Z. Phys. B - Condensed Matter 60, 205-213 (1985)

Condensed Matter
Zeitschrift für Physik B

Learning and Pattern Recognition in Spin Glass Models

W. Kinzel
Institut für Festkörperforschung der Kernforschungsanlage Jülich, Federal Republic of Germany

Received May 20, 1985

Dedicated to B. Mühlschlegel on the occasion of his 60th birthday

Spin glass models have a complex phase space which may be used to store information. By an asynchronous relaxational dynamics noisy patterns are recognized very fast. In particular the Hopfield model which may be a simple model for neural networks is analyzed in detail. By numerical simulations and analytical approximations the number of patterns to be stored and the amount of noise to be recognized is calculated. It is demonstrated that a random network can learn patterns by reducing frustrated bonds.

I. Introduction

Spin glasses are disordered magnetic materials with unusual properties [1]. In usual magnets, below a critical temperature T_c the magnetic moments are aligned in a periodic arrangement. However, in spin glasses, below a freezing temperature T_f the moments are frozen into *randomly* distributed directions. In usual magnets there is only one ordered state - up to rotations or reflections of the whole spin state. In spin glasses there are infinitely many stable and metastable spin configurations. This leads to a very broad spectrum of relaxation times which shows up in many experiments [1].

Spin glass behaviour is rather universal; it is found in many different magnetic materials. It turned out that only two properties are relevant for spin glasses: disorder and competition of the magnetic interactions. Therefore spin glasses are well described by a simple model introduced by Edwards and Anderson (EA) [2]. The Hamiltonian is given by

$$H = -\sum_{ij} J_{ij} S_i S_j - h \sum_i S_i. \tag{1}$$

The spins are modelled by Ising variables $S_i = \pm 1$ which may be located on a regular lattice. The interaction J_{ij} are randomly distributed by a Gaussian distribution $P(J)$ given by

$$P(J) \propto \exp(-J^2/2\Delta J^2) \tag{2}$$

h contains an external magnetic field. The dynamics of this model is defined by the single spin flip relaxational dynamics used in Monte Carlo simulations [3].

Computer simulations have shown that even the two-dimensional *short range* EA model describes many experiments in many details [2, 4, 5]. For instance the existence of metastable states with very long relaxation times can be seen in the decay of a remanent magnetization. While in thermal equilibrium the global magnetization M is zero, a spin glass is trapped in an metastable state with $M \neq 0$ after an external magnetic field h has been applied. This remanent magnetization decays very slowly with a power law in time t^{-a} to its equilibrium value $M = 0$. The exponent a depends on temperature T, the previously applied field h and on the history of the sample (field cooled vs zero field cooled) [4, 5].

In the case of *infinite range* couplings J_{ij} the EA model, Eqs. (1) and (2), has a sharp phase transition at a critical temperature T_c [6]. For $T < T_c$, there exist infinitely many thermodynamic states $\mathbf{m}^\alpha = (\langle S_1 \rangle_\alpha, \langle S_2 \rangle_\alpha, \ldots, \langle S_N \rangle_\alpha)$ where $\langle \ldots \rangle_\alpha$ means the thermal expectation value of the phase space α evolving from the spontaneously broken symmetry. If one defines an overlap $q_{\alpha\beta}$ between the states α and β by

$$q_{\alpha\beta} = \frac{1}{N} \sum_i m_i^\alpha m_i^\beta \tag{3}$$

then the thermodynamics of the system is given by the distribution $P(q)$ of overlaps q [7]. $P(q)$ can be calculated using the replica method with spontaneously broken replica symmetry [1, 7].
The overlap $q_{\alpha\beta}$ defines a distance $d_{\alpha\beta}$ between the states α and β by

$$d_{\alpha\beta}=(1-q_{\alpha\beta})/2. \tag{4}$$

With respect to such a distance the spin glass states have an ultrametric structure, i.e. any three states α, β, γ form an isoceles triangle ($d_{\alpha\beta}=d_{\beta\gamma}\geqq d_{\alpha\gamma}$ after relabeling the indices) [8].
This complex structure of spin glass states may be used to store and retrieve information. In fact, the EA model has been discussed by Hopfield as a simple model of neural networks [9]. Each neuron is modelled by just two states: $S_i=+1(-1)$ means that the neuron at site i is firing (not firing). The neuron i receives an electrical potential E_i from its connected neighbours j given by

$$E_i=\sum_j J_{ij}S_j. \tag{5}$$

Here the bonds J_{ij} model the synapses from neuron j to neuron i. Since there are excitatory ($J_{ij}>0$) and inhibitatory ($J_{ij}<0$) synapses and since the synapses can adjust the transmission of the potential, the bonds J_{ij} can vary in strength and sign. Each neuron fires only if the potential E_i is larger than a threshold value h_i, i.e. if one has $(E_i-h_i)S_i>0$. For zero thresholds $h_i=0$ and for symmetric bonds $J_{ij}=J_{ji}$ one recovers the spin glass model Eq. (1); the neural states are the local minima of Eq. (1) and the relaxation into these states is just the zero temperature dynamics of the spin glass.
The problem arises to choose the bonds J_{ij} such that many patterns are local minima of the energy H, Eq. (1), simultaneously. This problem has been solved by methods of simple linear algebra [10]. Other questions are: How many patterns can be stored and how much noise can be recognized by the network? Part of these problems have already been investigated by Hopfield [9].
In this paper the Hopfield model is investigated in more details. In Sect. II the model is defined and its thermodynamics is discussed. The following section presents a recursion relation which estimates the error made by retrieving learned random patterns. The distribution of retrieval error is calculated numerically as a function of the amount of *(i)* stored patterns and *(ii)* noise of the input in Sect. IV. In addition the Hopfield algorithm is illustrated for simple pictures. In Sect. V it is demonstrated that a random network can learn patterns by destroying all bonds which are not satisfied by one of the patterns (a property which is called frustration [11]). Section V summarizes the results and presents some possible extensions of the Hopfield mechanism.

II. Definition and Thermodynamics of the Hopfield Model

Consider the case where one wants to store M patterns $\mathbf{S}^{(1)}=(S_1^{(1)},\ldots,S_N^{(1)})$, $\mathbf{S}^{(2)}$, $\mathbf{S}^{(3)},\ldots,\mathbf{S}^{(M)}=(S_1^{(M)},\ldots,S_N^{(M)})$ in the network, Eq. (1), with N Ising variables $S_1,\ldots,S_N$. Here each pattern is given in a binary representation with $S_i^{(k)}=\pm 1$. For simplicity we chose M *random* patterns where each $S_i^{(k)}$ is randomly taken from a ± 1 distribution with equal weight of $+1$ and -1.
Storing M patterns means to choose the bonds J_{ij} such that for each pattern $\mathbf{S}^{(k)}$ each spin (or at least most of the spins) points into the direction of its internal field, i.e.

$$S_i^{(k)}\sum_j J_{ij}S_j^{(k)}>0. \tag{6}$$

For random patterns one may use the Cooper prescription [9]

$$J_{ij}=\sum_{k=1}^{M} S_i^{(k)}S_j^{(k)} \quad (i\neq j) \tag{7}$$

where we assume that each spin S_i interacts with z other spins S_j, only; all other bonds J_{ij} are set to zero. The dynamics of this model is introduced as an asynchronous relaxation of the energy, Eqs. (1) and (7) ($T=0$ Monte Carlo procedure). This means that one runs sequentially in any order through the system and flips each spin into the direction of its internal field if necessary.
Such models where the interaction J_{ij} is given by M "modes" $\mathbf{S}^{(k)}$ as in Eq. (7), have been discussed as models for spin glasses. For $M=1$ and zero external field this model is equivalent to a ferromagnet [12]. For a *finite* number M and for infinite range of bonds ($z=N$) this model can be solved exactly [5, 13–16]. One finds a phase transition into one of the M modes $\mathbf{S}^{(k)}$.
Moreover, the states $\mathbf{S}^{(k)}$ are the only groundstates [16]. However, the replica symmetry is not broken [14–16] and the decay into the equilibrium is exponentially fast [4, 15], hence the model with finite M does not contain typical spin glass properties [15].
For *infinite* number M of modes the thermodynamics of the Hopfield model is less clear. For $M\to\infty$ the bonds J_{ij} are Gaussian distributed according to Eq. (7). However, one still has correlations between different bonds, for instance the av-

erage frustration loop

$$\overline{J_{ij}J_{jk}J_{ki}}=M \tag{8}$$

is nonzero in contrast to the EA model, Eqs. (1) and (2). Nevertheless, in some cases for $M=z=N$ the Hopfield model seems to be equivalent to the infinite range EA model [14]. Hence, if $p_s=\frac{M}{N}$ is constant in the thermodynamic limit we expect spin glass behaviour for $p_s>1$ and ferromagnet - like behaviour for $p_s=0$ (M finite or maybe $M\propto N^x$ with $x<1$). Therefore the Hopfield model interpolates continuously between ferromagnets and spin glasses and it should be interesting to study the evolution of the complex spin glass phase if one approaches it from the well known ferromagnetic side.

III. Recursion Relations for the Retrieval Error

In this section we try to estimate the number of patterns to be stored and the amount of noise to be recognized, using a simple iteration approach. Consider again the Hopfield model with M random patterns as in Eq. (7) A distance d between the state $\mathbf{S}=(S_1,S_2,\ldots,S_N)$ of the network and one of the M patterns $\mathbf{S}^{(k)}$ is again defined by the overlap q between $\mathbf{S}$ and $\mathbf{S}^{(k)}$ as in Eqs. (3) and (4), one has

$$q(\mathbf{S},\mathbf{S}^{(k)})=\frac{1}{N}\mathbf{S}\,\mathbf{S}^{(k)};\quad d=(1-q)/2 \tag{9}$$

d is the Hamming distance of two informations given in binary form, i.e. d is the fraction of bits which are different in $\mathbf{S}$ and $\mathbf{S}^{(k)}$.

Now if the initial state $\mathbf{S}$ of the network is set to be $\mathbf{S}=\mathbf{S}^{(k)}$ for one pattern k, then we would like to know which is the distance $d(\mathbf{S}',\mathbf{S}^{(k)})$ between the final relaxed state $\mathbf{S}'$ and the pattern $\mathbf{S}^{(k)}$. If $d=0$ then this pattern has been learned completely.

On the other side, if we start with a noisy pattern $\mathbf{S}$ with $0\leqq p_n\equiv d(\mathbf{S},\mathbf{S}^{(k)})<0.5$ we would like to know how much the final state $\mathbf{S}'$ has approached the pure pattern $\mathbf{S}^{(k)}$. This retrieval error is again measured by $d(\mathbf{S}',\mathbf{S}^{(k)})$.

For simplicity we study a different dynamics where all spins are updated *at the same time* (in contrast to asynchronous single spin updating). Thus $\mathbf{S}(t)$ evolves in discrete time steps $t=0,1,2,\ldots$ and one has

$$S_i(t+1)=\operatorname{sgn}\sum_j J_{ij}S_j(t). \tag{10}$$

The distance $d(\mathbf{S}(t+1),\mathbf{S}^{(k)})$ is given by the fraction of spins with $S_i^{(k)}\cdot S_i(t+1)<0$. One has

$$\begin{aligned}S_i^{(k)}S_i(t+1)&=\operatorname{sgn}\Big[S_i^{(k)}S_i^{(k)}\sum_j S_j^{(k)}\cdot S_j(t)\\&\quad+\sum_{\nu\neq k}\sum_j S_i^{(k)}S_i^{(\nu)}S_j^{(\nu)}S_j(t)\Big]\\&=\operatorname{sgn}\left[z\,q(\mathbf{S}^{(k)},\mathbf{S}(t))+R_i\right].\end{aligned} \tag{11}$$

If correlations between $\mathbf{S}(t)$ and the pattern $\mathbf{S}^{(k)}$ are neglected R_i is a random variable with

$$\langle R_i\rangle=0,\quad \Delta^2\equiv\langle R_i^2\rangle\simeq z\cdot M \tag{12}$$

Here $\langle\ldots\rangle$ means the average over the distributions of random patterns. For large N values R_i is distributed by the Gaussian

$$P(R)=\frac{1}{\sqrt{2\pi}\,\Delta}\exp(-R^2/2\Delta^2). \tag{13}$$

Since $d(\mathbf{S}(t+1),\mathbf{S}^{(k)})$ is given by the fraction of sites i with $z\cdot q(\mathbf{S}(t)\cdot\mathbf{S}^{(k)})+R_i<0$, one obtains

$$q(t+1)=\Phi\left(\frac{z\cdot q(t)}{\Delta}\right) \tag{14}$$

with $\Phi(x)=\sqrt{\frac{2}{\pi}}\int_0^x\exp(-y^2/2)\,dy=\operatorname{erf}(x/\sqrt{2})$, $z/\Delta=\sqrt{z/M}$ and $q(t)=q(\mathbf{S}(t),\mathbf{S}^{(k)})$.

Eq. (14) is the main result of this approximation. Starting from a state which has some overlap $q(0)$ to one pattern $\mathbf{S}^{(k)}$ Eq. (14) describes how this overlap changes with time. The final overlap q^* is given by the fixed point

$$q^*=\Phi\left(\sqrt{\frac{z}{M}}\,q^*\right) \tag{15}$$

q^* describes the final stationary overlap and $d^*=(1-q^*)/2$ is the retrieval error.

Equation (15) has interesting consequences which are discussed in the following:

(1) A nonzero q^* exists only if

$$2/\pi\simeq 0.8>M/z. \tag{16}$$

For $M/z\simeq 2/\pi$ the retrieval overlap goes to zero. This means that for short range models with a small finite coordination number z only some few patterns can be stored in the network, even in the thermodynamic limit $N\to\infty$. In addition there is always a large retrieval error d^*.

(2) The retrieval error is independent of the initial overlap (we consider only cases where $\mathbf{S}(0)$ has overlap to only *one* of the M patterns). This means that the network recognizes each amount p_n of noise ($p_n=(1-d(0))/2$) with the *same* error.

(3) For the infinite range model $z=N$ one can store infinitely many patterns with some error d^* given by

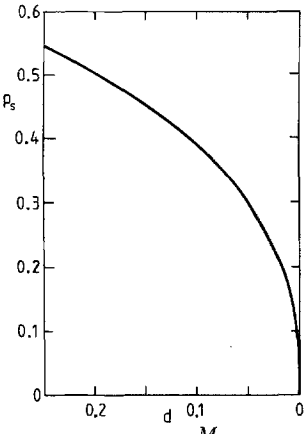

Fig. 1. Fraction $p_s=\frac{M}{N}$ of M random patterns learned by the bonds of an infinite range network of N spins as a function of the retrieval error d. For a given error tolerance d one can store up to $p_s N$ patterns

Eq. (15). d^* depends on the fraction $p_s=\frac{M}{N}$ of stored patterns. For a finite number M one has $p_s=0$ for $N\to\infty$, therefore Eq. (15) gives zero error ($q^*=1$); each pattern is completely stable in agreement with Ref. 16. Figure 1 shows p_s as a function of the retrievel error given by (15). For instance if one allows 0.5% error one can store up to about $M=0.15\,N$ patterns. Each pattern is recognized with up to $p_n=50\%$ noise.

Equation (15) involves two approximations: *(i)* Higher correlations between the states $\mathbf{S}(t)$ and the patterns $\mathbf{S}^{(k)}$ are neglected and *(ii)* All spins are updated at once. In particular *(ii)* leads to a wrong result in the limit of large values of p_s. In this case one obtains the infinite range EA model and q^* is just its saturated remanent magnetization at zero temperature (=overlap to the initial completely aligned state). Therefore we expect $q^*\simeq 0.15$ [17] while (15) gives $q^*=0$ for $p_s>2/\pi$. The reason for a nonzero q^* is that each spin flip stabilizes most of the other spins since the energy is decreased. However, there is still no theory of the remanent magnetization of spin glasses available.

IV. Numerical Results

In this section we investigate the Hopfield model numerically. We consider the infinite range model $z=N$ with N spins $S_i=\pm 1$ and store M random patterns $\mathbf{S}^{(k)}$, $k=1,\ldots,M$ in the bond network J_{ij} with the Cooper presciption of Eq. (7). The dynamics is a single spin flip relaxation at zero temperature as described in Sect. II. Thus the time step is one run through the spin system.

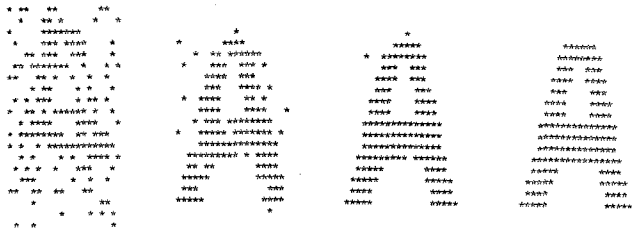

Fig. 2. In addition to 29 other random patterns the pattern "A" has been stored in a network of 20×20 spins. A star shows $S_i=+1$ while a variable $S_i=-1$ is not shown. A pattern with 30% noise relaxes in 4 steps to its pure learned pattern (from left to right)

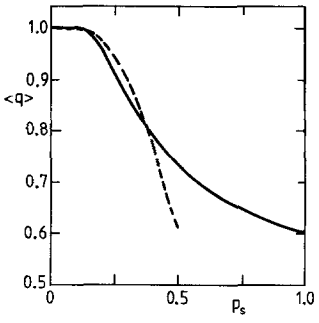

Fig. 3. Average retrieval overlap of the pure random patterns as a function of the fraction p_s of learned patterns. The solid curve shows the results for $N=400$ which is very close to the one of $N=100$, the dashed line shows the approximation of (15)

For an illustration we choose the first pattern $\mathbf{S}^{(1)}$ to be the pattern "A" as shown in Fig. 2. Here the spins S_i are arranged in a 20×20 lattice and a star indicates $S_i=1$ while $S_i=-1$ is not shown. The initial (left) state of Fig. 2 is the "A" with 30% spins being flipped randomly. From left to right it is shown that even such noisy inputs relax within few time steps to their pure learned patterns $\mathbf{S}^{(k)}$.

The simple estimate of (15) has already shown that for nonzero p_s and $N\to\infty$ the network does not learn the patterns $\mathbf{S}^{(k)}$ but some different patterns $\tilde{\mathbf{S}}^{(k)}$ with a retrieval overlap $q^*=\frac{1}{N}\mathbf{S}^{(k)}\cdot\tilde{\mathbf{S}}^{(k)}$. We have calculated this overlap as a function of p_s numerically. The system starts with the state $\mathbf{S}(0)=\mathbf{S}^{(k)}$ and relaxes to a state $\mathbf{S}(t=\infty)$ with an overlap $q=\mathbf{S}(0)\cdot\mathbf{S}(t=\infty)/N$ or an error $d=(1-q)/2$. Figure 3 shows this overlap as a function of p_s averaged over 10^4 random patterns. For $p_s\lesssim 0.14$ one observes a rather flat plateau at zero retrieval error ($q=1$) in agreement with Ref. 9. For comparison Fig. 3 also shows the approximation, Eq. (15). In this case one has a plateau, too; however, the function decreases more smoothly and for large p_s one obtains the wrong result $q\to 0$ as discussed in Sect. III.

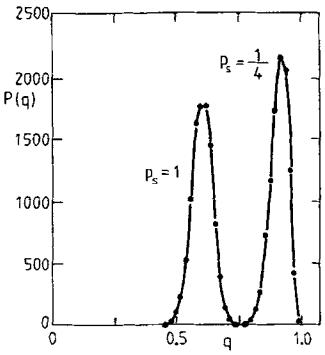

Fig. 4. Distribution $P(q)$ of the retrieval overlap of the pure patterns for two different p_s values and $N=400$. Since the width of $P(q)$ goes to zero with $1/\sqrt{N}$, q is a selfaveraging quantity

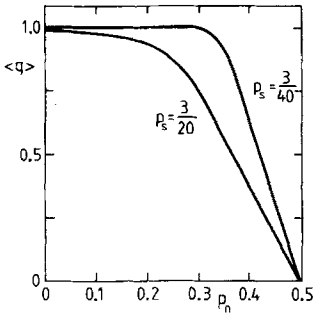

Fig. 5. Average retrieval overlap q as a function of noise p_n initially applied to one of the M learned patterns. The data are shown for $N=400$ and for two different values of $p_s=M/N$

For finite N values the retrieval overlap q is distributed by a function $P(q)$ as shown in Fig. 4. However, we have found that the width of $P(q)$ goes to zero with $1/\sqrt{N}$. Hence, in the thermodynamic limit q is selfaveraging; the distance between the pattern $\mathbf{S}^{(k)}$ and the stored pattern $\tilde{\mathbf{S}}^{(k)}$ is *the same* for all patterns $k=1,\ldots,M$. For large values of p_s this property shows that the remanent magnetization of spin glasses is a selfaveraging quantity [17]. Note that this statement is not trivial since the distribution $P(q)$ of overlaps between pure spin glass states is not selfaveraging [7]. The results of Fig. 4 are in disagreement with the corresponding data of Hopfield [9].

If p_s is small then the retrieval error of the pure patterns is almost zero. In this case it is interesting to see how much noise is recognized by the network. From Eq. (15) we obtained that the retrieval error was independent of the amount of initial noise. Figure 5 shows that this is not quite true.

If we consider the initial state $\mathbf{S}(0)$ to be a pattern $\mathbf{S}^{(k)}$ with a fraction p_n of spins being flipped then one has

$$\mathbf{S}(0)\cdot\mathbf{S}^{(k)}=(2p_n-1)\cdot N \tag{17}$$

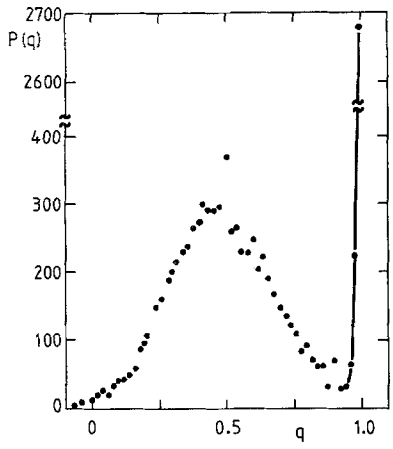

Fig 6. Distribution of retrieval overlap of noisy patterns with $p_s=3/40$ and $p_n=0.4$ ($N=400$)

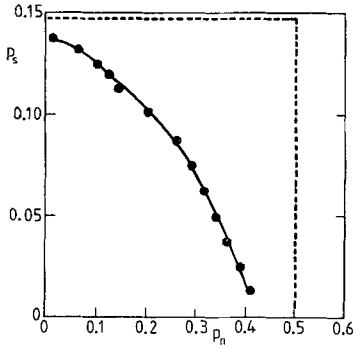

Fig. 7. "Phase boundary" of the fraction p_s of stored random patterns as a function of initial noise p_n. In the left region noisy patterns are recognized with less than 0.5% error, i.e. $\langle q\rangle \leqq 0.99$ ($N=400$)

for the pattern k and zero overlap to the other $M-1$ patterns. Then the system is relaxed and the final overlap $q=\mathbf{S}^{(k)}\cdot\mathbf{S}(t=\infty)/N$ is calculated. Figure 5 shows the average overlap q as a function of noise p_n. For 30 patterns ($N=400$) one can recognize up to 30% noise without remarkable error. For 60 patterns, however, even the pure patterns are retrieved with some error which increases with the amount of noise.

Figure 6 shows the distribution of overlaps q for $p_s=3/40$ and $p_n=0.4$. For such a large amount of noise sometimes the patterns are completely recognized ($q=1$) while in other cases the network relaxes to patterns which are still closer to the initially applied patterns than to the other ones. Note that at least for a finite number M of patterns ($p_s=0$) such metastable patterns do not exist if one introduces a thermodynamics to this model and goes to sufficiently high temperatures [16].

For a given maximal retrieval error one may ask how many patterns can be stored and how much

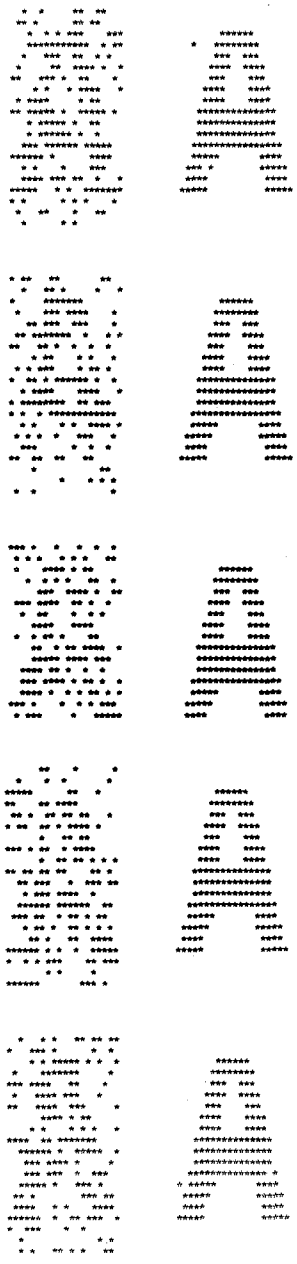

Fig. 8. Relaxation of noisy patterns in a network of 20×20 spins as in Fig. 2 with $p_s = 3/40$ and $p_n = 0.3$. The left noisy input relaxes to the right final state which, according to Figs. 5 and 7 has less than 0.5% error to the learned pattern "A"

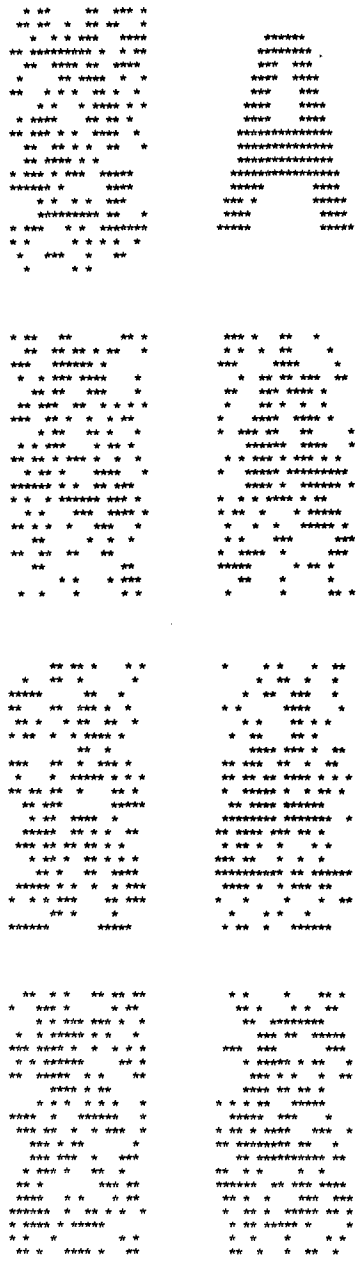

Fig. 9. Same as Fig. 8 but with $p_n = 40\%$ noise. As already seen from Fig. 6 some noisy patterns relax to their pure learned pattern while others relax to metastable final patterns

noise can be recognized by the Hopfield model. Figure 7 shows the answer for a retrieval error of 0.5%. One obtains two regions in the (p_s, p_n)-phase which are separated by a "phase boundary". In the region containing the origin (0,0) noisy patterns are recognized (on the average) with less than 0.5% error.

These quantitative findings are illustrated in Figs. 2, 8 and 9. In Fig. 8 the system is just below the phase boundary of Fig. 7, hence the network recognizes the noisy pattern "A" with only small error. However, in Fig. 9 the noise level is so large that only in some cases the network relaxes to its pure learned pattern, in agreement with Fig. 6.

Up to now we have considered random patterns $\mathbf{S}^{(k)}$, only; i.e. each variable $S_i^{(k)}$ is randomly taken from a symmetric ± 1 distribution. However, if the patterns have a nonzero overlap then the Cooper prescription, Eq. (7), does not work. This can already be seen from Eq. (11). With $\langle \mathbf{S}^{(\nu)} \cdot \mathbf{S}^{(\mu)} \rangle / N \neq 0$ one obtains

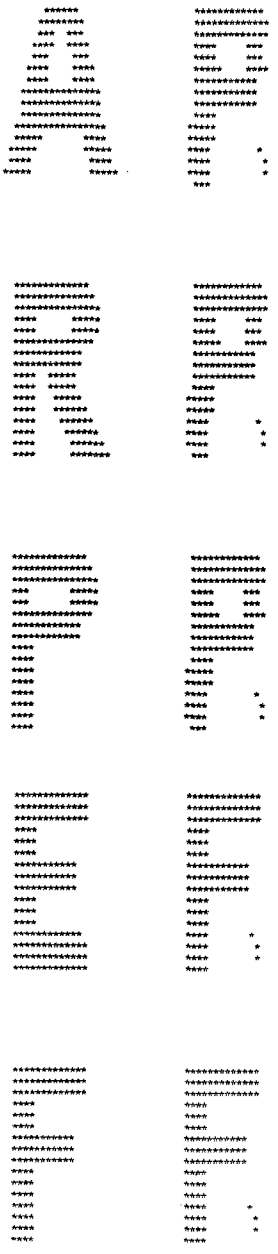

Fig. 10. The five patterns of the left side are stored in a 20×20 system using Eq. (7). The right side shows the corresponding relaxed states demonstrating that the Cooper algorithm cannot learn correlated patterns

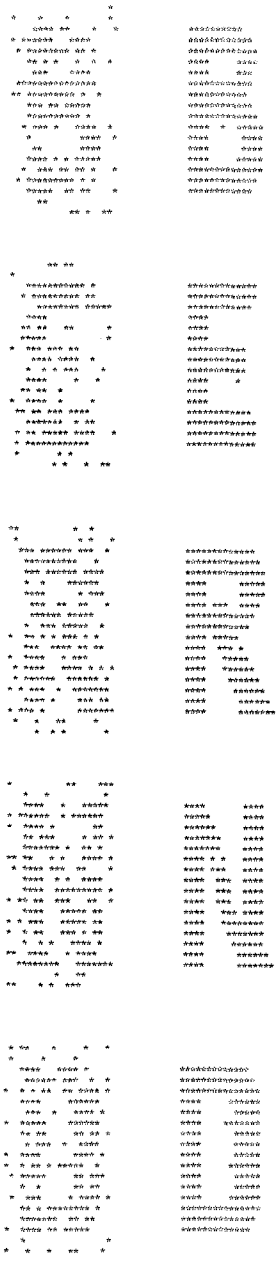

Fig. 11. An unorganized 20×20 network has learned five patterns simultaneously by killing all frustrated connections. A noisy input (left) relaxes to the learned pattern (right) within a few time steps. The fraction of noise of the input is $p_n = 20\%$

$$\langle R_i \rangle = 0 (z \cdot M). \tag{18}$$

Therefore the second term of Eq. (11) cancels the first one if $S_i^{(k)} S_i^{(\nu)} < 0$. This means that bits $S_i^{(k)}$ of a pattern k which are different to most of the other bits $S_i^{(\nu)}$ are not stable. This is demonstrated in Fig. 10 where five patterns A, R, P, E and F are stored in a 20×20 network. The legs of "R" and "E" are different from the other patterns, therefore they disappear in the final relaxed state. For such correlated patterns one can find the network J_{ij} using methods of linear algebra [10].

V. Learning by Regression of Frustrated Bonds

This Section is motivated by neurobiology. In the first few years of life many synapses of the brain are dying off although the amount of learned infor-

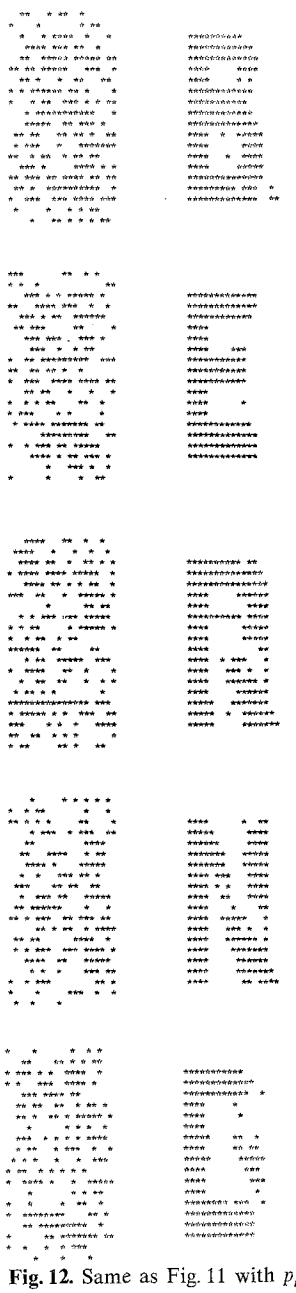

Fig. 12. Same as Fig. 11 with $p_n = 35\%$

mation is large and permanent [18]. We have studied this effect in the Hopfield model.

We start from a completely unorganized network, Eq. (1). We use Gaussian distributed infinite range couplings as in (2); i.e. we start from the EA model of spin glasses. Then we reduce all bonds which are frustrated in one of the patterns to be learned. This means if we want to store M patterns $\mathbf{S}^{(1)}, \ldots; \mathbf{S}^{(M)}$ we replace J_{ij} by zero if

$$J_{ij} S_i^{(k)} S_j^{(k)} < 0 \tag{19}$$

for any pattern k. Hence we obtain a model which has many stable states which are all *completely unfrustrated*. Of course this condition is much stronger than (6) since in this case (6) holds for any bond (not only for the internal field).

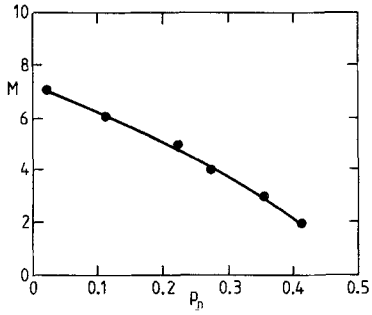

Fig. 13. "Phase boundary" of the 20×20 network which has learned M random pattern by killing frustated bonds. In the lower left region noisy patterns are recognized with less than 0.5% error

By definition each state is stable, i.e. the retrieval error for the pure patterns is zero. However, the problem remains how well the network recovers noisy patterns. To do so we expect that the number of nonzero bond should be at least of the order of the number N of spins; otherwise there is no sufficiently dense network of connections and almost any initial state is stable. If we use M random patterns then the number N_b of bonds which are not reduced is easily calculated from (19), on obtains

$$N_b/N^2 = 1/2^M. \tag{20}$$

If we demand $N_b > N$ we see that one can store at most $M = \ln N / \ln 2$ patterns. Hence the amount of patterns which can be stored is very small (note that $M \propto N$ for the Cooper algorithm, Eq. (7)).

If the patterns are correlated then less bonds are reduced by (19). Thus this algorithm is not very sensitive to the amount of correlations between learned patterns.

Figure 11 shows that even very noisy input is well recognized by this network. Five correlated patterns have been learned in a 20×20 system by reducing all frustrated bonds; only 16% of the initial bonds remained. For 20% noise all letters are well recovered while for 30% noise some larger errors are introduced as shown in Fig. 12. Again we expect that such errors may disappear with a stochastic dynamics by first heating the system up and then cooling it slowly down [16].

Figure 13 shows the "phase diagram" of this algorithm applied to random patterns. As in Fig. 7 noisy patterns are recognized with less than 0.5% errors below the indicated curve.

VI. Summary and Conclusions

It has been demonstrated in detail that the Hopfield model can store and recognize patterns. The infor-

mation is stored in the whole network of bonds simultaneously, hence the storage is very insensitive to defects. The information is retrieved in a few adaptions per site by an asynchronous and parallel relaxation of the energy of this model. Information is not retrieved by an address as in present day computers but by content; even very noisy patterns are recognized fast. Information retrieval is a spontaneous property of the network; it does not need any complex processing unit. As was suggested by Hopfield these properties may be compared with biological computation.

In this paper the retrieval error (or overlap) was discussed as a function of the number of patterns to be stored and the amount of noise of the input information. For random patterns and the Cooper definition of the bonds an approximate recursion relation for the dynamics of the retrieval overlap was derived. For a finite number of connections per site the patterns can be stored with a certain error, only. For the infinite range model a finite number of patterns can be stored completely. For a given maximum error the amount of stored patterns is proportional to the number of sites. Numerical simulations showed that for pure inputs the retrieval error is a self-averaging quantity. For a given retrieval error the amount of stored patterns and of input noise was calculated.

With an increasing number of stored patterns per site the Hopfield model interpolates continuously between ferromagnets and spin glasses. The retrieval overlap of spin glasses is just their saturated remanent magnetization. It will be interesting to see whether already with a small fraction of stored patterns the model has spin glass properties, which for instance show up in replica symmetry breaking.

It was demonstrated that spin models can learn information by killing frustrated bonds. Starting from a completely unorganized network with random bonds (=EA model) the system could store correlated pattern by reducing all bonds which are frustrated in one of the patterns. Even very noisy imputs could be recognized well. This learning process was motivated by biological behaviour: In the first few years of life many synapses of the brain are dying off although the amount of learned information is large and permanent. It should be interesting to study the statistical mechanics of this model since it has many stable states which are unfrustrated.

In the Hopfield model patterns are recognized by the relaxation of the energy. The model has several stable states which are attractors for parts of the phase space. One may consider this as classification or categorization of information. However, the basin of attraction is mainly determined by the overlap (or Hamming distance). Hence the Hopfield model cannot recognize for instance a small and a capital letter to be the same. This is different in other disrete dynamical models which reinforce differences in the patterns [19]. These models can be constructed such that two arbitrary patterns give either the same or a different output as requested. Is it possible to construct a spin model with the information being stored in bonds which has similar properties?

In fact the recursion relation of (14) indicates that this should be possible. This relation also describes a layered structure where the first layer is the input and the last layer is the output. The information is stored in a network of bonds between the layers. The bonds can be chosen such that two different inputs gives the same output. To what extend this works will be a subject of a future contribution.

The author would like to thank M. Kerszberg and P. Rujan for many useful discussions.

References

1. For a review see: Fischer, K.H.: Phys. Status Solidi **116**, 357 (1983); (**130**) (1985)
2. Edwards, S.F., Anderson, P.W.: J. Phys. F **5**, 968 (1975)
3. Binder, K. (ed.): Monte Carlo methods in statistical physics. In: Topics in Current Physics. Vol. 7. Berlin, Heidelberg, New York: Springer 1979
4. Binder, K., Schröder, K.: Phys. Rev. B **14**, 2142 (1976)
5. Kinzel, W.: Phys. Rev. B **19**, 4595 (1979)
6. Sherrington, D., Kirkpatrick, S.: Phys. Rev. Lett. **35**, 1792 (1975)
7. Parisi, G.: Phys. Rev. Lett. **43**, 1954 (1979); **50**, 1946 (1983)
8. Mezard, M., Parisi, G., Sourlas, N., Toulouse, G., Virasoro, M.: Phys. Rev. Lett. **52**, 1156 (1984)
9. Hopfield, J.J.: Proc. Natl. Acad. Sci. USA **79**, 2554 (1982)
10. Personnaz, L., Guyon, I., Dreyfus, G.: J. Phys. (Paris) Lett. **46**, L359 (1985)
11. Toulouse, G.: Comm. Phys. **2**, 155 (1977)
12. Mattis, D.C.: Phys. Lett. **56** A, 421 (1976)
13. Sherington, D.: Phys. Lett. **58** A, 36 (1976); Hemmen, J.L. van: Phys. Rev. Lett. **49**, 409 (1982); Provost, J.P., Vallee, G.: Phys. Rev. Lett. **50**, 598 (1983)
14. Benamira, F., Provost, J.P., Vallee, G.: (to be published)
15. Choy, T.C., Sherrington, D.: J. Phys. C **17**, 739 (1984)
16. Amit, D.J., Gutfreund, H., Sompolinski, H.: (to be published)
17. Kinzel, W.: (to be published)
18. Purves, D., Lichtman, J.W.: Science **210**, 153 (1980); I thank Kerszberg, M. for bringing this fact to my attention
19. Huberman, B.A., Hogg, T.: Phys. Rev. Lett. **52**, 1048 (1984); Hogg, T., Huberman, B.A.: J. Stat. Phys. (1985); Proc. Natl. Acad. Sci. USA **81**, 6871 (1984)

W. Kinzel
Institut für Festkörperforschung
Kernforschungsanlage Jülich GmbH
Postfach 1913
D-5170 Jülich 1
Federal Republic of Germany