Replica symmetry breaking and the nature of the spin glass phase

M. Mézard, G. Parisi (+), N. Sourlas, G. Toulouse (*) and M. Virasoro (++)

Laboratoire de Physique Théorique de l'Ecole Normale Supérieure (**)

(+) Universita di Roma II, Tor Vergata, Italy
(*) Laboratoire de Physique de l'Ecole Normale Supérieure, Paris, France

Résumé. — Récemment, l'un d'entre nous a proposé, comme paramètre d'ordre pour les verres de spin, une distribution de probabilité. Nous montrons que cette probabilité dépend de la réalisation particulière des couplages, même à la limite thermodynamique, et nous étudions sa distribution. Nous montrons aussi que l'espace des états est muni d'une topologie ultramétrique.

Abstract. — A probability distribution has been proposed recently by one of us as an order parameter for spin glasses. We show that this probability depends on the particular realization of the couplings even in the thermodynamic limit, and we study its distribution. We also show that the space of states has an ultrametric topology.

1. Introduction.

The usual approach to investigate the different phases of a physical system starts with the mean field approximation. In the case of spin glasses, even this first step has required a lot of effort [1-5].

The mean field approximation has been formulated by using the infinite range model [2]. The partition function is given by :

$$Z(\beta, J, h) = \sum_{\{\sigma\}} e^{-\beta H (J,h,\sigma)} \qquad (1)$$

$$H(J, h, \sigma) = - \sum_{i,j} J_{ij} \, \sigma_i \, \sigma_j - h \sum_i \sigma_i \qquad (2)$$

where the σ_i $i = 1, ..., N$ are Ising spins, the sum $\sum_{i,j}$ runs over all pairs of spins, and J_{ij} are random couplings obeying a given probability distribution, which we suppose to be symmetric, with variance $1/\sqrt{N}$.

In order to study the Hamiltonian (2), four main approaches have been used. The replica approach [1-2, 4-5] in which one considers n copies of the same

(**) Laboratoire Propre du Centre National de la Recherche Scientifique, associé à l'Ecole Normale Supérieure et à l'Université de Paris Sud. Postal Address : 24, rue Lhomond, 75231 Paris Cedex 05, France.
(++) Permanent address : Departimento di Fisica, Universita di Roma I, La Sapienza, Italy.

system, averages over the coupling distribution, and at the end takes the limit $n \to 0$. In this way one computes the averages $\overline{O_i(J)}$ of the physical observables $O_i(J)$ (free energy, magnetic susceptibility, correlation functions, etc.) over the coupling distribution. (In this paper we shall always denote the thermal averages by $\langle \ \rangle$ and the averages over the coupling distribution by $\overline{\quad}$.) The other three approaches are the self-consistent field approach [3, 6], the dynamical approach [7] and the numerical simulations [2, 8].

The picture which has emerged from the four above approaches is that the spin glass phase is characterized by the existence of a large number (infinite when $N \to \infty$) of equilibrium states $\alpha = 1, 2, ...$ almost degenerate (free energy valleys separated by free energy barriers becoming infinitely high in the thermodynamic limit).

In a recent paper [10] (to be referred to later as [I]), one of us has proposed an order parameter for the spin glass phase and has shown its connection and interpretation in terms of the many valley picture. In the present paper we further continue this investigation. A short version of the present work has been published elsewhere [11]. The Boltzmann-Gibbs measure is :

$$\langle \, 0(\sigma) \, \rangle = \frac{\sum\limits_{\{\sigma\}} 0(\sigma) \exp[- \beta H(\sigma)]}{\sum\limits_{\{\sigma\}} \exp[- \beta H(\sigma)]} \qquad (3)$$

which can be decomposed as a sum over the pure equilibrium (clustering) states :

$$\langle \cdots \rangle = \sum_{\alpha} P_{\alpha} \langle \cdots \rangle_{\alpha}, \qquad \sum_{\alpha} P_{\alpha} = 1. \qquad (4)$$

We may characterize a pure state α of a spin glass by the magnetization $m_i^{\alpha} = \langle \sigma_i \rangle_{\alpha}$ at each point i of the system. Following [I] we define the overlap $q^{\alpha\beta}$ of the two pure states α and β :

$$q^{\alpha\beta} = \frac{1}{N} \sum_{i=1}^{N} m_i^{\alpha} m_i^{\beta} \qquad (5)$$

($q^{\alpha\alpha}$ is the familiar Edwards-Anderson order parameter $q_{\text{E.A.}}^{\alpha} = \frac{1}{N} \sum_i \langle \sigma_i \rangle_{\alpha}^2$), and the probability $P_J(q)$ for a pair of states (α, β) to have an overlap q :

$$P_J(q) = \sum_{\alpha,\beta} P_{\alpha} P_{\beta} \, \delta(q - q^{\alpha\beta}). \qquad (6)$$

We shall denote by $P(q)$ the average of $P_J(q)$ over the coupling distribution :

$$P(q) = \overline{P_J(q)}. \qquad (7)$$

Let us call q_{Max} and q_{Min} the maximum and the minimum possible overlaps between two states at a given temperature T and magnetic field H. Obviously $-1 \leqslant q_{\text{Min}} \leqslant q_{\text{Max}} \leqslant 1$. As in [I] we define

$$x(q) = \int_{-1}^{q} dq' \, P(q') \qquad (8)$$

and $q(x)$ the inverse function of $x(q)$. Because of (6) and (8) : $x(q_{\text{Max}}) = 1$ and $x(q_{\text{Min}}) = 0$.

Let us also define :

$$y(q) = 1 - x(q) = \int_{q}^{1} dq' \, P(q') \qquad (9)$$

which is the probability of two pure states to have an overlap larger than q.

In the replica approach the order parameter is a $n \times n$ matrix Q_{ab}. In the limit $n \to 0$ (if we follow the Parisi pattern of replica symmetry breaking) the matrix Q is characterized by a function $Q(x)$ where $0 \leqslant x \leqslant 1$. It was shown in [I] that $q(x)$ (the inverse function of $x(q)$ defined in (8)), is identical to $Q(x)$, thus giving a physical interpretation to the replica symmetry breaking.

In the familiar case of an homogeneous ferromagnetic system (i.e. in the absence of any disorder), if we start from the high temperature phase and cool down below the Curie temperature, the probabilities P_+ and P_- of arriving to a state of magnetization $+ m(T)$ or $- m(T)$, are well known to depend on the boundary conditions we have imposed on the system. So, even in this simple case, P_{α} and therefore also $P_J(q)$ are not « good, extensive quanti-

ties ». This trivial observation makes one strongly suspect that the same is true in the more complicated spin glass case. This suspicion will become a certainty in chapter 2, since we shall show that $P_J(q)$ depends on the realization of the couplings, even in the thermodynamic limit : it is not a « self averaging » quantity, in the sense that, as far as $P_J(q)$ is concerned, an increasing size of the sample does not imply an average over all disorder configurations.

The same properties are true for

$$Y_J(q) = \int_{q}^{q_{\text{Max}}} dq' \, P_J(q'), \qquad \overline{Y_J(q)} = y(q). \qquad (10)$$

We show that the probability distribution of $Y_J(q)$ is calculable in the framework of the replica scheme. We find that this distribution is such that the most probable value of Y_J is 1, and therefore differs from its mean value y.

The possibility of such a behaviour for disordered systems has been suspected before [12], but it is the first time, to our knowledge, that this is demonstrated in the context of the S.K. model. The order parameter, far from being a parameter, was shown to be a function, interpreted as a probability law. And now, on top of that, there appears a probability law for this function, i.e. a probability law for a probability law.

In chapter 2 we show another remarkable property of the spin glasses which is the ultrametric topology of the space of states. Taking any three pure spin glass states α, β, γ and computing the three overlaps $q^{\alpha\beta}$, $q^{\beta\gamma}$, $q^{\gamma\alpha}$, we find that at least two of them are equal. Spaces with such a property are called ultrametric spaces [13]. From this property we show that for any value of q, by grouping together all the states with an overlap bigger than q, we separate the space of states into disjoint clusters. Each such cluster is again divided into smaller clusters, by grouping together the states with an overlap bigger than $q' > q$. This procedure can be repeated indefinitely. So we prove that the space of pure states has a hierarchical structure. This hierarchical structure is a characteristic property of ultrametric spaces.

In chapter 4 we compute the cluster distribution. In particular we show that for any value of y in the spin-glass phase (this means for $T < T_g$), the number of clusters is infinite.

Another remarkable property of all the probability distributions we consider is their universality : they depend on the different parameters of the problem (temperature, magnetic field, the particular value of q we are considering) only through the mean value y of $Y_J(q)$.

We should emphasize that all our results are obtained in the framework of the replica symmetry

breaking scheme of reference [5], that is in purely static terms.

Our paper is organized as follows. In chapter 2 we recall the replica symmetry breaking mechanism from which we then derive the ultrametric structure of the space of the pure spin glass states. We also show that $P_J(q)$ is not self-averaging in the thermodynamic limit.

In chapter 3 we compute the probability distribution of $Y_J(q)$. In chapter 4 we study the distribution properties of the clusters in the space of the spin glass states.

2. Hierarchical organization of the spin glass states.

We first recall the replica symmetry breaking (R.S.B.) mechanism because in the following we will make an explicit use of it.

The $n \times n$ matrix Q_{ab} is constructed through the following recursive algorithm of successive R.S.B.

a) No symmetry breaking : $Q_{ab}^{(0)} = Q_0$ for $a \neq b$, $Q_{aa} = 0$.

b) First R.S.B. : The $n \times n$ matrix $Q^{(0)}$ is broken into $\dfrac{n}{m_1} \times \dfrac{n}{m_1}$ blocks (submatrices) Q^{A_1,B_1}, A_1, $B_1 = 1$, ..., $\dfrac{n}{m_1}$ where the Q^{AB} are $m_1 \times m_1$ matrices. For the non-diagonal submatrices $A_1 \neq B_1$, $Q_{\alpha\beta}^{A_1 B_1} = Q_0$ and $Q_{aa}^{A_1 A_1} = 0$, $Q_{ab}^{A_1 A_1} = Q_1$, $a \neq b$ for the diagonal submatrices. This completes the construction of the matrix $Q^{(1)}$.

c) Second R.S.B. The same procedure of R.S.B. is repeated with the diagonal submatrices $Q^{A_1 A_1}$. They are broken into $\dfrac{m_1}{m_2} \times \dfrac{m_1}{m_2}$ submatrices $Q^{A_2 B_2}$, A_2, $B_2 = 1, ..., \dfrac{m_1}{m_2}$ and $Q_{ab}^{A_2 B_2} = Q_1$ for $A_2 \neq B_2$, $Q_{ab}^{A_2 A_2} = Q_2$ for $a \neq b$, $Q_{aa}^{A_2 A_2} = 0$.
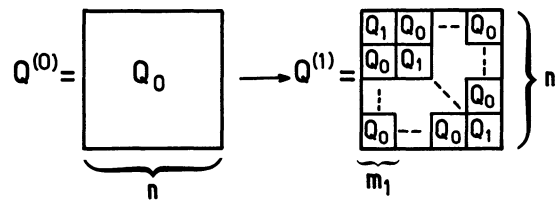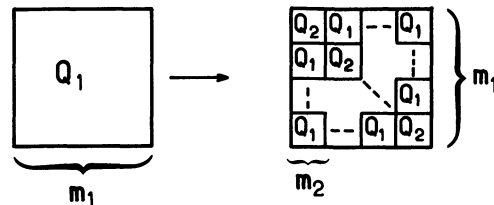
Fig. 1. — Iterative procedure for the construction of the matrix Q_{ab}.

d) The same procedure is repeated indefinitely. The whole process is illustrated in figure 1.

By construction $n \geqslant m_1 \geqslant \cdots \geqslant m_M \geqslant 1$. In the $n \to 0$ limit this becomes $1 \geqslant m_M \geqslant \cdots \geqslant m_1 \geqslant 0$, and in the limit $M \to \infty$, the m_i's become a continuous variable $m_i \to x$, $0 \leqslant x \leqslant 1$, $m_{i+1} \to x + \mathrm{d}x$. The Q_i's become the well-known function $Q(x)$.

It was shown in [I] that the characteristic function $g(y)$

$$g(y) = \int \mathrm{d}q \, \overline{P_J(q)} \, \mathrm{e}^{yq} = \int \mathrm{d}q \, P(q) \, \mathrm{e}^{yq} \qquad (11)$$

can be computed in the replica framework and is given by

$$g(y) = \frac{1}{n(n-1)} \sum_{\substack{a=1 \\ a \neq b}}^{n} \sum_{b=1}^{n} \mathrm{e}^{yQ_{ab}} \xrightarrow[n \to 0]{} \int_0^1 \mathrm{d}x \, \mathrm{e}^{yQ(x)} \qquad (12)$$

where the sum is over all pairs of distinct replica indices [14].

Let us now consider any three pure states α_1, α_2, α_3 and $P_J(q_1, q_2, q_3)$ the probability for them to have overlaps $q_1 = q^{\alpha_2 \alpha_3}$, $q_2 = q^{\alpha_3 \alpha_1}$, $q_3 = q^{\alpha_1 \alpha_2}$, respectively. $P_J(q_1, q_2, q_3)$ is obviously symmetric under permutations of its arguments. In order to compute $P_J(q_1, q_2, q_3)$ in the R.S.B. scheme, following [I], we consider the generalized Laplace transform

$$g_J(y_1, y_2, y_3) = \int \mathrm{d}q_1 \, \mathrm{d}q_2 \, \mathrm{d}q_3 \, \exp\left(\sum_{i=1}^{3} q_i \, y_i \right) P_J(q_1, q_2, q_3) \qquad (13)$$

and we take three identical copies of the systems, with spins σ_1, σ_2, σ_3 and Hamiltonian $H_3(\sigma_1, \sigma_2, \sigma_3) = H(\sigma_1) + H(\sigma_2) + H(\sigma_3)$. Then

$$g_J(y_1, y_2, y_3) = \left\langle \exp\left[\frac{1}{N} \sum_i \left(y_1 \, \sigma_2(i) \, \sigma_3(i) + y_2 \, \sigma_3(i) \, \sigma_1(i) + y_3 \, \sigma_1(i) \, \sigma_2(i) \right) \right] \right\rangle_3 \qquad (14)$$

where $\langle \quad \rangle_3$ means that the expectation value is taken with respect to the Hamiltonian H_3. It is then possible to compute the average over the J's

$$g(y_1, y_2, y_3) = \overline{g_J(y_1, y_2, y_3)}.$$ (15)

By introducing n replicas and letting $n \to 0$

$$g(y_1, y_2, y_3) = \frac{1}{n(n-1)(n-2)} \sum_{\substack{a=1 \\ a \neq b}}^{n} \sum_{\substack{b=1 \\ b \neq c}}^{n} \sum_{\substack{c=1 \\ a \neq c}}^{n} e^{y_1 Q_{ab}} e^{y_2 Q_{bc}} e^{y_3 Q_{ca}}.$$ (16)

As in [I] the sum in (16) must run over all replica indices because of the presence of R.S.B. (In this context see also [14].) Using the R.S.B. algorithm one constructs the matrices $A_{ab}(y) \equiv e^{yQ_{ab}}$ for $a \neq b$ and $A_{aa}(y) = 0$. Then

$$g(y_1, y_2, y_3) = \lim_{n \to 0} \frac{1}{n(n-1)(n-2)} \operatorname{Tr} \left[A(y_1) A(y_2) A(y_3) \right].$$ (17)

As usual, in the $n \to 0$ limit one has to replace $m_i \to x$ and $Q_i \to Q(x)$. Thus, after some algebra, we get : $(P(q_1, q_2, q_3) \equiv \overline{P_J(q_1, q_2, q_3)})$

$$P(q_1, q_2, q_3) = \tfrac{1}{2} P(q_1) x(q_1) \delta(q_1 - q_2) \delta(q_1 - q_3) +$$

$$+ \tfrac{1}{2} \left\{ P(q_1) P(q_2) \theta(q_1 - q_2) \delta(q_2 - q_3) + \text{permutations} \right\}.$$ (18)

This formula is quite interesting. It means that if we take any three states, at least two pairs of them will have the same overlap with probability one. This property is reminiscent of ultrametric spaces [13] [1].

Let's now consider two states γ and γ' such that $q^{\alpha\gamma} \geqslant q$, $q^{\alpha\gamma'} \geqslant q$. It follows from (18) that if $q^{\alpha\gamma} = q^{\alpha\gamma'}$ then $q^{\gamma\gamma'} \geqslant q$ and if $q^{\alpha\gamma} \neq q^{\alpha\gamma'}$ then $q^{\gamma\gamma'} = q^{\alpha\gamma}$ or $q^{\gamma\gamma'} = q^{\alpha\gamma'}$.

It follows from this property that the states are organized in non overlapping clusters.

Let α and β be two pure states of our system, $I_\alpha(q)$ the set of states γ which have an overlap with α bigger or equal to q ($q^{\alpha\gamma} \geqslant q$ for every γ) and similarly for $I_\beta(q)$ ($\delta \in I_\beta(q) \Rightarrow q^{\beta\delta} \geqslant q$). It follows from the previous remark that for any pair of states γ, γ' belonging to $I_\alpha(q)$ (or to $I_\beta(q)$), $q^{\gamma\gamma'} \geqslant q$. It is a consequence of (18) that two sets $I_\alpha(q)$ and $I_\beta(q)$ are always either identical or disjoint because if there existed a pure state γ, belonging to both $I_\alpha(q)$ and $I_\beta(q)$, i.e. $q^{\alpha\gamma} \geqslant q$ and $q^{\beta\gamma} \geqslant q$, then by equation (18) one would get $q^{\alpha\beta} \geqslant q$.

The following hierarchical structure of the states of the spin glass phase emerges from the previous remarks. For any q, $q_{\text{Min}} \leqslant q \leqslant q_{\text{Max}}$, the states are organized into disjoint clusters, such that any pair (α, β) of states inside the same cluster has an overlap $q^{\alpha\beta} \geqslant q$. Now one can again divide each of the previous clusters into disjoint smaller clusters by choosing a q', $q < q' \leqslant q_{\text{Max}}$ and grouping together the states with overlap bigger than q'. This procedure can be repeated indefinitely. We have proved that the space of pure states has a hierarchical structure (characteristic of ultrametric spaces).

In order to see whether the probabilities $P_J(q)$ approach a definite limit or fluctuate when the number N of spins becomes infinite, we will now compute $\overline{P_J(q_1) P_J(q_2)} - P(q_1) P(q_2)$. One way of computing $\overline{P_J(q_1) P_J(q_2)}$ is to consider four pure states α_1, α_2, α_3, α_4, compute the averaged over J probability $P(q_1, q_2)$ to have $q^{\alpha_1\alpha_2} = q_1$, $q^{\alpha_3\alpha_4} = q_2$. The computation of $P(q_1, q_2)$ is similar to the computation of $P(q_1, q_2, q_3)$: we consider the generalized Laplace transform $g(y_1, y_2)$ of $P(q_1, q_2)$ and use again the replica trick.

We finally get

$$g(y_1, y_2) = \int dq_1 \, dq_2 \, e^{y_1 q_1 + y_2 q_2} P(q_1, q_2)$$

$$= \int dq_1 \, dq_2 \, e^{y_1 q_1 + y_2 q_2} \overline{P_J(q_1) P_J(q_2)}$$

---

[1] The definition of an ultrametric space requires the definition of a distance. A natural choice is

$$d(\alpha, \beta) = \frac{1}{N} \sum_i (m_i^\alpha - m_i^\beta)^2.$$

As $q^{\alpha\alpha} = q_{\text{E.A.}}$ for almost all states α (see (47)), one has simply $d(\alpha, \beta) = 2(q_{\text{E.A.}} - q^{\alpha\beta})$. The characteristic property of ultrametric spaces is that if one takes any three points of the space, they form an isosceles triangle, with the two equal angles larger than or equal to the third one.

$$= \frac{1}{n(n-1)(n-2)(n-3)} \sum_{a=1}^{n} \sum_{b=1}^{n} \sum_{c=1}^{n} \sum_{d=1}^{n} e^{y_1 Q_{ab} + y_2 Q_{cd}} \tag{19}$$

where the sum is restricted to ensembles of replicas a, b, c, d which are all different.

Taking the $n \to 0$ limit we get :

$$\overline{P_J(q_1) \, P_J(q_2)} = \frac{1}{3} P(q_1) \, \delta(q_1 - q_2) + \frac{2}{3} P(q_1) \, P(q_2) \,. \tag{20}$$

Formula (20) is a direct and manifest proof that $P_J(q)$ (and therefore $x_J(q)$) fluctuates with J even after the thermodynamic limit is taken.

In the next section we will compute the probability distribution of $Y_J(q)$.

We should emphasize that all our results have been obtained within the framework of the replica symmetry breaking scheme of reference [5].

3. Reconstruction of the probability distribution.

As the structure of the ensemble of pure states depends on the realization of the couplings, the function $P_J(q)$ fluctuates and one would like to know the probability distribution of this function. This could, in principle, be studied from the moments $\overline{P_J(q_1) \, P_J(q_2) \dots P_J(q_k)}$, but it turns out that the direct computation of these moments with the method we have sketched before is a difficult task, as soon as k gets larger than 3 or 4.

We shall see in the following that it is much easier to calculate the moments $\overline{Y_J(q)^k}$ (all the $Y_J(q)$ taken at the same value of q). The second moment is already contained in equation (20)

$$\overline{Y_J(q)^2} = \int_q^{q_{Max}} dq_1 \int_q^{q_{Max}} dq_2 \, \overline{P_J(q_1) \, P_J(q_2)} = \frac{1}{3} y(q) + \frac{2}{3} y(q)^2 \,. \tag{21}$$

We have computed similarly

$$\overline{Y_J(q)^3} = \frac{1}{5 \, !!} \left(3 \, y(q) + 7 \, y(q)^2 + 5 \, y(q)^3 \right). \tag{22}$$

The remarkable property of these two equations is that the n'th moment of $Y_J(q)$ ($n = 1, 2, 3$) is expressed in terms only of its mean value $y(q)$ (it is in fact a n'th degree polynomial in y). In particular there is no coupling between different values of q. Those properties are true for any n in the R.S.B. scheme of reference [5] and are due to the ultrametric topology of the replica space itself, whose signature has been seen in other properties [18].

In fact we can interpret Q_{ab} as the overlap between replicas a and b :

— With the first symmetry breaking, the n replicas are organized in n/m_1 clusters of size m_1. The overlap between two replicas within the same cluster is Q_1 while the overlap between two replicas belonging to different clusters is Q_0. As $Q(x)$ is a monotonous function, Q_1 is larger than Q_0 and hence the overlap of the replicas within a cluster is larger than the overlap between replicas in different clusters, as it should be.

— Performing the second breaking, one sees that the replicas inside one cluster of size m_1 are themselves grouped in subclusters of size m_2, etc.

The structure of replica space can be pictorially described as in figure 2. It is an ultrametric space where the sizes of all the clusters at a given scale Q_k are the same and are equal to m_k [2].

Now let us suppose that we want to compute a quantity such as $\overline{Y_J(q)^p}$ which involves only one scale of distances q between pure states. As was shown in the previous section, this quantity can be obtained by using the replica formalism : in the precise case of $\overline{Y_J(q)^p}$, one must choose p pairs of replicas such that

— all the replicas be distinct

— the overlap between the two replicas of a pair be larger than q.

From the ultrametric topology of replica space, the choice of a scale of distances q naturally induces a partition of the n replicas into n/m disjoint clusters of size m, such that the overlap of replicas inside the same cluster be larger than q, while replicas in different clusters have an overlap smaller than q. So if we are interested in only one scale q, we can forget about all the other structures of the replica space. This means that we can compute $\overline{Y_J(q)^p}$ by applying only one replica symmetry breaking (i.e. from the matrix $Q_{ab}^{(1)}$ constructed in the

---

[2] As Q_{aa} is taken equal to zero, the ultrametric structure is true, strictly speaking, only if one doesn't consider the self overlap of a replica.
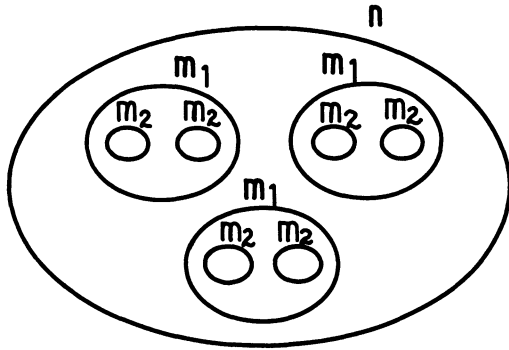
Fig. 2. — Representation of the ultrametric structure of replica space.

previous chapter) provided $Q_0 < q < Q_1$. With full R.S.B., it is known that the numbers m_k, in the limit $n \to 0$, are equal to the value of x for which $q(x) = q_k$. So the right value of m for this single R.S.B. is $m = x(q) = 1 - y(q)$.

With a single breaking of the replica symmetry, the moments can be computed much more easily. Let us for instance compute in detail the second one, defined as :

$$\overline{Y_J(q)^2} = \sum_{\alpha,\beta,\gamma,\delta} P_\alpha P_\beta \, \theta(q^{\alpha\beta} - q) \, P_\gamma P_\delta \, \theta(q^{\gamma\delta} - q) \, . \tag{23}$$

α, β, γ, δ label the pure states, and θ is the usual step function. As explained in section 2, one obtains, with the replica formalism :

$$\overline{Y_J(q)^2} = \frac{1}{n(n - 1)(n - 2)(n - 3)} \sum_{a,b,c,d}' \theta(Q_{ab} - q) \, \theta(Q_{cd} - q)$$

where the symbol $\sum'$ means that the four replica indices a, b, c, d, must be different.

At the scale q, the replicas are grouped into clusters of size m, and one has two possibilities :

— Either the two pairs of replicas are in the same cluster, which we describe as :
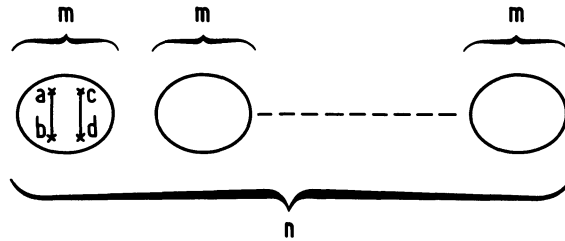
This contribution to the sum is $\dfrac{n(m - 1)(m - 2)(m - 3)}{n(n - 1)(n - 2)(n - 3)}$ which gives, as $n \to 0$, $m \to 1 - y(q)$: $\dfrac{1}{3!} y(1 + y)(2 + y)$.

— Or the two pairs are in different clusters :
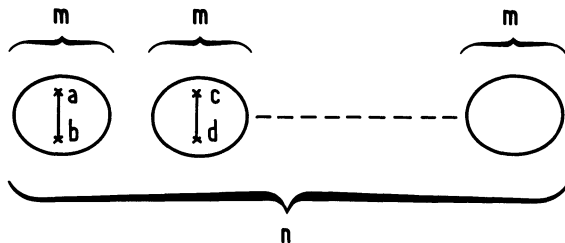
which gives a contribution : $\dfrac{n(m - 1)(n - m)(m - 1)}{n(n - 1)(n - 2)(n - 3)} = \dfrac{1}{3!} y^2(1 - y)$. So the final result for $\overline{Y_J(q)^2}$ is (we now omit the explicit reference to the scale q) :

$$\overline{Y_J^2} = \frac{1}{3}(y + 2 y^2) \, .$$

One can write down the rules which give the general moment $\overline{Y_J^p}$:

— One must take p pairs of replicas and distribute them in all the possible distinct ways between the clusters. The diagram obtained :
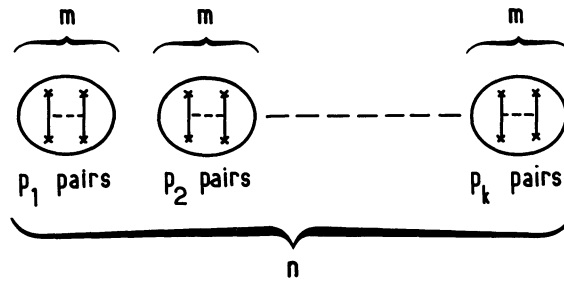
gives a contribution

$$S \frac{p\,!\,(k-1)\,!}{(2\,p-1)\,!\,p_1\,!\,...\,p_k\,!}(1-y)^{k-1}\,y^k[(1+y)(2+y)]^{N_2}\,[(3+y)(4+y)]^{N_3}\,... \tag{24}$$

where

— $k \equiv N_1$ is the number of clusters containing at least one pair
— N_r is the number of clusters containing at least r pairs
— S is the symmetry factor of the diagram, given by :

$$S = \frac{1}{\displaystyle\prod_{r=1}^{p}(N_r - N_{r+1})\,!}\,. \tag{25}$$

With these rules we have computed the first seven moments :

$$\overline{Y_J^3} = \frac{1}{5\,!!}\,[3\,y + 7\,y^2 + 5\,y^3]$$

$$\overline{Y_J^4} = \frac{1}{7\,!!}\,[15\,y + 39\,y^2 + 37\,y^3 + 14\,y^4]$$

$$\overline{Y_J^5} = \frac{1}{9\,!!}\,[105\,y + 296\,y^2 + 326\,y^3 + 176\,y^4 + 42\,y^5]$$

$$\overline{Y_J^6} = \frac{1}{11\,!!}\,[945\,y + 2\,838\,y^2 + 3\,458\,y^3 + 2\,228\,y^4 + 794\,y^5 + 132\,y^6]$$

$$\overline{Y_J^7} = \frac{1}{13\,!!}\,[10\,395\,y + 32\,859\,y^2 + 43\,191\,y^3 + 31\,235\,y^4 + 13\,553\,y^5 + 3\,473\,y^6 + 429\,y^7]\,(^3). \tag{26}$$

(^{3}) We have noticed that, up to the 7th moment, the following formulas are true :

$$\overline{Y_J^K} \sim \frac{1}{2\,K-1}\,y \qquad y \to 0$$

$$\overline{Y_J^K} \sim \frac{2^K}{(K+1)\,!}\,y^K \qquad y \to \infty$$

$$\overline{Y_J^K} = \frac{(-1)^K}{(2\,K-1)\,!!} \qquad y = -1$$

$$\overline{Y_J^K} = \frac{2.3^{K-1}(-1)^K}{(2\,K-1)\,!!} \qquad y = -2\,.$$

We have also shown that, up to $K = 3$, the polynomials $R_K(y)$ which give $\overline{Y_J^K}$:

$$\overline{Y_J^K} \equiv \frac{1}{(2\,K-1)\,!!}\,R_K(y)$$

are simply related to the number of diagrams of different types (planar or non planar) of a free field matrix theory : considering a free field theory for $N \times N$ real symmetric matrices M_{ab}, with the constraint that the diagonal elements $M_{aa} = 0$, one finds that, up to $K = 3$:

$$\frac{1}{N(N-1)}\,\langle\,\mathrm{Tr}\,(M^{2K})\,\rangle = R_K(y)\,, \quad \text{where} \quad y = N-1\,.$$

We have not been able to obtain a general formula for the coefficients of the polynomial giving $\overline{Y_j^p}$. Instead we have numerically reconstructed the probability distribution of the variable Y at a given $q : \Pi_y(Y)$ (the index y is here to remind us that this distribution depends on $y(q)$ only).

In order to compute $\Pi_y(Y)$ from its moments, it would be useful to control eventual singularities. In particular, the behaviour of $\Pi_y(Y)$ around $Y \simeq 1$ is reflected by the large p behaviour of the moments $\mu_p \equiv \overline{Y_j^p}$ (if $\Pi_y(Y) \sim (1 - Y)^{-\alpha}$, then $\mu_p \sim p^{-1+\alpha}$). From the seven first moments we tentatively conclude to a singularity of the type $\Pi_y(Y) \sim (1 - Y)^{-y}$. (We will give a different argument for this singularity in the next section.) We have inverted the moments μ_p by developing $\Pi_y(Y)$ on the basis of the Jacobí polynomials which are orthogonal on the interval [0, 1] with respect to the integration measure $(1 - Y)^{-y}$. This amounts to factorizing the $(1 - Y)^{-y}$ singularity into the integration measure and developing a smoother function on the polynomials. A good test of the method was its rapid convergence. We have also tried the maximum entropy method with similar results (see Fig. 3) [17].
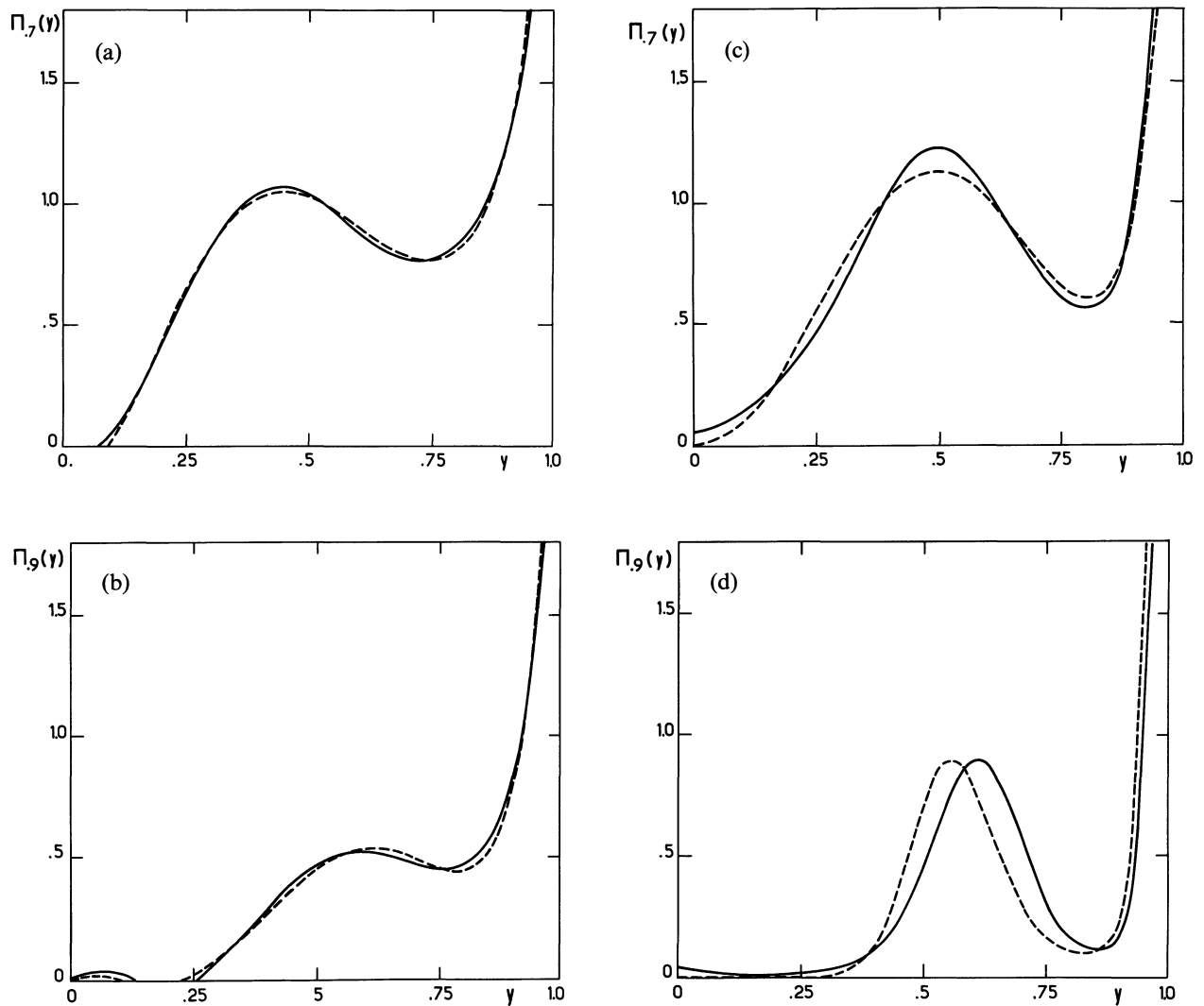

Fig. 3. — (a) Probability distribution of the variable Y, for the average value $y = 0.7$, reconstructed by projecting $\Pi(Y)$ on the Jacobi polynomials. The dashed curve is the probability obtained by inverting the first 6 moments while the full line is obtained from the first 7 moments.

(b) Probability distribution of the variable Y, for the average value $y = 0.9$, reconstructed by projecting $\Pi(Y)$ on the Jacobi polynomials. The dashed curve is the probability obtained by inverting the first 6 moments while the full line is obtained from the first 7 moments.

(c) Probability distribution of the variable Y, for the average value $y = 0.7$, reconstructed from the maximum entropy method. The dashed curve is the probability obtained by inverting the first 6 moments while the full line is obtained from the first 7 moments.

(d) Probability distribution of the variable Y, for the average value $y = 0.9$, reconstructed from the maximum entropy method. The dashed curve is the probability obtained by inverting the first 6 moments while the full line is obtained from the first 7 moments.

The Π_y we have obtained are plotted in figure 3 for various values of y. It is interesting to notice that the average value y of Y is very different from its most probable value $y_{\mathrm{M.P.}}$, which is in general equal to 1. In this case if one keeps to the most probable value, one gets a curve $y_{\mathrm{M.P.}}(q) = 1$ for every q, which is a replica symmetric behaviour, while the average value $y(q)$ obtained with the replica symmetry breaking is very different from $y_{\mathrm{M.P.}}$. It has been emphasized many times that this kind of behaviour often appears in disordered systems [12].

On the other hand the most important property of this probability distribution of Y at a given q is the fact that it does not depend on anything else than $y(q)$. This explains why the integrated probability y (or equivalently x) is so interesting : in terms of this variable, there exists a universal behaviour of the glassy phase in the following sense : for a given value of the parameters (temperature and magnetic field) T_1, H_1, and a given scale of overlap q_1, the probability distribution of Y is entirely determined by y. If one changes the parameters to values T_2, H_2, there exists a change of scale of overlaps $q_1 \rightarrow q_2$ such that the probability distribution of $Y(q_2)$ in this second system be exactly identical to the probability distribution of $Y(q_1)$ in the first system ; the rescaling in overlaps is given by : $y_{T_1,H_1}(q_1) = y_{T_2,H_2}(q_2)$. The universality comes from the fact that the whole dependence on H and T is through the function $y_{T,H}(q)$. But in order to compute the dependence of y on q, T, H, we must minimize the free energy and for this the entire sequence of R.S.B.'s is required.

Finally let us emphasize the crucial role played by the ultrametric topology of the replica space : because of this topology, a quantity which involves p scales of distances q_1, ..., q_p can be computed with p explicit breakings of the replica symmetry. This allowed us to compute the moments $\overline{Y_J(q)^r}$ to an arbitrarily large order, with a single replica symmetry breaking.

4. Distribution properties of the clusters of spin glass states.

We have shown that the space of pure states has a hierarchical structure that consists of clusters contained in clusters. In this section we will try to give more information on this structure, especially on the numbers and sizes of the clusters.

As in the previous section, given a certain scale q, we group into clusters all the states that have an overlap larger than q, defining in this way a partition of the ensemble of pure states into K clusters. Let us call W_I the weight of the I'th cluster :

$$W_I = \sum_{\alpha \in I} P_\alpha \tag{27}$$

obviously

$$\sum_I W_I = 1 \tag{28}$$

We are interested in the distribution of the weights W_I. (It is clear that the knowledge of all the W_I's for every scale q completely characterizes the space of states.) In order to take advantage of the universality demonstrated in the previous section we shall define the scale through the variable $y \equiv y(q)$.

Let $f_J(W, y)\, dW$ be the number of clusters that have weights W_I between W and $W + dW$, for a given distribution of couplings J :

$$f_J(W, y) = \sum_I \delta(W - W_I). \tag{29}$$

The average function $\overline{f_J(W, y)}$ can be computed in the following way : the moments $M_k = \displaystyle\int_0^1 W^k\, \overline{f_J(W, y)}$ are given by :

$$M_k = \overline{\sum_I W_I^k}$$

$$= \overline{\sum_I \left(\sum_{\alpha_1 \in I} \cdots \sum_{\alpha_k \in I} P_{\alpha_1} \cdots P_{\alpha_k} \right)}. \tag{30}$$

This is the total probability that k states are in the same cluster :

$$M_k = \overline{\sum_{\alpha_1} \cdots \sum_{\alpha_k} P_{\alpha_1} \cdots P_{\alpha_k}\, \theta(q^{\alpha_1 \alpha_2} - q)\, \theta(q^{\alpha_1 \alpha_3} - q) \dots \theta(q^{\alpha_1 \alpha_k} - q)}. \tag{31}$$

In the same way as in the previous section, this quantity can be calculated by going to replica space, with one R.S.B. at the scale q : M_k is equal to the probability of choosing k different replicas, all belonging to the same cluster (in replica space) :

$$M_k = \frac{1}{n(n-1)\dots(n-k+1)}\, n(m-1)\dots(m-k+1). \tag{32}$$

In the limit $n \to 0$, $m \to 1 - y$, we get :

$$M_k = \frac{\Gamma(k + y - 1)}{\Gamma(y) \times \Gamma(k)}.$$

(33)

It turns out that these moments can be inverted, yielding :

$$\overline{f_J(W, y)} = \frac{W^{y-2}(1 - W)^{-y}}{\Gamma(y)\,\Gamma(1 - y)}.$$

(34)

A number of physical conclusions can be obtained from this expression :

1) The average multiplicity of clusters is infinite for any value of y :

$$\int_0^1 \overline{f_J(W, y)}\, dW = \infty.$$

2) The infinite number of clusters is concentrated around $W = 0$, and their overall weight is infinitesimal : indeed, if one introduces a cut-off ε in the region of small weights, one gets an average number of clusters :

$$K_\varepsilon \equiv \int_\varepsilon^1 \overline{f_J(W, y)}\, dW \sim \frac{\varepsilon^{y-1}}{(1 - y)\,\Gamma(y)\,\Gamma(1 - y)} \qquad \begin{pmatrix} \varepsilon \ll 1 \\ y \neq 1 \end{pmatrix}$$

(35)

which diverges as $\varepsilon \to 0$, but their total weight

$$\int_\varepsilon^1 W\, \overline{f_J(W, y)}\, dW \sim 1 - \frac{\varepsilon^y}{\Gamma(1 + y)\,\Gamma(1 - y)} \qquad \begin{pmatrix} \varepsilon \ll 1 \\ y \neq 1 \end{pmatrix}$$

(36)

goes to one. For each J, the total number of clusters having a weight larger than ε is an integer between 0 and $1/\varepsilon$. But the average number K_ε is much smaller since it behaves like ε^{y-1}.

3) y is precisely the average size of clusters : this is in fact nothing but the definition of y, and this result is valid for each configuration of the couplings, since :

$$y_J = \sum_\alpha P_\alpha \sum_\beta P_\beta\, \theta(q^{\alpha\beta} - q)$$

$$= \sum_\alpha P_\alpha\, W_{I \ni \alpha} = \sum_I W_I^2.$$

(37)

4) The average probability, when choosing a state at random, that it be in a cluster of weight W is $W\, \overline{f_J(W, y)}$. When $y \to 1$, this probability is strongly peaked around $W = 1$. (The integrated probability in the interval $[0, 1/2]$, and in any finite interval that does not contain 1, goes to zero when $y \to 1$.) But for each J, there can be at most one cluster in the interval $]1/2, 1]$ (since $\sum_I W_I = 1$), and therefore in the limit $y \to 1$ there is, for each J, one large isolated cluster that dominates.

5) The same kind of argument explains the nature of the singularity at $Y = 1$ of the function $\Pi_y(Y)$ computed in chapter 3 : $\Pi_y(Y)$ is defined as :

$$\Pi_{y(q)}(Y) = \overline{\delta(y_J(q) - Y)} = \overline{\delta\left(\sum_I W_I^2 - Y\right)}.$$

(38)

For Y near to 1, the configurations which contribute are those for which one cluster dominates :

$$\Pi_{y(q)}(Y) = \overline{\delta(W_M^2 - Y)}, \qquad (Y \to 1)$$

(39)

where W_M is the weight of the largest cluster. Thus from (34) :

$$\Pi_y(Y) \sim \frac{2^{y-1}}{\Gamma(y)\,\Gamma(1 - y)}(1 - Y)^{-y}, \qquad (Y \to 1).$$

(40)

One can go further in this analysis by computing the fluctuations $\overline{f_J(W_1, y)\, f_J(W_2, y)}$:

$$\overline{f_J(W_1, y)\, f_J(W_2, y)} = \overline{\sum_I \delta(W_I - W_1) \sum_{I'} \delta(W_{I'} - W_2)}$$

$$= \delta(W_1 - W_2)\, \overline{f_J(W_1, y)} + \overline{\sum_{I \neq I'} \delta(W_I - W_1)\, \delta(W_{I'} - W_2)}.$$

(41)

We have to compute only the term in which $I \neq I'$. This can be done in the same way as before : one computes all the moments (in W_1 and W_2) by going to replica space. The result is :

$$\overline{\sum_{I \neq I'} \delta(W_I - W_1)\,\delta(W_{I'} - W_2)} = \frac{(1 - y)\,\theta(1 - W_1 - W_2)(W_1\,W_2)^{y-2}(1 - W_1 - W_2)^{1-2y}}{\Gamma(y)\,\Gamma(y)\,\Gamma(2 - 2\,y)}. \qquad (42)$$

This result can be generalized :

$$\overline{\sum_{I_1,...,I_k}' \prod_{l=1}^{k} \delta(W_{I_l} - W_l)} = \frac{(1 - y)^{k-1}\,\Gamma(k)}{\Gamma(y)^k\,\Gamma(k - ky)}\,\theta\left(1 - \sum_{l=1}^{k} W_l\right)\left(\prod_{l=1}^{k} W_l\right)^{y-2}\left(1 - \sum_{l=1}^{k} W_l\right)^{k-ky-1} \qquad (43)$$

(The Σ' means that all the indices $I_1, ..., I_k$ are different.)

One can deduce that the probability distribution of the number of clusters having a weight larger than ε is not peaked :

$$\frac{\overline{K_\varepsilon^2} - \overline{K_\varepsilon}^2}{\overline{K_\varepsilon}^2} = \frac{(1 - y)\,\Gamma(1 - y)^2}{\Gamma(2 - 2\,y)} - 1 \sim 1 \qquad \left(\begin{array}{c} \varepsilon \ll 1 \\ y \neq 1 \end{array}\right). \qquad (44)$$

We have a rather clear picture of the structure of the ensemble of pure states : there is an infinity of states in small clusters, carrying a total weight also small, and as soon as y goes near to one, there is one isolated cluster that dominates.

---

5. Conclusions.

In the previous sections we have shown how the information contained in Parisi's R.S.B. scheme can be decoded to derive the structure of the space of states of the system. We showed that the fluctuations with J_{ik} do not disappear even in the $N \to \infty$ limit. Henceforth we gave examples about how the distribution probability for these fluctuations can be estimated by inverting the moments.

Our calculations suggest a new set of numerical simulations whereby this dependence would be explicitly checked. This would be a crucial test of our results. It would be also interesting to see whether our predictions survive in a finite dimensional system with short range forces.

We again stress the « universality » of the results of our calculations. Through the change of variable

$$q \to y, \qquad y = \int_q^{q_{Max}} P(q')\,dq'. \qquad (45)$$

We were able to eliminate all references to the particular order parameter $q(x)$, the stationary point of Pari's free energy. As a consequence the two following probability distributions are equal :

(a) the probability distribution of the Boltzmann-Gibbs factor of the pure states $\alpha(P_\alpha)$ at $T_0 \leqslant T_g$;

(b) the probability distribution of the W_I of the clusters at $T_1 < T_0$ defined through the equation :

$$W_I = \sum_{\alpha \in I} P_\alpha \quad \text{where} \quad \alpha, \alpha' \in I \quad \text{if} \quad q^{\alpha\alpha'} \geqslant q_{E.A.}(T_0).$$

The additional information that the order parameter $q(x)$ has a plateau, i.e.

$$q(x, T) = q_{E.A.}(T) \qquad x \geqslant x_0(T)$$

implies that, up to sets of zero probability, all states must have

$$q^{\alpha\alpha}(T) \leqslant q_{E.A.}(T).$$

On the other hand, the choice of scale $q = q_{E.A.}(T)$ induces a partition of the states into clusters. Each cluster must contain at most one state α, and the only states which contribute are those with $q^{\alpha\alpha} \geqslant q_{E.A.}(T)$. But we have proven that, whatever the scale of the partition into clusters $I : \sum_I W_I = 1$. Hence we obtain for all α :

$$q^{\alpha\alpha} = q_{E.A.} \quad \text{(with probability one)}. \qquad (46)$$

Then from formula (37), one finds :

$$\sum_\alpha P_\alpha^2 = \lim_{q \to q_{E.A.}} y(q). \qquad (47)$$

If the function $q(x)$ has a plateau, as it is commonly believed, the right hand side of this equation is nothing but the length of this plateau. As $P_\alpha < 1$, one can then conclude that a few states α dominate the sum $\sum_\alpha P_\alpha^2$.

Several questions remain open :

a) are there physical observables such that their infinite volume limits do not fluctuate with J ?

b) what happens when one adds corrections to the mean field approximation ? In this context we remark that it has recently been proved [15] that the free energy is self-averaging in the thermodynamic limit for short range interactions. This has also been shown in the S.K. model, and the finite volume corrections have been computed [16].

The hierarchical structure (ultrametric topology) of states was demonstrated for the J average but it obviously applies for every realization of J. The

equality of the probability distributions mentioned above suggests the following picture : when heating a spin glass from zero temperature up to T_g we go through a series of micro phase transitions characterized by the melting of two or more states into one state at the higher temperature. Furthermore we should have

$$\frac{- NF_\alpha(T + \varepsilon)}{kT} = \ln \sum_{\alpha'} \exp \frac{- NF_{\alpha'}(T - \varepsilon)}{kT} + K$$

(48)

where K is in dependent of α, and F_α is the free energy density of the state α.

Finally we point out that the ultrametric structure allows the definition of a non ergodic Brownian motion. Indeed, if in an ultrametric space a point can jump up to a distance δ in a single step, after N steps it can arrive only at distance δ. (This follows from the well known fact that two overlapping spheres coincide in an ultrametric space.) We think that this fact will have interesting consequences for the dynamical approach to equilibrium of the spin glass.

References

[1] EDWARDS, S. F. and ANDERSON, P. W., J. Phys. F **5** (1975) 965.
[2] SHERRINGTON, D. and KIRKPATRICK, S., Phys. Rev. Lett. **32** (1975) 1792.
[3] THOULESS, D. J., ANDERSON, P. W. and PALMER, R., Philos. Mag. **35** (1977) 593.
[4] DE ALMEIDA, J. R. L. and THOULESS, D. J., J. Phys. A **11** (1978) 983.
[5] PARISI, G., Phys. Rev. Lett. **43** (1979) 1754 and J. Phys. A **13** (1980), L 117, 1101, 1887.
[6] DE DOMINICIS, C., GABAY, M., GAREL, T. and ORLAND, H., J. Physique **41** (1980) 923 ;
   BRAY, A. J. and MOORE, M. A., J. Phys. C **13** (1980) L 469, L 907.
[7] SOMPOLINSKY, H. and ZIPPELIUS, A., Phys. Rev. Lett. **47** (1981) 359 ; Phys. Rev. B **25** (1982) 6860.
[8] See for example : MACKENZIE, N. D. and YOUNG, A. P., Phys. Rev. Lett. **49** (1982) 301 ;
   YOUNG, A. P., Phys. Rev. Lett. **51** (1983) 1206.
[9] For recent reviews : PARISI, G., lectures given at the 1982 Les Houches Summer School; PARISI, G., in Disordered Systems and Localization (Springer Verlag) 1981 ; TOULOUSE G., contribution to the Heidelberg Colloquium on spin glasses, Lecture Notes in Physics (Springer Verlag) **192** (1983) 2 ; DE DOMINICIS, C., contribution to the Heidelberg Colloquium on spin glasses, Lecture Notes in Physics (Springer Verlag) **192** (1983) 103.
[10] PARISI, G., Phys. Rev. Lett. **50** (1983) 1946.
[11] MEZARD, M., PARISI, G., SOURLAS, N., TOULOUSE, G., VIRASORO, M., LPTENS preprint 83/39 (1983).
[12] ANDERSON, P. W., in Les Houches 1978, Ill condensed matter (North Holland) ;
   DERRIDA, B., Phys. Rep. **67** (1980) 29.
[13] BOURBAKI, N., Espaces Vectoriels Topologiques, Paris 1966. We thank R. Rammal for pointing this to us.
[14] DE DOMINICIS, C. and YOUNG, A. P., J. Phys. A **16** (1983) 2063.
[15] KHANIN, K. M. and SINAI, Y. G., J. Stat. Phys. **20** (1979) 573.
[16] KONDOR, I., Saclay preprint (1983).
[17] MEAD, L. R. and PAPANICOLAOU, N., Washington University, St Louis preprint 1983.
[18] DE DOMINICIS, C. and KONDOR, I., Phys. Rev. B **27** (1983) 606.