J. Phys. A: Math. Gen. **21** (1988) 2995-3011. Printed in the UK

The roles of stability and symmetry in the dynamics of neural networks

Werner Krauth†, Jean-Pierre Nadal‡ and Marc Mézard§

† Département de Physique, Ecole Normale Supérieure, 24 rue Lhomond, 75231 Paris Cedex 05, France

‡ Groupe de Physique des Solides, Ecole Normale Supérieure, 24 rue Lhomond, 75231 Paris Cedex 05, France

§ Laboratoire de Physique Théorique, Ecole Normale Supérieure, 24 rue Lhomond, 75231 Paris Cedex 05, France

Received 28 March 1988

Abstract. In this paper we study the retrieval phase of spin-glass-like neural networks. Considering that the dynamics should depend only on gauge-invariant quantities, we propose that two such parameters, characterising the symmetry of the neural net's connections and the stabilities of the patterns, are responsible for most of the dynamical effects. This is supported by a numerical study of the shape of the basins of attraction for a one-pattern neural network (OPN) model. The effects of stability and symmetry on the short-time dynamics of this model are studied analytically, and the full dynamics for vanishing symmetry is shown to be exactly solvable.

1. Introduction

During the past few years there has been tremendous interest in the theory of neural networks. Outstanding popularity was gained by the Hopfield model (Hopfield 1982, Amit 1987) which, by the symmetry of its interactions and by its Monte Carlo dynamics, has had great appeal for physicists trained in statistical mechanics. Examples have shown, however, that similar behaviour may be reached with quite different architectures and philosophies, such as the asymmetric strict-stability models (Kohonen 1984, Personnaz *et al* 1985, Kanter and Sompolinsky 1987). At present, many questions still appear to be open, concerning architectures, learning rules, storage prescriptions, etc.

Despite the many open issues, all neural networks will, in the retrieval phase, function basically in the same way: started close to a memory state, a given initial configuration will flow towards it, and there will be interference effects due to other memory states which will slow down and possibly inhibit convergence.

In this paper we will exclusively be interested in this retrieval phase. We will present a particular model which allows us to study retrieval without having to specify the details of the learning phase of the system. More specifically, we propose that the symmetry of the net's connections and the stabilities of the memories are responsible for most of the dynamical effects in spin-glass-like neural networks. Our one-pattern neural network (OPN) model highlights the aspects of the stability and of the symmetry, quantities whose definition will be given in the next section.

Working with the OPN model instead of with a complete neural network (with a given set of patterns and a specified learning rule) allows us to make quite general statements, but it involves approximations. It is therefore a crucial step in our argument

0305-4470/88/132995+17$02.50

that it is possible to step back from the complete network to the OPN model. We show in § 3 that, at least for the two learning rules we checked, the OPN approximation preserves the shape of the basins of attraction.

In the large-N (number of neurons) limit, the OPN model becomes quite simple, and independent of many details of the underlying net. We will present analytical calculations (§ 4) which will allow us to gain further insight into the dynamics and, hopefully, into the workings of neural nets. Our conclusions will then be summarised in § 5.

2. Stability, symmetry and gauge transformations

Let us first recall a number of well established facts and fix our notations. In all situations below we regard a network with spins $S_i = \pm 1$ ($i = 1, \ldots, N$) and an ($N \times N$) matrix (J_{ij}) of synaptic couplings J_{ij} with $J_{ii} = 0$. For simplicity we restrict ourselves to parallel dynamics at zero temperature

$$S_i(t+1) = \operatorname{sgn}\left(\sum_j J_{ij} S_j(t)\right) \qquad i = 1, \ldots, N. \tag{2.1}$$

We denote the patterns as $\boldsymbol{\xi}^\mu = (\xi_1^\mu, \xi_2^\mu, \ldots, \xi_N^\mu)$, $\mu = 1, \ldots, p$, and the storing ratio as $\alpha = p/N$.

The 'stabilities' Δ_i, Δ_i^μ, are quantities defined on all rows of the matrix (J_{ij}) as

$$\Delta_i = \min_\mu \Delta_i^\mu \qquad \Delta_i^\mu = \sum_j \xi_i^\mu J_{ij} \xi_j^\mu \left(\sum_j J_{ij}^2\right)^{-1/2}. \tag{2.2}$$

If $\Delta_i^\mu > 0$, $i = 1, \ldots, N$, the pattern $\boldsymbol{\xi}^\mu$ is a fixpoint of the dynamics (2.1). Many algorithms have been proposed to compute a matrix (J_{ij}) such that this condition is fulfilled (Minski and Papert 1969, Gardner *et al* 1987b, Pöppel and Krey 1987, Diederich and Opper 1987), and an optimal stability algorithm was published recently (Krauth and Mézard 1987). For random patterns ($\xi_i^\mu = \pm 1$ with probability $\frac{1}{2}$) it is not possible to store more than $2N$ patterns with $\Delta_i > 0$, $i = 1, \ldots, N$ (Venkatesh 1986, Gardner 1987). For fixed α, Δ_i will be constant as $N \to \infty$.

Recently Forrest has shown the importance of the stability for the dynamics of a symmetric neural network, obtained with the learning algorithm of Gardner *et al* (1987b): the greater the stability, the larger the basin of attraction (Forrest 1987). We have shown in addition that, on an asymmetrically diluted network (in the limit of extreme dilution), the typical stability is the only dynamically relevant parameter if the elements of the matrix (J_{ij}) are all of the same order (Mézard *et al* 1988).

We now define the 'symmetry' of the matrix (J_{ij}) as

$$\eta = \sum_{i \neq j} J_{ij} J_{ji} \left(\sum_{i \neq j} J_{ij}^2\right)^{-1} \tag{2.3}$$

η measures the relative weights of the symmetric ($2J_{ij}^{\mathrm{s}} = J_{ij} + J_{ji}$) and the antisymmetric ($2J_{ij}^{\mathrm{a}} = J_{ij} - J_{ji}$) parts of ($J_{ij}$):

$$\eta = \left(\sum_{i \neq j} (J_{ij}^{\mathrm{s}})^2 - \sum_{i \neq j} (J_{ij}^{\mathrm{a}})^2\right)\left(\sum_{i \neq j} J_{ij}^2\right)^{-1}. \tag{2.4}$$

In particular, $\eta = 1$ (respectively -1) for a fully symmetric (respectively fully antisymmetric) matrix. $\eta = 0$ means that symmetric and antisymmetric parts have the same weights, which is, for example, the case if, for $i < j$, J_{ij}^{s} and J_{ij}^{a} are random variables

of the same distribution. The relevance of η to the dynamics of spin-glass-like systems has been noted by several authors (Toulouse 1988, Gutfreund *et al* 1987, Rieger *et al* 1988). For convenience we mention the relation of η with the parameter k of Gutfreund *et al* (1987):

$$\eta = \frac{1-k^2}{1+k^2}.$$

We will several times consider gauge transformations about a state $\boldsymbol{S}$, defined by

$$J_{ij} \to J_{ij}S_iS_j \qquad \xi_i^\mu \to \xi_i^\mu S_i. \tag{2.5}$$

The transformation (2.5) leaves the dynamics of the network unchanged. Conversely, only gauge-invariant quantities can be important for the dynamics of the neural network. The symmetry η and the stabilities Δ_i are gauge invariant. As pointed out by Derrida (1988), other such quantities are, for example, the higher-order correlations

$$\sum_{j,k,\ldots,m} J_{ij}J_{jk}\ldots J_{mi} \qquad \text{or} \qquad \sum_{j,k,\ldots,m} \xi_i^\mu J_{ij}J_{jk}\ldots J_{mn}\xi_n^\mu \tag{2.6}$$

whose influence on the dynamics of the net cannot be excluded *a priori*. We excluded another such quantity, J_{ii}, from our considerations by setting it equal to zero. On a fully connected network a diagonal term of order 1 (for non-diagonal terms of order $1/\sqrt{N}$) can be shown to influence the convergence properties favourably (Mézard *et al* 1988).

The gauge transformation (2.5) thus has a certain fundamental importance in our context. In addition, we will make use of the transformation (2.5) about pattern $\boldsymbol{\xi}^\mu$, whenever treating $\boldsymbol{\xi}^\mu$ explicitly, as a simple change of coordinates. $\boldsymbol{\xi}^\mu$ then conveniently transforms into $\mathbf{1} := (1, 1, \ldots, 1)$ and the definitions of magnetisation $q = \Sigma_i S_i/N$ and overlap with pattern $\boldsymbol{\xi}^\mu (\Sigma_i S_i \xi_i^\mu / N)$ become equivalent.

3. One-pattern model

We define a one-pattern neural network (OPN) model to consist of a neural network with one single pattern, $\boldsymbol{\xi}^1 = \mathbf{1}$. The coupling matrix (J_{ij}) does not result from a learning procedure; (J_{ij}) is a random matrix, with elements J_{ij} taken from a given distribution (such as $J_{ij} = \pm 1$), which are subjected to two types of constraints. First, the pattern $\mathbf{1}$ is to be stable with a fixed stability $\Delta_i^1 > 0$, $i = 1, \ldots, N$, so that

$$\sum_j J_{ij} = \Delta_i^1 \left(\sum_j J_{ij}^2 \right)^{1/2} \qquad i = 1, \ldots, N. \tag{3.1}$$

Second, the matrix (J_{ij}) is to have a certain degree of symmetry η, as defined in equation (2.4).

While $\mathbf{1}$ is thus the only pattern which is retained explicitly in the OPN model, the quantities η and Δ are meant to represent of the order of N other patterns, and the details of the learning algorithm. By taking the values of all the stabilities greater than zero, we mimicked a strict-stability learning algorithm. Had we allowed fluctuations of the stabilities of the same order as their mean value, for example by putting

$$\sum_j J_{ij} = (\Delta + z_i) \left(\sum_j J_{ij}^2 \right)^{1/2} \qquad i = 1, \ldots, N \tag{3.2}$$

with z_i a normalised random variable, a model closely related to the Hopfield model would have resulted. Note that a formula like equation (3.2) with $\Delta = 1/\sqrt{\alpha}$ holds in the Hopfield model. We will elaborate some differences between the strict-stability OPN model and the OPN model with average stability in § 4.2.

It is a key point in this paper that the dynamics of a neural network (with $p = \alpha N$ patterns, and a specific learning rule) close to a given pattern $\boldsymbol{\xi}^\mu$ depends mainly on Δ_i^μ and on η, i.e. is similar to the dynamics of the corresponding OPN model (whose matrix (J_{ij}) has unchanged row sums (3.1), identical symmetry η and the same distribution function for the J_{ij}).

3.1. Numerical testing of the OPN model

There are many aspects of the dynamics one could compare, among which we choose to restrict ourselves to a detailed study of the shape of the basins of attraction. Our numerical studies indicate, however, that this comparison may bear also on other points, such as convergence times or short-time dynamics.

We performed computer simulations, making use of two different learning algorithms (see below), and of a randomising procedure which transformed a given matrix (J_{ij}) into a corresponding OPN matrix: this procedure RANDOMISE, given in table 1, scrambles the matrix, keeping the stability of the chosen pattern and establishing the symmetry at a prescribed value.

A numerical simulation run would now proceed as follows.

(i) For a certain number of random patterns, we compute a coupling matrix (J_{ij}) using one specific learning algorithm (see later) and determine its symmetry η.

(ii) We pick an index μ with $1 \leqslant \mu \leqslant p$ at random and transform $\boldsymbol{\xi}^\mu$ into $\mathbf{1}$ using equation (2.5).

(iii) For typically 200 arbitrary initial states per magnetisation q_0 we follow the dynamics (equation (2.1)) during 50 time steps, if the state does not get trapped into

Table 1. Symbolic listing of subroutine RANDOMISE (see, for example, Papadimitriou and Steiglitz (1982) for the symbolic programming language used). On output the $(N \times N)$ matrix (J_{ij}) will be a scrambled version of the input matrix, with unchanged row sums and symmetry $= \eta'$ (η' may be positive or negative). The first row of (J_{ij}) remains unchanged in order to eliminate a permutative degree of freedom.

```
procedure RANDOMISE
for i := 2, 3, ..., N do (comment: random permutations)
begin
                J(i, i) ↔ J(i, N)
                for k := N − 1, N − 2, ..., 2 do J(i, k) ↔ J(i, ran(k))
                (comment: ran(k) produces random integers between 1 and k)
                J(i, i) ↔ J(i, N)
end
compute η (comment: see equation (2.3)) and norm := Σ_{i,k} J(i, k)^2
while |η/η'| < 1 do (comment: Monte Carlo step)
begin
                pick a triplet (i, k, l) of mutually different random integers with 2 ≤ i ≤ N,
                1 ≤ k ≤ N, 1 ≤ l ≤ N
                del := (J(i, l) − J(i, k))J(k, i) + (J(i, k) − J(i, l))J(l, i)
                if del sgn(η') > 0 then J(i, k) ↔ J(i, l), η' := η' + del/norm
end
```

a fixed point or a cycle before this point. We record, among other quantities, the probability of perfect recall as a function of q_0 (this way of doing the simulations is due to Forrest (1987)).

(iv) Using the procedure RANDOMISE (see table 1) we scramble the matrix (J_{ij}) and create random matrices with unchanged row sums and various values of the symmetry η', in particular with the original degree of symmetry $\eta' = \eta$.

(v) We repeat (iii) for each value of η'.

3.1.1. Optimal stability algorithm. The first algorithm we considered was the optimal stability learning rule (Krauth and Mézard 1987) which, for a given set of patterns, has been proven to find the matrix (J_{ij}) with maximal values of the Δ_i (see equation (2.2); in the present context this fact is, however, of no importance). We took values of N between 100 and 400, and varied α between $\frac{1}{4}$ and $\frac{1}{2}$. Results of one such run with $N = 100$ and $p = 50$ are shown in figure 1, which shows the probability of perfect recall p_{perf} as a function of the initial overlap q_0. The data thus give the probability of flowing towards pattern **1** when starting with a random state of magnetisation q_0. The two curves result from least-squares fits to functions

$$p_{\text{perf}} \approx \tfrac{1}{2}\{\tanh[a(q_0 - q_c)] + 1\} \tag{3.3}$$

with q_c the critical overlap. In this particular example $0.82 \leqslant \Delta_i^\mu$, $\Sigma_i \Delta_i^\mu / N = 1.15$; the matrix is almost symmetric: $\eta = 0.97$. We see in figure 1 that the OPN model is capable of reproducing qualitatively the dynamics of the optimal stability matrix. The values of a (in equation (3.3)) correspond closely, and the critical overlap q_c differs slightly. OPN matrices with $\eta' = 0.5$ and $\eta' = 0$ give a poorer description of the original dynamical behaviour.

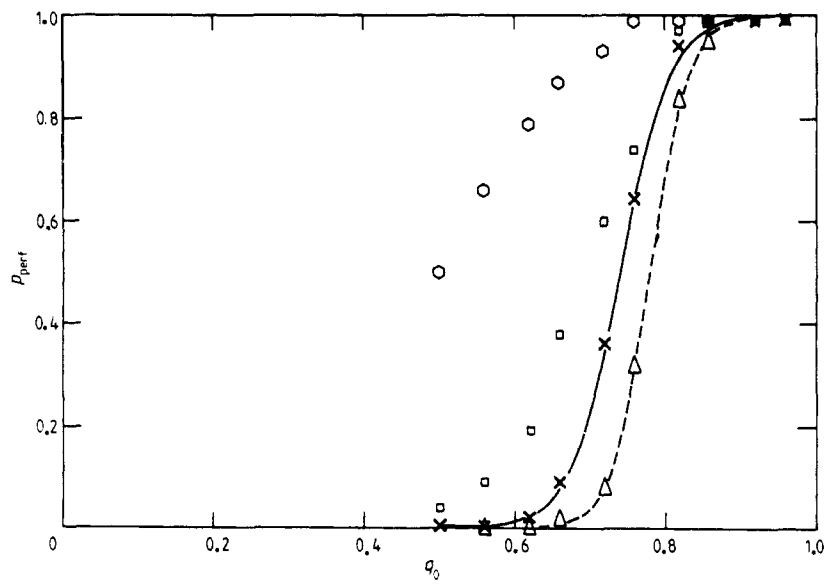

Figure 1. Comparison of the basins of attraction between the optimal stability net (×, $\eta = 0.97$, $\Sigma_i \Delta_i / N = 1.15$) and the corresponding OPN models (△, $\eta' = \eta$; □, $\eta' = 0.5$, ○, $\eta' = 0$) for $N = 100$ and $p = 50$ random patterns. Shown is the probability of perfect recall p_{perf} as a function of the initial overlap q_0 with pattern 1. The full curve is a least-squares fit with function equation (3.3) for the original dynamics and the broken curve corresponds to the OPN model dynamics with $\eta' = \eta$.

Figure 1 is typical of what we have found with the minimum overlap algorithm: there are finite differences between the original and the scrambled matrix dynamics with $\eta' = \eta$. q_c is smaller for the original matrix than for the OPN matrix if α is large, and smaller if α is small, the two dynamics being identical for α of order $\frac{1}{3}$. These differences seem to remain finite as $N \to \infty$, and do not vary much with N for the values of N we considered. Choosing values of η' different from η tends to give larger variances with the original dynamical behaviour; the dependence on η is more pronounced for larger values of the stabilities. In any case, the slope a in equation (3.3) seems to go to infinity with N, indicating the convergence towards a step function ($p_{\text{perf}} = 0$ for $q_0 < q_c$; $p_{\text{perf}} = 1$ for $q_0 > q_c$), in agreement with the results of Forrest (1987).

3.1.2. Simplex-based algorithm. We repeated the same simulations as in § 3.1 for the simplex-based learning rule (Krauth and Mézard 1987). This rule differs from the preceding algorithm in that the quantities

$$\Delta_i' = \sum_j \xi_i^\mu J_{ij} \xi_j^\mu \left(\max_j |J_{ij}| \right)^{-1} \tag{3.4}$$

are optimised. In view of the larger algorithmic complexity of the simplex-based algorithm we restricted ourselves to $20 \leqslant N \leqslant 80$ and treated again cases with $\frac{1}{4} \leqslant \alpha \leqslant \frac{1}{2}$. Results of two runs are given in figure 2. Surprisingly, there is hardly any difference between the original matrix and the OPN model dynamics (with $\eta' = \eta$), for any value of α. Again, the dynamical properties of the OPN system differ widely for large values of Δ_i, less so for small stability.

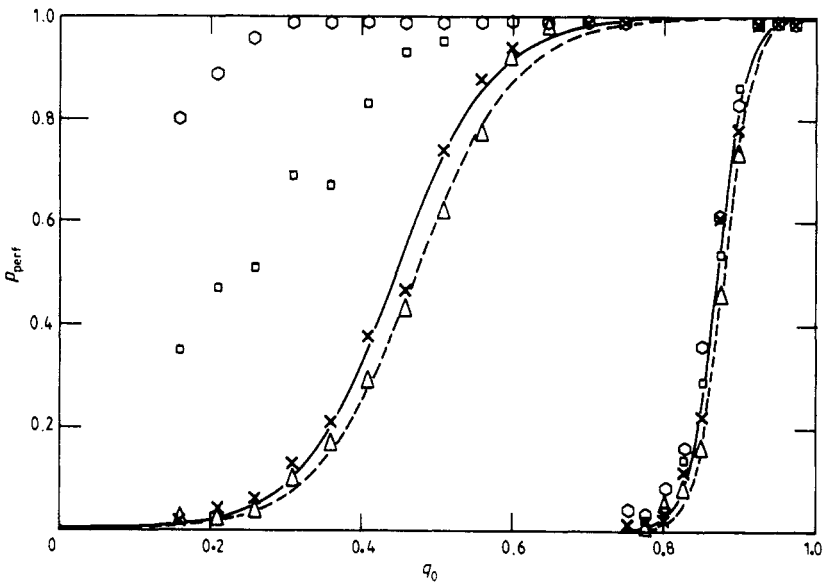

Figure 2. Comparison of neural network dynamics (×, simplex net) for $N = 80$ with $p = 20$ (left: $\eta = 0.78$, $\Sigma_i \Delta_i / N = 1.44$) and $p = 40$ (right: $\eta = 0.76$, $\Sigma_i \Delta_i / N = 0.90$) random patterns with the corresponding OPN model dynamics (△, $\eta' = \eta$; □, $\eta' = 0.5$; ○, $\eta' = 0$). Shown is the probability of perfect recall p_{perf} as a function of initial overlap q_0. The full curve represents the full neural network and the broken curve corresponds to the OPN model dynamics with $\eta' = \eta$. Note the excellent agreement of the two systems for $\eta' = \eta$ and the poor agreement for differing values of the symmetry.

3.2. Comments

In the OPN model we search for a description of the dynamical behaviour close to $\boldsymbol{\xi}^\mu$ *not* as a function of all the input patterns and of the learning rule, but as a function of a few gauge-invariant combinations of $\boldsymbol{\xi}^\mu$ and of (J_{ij}). It seems plausible that not all of the higher-order correlations in equation (2.6) are relevant. The procedure RANDOMISE allows us to test the hypothesis that η and Δ_i are the most important such quantities.

With a judicious choice of parameters, qualitative agreement between the full matrix model and the OPN model was perhaps to be expected. We are surprised by the excellent agreement for the simplex-based rule, and we have not yet tried to trace the origins of the differences in the case of the minimum overlap algorithm.

Note that the choice of uncorrelated patterns $\boldsymbol{\xi}^\mu$ and of the two algorithms is somewhat accidental and of no real significance in the context of this paper.

4. OPN model in the large-N limit

The results of § 3 seem to us to justify an attempt to understand the OPN model in somewhat greater detail. We will present therefore our analytical calculations for a particular OPN model for which we have calculated exactly the dynamics for short times. For the first time step $t = 1$, the dynamics depends solely on the stability since the sites are uncoupled. The effects of the symmetry are visible starting from $t = 2$. For large t, we expect there to be a sharp transition between a region of strong recall and a region of weak recall. Our calculation up to $t = 4$ provides already a good indication of this transition.

The precise OPN model we consider consists of a matrix (J_{ij}) with $J_{ij} = \pm 1$ $(J_{ii} = 0)$, for which there is the same ('typical') stability on all rows:

$$\sum_j J_{ij} = \Delta\sqrt{N} \qquad i = 1, \ldots, N. \tag{4.1}$$

As before, we impose a certain symmetry η. Beyond these conditions (J_{ij}) is random. This model differs only slightly from the one considered in § 3: the choice of integer values for J_{ij} will be seen to be of no significance in the limit $N \to \infty$ (only the second moment of the J_{ij} distribution enters the calculation), and equation (4.1) differs from equation (3.1) only in that the stabilities Δ_i are replaced by a typical stability Δ.

The symmetry η can be imposed on the matrix (J_{ij}) by splitting J_{ij} into symmetric and antisymmetric parts (see above equation (2.4)) and by choosing independently

$$\begin{aligned} &(J^{\mathrm{s}}_{ij}, J^{\mathrm{a}}_{ij}) = (\pm 1, 0) && \text{with probability } \tfrac{1}{4}(1+\eta) \text{ each} \\ &(J^{\mathrm{s}}_{ij}, J^{\mathrm{a}}_{ij}) = (0, \pm 1) && \text{with probability } \tfrac{1}{4}(1-\eta) \text{ each} \end{aligned} \tag{4.2}$$

while at the same time keeping the constraints (4.1) enforced by a product of δ functions.

In the limit $N \to \infty$ we determined the dynamics of this model for four time steps, i.e. we calculated the expectation values

$$q_1 = \langle S^{t=1} \rangle, \ldots, q_4 = \langle S^{t=4} \rangle \tag{4.3}$$

and correlation functions such as

$$q_{02} = \langle S^{t=0} S^{t=2} \rangle. \tag{4.4}$$

The mean $\langle\,\rangle$ denotes the trace taken over the couplings (4.2) *and* over initial states with $q_0 = q^{t=0}$ fixed.

4.1. Principle of the calculation, $t = 1$

Our calculations are quite similar to those on the spin-glass model by Gardner *et al* (1987a). They are as cumbersome as well. We choose to sketch only the calculation of q^1 in the main text (using the systematic way that works also for q^t, $t > 1$) in order to make plausible where the problem resides, and we relegate the more involved calculations for $t > 1$ to the appendix. As a matter of fact, the calculation at $t = 1$ can be phrased much more simply, but without any possibility of generalising it (Mézard *et al* 1988).

To determine q_1 we calculate the averaged 'partition function'

$$\bar{Z} = \mathrm{Tr}_J \, \mathrm{Tr}_{S^1} \, \delta\left(\sum_j J_{ij} - \Delta\sqrt{N}\right) \prod_i \left(S_i^1 \sum_j J_{ij} S_j^0\right) \tag{4.5}$$

where Tr_J goes over the four cases in formula (4.2). Introducing integral representations for the Kronecker δ

$$\begin{aligned}\delta(J_{ij} - \Delta\sqrt{N}) &= \int_{-\pi}^{\pi} \frac{\mathrm{d}\gamma_i}{2\pi} \exp\left[\mathrm{i}\gamma_i\left(\sum_j J_{ij} - \Delta\sqrt{N}\right)\right] \\ &= \int_{-\infty}^{\infty} \frac{\mathrm{d}\gamma_i}{2\pi\sqrt{N}} \exp\left[\mathrm{i}\gamma_i\left(\frac{1}{\sqrt{N}}\sum_j J_{ij} - \Delta\right)\right]\end{aligned} \tag{4.6}$$

and for the Heaviside function θ

$$\theta\left(S_i^1 \sum_j J_{ij} S_j^0\right) = \int_{-\infty}^{\infty} \mathrm{d}x_i \int_0^{\infty} \frac{\mathrm{d}\lambda_i}{2\pi} \exp\left[\mathrm{i}x_i\left(\frac{1}{\sqrt{N}}\sum_j J_{ij} S_j^0 - \lambda_i S_i^1\right)\right] \tag{4.7}$$

we arrive at

$$\begin{aligned}\bar{Z} = \mathrm{Tr}_J \prod_i \mathrm{Tr}_{S_i^1} \int_{-\infty}^{\infty} \frac{\mathrm{d}\gamma_i}{\sqrt{N}2\pi} \int_0^{\infty} \frac{\mathrm{d}\lambda_i}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \frac{\mathrm{d}x_i}{\sqrt{2\pi}} \exp(-\mathrm{i}\gamma_i\Delta - \mathrm{i}x_i\lambda_i S_i^1) \\ + \sum_{i<j}\left(\frac{\mathrm{i}}{\sqrt{N}} J_{ij}^{\mathrm{s}}(\gamma_i + \gamma_j + x_i S_j^0 + x_j S_i^0) + \frac{\mathrm{i}}{\sqrt{N}} J_{ij}^{\mathrm{a}}(\gamma_i - \gamma_j + x_i S_j^0 - x_j S_i^0)\right).\end{aligned} \tag{4.8}$$

In order to calculate any useful macroscopic variables (such as q_1), one will have to introduce an additional *ad hoc* source term (e.g. $\exp(h \sum_i \sigma_i^1)$; q_1 will then be given by $\partial \log Z(h)/\partial h$ at $h = 0$).

Taking the trace over J_{ij}^{s} and J_{ij}^{a} turns the term inside large round brackets in equation (4.8) into

$$\begin{aligned}(\,) \to \log\left[\frac{1+\eta}{2} \cos\left(\frac{\gamma_i + \gamma_j + x_i S_j^0 + x_j S_i^0}{\sqrt{N}}\right)\right. \\ \left. + \frac{1-\eta}{2} \cos\left(\frac{\gamma_i - \gamma_j + x_i S_j^0 - x_j S_i^0}{\sqrt{N}}\right)\right]\end{aligned} \tag{4.9}$$

which after expanding the trigonometric functions gives

$$\exp \sum_{i<j} (\,) \to \prod_{i \neq j} \exp\left(-\frac{1}{2N}[(\gamma_i + x_i S_j^0)^2 + \eta(\gamma_i + x_i S_j^0)(\gamma_j + x_j S_i^0)]\right). \tag{4.10}$$

All the tediousness of the calculation now stems from the term in η which forces us to introduce Gaussian integrations in order to decouple sites $1, \ldots, N$, and afterwards

to use saddle-point integrations to eliminate the extra variables (order parameters). The results for $t=1$ are

$$q_1=\left(\frac{2}{\pi}\right)^{1/2}\int_0^{\Delta q_0/(1-q_0^2)^{1/2}} \mathrm{d}z\, \mathrm{e}^{-z^2/2}=:\operatorname{erf}\left(\frac{\Delta q_0}{[2(1-q_0^2)]^{1/2}}\right) \tag{4.11}$$

$$q_{01}=q_0 q_1 . \tag{4.12}$$

As expected (Mézard *et al* 1988), q_1 does not depend on the symmetry at all, but only on Δ. There are two cases. For $\Delta \leqslant \sqrt{\pi/2}$ the function q_1 is partly below and partly above the diagonal, while for $\Delta \geqslant \sqrt{\pi/2}$ the mean magnetisation at $t=1$ is larger than the magnetisation at $t=0$ for all $q_0>0$. The connected correlation function $\langle S^0 S^1\rangle_c := q_{01}-q_0 q_1$ is always zero. It is for this reason that simpler derivations of equation (4.11) are possible.

4.2. Solution for $\eta=0$, all times

For $t>1$ the calculation shown in § 4.1 has to be generalised. The changes consist in additional θ functions and in traces over spin configurations at later times. Gaussian integrations now become inevitable; they serve to decouple a growing number of terms in η (cf equation (4.10)) and, physically, to fix a growing number of order parameters, the connected correlation functions (see the appendix).

There is, however, the special case $\eta=0$, in which the troublesome term in equation (4.10) is multiplied by 0. In this case of 'zero symmetry', it may be shown by recursion that we can solve the dynamics explicitly for all times: $q^{(1)}$ is an iteration graph and $q^{(t+1)}=q^{(1)}(q^{(t)})$:

$$q_{t+1}=\operatorname{erf}\left(\frac{\Delta q_t}{[2(1-q_t^2)]^{1/2}}\right) \qquad \eta=0 \qquad t=0,1,\ldots \tag{4.13}$$

(cf equation (4.11)). Thus, the case $\eta=0$ is equivalent to the dynamics on the asymmetrically diluted lattice, which was introduced by Derrida *et al* (1987). The fact that simplifications arise for vanishing symmetry in spin-glass models has been noted by several authors (Gutfreund *et al* 1987, Toulouse 1988, Rieger *et al* 1988, Crisanti and Sompolinsky 1988).

For zero symmetry all the connected correlation functions are zero; the system has no dynamical memory: S^{t_1} and S^{t_2} have just the same overlap as two randomly chosen states with magnetisation q^{t_1} and q^{t_2}. In this case we are able to exactly calculate the critical overlap q_c, which corresponds to the unstable fixpoint of the iteration graph. For $\Delta<\sqrt{\pi/2}$, $0<q_c<1$ ($q=0$, $q=1$ are stable fixpoints) and for $\Delta>\sqrt{\pi/2}$, $q_c=0$ (the two smaller fixpoints have merged). There remains a single stable fixpoint of the iteration (4.13), the pattern **1**.

It is worth mentioning here that, if the stabilities Δ_i fluctuate, with for example

$$P(\{\Delta_i\},\ i=1,\ldots,N)=\prod_i P(\Delta_i)$$

which leads to equation (3.2), the dynamics is again solvable at $\eta=0$. By a similar calculation, one gets that, for $\eta=0$ and $t\geqslant 1$,

$$q_{t+1}=\int \mathrm{d}\Delta\, P(\Delta)\operatorname{erf}\left(\frac{\Delta q_t}{[2(1-q_t^2)]^{1/2}}\right). \tag{4.14}$$

A Hopfield-type case would correspond to a Gaussian distribution:

$$P(\Delta_i) = (2\pi)^{-1/2} \exp[-\tfrac{1}{2}(\Delta - \Delta^*)^2] \tag{4.15}$$

with

$$\Delta^* = 1/\sqrt{\alpha} \tag{4.16}$$

and this gives

$$q_{t+1} = \mathrm{erf}(\Delta^* q_t/\sqrt{2}). \tag{4.17}$$

As expected, $q_\infty < 1$. Interestingly, not all connected correlations vanish. We do have, for any $t > 0$, $q_{t0} = q_t q_0$, but for $0 < t' < t$, $q_{t't}$ is different from $q_t q_{t'}$ and is a function of $\{q_{t'-1,t-1}, q_{t-1}, q_{t'-1}\}$.

4.3. Results for $2 \leqslant t \leqslant 4$, $\eta \neq 0$

In the case of non-zero symmetry we are not allowed to iterate q^1: the correlations between J_{ij} and J_{ji} induce non-trivial correlations between $\boldsymbol{S}(t_1)$ and $\boldsymbol{S}(t_2)$; however, all connected correlation functions between times differing by an odd number of steps remain zero (as in equation (4.11)). This fact is what makes our calculation up to $t = 4$ feasible. Details of the calculation may be found in the appendix and some results are shown in figures 3-6.

For $t = 2$ we find in particular

$$q_2 = \sum_{S_0} \frac{1 + q_0 S_0}{2} \mathrm{erf}\left(\frac{\Delta q_1 + \mathrm{i}\eta V_{01}(S_0 - q_0)}{[2(1 - q_1^2)]^{1/2}} \right) \tag{4.18}$$

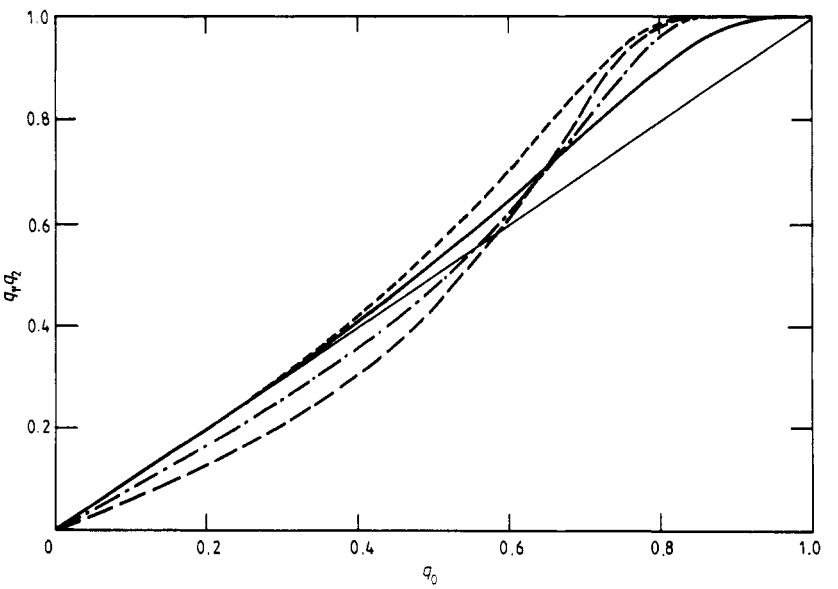

Figure 3. One-step overlap q_1 (——) and two-step overlaps q_2 as a function of initial overlap q_0 for stability $\Delta = \sqrt{\pi/2}$ and symmetry $\eta = -1$ (— — —), 0 (- - -), 1 (— - — -) in the OPN model (q_1 is independent of η).

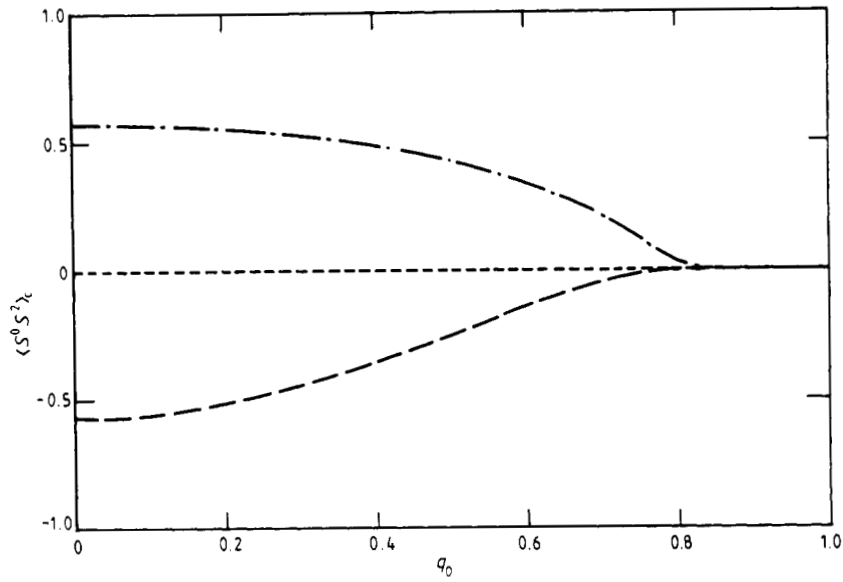

Figure 4. Connected correlation function $\langle S^0 S^2 \rangle_c = q_{02} - q_0 q_2$ for stability $\Delta = \sqrt{\pi/2}$ as a function of initial overlap q_0 for symmetry $\eta = -1$ (— — —), 0 (- - -), 1 (— - — -) in the OPN model.

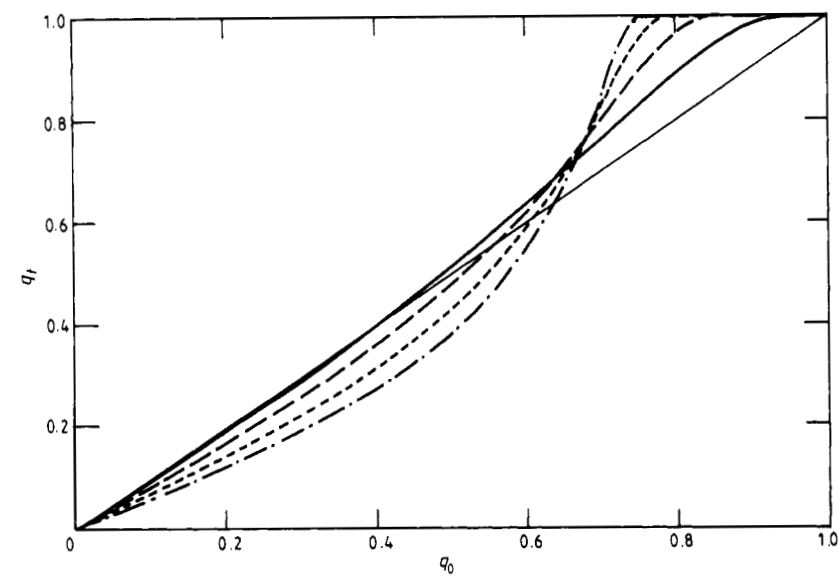

Figure 5. 1, . . . , 4-step overlap $q_1, \ldots, q_4$ as a function of initial overlap q_0 for stability $\Delta = 1.2$ and symmetry $\eta = 0.75$. ——, q_1; — — —, q_2, - - -, q_3. — · — ·, q_4.

where

$$iV_{01} = \left(\frac{2}{\pi(1-q_0^2)}\right)^{1/2} \exp\left(\frac{-\Delta^2 q_0^2}{2(1-q_0^2)}\right). \tag{4.19}$$

The only difference with q_1, equation (4.11), is a term which depends on η. It introduces correlations between the inital state and the one at $t=2$. Figure 3 illustrates the influence of the symmetry η at $t=2$ for $\Delta=\sqrt{\pi/2}$ (cf preceding section), and figure 4 gives the connected correlation functions for the same values of Δ and η. The high degree of correlation (anticorrelation) for large positive (negative) symmetry translates into a large probability to end up in a cycle of length 2 (of length 4 with inverted second and fourth step), as we have been able to observe numerically.

Our formulae, given in the appendix up to $t=4$, are shown in figure 5 for symmetry $\eta=0.75$ and $\Delta=1.2$. There we see in fact a step function forming, indicating a critical overlap $q_c \approx 0.7$. For a more concise representation of our formulae see figure 6. There we plotted the values of q_0, for which $q_1=q_0, \ldots, q_4=q_0$, for $\eta=0.5$, as a function of Δ. In addition, we included the numerically determined critical overlaps (at time $t=50$, using procedure RANDOMISE).

Figure 6 demonstrates that our calculation up to $t=4$ already provides a fair approximation for the critical overlap and allows at least a qualitative discussion of the OPN model also for large times. There is little influence of the symmetry for small stability, as has already been noted in § 3, and important influence for larger values of the stability. At $\Delta \approx 1.6$, the critical value of the magnetisation q_c becomes zero. This, however, has not to be taken to mean that half of the phase space flows towards **1**. It should be realised that in a space of high dimensionality (N) almost all configurations have overlap with **1** which is of the order of $1/\sqrt{N}$. For $q_0=0$, the OPN model

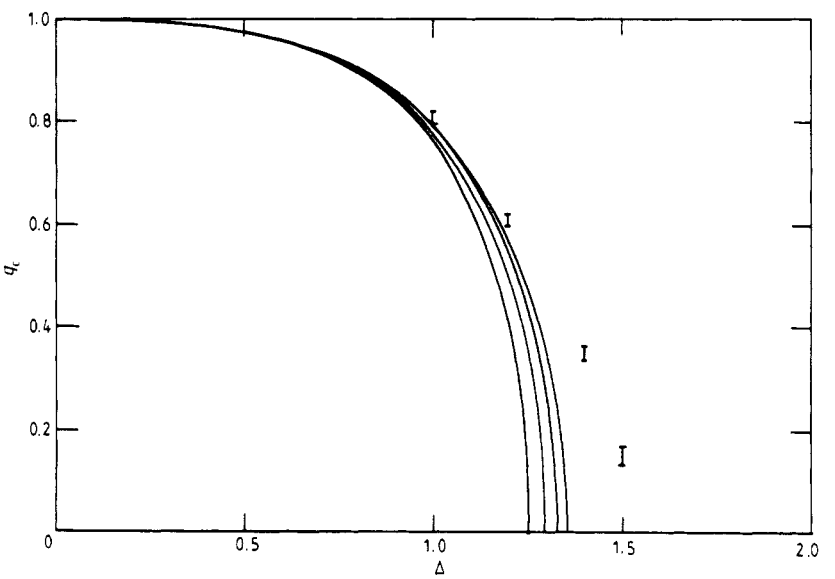

Figure 6. The curves give, from left to right, the analytical 1, 2, ..., 4-step approximations at $\eta=0.5$ for the critical overlap q_c (values of q_0 for which $q_1=q_0, \ldots, q_4=q_0$) as a function of q_0. The points with error bars are numerical values of q_c determined using procedure RANDOMISE. The values of q_c obtained here have to be compared with those at $\eta'=0.5$ shown on figures 1 and 2.

behaves as a spin glass, and our formulae for $q_0 = 0$ and $\eta = 1$ are equivalent to those of Gardner *et al* (1987a).

5. Conclusion

In this paper we presented a one-pattern neural network model with which to describe the retrieval phase of a neural network. This model stresses the importance of the stability and symmetry concepts. Using numerical methods we found it to provide in general a good qualitative description of real neural networks, and even a good quantitative agreement in some cases.

Analytical calculations enabled us to obtain the dynamics at short times, and to demonstrate that for 'zero symmetry' ($\eta = 0$) the dynamics is solvable and simple: the dynamics is equivalent to the one on the extremely diluted asymmetric network. We expect this last result to be quite general. The equivalence of the dynamics of the OPN model for zero initial overlap with the pattern and that of a spin glass provides a rather nice illustration of the similarities between spin glasses and neural networks. Without partial information about the patterns, the latter will, for all finite times, behave as the former.

In the OPN model we have shown that the critical overlap q_c (the border of the basin of attraction) depends on a function of the stability and the symmetry: typically, the larger the stability Δ and the smaller the symmetry η, the smaller q_c. We expect this to be true in a real neural network, where, of course, the two variables are no longer free, and depend on the learning rule and on the number of stored patterns.

To study the influence of the stability and the symmetry on real neural networks, we envisage several possibilities. As an example, one can consider iterative learning procedures with a *tabula non rasa* scheme (Toulouse *et al* 1986, Personnaz *et al* 1986), modulating the values of the stability and of the symmetry by proper choices of the initial matrix. Then one can look, at a given value of α, for the values leading to the largest basins of attraction. As another possibility, one could study relatively sparse networks, in which the degree of symmetry can be reduced by setting an important number of links J_{ij} equal to zero. More generally, it may be that one will have to look for learning schemes which try to reach reasonable values of the stability while trying to keep the symmetry low.

Acknowledgments

We are very grateful to B Derrida for helpful discussions, and especially for pointing out the relevance of gauge-invariant quantities. Discussions with G Toulouse are gratefully acknowledged. WK acknowledges financial support by Carl Duisberg Gesellschaft and by Studienstiftung des deutschen Volkes. The numerical simulations, done on a VAX 750 and a CONVEX C1, were supported by GRECO 70 'Expérimentation Numérique'.

Appendix

We determine the dynamics of the OPN model up to four time steps. To do so, we

calculate the 'partition function' $Z(t)$:

$$Z = \mathrm{Tr}_{S^1,S^2,\ldots,S^t} \prod_{i=1,N} \prod_{l=0,t-1} \theta\left(S_i^{l+1} \sum_j J_{ij} S_j^l\right) \tag{A1}$$

averaged over the distribution of the J_{ij}:

$$\bar{Z} = \mathrm{Tr}_J \prod_i \delta\left(\sum_j J_{ij} - \Delta\sqrt{N}\right) Z/V \tag{A2a}$$

$$V = \mathrm{Tr}_J \prod_i \delta\left(\sum_j J_{ij} - \Delta\sqrt{N}\right) \tag{A2b}$$

where Tr_J denotes the trace equation (4.2). Using the representations (4.6) and (4.7) for the δ and θ functions, respectively, we get

$$\bar{Z} = \frac{1}{V} \mathrm{Tr}_J \prod_i \int \frac{\mathrm{d}\gamma_i}{2\pi\sqrt{N}} \prod_i \prod_{l=0,t-1} \int \mathrm{d}x_i^l \int_0^\infty \frac{\mathrm{d}\lambda_i^l}{2\pi} \exp\left[\mathrm{i} \sum_i \gamma_i \left(\sum_j \frac{J_{ij}}{\sqrt{N}} - \Delta\right)\right]$$
$$\times \exp\left[\mathrm{i} \sum_{l=0,t-1} \sum_{i=1,N} x_i^l \left(\frac{1}{\sqrt{N}} \sum_j J_{ij} S_j^l - \lambda_i^l S_i^{l+1}\right)\right]. \tag{A3}$$

This leads to the following generalisation of equation (4.9) for t time steps:

$$\bar{Z} = \frac{1}{V} \mathrm{Tr}_J \prod_i \int \frac{\mathrm{d}\gamma_i}{2\pi\sqrt{N}} \prod_{i,l} \int \mathrm{d}x_i^l \int_0^\infty \frac{\mathrm{d}\lambda_i^l}{2\pi} \exp\left(-\mathrm{i} \sum_i \gamma_i \Delta - \mathrm{i} \sum_{i,l} x_i^l \lambda_i^l S_i^{l+1}\right)$$
$$\times \exp\left(\frac{\mathrm{i}}{\sqrt{N}} \sum_{i<j} J_{ij}^{\mathrm{s}} (a_{ij} + a_{ji}) + \frac{\mathrm{i}}{\sqrt{N}} \sum_{i<j} J_{ij}^{\mathrm{a}} (a_{ij} - a_{ji})\right) \tag{A4}$$

with

$$a_{ij} = \gamma_i + \sum_{l=0,t-1} x_i^l S_j^l. \tag{A5}$$

Taking the trace on the $J_{ij}^{\mathrm{s,a}}$, and keeping only the dominant terms in N, we get

$$\bar{Z} = \frac{1}{V} \prod_i \frac{\mathrm{d}\gamma_i}{2\pi\sqrt{N}} \prod_l \int \mathrm{d}x_i^l \int_0^\infty \frac{\mathrm{d}\lambda_i^l}{2\pi} \exp\left(-\mathrm{i}\Delta \sum_i \gamma_i - \mathrm{i} \sum_{i,l} x_i^l \lambda_i^l S_i^{l+1}\right)$$
$$\times \exp\left(-\frac{1}{2N} \sum_{i \neq j} (a_{ij}^2 + \eta a_{ij} a_{ji})\right). \tag{A6}$$

After replacing a_{ij} in (A6) by the RHS of (A5), one can perform the integration on the γ_i. Then it is convenient to introduce the macroscopic parameters

$$q_l = \frac{1}{N} \sum_j S_j^l \qquad 1 \leqslant l \leqslant t-1 \tag{A7a}$$

$$q_{ll'} = \frac{1}{N} \sum_j S_j^l S_j^{l'} \qquad 0 \leqslant l' < l \leqslant t-1 \tag{A7b}$$

$$V_{ll'} = \frac{1}{N} \sum_j x_j^l S_j^{l'} \qquad 0 \leqslant l' \neq l \leqslant t-1 \tag{A7c}$$

$$U_l = \frac{1}{N} \sum_j x_j^l S_j^l \qquad 0 \leqslant l \leqslant t-1 \tag{A7d}$$

$$T_l = \frac{1}{N} \sum_j x_j^l \qquad 0 \leqslant l \leqslant t-1 \tag{A7e}$$

by means of conjugate variables $\hat{q}_l$, $\hat{q}_{ll'}$, $\hat{V}_{ll'}$, $\hat{U}_l$, $\hat{T}_l$, and one obtains

$$\bar{Z} = \int \prod_{l=1}^{t-1} \mathrm{d}q_l \frac{\mathrm{d}\hat{q}_l}{2\pi} \prod_{l=0}^{t-1} \mathrm{d}U_l \frac{\mathrm{d}\hat{U}_l}{2\pi} \mathrm{d}T_l \frac{\mathrm{d}\hat{T}_l}{2\pi} \prod_{l \neq l'} \mathrm{d}V_{ll'} \frac{\mathrm{d}\hat{V}_{ll'}}{2\pi} \prod_{l<l'} \mathrm{d}q_{ll'} \frac{\mathrm{d}\hat{q}_{ll'}}{2\pi} \exp[N(G+f)] \tag{A8}$$

with

$$\begin{aligned} G = & \sum_{l=1}^{t-1} q_l \hat{q}_l + \sum_{l=0}^{t-1} (U_l \hat{U}_l + T_l \hat{T}_l) + \sum_{l<l'} q_{ll'} \hat{q}_{ll'} + \sum_{l \neq l'} V_{ll'} \hat{V}_{ll'} + \mathrm{i}\Delta \sum_l q_l T_l \\ & + \eta \left(\tfrac{1}{2} \sum_l [\eta - q_l^2(1+\eta)] t_l^2 + \frac{1}{l} \sum_{l \neq l'} T_l T_{l'} [\eta(q_{ll'} - q_l q_{l'}) - q_l q_{l'}] \right. \\ & \left. - \tfrac{1}{2} \sum_l U_l^2 - \tfrac{1}{2} \sum_{l \neq l'} V_{ll'} V_{l'l} + \sum_l q_l U_l T_l + \sum_{l \neq l'} q_l T_{l'} V_{ll'} \right) \end{aligned} \tag{A9}$$

and

$$\begin{aligned} f = \sum_{S_0 = \pm 1} \frac{1+q_0 S_0}{2} \log \left[\sum_{\substack{S_l = \pm 1 \\ l=1,t}} \exp\left(- \sum_{l=1}^{t-1} \hat{q}_l S_l - \sum_{0 \leq l < l' < t-1} \hat{q}_{ll'} S_l S_{l'} \right) \right. \\ \left. \times \prod_{l=0}^{t-1} \int_{-\infty}^{\infty} \mathrm{d}x_l \int_0^{\infty} \frac{\mathrm{d}\lambda_l}{2\pi} \exp\left(\sum_l \Psi_l \right) \right] \end{aligned} \tag{A10}$$

and, finally,

$$\Psi_l = -\frac{1-q_l^2}{2} x_l^2 - \sum_{l'<l} x_l x_{l'} (q_{ll'} - q_l q_{l'}) - \mathrm{i} x_l \lambda_l S_{l+1} - \hat{T}_l x_l - \hat{U}_l x_l S_l - \sum_{l' \neq l} \hat{V}_{ll'} x_l S_{l'}. \tag{A11}$$

We then take the saddle points. As in Gardner *et al* (1987a), q_l, $q_{l'l}$, $V_{l'l}$ are functions of the time steps $l'' < l$, and many order parameters are zero. Noting that the sum over S_t together with the integral over $\mathrm{d}\lambda_{t-1}$ leads to $\delta(x_{t-1})$, one deduces

$$\begin{aligned} & U_l = T_l = \hat{U}_l = \hat{q}_l = 0 \qquad l \leq t-1 \\ & V_{ll'} = 0;\ \hat{q}_{ll'} = 0;\ \hat{V}_{l'l} = 0 \qquad l' < l \end{aligned} \tag{A12}$$

and

$$\begin{aligned} & \hat{T}_l = -\mathrm{i}\Delta q_l - \eta \sum_{l'<l} V_{l'l} q_{l'} \\ & \hat{V}_{ll'} = \eta V_{l'l} \qquad l' < l. \end{aligned} \tag{A13}$$

At the saddle point, $G = 0$, and f can be written as

$$\begin{aligned} & f = \sum_{S_0} \frac{1+q_0 S_0}{2} \log \sum_{S_1, \ldots, S_t} \prod_{l=0}^{t-1} \int \mathrm{d}x_l \int \frac{\mathrm{d}\lambda_l}{2\pi} \exp\left(\sum_{l=0}^{t-1} \Psi_l \right) \\ & \Psi_l = -\frac{1-q_l^2}{2} x_l^2 - \sum_{0 \leq l' < l} x_l x_{l'} (q_{ll'} - q_l q_{l'}) - \mathrm{i} x_l \lambda_l S_{l+1} + \mathrm{i}\Delta q_l x_l - \eta \sum_{0 \leq l' < l} V_{l'l} (S_{l'} - q_{l'}) x_l \end{aligned} \tag{A14}$$

and q_l, $q_{ll'}$, $V_{ll'}$ are given by

$$\begin{aligned} & q_l = \langle\langle S_l \rangle\rangle \\ & q_{l'l} = \langle\langle S_l S_{l'} \rangle\rangle \qquad l' < l \\ & V_{l'l} = \langle\langle x_{l'} S_l \rangle\rangle \qquad l' < l \end{aligned} \tag{A15}$$

where $\langle\langle\ \rangle\rangle$ is the average taken with (A14).

Once (A15) has been obtained for $l \leqslant t-1$, we can calculate q_t, q_{lt} and V_{lt} for $l < t$ by introducing *ad hoc* source terms $\exp(hS_t)$, $\exp(hS_lS_t)$, and $\exp(hx_lS_t)$, respectively, and by taking the derivative with respect to h at $h=0$.

One can see also that $q_{ll'} - q_lq_{l'}$ is zero if l and l' are of differing parity, and that $V_{l'l}$ is zero if l and l' agree in parity. For $\eta = 0$, $q_{ll'} - q_lq_{l'}$ is always zero.

This leads to the following results for $t \leqslant 4$:

$t=1$

$$q_1 = \left(\frac{2}{\pi}\right)^{1/2} \int_0^{\Delta q_0/(1-q_0^2)^{1/2}} \mathrm{d}z\, \mathrm{e}^{-z^2/2}$$

$$\mathrm{i}V_{01} = \left(\frac{2}{\pi(1-q_0^2)}\right)^{1/2} \exp\left(\frac{-\Delta^2 q_0^2}{2(1-q_0^2)}\right)$$

$t=2$

$$q_2 = \sum_{S_0} \frac{1+q_0S_0}{2} \left(\frac{2}{\pi}\right)^{1/2} \int_0^{x(S_0)} \mathrm{d}z\, \mathrm{e}^{-z^2/2}$$

$$q_{02} = \sum_{S_0} \frac{S_0+q_0}{2} \left(\frac{2}{\pi}\right)^{1/2} \int_0^{x(S_0)} \mathrm{d}z\, \mathrm{e}^{-z^2/2}$$

$$\mathrm{i}V_{12} = \sum_{S_0} \frac{1+q_0S_0}{2} \left(\frac{2}{\pi(1-q_1^2)}\right)^{1/2} \exp\left(-\frac{1}{2(1-q_1^2)}[\Delta q_1 + \mathrm{i}V_{01}\eta(S_0-q_0)]^2\right)$$

with

$$x(S_0) = \frac{\Delta q_1 + \eta \mathrm{i}V_{01}(S_0-q_0)}{(1-q_1^2)^{1/2}}$$

$t=3$

$$q_3 = \sum_{S_1,S_3} S_3 I(S_1, S_3)$$

$$q_{13} = \sum_{S_1,S_3} S_1S_3 I(S_1, S_3)$$

$$\mathrm{i}V_{03} = \sum_{S_1,S_3} S_1S_3 \frac{\exp[-\frac{1}{2}(1-\omega_{02}^2)x_0^2]}{[2\pi(1-q_0^2)]^{1/2}} \int_{x_2-\omega_{02}x_0S_1S_3}^{\infty} \frac{\mathrm{d}y}{\sqrt{2\pi}}\, \mathrm{e}^{-y^2/2}$$

$$\mathrm{i}V_{23} = \sum_{S_1,S_3} S_1S_3 \frac{\exp[-\frac{1}{2}(1-\omega_{02}^2)x_2^2]}{[2\pi(1-q_2^2)]^{1/2}} \int_{x_0-\omega_{02}x_2S_1S_3}^{\infty} \frac{\mathrm{d}y}{\sqrt{2\pi}}\, \mathrm{e}^{-y^2/2}$$

with

$$I(S_1, S_3) = \int_{x_0}^{\infty} \frac{\mathrm{d}y}{\sqrt{2\pi}} \int_{x_2}^{\infty} \frac{\mathrm{d}z}{\sqrt{2\pi}} \left(\frac{D_{02}}{(1-q_0^2)(1-q_2^2)}\right)^{1/2} \exp[-\tfrac{1}{2}(y^2+z^2-2\omega_{02}yzS_1S_3)]$$

where

$$D_{02} = (1-q_0^2)(1-q_2^2) - (q_{02}-q_0q_2)^2$$

$$\omega_{02} = (q_{02}-q_0q_2)/[(1-q_0^2)(1-q_2^2)]^{1/2}$$

$$x_0 = -\Delta q_0 S_1 \left(\frac{1-q_2^2}{D_{02}}\right)^{1/2}$$

$$x_2 = -S_3[\Delta q_2 + \eta \mathrm{i}V_{12}(S_1-q_1)]\left(\frac{1-q_0^2}{D_{02}}\right)^{1/2}$$

$t = 4$

$$q_4 = \sum_{S_0, S_2, S_4} \frac{1 + q_0 S_2}{2} S_4 J(S_0, S_2, S_4)$$

$$q_{l,4} = \sum_{S_0, S_2, S_4} \frac{1 + q_0 S_0}{2} S_l S_4 J(S_0, S_2, S_4)$$

for $l = 0, 2$, with

$$J(S_0, S_2, S_4) = \int_{x_1} \frac{dy}{\sqrt{2\pi}} \int_{x_3} \frac{dz}{\sqrt{2\pi}} \left(\frac{D_{13}}{(1 - q_1^2)(1 - q_3^2)} \right)^{1/2} \exp[-\tfrac{1}{2}(y^2 + z^2 - 2\omega_{13} yz S_2 S_4)]$$

where

$$D_{13} = (1 - q_1^2)(1 - q_3^2) - (q_{13} - q_1 q_3)^2$$

$$\omega_{13} = (q_{13} - q_1 q_3)/[(1 - q_1^2)(1 - q_3^2)]^{1/2}$$

$$x_1 = -\left(\frac{1 - q_3^2}{D_{13}} \right)^{1/2} S_2[(\Delta q_1 + \eta i V_{01}(S_0 - q_0)]$$

$$x_3 = -\left(\frac{1 - q_1^2}{D_{13}} \right)^{1/2} S_4[\Delta q_3 + \eta i V_{03}(S_0 - q_0) + \eta i V_{23}(S_2 - q_2)].$$

References

Amit D J 1987 *Heidelberg Colloq. on Glassy Dynamics* ed J L van Hemmen and I Morgenstern (Berlin: Springer) p 430

Crisanti A and Sompolinsky H 1988 *Preprint* Jerusalem

Derrida B 1988 private communication

Derrida B, Gardner E and Zippelius A 1987 *Europhys. Lett.* **4** 167

Diederich S and Opper M 1987 *Phys. Rev. Lett.* **58** 949

Forrest B M 1988 *J. Phys. A: Math. Gen.* **21** 245

Gardner E 1987 *Europhys. Lett.* **4** 481

Gardner E, Derrida B and Mottishaw P 1987a *J. Physique* **48** 741

Gardner E, Stroud N and Wallace D J 1987b *preprint Edinburgh 87/394*

Gutfreund H, Reger J D and Young A P 1988 *J. Phys. A: Math. Gen.* **21** 2775

Hopfield J J 1982 *Proc. Natl Acad. Sci. USA* **79** 2554

Kanter I and Sompolinsky H 1987 *Phys. Rev.* A **35** 380

Kohonen T 1984 *Self Organization and Associative Memory* (Berlin: Springer)

Krauth W and Mézard M 1987 *J. Phys. A: Math. Gen.* **20** L745

Mézard M, Nadal J P and Krauth W 1988 in preparation

Minsky M and Papert S 1969 *Perceptrons* (Cambridge, MA: MIT Press)

Papadimitriou D and Steiglitz K 1982 *Combinatorial Optimization: Algorithms and Complexity* (Engelwood Cliffs, NJ: Prentice-Hall)

Personnaz L, Guyon I and Dreyfus G 1985 *J. Physique* **16** L359

Personnaz L, Guyon I, Dreyfus G and Toulouse G 1986 *J. Stat. Phys.* **43** 411

Pöppel G and Krey U 1987 *Europhys. Lett.* **4** 481

Rieger H, Schreckenberg M and Zittartz J 1988 *J. Phys. A: Math. Gen.* **21** L263

Toulouse G 1988 Private communication

Toulouse G, Dehaene S and Changeux J P 1986 *Proc. Natl Acad. Sci. USA* **83** 1695

Venkatesh S 1986 *Proc. Conf. on Neural Networks for Computing, Snowbird, Utah* (*AIP Conf. Proc.* **151**) ed J S Denker