The P$\to$SG transition is unaffected by having J_0 non-zero (because J_0 couples only to m which is zero near the transition), so the same phenomena are to be expected here as for the case above. By considering the replica symmetric saddle-point conditions

$$m = \left[\int \mathrm{D}x\ \cosh^n(\Xi)\tanh(\Xi)\right]\left[\int \mathrm{D}x\ \cosh^n(\Xi)\right]^{-1} \tag{4.3}$$

$$q = \left[\int \mathrm{D}x\ \cosh^n(\Xi)\tanh^2(\Xi)\right]\left[\int \mathrm{D}x\ \cosh^n(\Xi)\right]^{-1} \tag{4.4}$$

for small m and q, following Sherrington (1980), one may determine the orders of the various phase transitions, and where these are second order, give expressions for the relevant transition temperatures. For the P $\to$ F transition, if $\beta^{-2}\tilde{J}^{-1} > 3n - 2$ this phase change is second order, and occurs for $\beta J_0 = 1$. Again, comparison of this condition with that for the SK model shows the transition temperature to be unaffected, for this choice of n. Only if $\beta^{-2}\tilde{J}^{-1} < 3n - 2$ do the dynamics of the weights influence this transition.

Regarding the boundary separating SG and F phases, we focus on the point where this changes from a first- to second-order line. Even though the boundary is always distinct from that of the SK model, we would argue that if the order of the transition is unaffected then the influence of the weight dynamics must be marginal. We emphasize that the different relation between β and q in the SG phase, as compared with that in the SK model, means that the weight dynamics result in a repositioning of this phase boundary under all conditions, in contrast to the boundaries treated above. This effect follows from the mutual correlation of weights (3.6), given that, within the SG phase, the mean and variance of individual weights (3.5) match those of the SK model.

Sherrington (1980) observed that for $n \geqslant 4$ (and integral), the SG$\to$F transition was always second order, whilst for $n \in \{2, 3\}$ both orders were possible. We find that both orders persist beyond $n = 3$, but, owing to numerical difficulties, the limit was only inferred using an extrapolation procedure. Given that the temperature at which the SG → F transition-order changes is bounded above by the temperature of the P → SG transition, and that the numerical difficulties stem from these two becoming proximal beyond $n \sim 3$, we have extrapolated these two temperatures to coincidence (as functions of n) in order to deduce the limit beyond which the SG → F transition is uniquely second order. (The two temperature curves involved are those shown in figure 3.) It thus appears that for $n > 3.32$ ($T > 3.32\tilde{T}$) the weight-dynamics leave the order of this transition unchanged from that of a spin-system with suitably random interactions. We give our calculated values of the temperature of the order change in figure 3 (the value J_0 is uniquely determined on this line).

Examining the domain within which replica symmetry breaking occurs is a computationally intensive task. Rather than consider a one-step breaking scheme and explore where this deviates from a replica symmetric approach (as was viable for the case $J_0 = 0$), we will consider only the limit of local stability of the replica symmetric solution (i.e., the AT surface). Given that figure 2 would suggest that the AT line does indeed signal the onset of RSB, we believe that this surface should hold physical significance (again, this is not always the case, cf Krauth and Mézard 1989).

For large β (such that $m, q \to 1$), one may generalize Kondor's expression for the stability limit, such that

$$n_{\mathrm{AT}} \sim \beta^{-1} \tilde{J}^{-1/2} \left\{ |J_0| \tilde{J}^{-1/2} + \sqrt{J_0^2 \tilde{J}^{-1} + 2 \ln(\beta \tilde{J}^{1/2})} \right\}. \tag{4.5}$$

For smaller β the AT surface moves from the ferromagnetic phase (where it resides for the SK model) at lower temperatures, into the spin-glass phase for higher temperatures. A section of this surface is illustrated in figure 4. Again, only for n below about 0.3 does replica symmetry appear to break, i.e. for $T > 0.3\tilde{T}$ the dynamics of the spins should remain ergodic, with partial freezing of the system expected only for smaller temperature ratios.

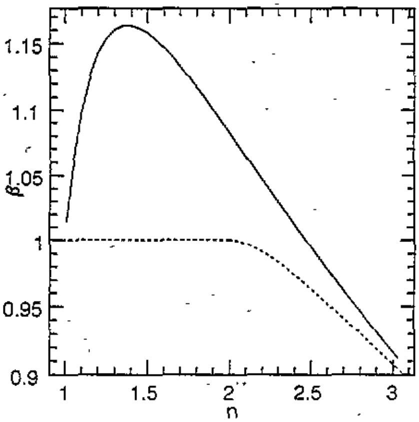

Figure 3. Inverse temperatures (β) at which SG → F transition changes order (full curve) as compared with P → SG transition (dotted curve). Both curves have $\tilde{J} = 1$.

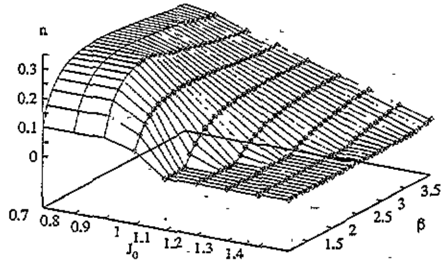

Figure 4. The de Almeida-Thouless surface for various J_0 and β with $\tilde{J} = 1$. The surface lies entirely above $n = 0$. Grid intersections represent calculated values of n; where ferromagnetic order is exhibited ($m \neq 0$), these nodes are marked with points.

5. Simulations of the dynamical system

There are two possible motivations for performing simulations for this problem, and they would lead to rather different computational strategies. Firstly, there is the question of whether the replica-method is successful in averaging an arbitrary positive power of a spin partition function. Alternatively, one might simulate the dynamical laws (2.2) and (2.3) directly and compare their behaviour with the predictions of the replica method.

Seeking to establish the viability of the replica method would focus one's simulations on (2.8); possibly with Monte Carlo integration being used for the trace over the couplings, within which the spin partition function would be evaluated for each choice of interactions. The calculation of Z_β might then be facilitated by using optimization procedures, such as were used by Kirpatrick and Sherrington (1978) in order to identify the dominating low-energy spin configurations, without requiring all the numerous spin states to be addressed individually. However, for the SK model, in which the weights J_{ij} are quenched random variables, it is argued that the free energy of the spin system ($-\beta^{-1}N^{-1}\ln Z_\beta$) should be self-averaging in the thermodynamic limit, i.e., that this quantity should be the same for all likely realizations of the weights. Therefore, if one sought to check the replica analysis by averaging $[Z_\beta]^n = \exp(n \ln Z_\beta)$, as suggested by (2.8), then one would be immediately in difficulty. That the spin free energy is self averaging implies that the choice of weights alters $\ln Z_\beta$ at a lower order than the dominant N^1, and hence that the variation in $[Z_\beta]^n$ that non-trivially weights the distribution of the J_{ij} (as implied by (3.5), (3.6)), is due to these marginal corrections. So, in order to numerically examine $\tilde{Z}_{\tilde{\beta}}$, and allied quantities such as q and m, it is likely that great care would be needed in evaluating Z_β so as to correctly weight different choices of J_{ij} in (2.8). Both because of these likely difficulties, and also because of the considerable successes of the replica method for $n \to 0$ in addition to its more rigorous application for $n \in \mathbb{Z}$, we have chosen to simulate the dynamics of the spins and interactions directly.

The modelling of (2.2), and particularly (2.3), has its own complications. The need to allow the spin system to reach equilibrium before updating the interactions (in accordance with (2.3)) may demand long equilibriation times, given the spin-glass character of this system. Further, this, potentially slow, relaxation must be endured repeatedly until the interactions have themselves been allowed sufficient time to reach equilibrium. The longevity of this two-stage process has limited us to examination of quite modest system sizes. Nevertheless some confirmation of the analytic results has been obtained, thereby, implicitly, lending support to the replica method that underlies them.

We have concentrated on $n \geqslant 2$, this being the region where the weight dynamics seem most significant, and have taken $J_0 = 0$ for simplicity. We proceed as follows: for a given set of couplings, the whole spin system is updated R_1 times, using the Metropolis algorithm, with individual spins being addressed in random order. These iterations are intended to allow the spin system to come to thermal equilibrium. The next R_2 spin-system updates are used to measure the order parameters $\langle S_i \rangle$ and the correlation functions $\langle S_i S_j \rangle$. The latter quantities are then used to evolve the weights, using a time-step $\Delta t = \tau\,\delta$ (the choice of which influences the Gaussian noise term in (2.3), such that $\mathrm{var}(\eta) \sim \delta^{-1/2}$). This whole process is repeated R_3 times (to allow for overall thermal equilibriation), after which, over the next R_4 such iterations, the order parameters m and q are obtained by averaging the values of $N^{-1}\sum_i \langle S_i \rangle$ and $N^{-1}\sum_i \langle S_i \rangle^2$ respectively. (The error bars in figures 5 and 6 are calculated, over these R_4 weight-updates, as standard deviations of the quantity $N^{-1}\sum_i \langle S_i \rangle^2$. Theory would predict these variations to vanish in the thermodynamic limit.) Thereafter the spin temperature is reduced (holding $n = \tilde{\beta}/\beta$ fixed), and the above steps

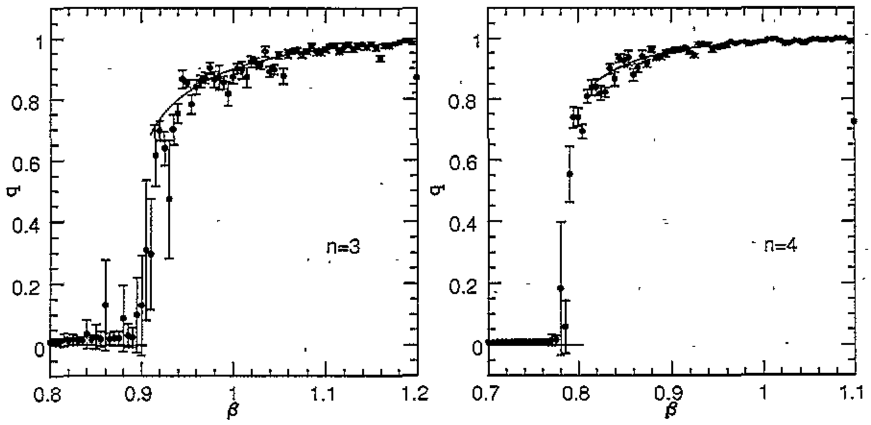

Figure 5. Comparison of theoretical order parameter, q, (full curve) with simulation results (points), for two temperature ratios, $n = 3$ and $n = 4$, both for $N = 80$, $J_0 = 0$, $\tilde{J} = 1$.

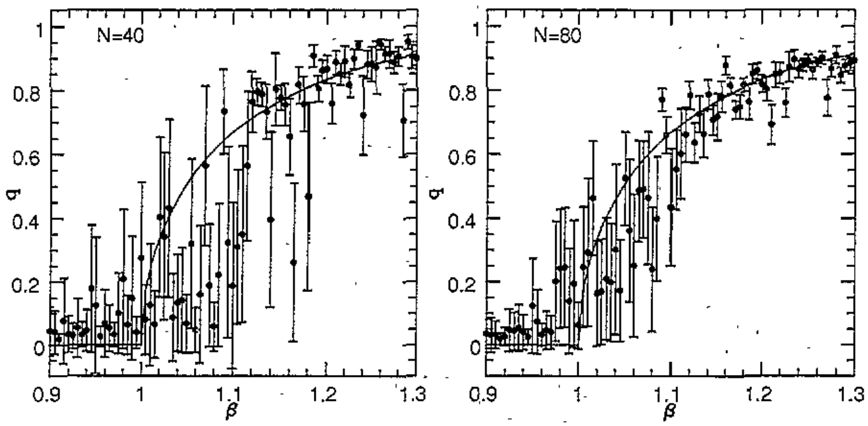

Figure 6. Theoretical order parameter, q, compared with simulation results, for $n = 2$, for two system sizes, $N = 40$ and $N = 80$, both at $J_0 = 0$, $\tilde{J} = 1$.

retraced. Typically $R_1, R_2 \sim 250$, $R_3, R_4 \sim 500$ and $\delta \sim 0.01$. That the results obtained by doubling the equilibriation and averaging times are not qualitatively different from those presented suggests that thermal equilibrium has been approached, although no more sophisticated tests have been undertaken. For reasonable sizes of system, broader time-spans would have required inordinate amounts of computer time.

Our initial conditions involve $J_{ij} = 0$ together with a random choice of S_i. A small field is applied to the spin system until the first weight update, in order to encourage ordering, but is thereafter removed. It was found rare for the spin system to jump reliably from the paramagnetic phase to the more ordered spin-glass phase (at least when a first order transition was expected), so our experiments always proceed from low to high temperature. In this way our system is likely always to start in the 'correct' phase.

Our results for $n = 3, 4$†, shown in figure 5, are very suggestive of the first-order transition expected on the basis of our calculations (and those of Sherrington 1980). For the SK model, in which the weights are quenched random variables, a second order transition at $\beta \tilde{J}^{1/2} = 1$ would be expected. That there is a plausible agreement between our replica calculations and the simulations of the dynamics that yielded them, give some credibility to the replica method for $n > 0$, as applied here.

Simulations of the dynamics for $n = 2$, where a second-order transition is expected at $\beta \tilde{J}^{1/2} = 1$, also show fair agreement with theory (figure 6), although finite-size effects appear to be more significant than for larger n. Examination of various increasing choices of system size, N, indicates definite reduction in the noise in the measured q, with improved agreement between theory and experiment towards larger N, although any finite-size scaling analysis would not seem justifiable for the data available.

6. Conclusion

We have investigated a spin model with two levels of dynamics, in which both spins and their interactions are dynamical variables. Mathematical analysis using the replica method, but for finite positive replica number, n, has indicated the existence of phase transitions in the system, whose nature can differ markedly from that of the underlying spin system in isolation. Instances where the influence of the spin system on the couplings leads to changes in both the order and location of phase transitions, relative to those for suitably random choices of interactions, have been predicted in addition to cases where the interdependence of the two sets of dynamical laws has insignificant effects. The pertinence of the relative noise levels of the spin and weight dynamics (i.e. $\tilde{\beta}/\beta$), has been indicated. Simulation results have shown good correspondence with the theoretical predictions.

In addition to offering some insight into likely properties of a class of coupled dynamical systems, the similarity of simulations to our predictions based on the replica method, in a less familiar realization, provides some evidence that the replica method is trustworthy over a wide range of positive n, at least where replica symmetry is preserved.

Towards the closing stages of our work we have learned that the replica method for $n < 0$ was being investigated by Mézard *et al* (private communication). In a dynamical context, this would correspond to an anti-Hebbian law for the interaction dynamics.

Acknowledgments

We would like to thank Dominic O'Kane for advice while this work was in progress. RWP would like to thank Jesus College, Oxford, for its kind award of a scholarship. Financial support from the SERC (under grant number 9130068X) and from the EEC (under grant number S/SC1*915121), is greatly appreciated.

References

de Almeida J R L and Thouless D J 1978 Stability of the SK solution of a spin glass model *J. Phys. A: Math. Gen.* **11** 983

† That n is here an integer is of no particular significance—see figure 1.

Amit D J, Gutfreund H and Sompolinsky H 1987 Statistical mechanics of neural networks near saturation *Ann. Phys., NY* **173** 30

Coolen A C C and Ruijgrok Th W 1988 Image evolution in Hopfield networks *Phys. Rev.* A **38** 4253

de Dominicis C and Kondor I 1983 Eigenvalues of the stability matrix for the Parisi solution of the long-range spin glass *Phys. Rev.* B **27** 606

Jonker H J J and Coolen A C C 1991 Unsupervised dynamic learning in layered neural networks *J. Phys. A: Math. Gen.* **24** 4219

Gardner E J 1988 The space of interactions in neural network models *J. Phys. A: Math. Gen.* **21** 257

Kirkpatrick S and Sherrington D 1978 Infinite-ranged models of spin glasses *Phys. Rev.* B **17** 4384

Kondor I 1983 Parisi's mean-field solution of spin glasses as an analytic continuation in the replica number *J. Phys. A: Math. Gen.* **16** L127

Krauth W and Mézard M 1989 Storage capacity of memory with binary couplings *J. Physique* **50** 3057

Mézard M, Parisi G and Virasoro M 1987 *Spin Glass Theory and Beyond* (Singapore: World Scientific)

Parisi G 1980a The order parameter for spin glasses: a function on the interval 0–1 *J. Phys. A: Math. Gen.* **13** 1101

—— 1980b Magnetic properties of spin glasses in a new mean field theory *J. Phys. A: Math. Gen.* **13** 1887

Sherrington D 1980 Ising replica magnets *J. Phys. A: Math. Gen.* **13** 637

Sherrington D and Kirkpatrick S 1975 Solvable model of a spin glass *Phys. Rev. Lett.* **35** 1792

Shinomoto S 1987 Memory maintenance in neural networks *J. Phys. A: Math. Gen.* **20** L1305

van Hemmen J L and Palmer R G 1979 The replica method and a solvable spin glass model *J. Phys. A: Math. Gen.* **12** 563