ON THE PROOF OF THE PARISI FORMULA
BY GUERRA AND TALAGRAND

by Erwin BOLTHAUSEN

1. THE SHERRINGTON-KIRKPATRICK MODEL

We consider "Ising spins" $\sigma_i \in \{-1, 1\}$, $i = 1, \ldots, N$. Spin configurations will be denoted by $\sigma = (\sigma_i)_{i=1,\ldots,N} \in \Sigma_N \overset{\text{def}}{=} \{-1, 1\}^N$. As the Sherrington-Kirkpatrick model (SK-model for short) is a mean-field model, there is no geometric structure of $\{1, \ldots, N\}$ assumed.

Let further J_{ij}, $1 \leqslant i < j \leqslant N$, be i.i.d. standard Gaussian random variables, defined on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$. These random variables form the "random environment". The **Hamiltonian** is the following random function $\Sigma_N \to \mathbb{R}$:

$$(1) \qquad H_{N,\omega}(\sigma) \overset{\text{def}}{=} \frac{1}{\sqrt{N}} \sum_{1 \leqslant i < j \leqslant N} J_{ij}(\omega) \sigma_i \sigma_j, \ \omega \in \Omega,$$

We will often drop ω and N in such expressions. Remark that for any σ, this is a random variable, and indeed a centered Gaussian one. The covariances are given by:

$$(2) \qquad \mathbb{E}(H_N(\sigma) H_N(\sigma')) = \frac{1}{N} \sum_{1 \leqslant i < j \leqslant N} \sigma_i \sigma_j \sigma_i' \sigma_j' = \frac{1}{2N} \sum_{i,j=1}^N \sigma_i \sigma_j \sigma_i' \sigma_j' - \frac{1}{2}$$

$$= \frac{N}{2} \left(\frac{1}{N} \sum_{i=1}^N \sigma_i \sigma_i' \right)^2 - \frac{1}{2}.$$

The quantity in brackets is the so-called **overlap** of the two spin configurations

$$R_N(\sigma, \sigma') \overset{\text{def}}{=} \frac{1}{N} \sum_{i=1}^N \sigma_i \sigma_i'.$$

The (random) Gibbs distribution $\mathcal{G}_{N,\beta,h,\omega}$ with inverse temperature $\beta > 0$, and external field with strength $h \in \mathbb{R}$ on Σ_N is defined by

$$\mathcal{G}_{N,\beta,h,\omega}(\sigma) \overset{\text{def}}{=} \frac{1}{Z_{N,\beta,h,\omega}} \exp\left[\beta H_{N,\omega}(\sigma) + h \sum_{i=1}^N \sigma_i \right],$$

where

$$Z_{N,\beta,h,\omega} \stackrel{\text{def}}{=} \sum_{\sigma} \exp\left[\beta H_{N,\omega}(\sigma) + h \sum_{i=1}^{N} \sigma_i\right]$$

is the appropriate norming constant[1]. h here is non-random. There exist also variants where h is a random variable or where $h \sum_{i=1}^{N} \sigma_i$ is replaced by $\sum_{i=1}^{N} h_i \sigma_i$, where the h_i are random variables, e.g. $h_i = \gamma g_i + h$, $\gamma > 0$, $h \in \mathbb{R}$, and the g_i again being independent standard Gaussian random variables. It has some advantages to include such a Gaussian external field, as we will see later, but for the moment, we do not consider this possibility.

We write F_N for the finite volume free energy

$$F_N(\beta, h) \stackrel{\text{def}}{=} \frac{1}{N} \log Z_{N,\beta,h},$$

which is a random variable, defined on Ω, and

$$f_N(\beta, h) \stackrel{\text{def}}{=} \mathbb{E} F_N(\beta, h)$$

its expectation, the so-called "quenched" free energy. Sometimes, "quenched" refers to the random quantity only, but there is not much difference, as we will explain. In contrast, the so-called "annealed" free energy is obtained by taking the expectation inside the logarithm. By Jensen's inequality, f_N is dominated by the annealed free energy.

Before proceeding with the discussion of the model, we try to explain why it is interesting.

The usual models of (non-random) Ising type are defined as follows. Consider a finite set Λ, and let $\Sigma_\Lambda \stackrel{\text{def}}{=} \{-1, 1\}^\Lambda$. Let further $A = (a_{ij})_{i,j\in\Lambda}$ be a real symmetric matrix, and $\mathbf{h} = (h_i)_{i\in\Lambda}$ be a real vector. Then the Gibbs measure $\mathcal{G}_{A,\mathbf{h}}$ on Σ_Λ is defined by

$$\mathcal{G}_{\Lambda,A,\mathbf{h}}(\sigma) \stackrel{\text{def}}{=} \frac{1}{Z_{\Lambda,A,\mathbf{h}}} \exp\left[\frac{1}{2} \sum_{i,j\in\Lambda} a_{ij}\sigma_i\sigma_j + \sum_{i\in\Lambda} h_i\sigma_i\right],$$

where of course

$$Z_{\Lambda,A,\mathbf{h}} \stackrel{\text{def}}{=} \sum_{\sigma} \exp\left[\frac{1}{2} \sum_{i,j\in\Lambda} a_{ij}\sigma_i\sigma_j + \sum_{i\in\Lambda} h_i\sigma_i\right].$$

Of great importance is the (finite volume) free energy, defined by

$$F_\Lambda(A, \mathbf{h}) = \frac{1}{|\Lambda|} \log Z_{A,\mathbf{h}}.$$

The importance of this quantity is coming from the fact that most of the physical interesting quantities can be expressed through it, like mean magnetization, entropy, etc.

---

[1] In contrast to the habit in physics, we do not take a minus in front of β, and we also do not apply β to h.

The best known example is the Ising model where Λ is a finite (large) box in $\mathbb{Z}^d$, and

$$a_{ij} \stackrel{\text{def}}{=} \begin{cases} \beta & \text{if } |i - j| = 1 \\ 0 & \text{otherwise.} \end{cases}$$

Short range models are usually rather difficult to analyze, and often a qualitatively good approximation is obtained from *mean field models* where every spin interacts with any other one on equal footing. The simplest mean-field model is the *Curie-Weiss-model*. Here

$$a_{ij} \stackrel{\text{def}}{=} \beta/\left|\Lambda\right|, \ \forall i, j \in \Lambda.$$

In that case one has with $N \stackrel{\text{def}}{=} |\Lambda|$

$$\frac{1}{2} \sum_{i,j \in \Lambda} a_{ij} \sigma_i \sigma_j = \frac{\beta}{2N} \left\{ \sum_{i \in \Lambda} \sigma_i \right\}^2,$$

and anything one wants to know can be derived from the Stirling approximation, and it becomes an easy exercise in elementary probability.

Spin glasses are models where the interactions are "disordered", which typically means that they are obtained as a random object. A topic which is still very poorly understood is the case of *short range* random interactions, for instance when $\Lambda = \{-n, \ldots, n\}^d$, and the a_{ij} are independent Gaussians for $|i - j| = 1$, and 0 otherwise. This is the *Edwards-Anderson model* on which there are ongoing controversial discussions in the physics community, the more so as it is very difficult to simulate on computers with a reasonably large box and in interesting dimensions. The SK-model is a mean-field model of this random interaction type, and it was invented in [18] certainly with the aim to have a simple model with disordered interaction. The $1/\sqrt{N}$ factor is easy to understand. In the Curie-Weiss model, each spin variable interacts with the other ones with a total interaction strength of order 1. Due to the cancellations between positive and negative J's, the situation is essentially the same for the SK-model.

The model is evidently closely connected with questions probabilists have been interested in for a long time, namely maxima (or minima) of (Gaussian) random vectors. For instance, $\lim_{\beta \to \infty} (1/\beta) \log Z_{N,\beta,0}$ is simply $\max_\sigma H_N(\sigma)$, which is just the maximum of a family of correlated Gaussians with a simple covariance structure. Probabilists have developed methods to investigate such questions for a long time, e.g. Dudley, Fernique, Talagrand, and many others. It is not difficult to see that $\max_\sigma H_N(\sigma)$ is of order N and to prove that there are constants $0 < C_1 < C_2$ satisfying

$$\lim_{N \to \infty} \mathbb{P}\left(C_1 N \leqslant \max_\sigma H_N(\sigma) \leqslant C_2 N\right) = 1.$$

However, the standard probabilistic techniques cannot derive the exact constant, which the Parisi-theory does, revealing a marvelous mathematical structure behind the problem, which is still *very* poorly understood, to this day.

The Parisi-theory applies to many other problems besides to the SK-model, e.g. to the assignment problem from combinatorial optimization, to the perceptron and the Hopfield net from neural networks, to coding theory, and to others. For some of these applications, see Nishimori [14].

Back to the SK-model, the first question one typically answers is the existence of the free energy in the thermodynamical limit (here just $N \to \infty$). It is however not at all clear that the free energy

$$\lim_{N \to \infty} F_N(\beta, h)$$

exists. In principle, even if the limit exists, it could be a random variable. This possibility is however ruled out by Gaussian concentration inequalities. One says that the free energy is "self-averaging", meaning that no randomness remains in the $N \to \infty$ limit. For a proof of the following inequality, see for instance [12].

PROPOSITION 1.1. — *Let γ_n be the standard Gaussian distribution on $\mathbb{R}^n$. Let $f : \mathbb{R}^n \to \mathbb{R}$ be a Lipshitz continuous function with Lipshitz constant L. Then for any $u > 0$*

$$\gamma_n \left(f > \int f d\gamma_n + u \right) \leqslant \exp\left[-u^2 / 2L^2 \right].$$

If we apply this inequality to $F_N(\beta, h)$, regarded as a function of the standard Gaussian vector $(J_{ij})_{1 \leqslant i < j \leqslant N}$, then one gets

$$\mathbb{P}\left(\left| \frac{1}{N} \log Z_{N,\beta,h} - \frac{1}{N} \mathbb{E} \log Z_{N,\beta,h} \right| \geqslant N^{-1/4} \right) \leqslant 2 \exp\left[-\frac{N^{1/2}}{\beta^2} \right].$$

It is therefore clear that instead of investigating $\lim_{N \to \infty} F_N(\beta, h)$, one can as well investigate the non-random object $\lim_{N \to \infty} f_N(\beta, h)$. The existence of this limit had been open for a long time, until Guerra and Toninelli [11] found a very nice, and not so obvious superadditivity property:

$$(3) \qquad\qquad \mathbb{E} \log Z_{N_1 + N_2} \geqslant \mathbb{E} \log Z_{N_1} + \mathbb{E} \log Z_{N_2},$$

from which one easily derives that

$$f(\beta, h) = \lim_{N \to \infty} f_N(\beta, h)$$

exists.

For the SK-model, the inequality came somewhat as a surprise. The proof is by a simple but very clever interpolation scheme which interpolates between the $(N_1 + N_2)$-system, and the two independent smaller systems. Such interpolation schemes are at the very base of the recent progress in the understanding of the SK model, as we will see later. I am not going to prove the inequality here, but I will explain the interpolation method in another case below (Section 3).

There are many quantities in the SK-model which are *not* self-averaging in the $N \to \infty$ limit, i.e. which stay random (or at least are believed to be so). An

example is the overlap of two independent "replicas": Take σ, σ' to be two independent realizations under $\mathcal{G}_{N,\beta,h,\omega}$ for a fixed ω, and calculate $R_N(\sigma, \sigma')$, and then take the Gibbs-expectation. This is still a random variable (being a function of the interaction strengths). For small β, these random variables have a non-random limit for $N \to \infty$, but the limit stays random for large β.

The case $h = 0$ has some evident symmetry properties which make life easier, particularly in the high-temperature region. This case is however somewhat misleading. In particular, the high temperature behavior for $h \neq 0$ is far from trivial, and it is actually of crucial importance for the understanding of the low temperature region.

For $h = 0$ and small enough β, the ("quenched") free energy equals the "annealed" free energy.

THEOREM 1.2 (Aizenman-Lebowitz-Ruelle). — *For $h = 0$, and $\beta \leqslant 1$, one has*

$$(4) \qquad f(\beta) = \lim_{N \to \infty} \frac{1}{N} \log \mathbb{E} Z_{N,\beta} = \frac{\beta^2}{4} + \log 2.$$

The second equation is evident:

$$\mathbb{E} Z_{N,\beta} = \sum_{\sigma} \mathbb{E} \exp\left[\beta H_N(\sigma)\right] = \sum_{\sigma} \exp\left[\frac{\beta^2}{2} \operatorname{var}\left(H_N(\sigma)\right)\right]$$
$$= 2^N \exp\left[\frac{\beta^2}{2} \operatorname{var}\left(H_N(\sigma)\right)\right] = 2^N \exp\left[\frac{\beta^2}{2}\left(\frac{N}{2} - \frac{1}{2}\right)\right]$$

from which the claim follows. The somewhat astonishing fact is that one can interchange the expectation with the logarithm. Of course, by Jensen, one always has

$$(5) \qquad \mathbb{E} \log Z_{N,\beta} \leqslant \log \mathbb{E} Z_{N,\beta},$$

and therefore $f(\beta) \leqslant \beta^2/4 + \log 2$. We will indeed show later that $f(\beta) < \beta^2/4 + \log 2$ for $\beta > 1$. The proof of the above result is surprisingly simple and can be done by a second moment computation, proving that $\mathbb{E} Z^2 \leqslant \text{const} \times (\mathbf{E} Z)^2$ for $\beta < 1$, which is easy. Together with Gaussian isoperimetry (Proposition 1.1), this proves (4). The original proof in [1] was more complicated, but it derived also a much more detailed picture of the remaining fluctuations of $\log Z_N$.

There are other models like directed polymers for which one can prove that the quenched free energy equals the annealed one in certain regions, but typically, this is not possible by a simple second moment method in the full region where it is true. The fact that a second moment computation gives the result in the SK-model up to the correct critical value (for $h = 0$) is rather surprising.

For $h \neq 0$, "quenched=annealed" is never true, which reveals that this is a much more interesting situation, even where β is small.

2. THE REPLICA COMPUTATION AND THE PARISI FORMULA

The first evaluation of the free energy $f(\beta, h)$ was by Sherrington and Kirkpatrick [18], who applied the so-called "replica trick". This is based on the observation that for a positive number x, one has $\log x = \lim_{n \downarrow 0} (x^n - 1)/n$. If X is positive random variable, one therefore has, provided the interchange of limits with the expectation is justified,

$$\mathbb{E}\log X = \lim_{n \downarrow 0} \frac{\mathbb{E}X^n - 1}{n}.$$

As integer moments are often easier to evaluate than non-integer ones, the "trick" is to evaluate $\mathbb{E}X^n$ for integer n, then somehow extend things analytically, and perform the above limit. This is the folk tale how the replica trick works, but for the SK-model, this is not quite the way it is done. In fact, one just *starts* the computation of $\mathbb{E}Z_N^n$ assuming that n is an integer, but as soon as convenient, one gives up this illusion and lets $n \to 0$, *before* really finishing the computation. From computations of the integer moments (in the $N \to \infty$ limit), one cannot derive the value of $f(\beta, h)$. I do not repeat the computation here, as it is done in many textbooks (see e.g. [14]), and the most interesting issue starts *after* the (non-rigorous) replica computation. The variational formula one obtains from the replica trick is

$$(6) \quad f(\beta, h) = \inf_{\mathbf{q}} \lim_{n \to 0} \left\{ -\frac{\beta^2}{2n} \sum_{\alpha < \beta \leqslant n} q_{\alpha\beta}^2 + \frac{1}{n} \log \operatorname{tr}_\sigma e^{L(\mathbf{q}, \sigma)} + \frac{\beta^2}{4} \right\} + \log 2,$$

where

$$L(\mathbf{q}, \sigma) \stackrel{\text{def}}{=} \beta^2 \sum_{\alpha < \beta} q_{\alpha\beta} \sigma^\alpha \sigma^\beta + h \sum_\alpha \sigma^\alpha,$$

and where tr_σ means taking the average over an n-fold "replicated"spin $(\sigma^1, \ldots, \sigma^n)$, and the infimum taken over $n \times n$-matrices. The above expression (6) lacks a proper mathematical meaning as it is not clear what it should mean to take the $n \to 0$ limit, and what an $n \times n$-matrix is for $n \sim 0$: What one does is to make an appropriate ansatz for the minimizing q-matrix, and then take a (formal) $n \to 0$ limit. Sherrington-Kirkpatrick made short work of the problem and assumed that there is no sufficient reason why the n replicas should behave "asymmetric", and put $q_{\alpha\beta} = q$, and then take the infimum only over $q \geqslant 0$. This leads to the so-called "replica symmetric solution". The computation is easy: We have that $\sum_{\alpha < \beta} q_{\alpha\beta}^2 = n(n-1)q^2/2$, and an n cancels out. Furthermore, the $n \to 0$ limit in this part is no longer particularly demanding: $\lim_{n \to 0}(n-1) = -1$. Taking the $n \to 0$ limit in the other part is only a bit more tricky. A simple computation yields

$$\frac{1}{n} \log \operatorname{tr}_\sigma e^{L(\mathbf{q}, \sigma)} = -q\beta^2/2 + \frac{1}{n} \log E\left(1 + n \log \cosh\left(g\beta\sqrt{q} + h\right)\right) + o(1),$$

where g is a standard Gaussian variable, and E the expectation for this Gaussian. Expanding in n and letting $n \to 0$, one obtains for this $-q\beta^2/2 + E \log \cosh\left(g\beta\sqrt{q} + h\right)$.

Summing things together, we get for this "replica symmetric solution"

$$(7) \quad RS\left(\beta, h\right) \stackrel{\text{def}}{=} \inf_{q} \left\{ \frac{\beta^2}{4} \left(1 - q\right)^2 + \int \log \cosh\left(x\beta\sqrt{q} + h\right) \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-x^2/2} dx \right\} + \log 2.$$

For later use, we find the minimizing q. A simple computation leads to the following fixed point equation for the optimal q :

$$(8) \qquad\qquad q = \int \tanh^2\left(h + \beta\sqrt{q}x\right) \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-x^2/2} dx.$$

For $h = 0$, $q = 0$ is always a solution, and for $\beta \leqslant 1$, it is the only one, as one can easily check. Therefore $RS\left(\beta, 0\right) = \beta^2/4$ for $\beta \leqslant 1$. For $\beta > 1$, there is however at least one other solution of this fixed point equation, which follows easily by calculating the derivative of the expression on the right hand side at $q = 0$. In fact, there is just *one* other solution $q\left(\beta\right) > 0$ which gives the minimum, and therefore $RS\left(\beta, 0\right) < \beta^2/4 + \log 2$ for $\beta > 1$. As we will prove $f\left(\beta, h\right) \leqslant RS\left(\beta, h\right)$ for all β, h, this proves that value in Theorem 1.2 is never correct for $\beta > 1$.

For $h > 0$, the equation (8) does have a unique positive solution:

LEMMA 2.1. — *Let $\beta, h > 0$ be arbitrary. Then* (8) *has a unique solution $q\left(\beta, h\right)$.*

The proof is due to Guerra and is short but a bit tricky. Talagrand has it in his book ([19]).

The main question is whether $f\left(\beta, h\right) = RS\left(\beta, h\right)$. It is certainly correct for $h = 0$ and $\beta \leqslant 1$, as we have seen before. However, for $\beta > 1$, it is not correct. This is far from trivial to see. It will however turn out that for $h \neq 0$, the formula is correct again for small β, but not for large ones. Even the small β case is highly non-trivial. That the solution cannot be correct for large β was already realized by Sherrington and Kirkpatrick by calculating the entropy, which has to be positive, but it can also be computed from the free energy, and if one uses RS, it becomes negative for large β. So already Sherrington and Kirkpatrick concluded that their own solution is not correct for large β.

The RS-solution is supposed to be correct for β up to the celebrated AT-line (de Almayda-Thouless line [3]), i.e. for β satisfying

$$(9) \qquad\qquad \beta^2 \int \frac{1}{\cosh^4\left(h + \beta\sqrt{q\left(\beta, h\right)}x\right)} \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-x^2/2} dx < 1,$$

but this is not yet proved; it is now simply a nasty analytical problem, as the Parisi-formula for $f\left(\beta, h\right)$ is proved for the whole temperature region. (The above condition comes up through a local stability computation.)

In order to overcome the problem with the replica symmetric solution for large β, there had been various proposals for a different ansatz for the minimizing problem in (6), no longer assuming that all the $q_{\alpha\beta}$ are equal. This is the famous "replica symmetry breaking". A particular ansatz for this is due to Parisi [15]. The ansatz

makes a *very special* assumption on the matrix $Q = (q_{\alpha\beta})$, namely that it has a kind of hierarchical organization. The question then remained if there could not be a better choice not satisfying the Parisi-ansatz. A justification of the Parisi-ansatz before Talagrand's proof was the proof that it is in a sense locally stable, by computing Hessians, and that it was the only one found having this property, but the really convincing argument was that the outcome had interesting consequences also outside the "replica formulation". Very nice explanations of these issues can be found in [14]. Here just a cursory explanation of what is going on.

The replica symmetric ansatz fixes the matrix Q to be of the following form

$$Q = \begin{pmatrix} 0 & q & q & \cdots & \cdots & q \\ & 0 & q & \cdots & \cdots & q \\ & & 0 & q & \cdots & q \\ & & & \ddots & & \vdots \\ & & & & 0 & q \\ & & & & & 0 \end{pmatrix}.$$

In the Parisi ansatz, one uses more complicated matrices. There are a number of levels. In the end, this number has to go to infinity, but let us first look at the simplest case, the case with one level of replica symmetry breaking. Here one takes a matrix of the form:

$$\begin{pmatrix} \begin{pmatrix} 0 & q_2 & q_2 \\ & 0 & q_2 \\ & & 0 \end{pmatrix} & \begin{pmatrix} q_1 & q_1 & q_1 \\ q_1 & q_1 & q_1 \\ q_1 & q_1 & q_1 \end{pmatrix} \\ & \begin{pmatrix} 0 & q_2 & q_2 \\ & 0 & q_2 \\ & & 0 \end{pmatrix} \end{pmatrix}$$

The rule is that one divides the $n \times n$-matrix by choosing $n_1 \leqslant n$ such that n/n_1 is an integer, and then one divides the matrix into $(n/n_1)^2$ submatrices of the form $n_1 \times n_1$. The diagonal blocks get q_2 above the diagonal, and the off-diagonal blocks all get q_1. In the above example, one has $n = 6$ and $n_1 = 3$. Then one does the computation analogously as above, keeps $m_1 \stackrel{\text{def}}{=} n_1/n$ fixed, and lets formally $n \to 0$. This leads to a variational problem. One can check that one can always assume that $0 \leqslant q_1 \leqslant q_2 \leqslant 1$. For β small it turns out, that nothing new is achieved: The optimal choice for the q's is $q_1 = q_2$, but for large β, some $q_1 < q_2$ give a lower value. This is the "one level symmetry breaking", but one can proceed by dividing the q_2 blocks in a similar fashion, which leads to a "two level symmetry breaking", and one can go on in this way with an arbitrary number of symmetry breakings. The calculations are somewhat lengthy but not difficult. Here is the outcome:

Let $K \in \mathbb{N}$ (the number of symmetry breakings), and then we choose parameters

(10) $$0 = m_0 < m_1 < \cdots < m_{K-1} < m_K = 1,$$

(11) $$0 = q_0 \leqslant q_1 < \cdots < q_K < q_{K+1} = 1.$$

For $i = 0, \ldots, K$ let g_i be Gaussian with variance $\beta^2 (q_{i+1} - q_i)$, and set $Y_{K+1} \overset{\text{def}}{=} \cosh\left(h + \sum_{i=0}^{K} g_i\right)$. Then one defines

$$(12) \qquad Y_K \overset{\text{def}}{=} \left[E_K\left(Y_{K+1}^{m_K}\right)\right]^{1/m_K} = E_K\left(Y_{K+1}\right),$$

where E_K means that one integrates out g_K, so that Y_K still depends on $g_0, \ldots, g_{K-1}$. Then one defines

$$Y_{K-1} \overset{\text{def}}{=} \left[E_{K-1}\left(Y_K^{m_{K-1}}\right)\right]^{1/m_{K-1}}$$

and so on, until one gets Y_1. Y_1 is still a random variable as it depends on g_0. Remark however, that in case $q_1 = 0$ which we do not exclude, there is no randomness left. In any case, we set

$$(13) \qquad \mathcal{P}_K\left(m, q; \beta, h\right) \overset{\text{def}}{=} E \log Y_1 - \frac{\beta^2}{4} \sum_{i=1}^{K} m_i\left(q_{i+1}^2 - q_i^2\right) + \log 2.$$

Then $\inf_{m,q} \mathcal{P}_K\left(m, q\right)$ is the value one obtains by optimizing (6) with the Parisi-ansatz at K levels of replica symmetry breaking, and therefore, believing that first of all the replica trick works, and secondly that the ansatz of Parisi finds the minimum, we get

$$(14) \qquad f\left(\beta, h\right) = \inf_{K,m,q} \mathcal{P}_K\left(m, q\right) = \lim_{K \to \infty} \inf_{m,q} \mathcal{P}_K\left(m, q\right).$$

THEOREM 2.2 (Parisi Formula). — *The Parisi-formula (14) is correct for all β, h.*

The proof is due to Guerra [9] who proved the upper bound, and Talagrand [20] who then finished the proof.

In the case of the SK-model, either one has $K = 1$, which gives the true value in the region where the replica-symmetric solution is correct, or one has to take $K \to \infty$, and therefore one has "replica symmetry breaking" at infinitely many levels. There are other models, with the minimum assumed at one level of symmetry breaking, i.e. $K = 2$. One can artificially cook up cases with arbitrary K, but $K = 1, 2, \infty$ seem to be the only ones coming up "naturally". In the case $K = \infty$, one can phrase the limit $K \to \infty$ directly as a variational problem involving continuous functions $q \to x\left(q\right)$. The finite K case then corresponds to taking step functions $x\left(q\right) = m_i$ for $q \in [q_i, q_{i+1})$.

Here is an outline of what the physicists believe to be the picture behind the RS-solution ($K = 1$), and the replica symmetry breaking ($K > 1$). This picture emerged partly from another non-rigorous approach, the so-called "cavity method" which led to the same formula for the free energy, and gave a clearer picture about the Gibbs distribution (see [13]).

The region where the RS-solution (7) is valid is characterized by the property that the σ_i under the Gibbs measure are still "fairly independent". The $h = 0$ case is simple because, due to symmetry, the expectation under the Gibbs measure is 0. For $h \neq 0$, the expectation of σ_i under the Gibbs measure $\mathcal{G}_{N,\beta,h,\omega}$ is $m_i \overset{\text{def}}{=} \mathcal{G}\left(\sigma_i\right)$ which satisfies

$\mathbb{E}m_i^2 = q(\beta, h)$, q being the solution of (8), equality in the $N \to \infty$ limit. The m_i are themselves approximately independent under the measure $\mathbb{P}$. One therefore has the following picture (for large N): The randomness of the disorder (i.e. the J_{ij}) produces the nearly i.i.d. random variables m_i, and given the disorder, the Gibbs measure has approximately independent spin variables σ_i with mean m_i. The property that the σ_i are approximately independent is reflected in the physics community saying that there is just "one pure state".

Given this picture, $q(\beta, h)$ has a precise mathematical interpretation in terms of the Gibbs measure. It is the almost sure limit (as $N \to \infty$) of the overlaps of two independent realizations of the spin variables:

$$(15) \qquad R_N(\sigma, \sigma') = \frac{1}{N} \sum_{i=1}^{N} \sigma_i \sigma_i' \simeq \frac{1}{N} \sum_{i=1}^{N} m_i^2 \simeq q(\beta, h)$$

by the law of large numbers. The precise statement is as follows: Let $\nu_N^{(2)}$ be the measure on $\Sigma_N \times \Sigma_N$ defined by

$$(16) \qquad \nu_N^{(2)}(\sigma, \sigma') \stackrel{\text{def}}{=} \int \mathbb{P}(d\omega)\, \mathcal{G}_{N,\omega}^{\otimes 2}(\sigma, \sigma'),$$

where $\mathcal{G}^{\otimes 2}$ denotes the twofold product Gibbs measure. Then for small enough β

$$\lim_{N \to \infty} \nu_N^{(2)} \left(|R_N(\sigma, \sigma') - q(\beta, h)| \geqslant \varepsilon \right) = 0, \ \forall \varepsilon > 0.$$

This means that the overlap of independent replicas is self-averaging. The high-temperature regime is now mathematically very well understood, mainly through the work of Michel Talagrand (see [19], Chapter 2).

In the low temperature regime things become much more complicated. First of all, the RS-solution is no longer correct, but this is only one aspect. The overlaps are no longer self-averaging but stay random. The Gibbs distribution splits into a "countable number of pure states", a statement made in the physics literature which is difficult to make mathematically precise. Essentially the "pure states" under the Gibbs-distribution should be organized in a hierarchical way. This hierarchy somehow reflects the hierarchical ansatz in the Parisi-matrices above. Nothing of this has been proved mathematically, and probably not all statements made in the physics literature should be taken (mathematically) too literally. One important aspect is "ultrametricity", which has the following precise mathematical meaning: Take three independent realizations $\sigma, \sigma', \sigma''$ under the Gibbs measure. Define $\nu_N^{(3)}$ for the three replicas, similarly as $\nu_N^{(2)}$ defined above. The claim is that for any $\varepsilon > 0$

$$(17) \qquad \lim_{N \to \infty} \nu_N^{(3)} \left[R_N(\sigma, \sigma'') \geqslant \min(R_N(\sigma, \sigma'), R_N(\sigma', \sigma'')) - \varepsilon \right] = 0.$$

There is no proof of this. Together with the so-called Ghirlanda-Guerra identities, which I am not discussing here, this would essentially characterize the model completely.

Despite the recent progress on the SK-model in the full temperature regime, which is explained below, much of the above picture is mathematically not understood. In a case with one level of symmetry breaking only, the so-called p-spin SK model, there are results by Michel Talagrand, which confirm the physicists predictions, if properly formulated (see [19] Chapter 6).

3. GUERRA'S INTERPOLATION SCHEME: THE REPLICA SYMMETRIC BOUND

Much of the recent progress on the SK-model is based on a very clever argument invented by Guerra, which leads to bounds on the free energy. These bounds are obtained by interpolating continuously between the system one is interested in, and a much easier one.

I will explain this in the simplest case, where one proves that the replica symmetric solution is a strict bound for the free energy, for all N and in the full region of parameters.

THEOREM 3.1. — *For all $\beta > 0$, $h \in \mathbb{R}$, and $N \in \mathbb{N}$ one has*

$$f_N(\beta, h) \leqslant RS(\beta, h),$$

where $RS(\beta, h)$ is defined by (7).

Proof. — The proof is by interpolation. Let for an arbitrary number $q \geqslant 0$, and $t \in [0, 1]$

$$(18) \qquad H(t, \sigma) \stackrel{\text{def}}{=} \sqrt{\frac{t}{N}} \sum_{1 \leqslant i < j \leqslant N} J_{ij} \sigma_i \sigma_j + \sqrt{1-t} \sum_{i=1}^{N} \sqrt{q} g_i \sigma_i,$$

$$\Phi(t, \sigma; \beta, h) \stackrel{\text{def}}{=} \beta H(t, \sigma) + h \sum_{i=1}^{N} \sigma_i,$$

where g_i is a set of standard Gaussian variables, independent of the J's. The basic idea of this interpolation is to relate the Hamiltonian we are interested in with a much simpler one with independent σ_i, which however have the right overlap structure: For $t = 0$, we have, conditioned on the g_i, independent spins with mean $\tanh(h + \beta\sqrt{q}g_i)$. Therefore, if we take two independent realizations σ, σ' (still conditioned on the g_i), we get

$$\frac{1}{N} \sum_{i=1}^{N} \sigma_i \sigma_i' \simeq \frac{1}{N} \sum_{i=1}^{N} \tanh^2(h + \beta\sqrt{q}g_i)$$

$$\simeq \int \tanh^2(h + \beta\sqrt{q}x) \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx$$

which equals q, if we take for q the solution of (8). The clever idea by Guerra is that one can control what happens along the path from $t = 0$ to $t = 1$. For the moment,

we have not even to assume that q is the right one, and we can just take it arbitrary $\geqslant 0$. We again define the partition function

$$\xi\left(t\right) \overset{\text{def}}{=} \sum_\sigma e^{\Phi\left(t,\sigma;\beta,h\right)},$$

and we write $\mathcal{G}_t\left(\sigma\right)$ for the corresponding Gibbs measure. Let

(19) $$\phi\left(t\right) \overset{\text{def}}{=} \frac{1}{N}\mathbb{E}\log\xi\left(t\right).$$

Remark that

(20) $$\phi\left(0\right) = \int \log\cosh\left(\beta\sqrt{q}x + h\right)\frac{1}{\sqrt{2\pi}}e^{-x^2/2}dx + \log 2$$

$$\phi\left(1\right) = f_N\left(\beta, h\right).$$

We compute the derivative of $\phi\left(t\right)$ with respect to t.

$$\phi'\left(t\right) =$$

$$\frac{\beta}{2\sqrt{t}N^{3/2}}\sum_\sigma \sigma_i\sigma_j\sum_{i<j}\mathbb{E}J_{ij}\frac{\exp\left[\Phi\left(t,\sigma\right)\right]}{\xi\left(t\right)} - \frac{\sqrt{q}}{2\sqrt{1-t}}\sum_\sigma \sigma_i\sum_i \mathbb{E}g_i\frac{\exp\left[\Phi\left(t,\sigma\right)\right]}{\xi\left(t\right)}$$

$$= A\left(t\right) - B\left(t\right), \text{ say.}$$

We use Gaussian partial integration to remove the J_{ij}, g_i: If Z is a standard Gaussian, and $F: \mathbb{R} \to \mathbb{R}$ any decent function, then $\mathbb{E}\left(ZF\left(Z\right)\right) = \mathbb{E}F'\left(Z\right)$. Applying this first to $Z = J_{ij}$, we get

$$\mathbb{E}J_{ij}\frac{\exp\left[\Phi\left(t,\sigma\right)\right]}{\xi\left(t\right)} = \beta\sqrt{\frac{t}{N}}\left\{\mathbb{E}\frac{\sigma_i\sigma_j\exp\left[\Phi\left(t,\sigma\right)\right]}{\xi\left(t\right)} - \sum_\tau \mathbb{E}\frac{\exp\left[\Phi\left(t,\sigma\right) + \Phi\left(t,\tau\right)\right]}{\xi\left(t\right)^2}\tau_i\tau_j\right\}.$$

After some elementary manipulations, we obtain

$$A\left(t\right) = \frac{\beta^2}{4}\mathbb{E}\sum_\sigma \frac{\exp\left[\Phi\left(t,\sigma\right)\right]}{\xi\left(t\right)} - \frac{\beta^2}{4N^2}\mathbb{E}\sum_{\sigma,\tau}\sum_{i,j}\sigma_i\sigma_j\tau_i\tau_j\frac{\exp\left[\Phi\left(t,\sigma\right) + \Phi\left(t,\tau\right)\right]}{\xi\left(t\right)^2}.$$

The first summand is simply $\beta^2/4$. In the second, we remark that $N^{-2}\sum_{i,j}\sigma_i\sigma_j\tau_i\tau_j$ is $R_N\left(\sigma,\tau\right)^2$. The summation over σ,τ is therefore just the expectation of $R_N\left(\sigma,\tau\right)^2$ under the product Gibbs measure at parameter t. We therefore have

$$A\left(t\right) = \frac{\beta^2}{4}\left(1 - \nu_{N,t}^{(2)}\left(R_N^2\left(\sigma,\tau\right)\right)\right),$$

where $\nu_{N,t}^{(2)}$ is the measure (16), but with the interpolated Hamiltonian at t. (We sometimes do not clearly distinguish between a measure and taking the expectation with respect to it.) Similarly, we get

$$B\left(t\right) = \frac{\beta^2 q}{2}\left(1 - \nu_{N,t}^{(2)}\left(R_N\left(\sigma,\tau\right)\right)\right),$$

and therefore

$$\frac{d\phi}{dt} = \frac{\beta^2}{4} \nu_{N,t}^{(2)} \left(1 - R_N^2\left(\sigma, \tau\right) - 2q\left(1 - R_N\left(\sigma, \tau\right)\right)\right)$$

(21)
$$= \frac{\beta^2}{4} \left\{ (1-q)^2 - \nu_{N,t}^{(2)} \left(\left(R_N\left(\sigma, \tau\right) - q\right)^2\right) \right\}$$

$$\leqslant \frac{\beta^2}{4} \left(1 - q\right)^2.$$

$$\phi\left(1\right) - \phi\left(0\right) \leqslant \frac{\beta^2}{4} \left(1 - q\right)^2.$$

We therefore get from (20) for *any* N and *any* $q \geqslant 0$

$$f_N\left(\beta, h\right) \leqslant \frac{\beta^2}{4} \left(1-q\right)^2 + \int \log \cosh\left(h + \beta\sqrt{q}x\right) \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx + \log 2.$$

Taking the infimum over q of the right hand side implies the theorem.   □

The proof does not only give the desired result, but gives also an expression of the difference, namely

(22)
$$RS\left(\beta, h\right) - f_N\left(\beta, h\right) = \frac{\beta^2}{4} \int_0^1 \nu_{N,t}^{(2)} \left(\left(R_N\left(\sigma, \tau\right) - q\right)^2\right) dt.$$

In order to prove that $f\left(\beta, h\right) = RS\left(\beta, h\right)$, one therefore "only" has to show that for the optimal q (i.e. the one given by (8)), one has $R_N\left(\sigma, \tau\right) \simeq q$ with large $\nu_{N,t}^{(2)}$-probability, at least in the t-average. This is not true for large β, but it is true for small β, as we will discuss later.

It should also be emphasized that the Gibbs measure $\mathcal{G}_t$ structurally is not much different from the original measure. In fact it is of the form

(23)
$$\frac{1}{Z} \exp\left[\beta' H_N\left(\sigma\right) + \gamma \sum_i g_i \sigma_i + h \sum_i \sigma_i\right],$$

where the g_i are new independent Gaussians, and γ is an additional parameter.

4. DERRIDA'S RANDOM ENERGY MODEL AND RUELLE'S CASCADES

This section is a deviation from the SK model and introduces a class of simple models invented by Derrida which in a certain vague sense are supposed to be "universal attractors" of much more complicated models like SK. On a mathematical level this is very far from being understood. Nonetheless, computations on Derrida's model (in Ruelle's asymptotic version) can be used to give a transparent proof of Guerra's bound of the free energy by the Parisi expression. We will explain this in the next section. This application is not the only reason that we spend some space explaining Derrida's models. At the core lies one of the main open problems in spin glass theory,

namely the claim of universal ultrametricity. Derrida's models are ultrametric by construction, and as Parisi's formula is closely connected with them, as we will see, it appears natural to conclude, that ultrametricity should hold. However, there seems to be no promising idea around how to prove that.

The basic difficulty of the SK-model is coming from the fact that the "energies" $H(\sigma)$ are correlated random variables. Derrida [7] realized that already something interesting is happening assuming that they are just independent random variables having (about) the correct variances. Therefore, we consider independent Gaussian random variables $(H_N(\alpha))_{\alpha \in \Sigma_N}$. Σ_N does not need to have any structure here, so we just let $1 \leqslant \alpha \leqslant 2^N$. In order to match the variance of the Hamiltonian in the SK-case, we should take $N/2$, but for convenience, we take variance N, and define the partition function, the free energy, and the Gibbs measure in the usual way

$$(24) \qquad Z_N(\beta) \overset{\text{def}}{=} \sum_\alpha \mathrm{e}^{\beta H_N(\alpha)}$$

$$(25) \qquad f(\beta) \overset{\text{def}}{=} \lim_{N \to \infty} \frac{1}{N} \log Z_N(\beta), \ \mathcal{G}_{N,\beta}(\alpha) \overset{\text{def}}{=} Z_N(\beta)^{-1} \mathrm{e}^{\beta H_N(\alpha)}$$

It is easy to see that the free energy is self-averaging, so that $f(\beta)$ is also the limit of the expectations, and therefore non-random. The Gibbs measure is again a *random* probability distribution on Σ_N, as the $H(\sigma)$ are random variables. The limiting free energy is not difficult to determine and is given by

$$(26) \qquad f(\beta) = \begin{cases} \beta^2/2 + \log 2 & \text{for } \beta \leqslant \beta_{\mathrm{cr}} = \sqrt{2\log 2} \\ \sqrt{2\log 2}\,\beta & \text{for } \beta \geqslant \beta_{\mathrm{cr}} = \sqrt{2\log 2}. \end{cases}$$

Much more interesting is the Gibbs distribution in the $N \to \infty$ limit. This can be derived from a well known probabilistic result on extreme values of i.i.d. Gaussian random variables. There exists a sequence $a_N \uparrow \infty$ (the exact value is of no importance, they are of order $\sqrt{2\log 2}N$) such that the random measure

$$\sum_\alpha \delta_{H_N(\alpha) - a_N}$$

converges weakly to a Poisson point process on $\mathbb{R}$ with intensity measure

$$\sqrt{2\log 2}\,\mathrm{e}^{-\sqrt{2\log 2}\,t}dt.$$

We write $\mathrm{PPP}(t \to a\mathrm{e}^{-at})$ for a Poisson point process with such a density. Remark that there is a largest point, simply because $a\mathrm{e}^{-at}$ is integrable at $+\infty$. In contrast, there is no smallest point, and the points are lying dense and denser the further one goes down the negative real axis. We can represent such a point process as $\sum_{i=0}^\infty \delta_{\xi_i}$, where $\xi_1 > \xi_2 > \cdots$ are real-valued random variables. We also just talk of the "point process (ξ_i)", meaning $\sum_{i=0}^\infty \delta_{\xi_i}$, but we tacitly always assume that the points are ordered downwards. (The point processes we consider will always have a largest point.) We are not really interested in the energy levels, but rather in the Gibbs weights, which are given as $\exp[\beta H_N(\alpha)]$. As we are interested only in the relative

weights, we can as well consider $\exp\left[\beta\left(H_N\left(\alpha\right) - a_N\right)\right]$. Of course, we could normalize the weights to a (random) probability distribution, but it turns out to better be not too hasty with that, and to consider first the limiting point process of these points which evidently converges in distribution to the transformation of the point process $\mathrm{PPP}\left(t \to \sqrt{2\log 2}\mathrm{e}^{-\sqrt{2\log 2}t}\right)$ obtained by applying the mapping $\xi \to \eta \overset{\mathrm{def}}{=} \mathrm{e}^{\beta\xi}$ to the points. This is a $\mathrm{PPP}\left(t \to xt^{-x-1}\right)$, with the parameter $x = x\left(\beta\right) = \sqrt{2\log 2}/\beta$, i.e. we have

$$(27) \qquad\qquad \sum_\alpha \delta_{\exp[\beta(H_N(\alpha)-a_N)]} \to \mathrm{PPP}\left(t \to x\left(\beta\right)t^{-x(\beta)-1}\right)$$

in distribution. The $\mathrm{PPP}\left(t \to xt^{-x-1}\right)$ (which of course are point processes on the positive real line) have a number of remarkable properties which are absolutely crucial for their appearance in the Parisi picture.

PROPOSITION 4.1. — *Assume (η_i) are the points of a $\mathrm{PPP}\left(t \to xt^{-x-1}\right)$, and let $Y_1, Y_2, \ldots$ be i.i.d. positive real random variables satisfying $EY^x < \infty$, being also independent of the point process. Set $\psi\left(x\right) \overset{\mathrm{def}}{=} \left(EY^x\right)^{1/x}$. Then $\sum_i \delta_{\psi(x)^{-1}Y_i\eta_i}$ is also a $\mathrm{PPP}\left(t \to xt^{-x-1}\right)$. In plain words, multiplying the points η_i by Y_i amounts to the same (when regarded as a point process) than multiplying the points with the constant $\psi\left(x\right)$. (We will see that this property is at the core of the Parisi formula).*

The proof is an easy exercise and I do not give it here. Note that the properties crucially depend on the special form of the intensity measure of the Poisson process. The property actually characterizes $\mathrm{PPP}\left(t \to xt^{-x-1}\right)$ as has recently been shown by Ruzmaikina and Aizenman [17].

In order to describe the limiting Gibbs distribution, one still has to apply a normalization, and it is plausible that we can interchange the normalizing operation with taking the limit in (27), i.e. we would like to conclude that the point process $\sum_\alpha \delta_{\mathcal{G}_{N,\beta}(\alpha)}$ converges weakly to the proper normalization of $\mathrm{PPP}\left(t \to xt^{-x-1}\right)$. There is however a difficulty. Let $\eta_1 > \eta_2 > \cdots > 0$ be the ordered (random) points of a $\mathrm{PPP}\left(t \to xt^{-x-1}\right)$. We would like to apply a normalization procedure by normalizing the weights η_i, setting

$$\overline{\eta}_i = \eta_i \Big/ \sum_j \eta_j \ .$$

This we can only do if the sum converges. One easily proves the following statement for the points of a $\mathrm{PPP}\left(t \to xt^{-x-1}\right)$

$$\sum_j \eta_j < \infty \text{ a.s.} \Longleftrightarrow x < 1.$$

If $x < 1$, we can therefore define the normalization procedure, obtaining the point process $\sum_i \delta_{\overline{\eta}_i}$ which we denote by $\mathcal{N}\left(\mathrm{PPP}\left(t \to xt^{-x-1}\right)\right)$. This is no longer a Poisson point process as is evident from the fact that the points sum up to 1. The following

result is plausible, but its proof still requires some work as the above normalization is not a continuous operation.

PROPOSITION 4.2. — *Assume $\beta > \sqrt{2\log 2}$. Then $\sum_\alpha \delta_{\mathcal{G}_{N,\beta}(\alpha)}$ converges weakly to $\mathcal{N}\left(PPP\left(t \to xt^{-x-1}\right)\right)$, where $x(\beta) = \sqrt{2\log 2}/\beta$.*

For a proof (in a more general setting), see [6] or [19], Chap 1. The result states that for low temperature, there are configurations α which have Gibbs weight of order 1 in the $N \to \infty$ limit, but these Gibbs weights stay random. So the limiting Gibbs distribution is not "self-averaging". Furthermore, there is a "countable" number of such configurations in the limit. More precisely: For any $\varepsilon > 0$ there exists a number $K(\varepsilon)$ such that the total Gibbs weight of the $K(\varepsilon)$ configurations with the largest weight is $\geqslant 1 - \varepsilon$, with $\mathbb{P}$-probability larger than $1 - \varepsilon$, and that uniformly in N. Furthermore $K(\varepsilon)$ has to go to ∞ for $\varepsilon \to 0$. The situation is easy to understand: For $\beta > \sqrt{2\log 2}$, the Gibbs weights concentrate on the configurations α for which the energies $H_N(\alpha)$ are maximal or close to the maximum. These energies (near the maximum) are spaced at a distance of order 1: The second largest is below the largest by a random distance which stays stochastically of order 1 in the $N \to \infty$ limit. The maximum energy is approximately at $\sqrt{2\log 2}N$, with some correction of order $\log N$.

If $\beta < \sqrt{2\log 2}$, the situation is completely different. The main contribution comes from energies approximately at a level aN, where $a < \sqrt{2\log 2}$ (actually $a = \beta$, by accident). At this level, the energies are lying tightly, with exponentially small typical spacings. Therefore, the maximum Gibbs weight for $\beta < \sqrt{2\log 2}$ is exponentially small in N, and in order to catch a macroscopic weight one has to sum over exponentially many individual configurations. Therefore, in the limit, "uncountable" many configurations contribute to the Gibbs measure.

A prediction of the Parisi theory is that the point process described above is a universal object in spin glass theory and appears as the distribution of the "pure states" in essentially all systems exhibiting "spin glass behavior", in particular in the SK model. It is difficult to give the notion of a "pure state", which is often appearing in the physics literature, a precise mathematical sense. This has been achieved only for the p-spin SK model which has a simpler structure than the regular SK model, by Talagrand (see [19], Chap 6).

The above model is called the "random energy model", REM for short. It is certainly an oversimplification, and Derrida [8] a bit later introduced a model which has hierarchical organized correlations. Shortly afterwards, Ruelle [16] in an attempt to get a clearer mathematical picture of the physicists predictions in spin glass theory introduced a point process version, which is the limiting object of Derrida's model. This model was then further investigated in [5] and elsewhere. These models are now called "generalized random energy models", or GREM for short. In contrast to the random energy model, they have a non-trivial notion of "overlaps".

Here is Derrida's version. We consider a tree with 2^N leaves and K branching levels, where K stays fixed (for the moment), and we let then $N \to \infty$. We write the elements of the tree as $\alpha = (\alpha_1, \ldots, \alpha_K)$ where $\alpha_i \in \{1, \ldots, 2^{N/K}\}$. For convenience, we always assume that N/K is an integer. We again write Σ_N for the collection of such α's. Evidently, we have 2^N elements in Σ_N. For $i \leqslant K$, we identify $(\alpha_1, \ldots, \alpha_i)$ with the "bond" from node $(\alpha_1, \ldots, \alpha_{i-1})$ to $(\alpha_1, \ldots, \alpha_i)$. To the bonds of the tree, we attach Gaussian random variables with variances proportional to N, but depending on the level inside the tree. We choose parameters $\sigma_1^2, \ldots, \sigma_K^2 > 0$ with $\sum_i \sigma_i^2 = 1$, and for $i \leqslant K$, $(\alpha_1, \ldots, \alpha_i)$ as above, we choose Gaussian random variables $X_{\alpha_1, \ldots, \alpha_i}^{(i)}$ which have variance $\sigma_i^2 N$. All these variables are independent. Then we define the random Hamiltonian

$$(28) \qquad H_N(\alpha) \overset{\text{def}}{=} \sum_{i=1}^{K} X_{\alpha_1, \ldots, \alpha_i}^{(i)},$$

i.e. for any leaf of the tree we sum the independent Gaussian variables attached to the bonds along the path from the root to this leaf.

The $H_N(\alpha)$ are evidently Gaussians with variance N, like in the REM case, but there are now correlations. Defining for $\alpha, \alpha' \in \Sigma_N$ the overlap

$$R(\alpha, \alpha') \overset{\text{def}}{=} \max\{i : (\alpha_1, \ldots, \alpha_i) = (\alpha_1', \ldots, \alpha_i')\}$$

one has

$$\mathbb{E} H_N(\alpha) H_N(\alpha') = N \sum_{i=1}^{R(\alpha, \alpha')} \sigma_i^2.$$

We impose the following condition

$$(29) \qquad \sigma_1^2 > \sigma_2^2 > \cdots > \sigma_K^2 > 0.$$

If this is not satisfied, just some levels disappear in the $N \to \infty$ limit, so we can as well make this assumption[2]. The partition function and the Gibbs weights are defined as before in (24), (25). Despite its simplicity the model has a number of surprising properties which can be summarized as follows:

– The limiting Gibbs measure is always that of random energy model if "properly interpreted". This is true despite the fact that the limiting point process of the energy levels is *not* that of a random energy model, but by normalizing, the "not-REM" part cancels out.

– The model keeps a non-trivial overlap structure of the configurations, even in the $N \to \infty$ limit, which also stays random (i.e. non-selfaveraging). Surprisingly however, the overlap structure becomes stochastically independent of the Gibbs-weights in the limit (which is not true for finite N). The overlap structure has a simple Markovian structure as a coalescent with explicitly defined transition probabilities, a fact worked out in [5].

---

[2] There is a delicate issue in case of equalities in (29) which we do not address here.

– The GREM overlap structure is of direct relevance for the Sherrington-Kirkpatrick model. It turns out that the Parisi-formula for the free energy of the SK shows up through a "one spin perturbation" of GREMs. This aspect can be used to prove that the Parisi-expression is an upper bound for $f_{SK}(\beta, h)$. This will be explained in the next section.

The free energy of the GREM can be computed (see [8]), but it is not of great relevance for the aspects discussed here.

We now describe Ruelle's limiting point process versions. One can perform the $N \to \infty$ separately on each level. On the first level, one simply has $2^{N/K}$ independent Gaussians with variance $\sigma_1^2 N$. After subtracting a suitable constant from these "energies", one arrives in the limit $N \to \infty$, using the same argument as for the REM, at a PPP $(t \to a e^{-at})$, with $a = \sqrt{2 \log 2} / \sqrt{K \sigma_1^2}$. As we are interested in the Gibbs distribution we can as well consider the point process where one maps the points ξ of the above point process to $\eta = \exp[\beta \xi] \in \mathbb{R}^+$. This leads to a PPP $(t \to x_1 t^{-x_1 - 1})$ on the positive real line where $x_1 = \sqrt{2 \log 2} / \sqrt{K} \beta \sigma_1$.

In this way, one proceeds along the tree and arrives at the following object. Set

$$x_i(\beta) \overset{\text{def}}{=} \frac{\sqrt{2 \log 2}}{\sqrt{K} \beta \sigma_i}.$$

The point process $\text{PPP}(t \to x_1 t^{-x_1 - 1})$ for the first level consists of countably many random points which we can assume to be ordered downwards. Call them $\eta_1^1 > \eta_2^1 > \cdots > 0$. For any $i \in \mathbb{N}$, i.e. for any point from the first level, we choose independent point processes $\text{PPP}(t \to x_2 t^{-x_2 - 1})$, whose points we again order downwards: $\eta_{i,1}^2 > \eta_{i,2}^2 > \cdots > 0$, and in this way we proceed: For $j \leqslant K$ and $i_1, \ldots, i_{j-1}$ fixed, $\left(\eta_{i_1 i_2 \ldots i_j}^j \right)_{i_j \in \mathbb{N}}$ are the points of a $\text{PPP}(t \to x_j t^{-x_j - 1})$. They are independent for different $i_1, \ldots, i_{j-1}$ and also for different levels j, and we again assume that for each of these point processes, the points are ordered downwards. This is essentially Ruelle's cascade construction.

We can compose these point processes of the individual levels by just multiplying the "abstract Gibbs weights" along the tree (which corresponds to summing the energy levels of Derrida's GREM along the tree). We therefore arrive at a point process with random points indexed by $\mathbf{i} = (i_1, \ldots, i_K)$, $i_j \in \mathbb{N}$,

(30) $$\eta_{\mathbf{i}} = \eta_{i_1}^1 \eta_{i_1, i_2}^2 \cdots \cdots \eta_{i_1, i_2, \ldots, i_K}^K.$$

This is not a Poisson point process, but after normalization, surprisingly, it is simply a normalized REM. For convenience assume $\beta > \beta_K^{\text{cr}} \overset{\text{def}}{=} \sqrt{2 \log 2} / \sqrt{K} \sigma_K$, so that $x_K(\beta) < 1$. (If this is not satisfied, one has to collapse some of the latter levels and one arrives essentially to the same conclusion for the remaining ones.) In that case

$$\sum_{\mathbf{i}} \eta_{\mathbf{i}} < \infty$$

with probability one, and so one can normalize the point process, defining

$$\overline{\eta}_{\mathbf{i}} \overset{\text{def}}{=} \frac{\eta_{\mathbf{i}}}{\sum_{\mathbf{j}} \eta_{\mathbf{j}}}.$$

Then one has the following properties:

– The point process

$$\Xi = \sum_{\mathbf{i}} \delta_{\overline{\eta}_{\mathbf{i}}}$$

is the normalization of a PPP $(t \to x_K t^{-x_K - 1})$.

– The point process of the Gibbs-distributions of Derrida's GREM converges weakly to the point process of Ruelle's GREM:

$$\sum_{\alpha} \delta_{\mathcal{G}_{\beta}(\alpha)} \to \Xi,$$

in distribution, as $N \to \infty$. For a proof, see [6].

The point process Ξ does not keep track of the way the points were produced through the tree, so it "forgets" the tree structure. This structure is however important for the Parisi picture. The tree structure can be retained in the following way. As usual we order the energy levels $\eta_{\mathbf{i}}$ downwards, i.e. we define a (random) bijection $\pi : \mathbb{N} \to \mathbb{N}^K$ such that $\eta_{\pi(k)}$ is the k-th largest element in the set $\{\eta_{\mathbf{i}}\}$. This leads to an overlap structure on $\mathbb{N}$, by measuring the hierarchical distance between $\pi(i)$ and $\pi(i')$, i.e. we set for $i, i' \in \mathbb{N}$

$$q(i, i') \overset{\text{def}}{=} \max \{r : \pi(i)_1 = \pi(i')_1, \ldots, \pi(i)_r = \pi(i')_r\}.$$

This leads to a sequence of (random) partitions of $\mathbb{N}$, which for $k \leqslant K - 1$ clumps together points in $\mathbb{N}$ whose π-value agrees on level k, i.e. we introduce the equivalence relation

$$i \sim_k i' \Longleftrightarrow q(i, i') \geqslant k,$$

which leads to a partition of $\mathbb{N}$ in the equivalence classes of $\sim_k$. If k decreases, the partitions become coarser. For $k = 0$, evidently all of $\mathbb{N}$ is clumped into one set. Remarkably, this sequence of random partitions is stochastically independent of Ξ itself. Furthermore the sequence of clustering has a very simple Markovian structure, of viewed backwards in k (see [5]).

5. GUERRA'S REPLICA SYMMETRY BREAKING BOUND: THE AIZENMAN-SIMS-STARR PROOF

In a remarkable paper [9], Guerra extended the bound derived in Section 3 to a bound of $f(\beta, h)$ by the Parisi solution. The proof is not very complicated, but hard to understand without knowledge of the cascade picture introduced in the last section. A bit later, Aizenman, Sims, and Starr [2] reproved the bound, and generalized it by

introducing what they call "random overlap structures", which serve as an abstract model for measures on a countable set which have a notion of "overlaps".

DEFINITION 5.1. — *A **random overlap structure** $\mathcal{R}$ (ROSt for short) consists of a finite or countable set A, a probability space $(\Gamma, \mathcal{G}, \mathbb{P})$, and random variables $\eta_\alpha \geqslant 0$, $q_{\alpha, \alpha'}$, $\alpha, \alpha' \in A$, satisfying the following properties*

(1) $\sum_\alpha \eta_\alpha < \infty$
(2) $(q_{\alpha, \alpha'})$ *is positive definite and satisfies* $q_{\alpha, \alpha} = 1$.

The η_α play the rôle of (unnormalized) Gibbs weights, and the q's are the abstract overlaps.

Example 5.2. — As an example take $A = \Sigma_N \overset{\text{def}}{=} \{-1, 1\}^N$. The η_σ, $\sigma \in \Sigma_N$, can be arbitrary. For $q_{\sigma, \sigma'}$ we take the standard overlap $R_N(\sigma, \sigma')$, as introduced before. We write $\mathcal{R}_N^{\text{SK}}$ for this overlap structure. The q here are nonrandom. On the other hand, we can use a (random) reordering of the set A by ordering the η_σ downwards: $\eta_1 > \eta_2 > \cdots > \eta_{2^N}$. After this random reordering, the q become random: $q_{1,2}$ for instance is the overlap of the two indices with the largest η-weight.

Example 5.3. — Another overlap structure is defined by Ruelle's probability cascades (30) introduced in the last section. Fix $0 = m_0 < m_1 < \cdots < m_K = 1$. We take $A = \mathbb{N}^K$, and the η are the (unnormalized) weights $\eta_{\mathbf{i}}$ as in the last section with $x_i \overset{\text{def}}{=} m_i$, $1 \leqslant i \leqslant K$ (see (30)). There is a slight problem because we have to take the last parameter $x_K = 1$, which implies that $\sum_{\mathbf{i}} \eta_{\mathbf{i}} = \infty$. This will not cause any difficulties for what we do below. The overlaps are defined in the following way. Fix a sequence $0 \leqslant q(1) < q(2) < \cdots < q(K) < q(K+1) = 1$, and we set

$$q_{\mathbf{i}, \mathbf{i}'} = q\left(\max\{k : (i_1, \ldots, i_k) = (i'_1, \ldots, i'_k)\} + 1\right),$$

i.e. we measure the hierarchical distance on the tree, and weight it with the function q. For this random overlap structure, we write $\mathcal{R}_K^{\text{Ruelle}}$.

Given any ROSt, we attach to it families of Gaussian random variables $(y_{\alpha, i})_{\alpha \in A, \, i \in \mathbb{N}}$, $(\kappa_\alpha)_{\alpha \in A}$ by requiring

$$(31) \qquad\qquad \mathbb{E}(\kappa_\alpha \kappa_{\alpha'}) = q_{\alpha, \alpha'}^2 / 2,$$

and the "cavity field" by

$$(32) \qquad\qquad \mathbb{E}(y_{\alpha, j} y_{\alpha', j'}) = q_{\alpha, \alpha'} \delta_{j, j'}.$$

The κ and the y are independent. In case, the q's themselves are random variables, these are just the conditional distributions, given (ξ, q). It is not difficult to see that such random variables exist. By an extension of the probability space, we can assume that all the random variables are defined on a single probability space.

For later use, we give the construction of the cavity variables for $\mathcal{R}_K^{\text{Ruelle}}$. We simply write

$$
(33) \qquad y_{\mathbf{i}} = g^{(0)} + \sum_{k=1}^{K} g_{i_1,\dots,i_k}^{(k)},
$$

where the g's are independent centered Gaussians, with $\operatorname{var}\left(g^{(0)}\right) = q\left(1\right)$, $\operatorname{var}\left(g^{(k)}\right) = q\left(k+1\right) - q\left(k\right)$. Furthermore, the $y_{\mathbf{i},j}$, $j \in \mathbb{N}$, are independent copies of $y_{\mathbf{i}}$. The $\kappa_{\mathbf{i}}$ are constructed in a similar way.

The above notion of a ROSt needs some explanation. The basic idea comes from what in the physics literature is called the "cavity method". We consider the standard SK-Hamiltonian, but now with $N+M$ spins, where one should think of N being much larger than M. We then try to write the Hamiltonian in terms of the Hamiltonian on N spin variables acting on the M "newcomers". We write $\tau_i = \sigma_{N+i}$ for the newcomers.

$$
\frac{\beta}{\sqrt{N+M}} \sum_{i<j \leqslant N+M} J_{ij}\sigma_i\sigma_j + h \sum_{i=1}^{N+M} \sigma_i
$$
$$
= \frac{\sqrt{N}}{\sqrt{N+M}} \frac{\beta}{\sqrt{N}} \sum_{i<j \leqslant N} J_{ij}\sigma_i\sigma_j + h \sum_{i=1}^{N} \sigma_i + \frac{\beta}{\sqrt{N+M}} \sum_{j=1}^{M} \left(\sum_{i=1}^{N} J_{i,N+j}\sigma_i \right) \tau_j
$$
$$
+ \frac{\beta}{\sqrt{N+M}} \sum_{i<j \leqslant M} J_{N+i,N+j}\tau_i\tau_j + h \sum_{j=1}^{M} \tau_j.
$$

If $M \ll N$, then we can neglect the interaction among the newcomers, i.e. we can drop the fourth summand on the right hand side above. Furthermore, we may as well replace $\sqrt{N+M}$ by $\sqrt{N}$ in the third summand. Defining the cavity variables

$$
y_{\sigma,j} \overset{\text{def}}{=} \frac{1}{\sqrt{N}} \sum_{i=1}^{N} J_{i,N+j}\sigma_i,
$$

we see that they have exactly the right distribution as required in (32), with respect to the random overlap structure $\mathcal{R}_N^{\text{SK}}$ coming from the N system. In the first summand, we have to be more careful: $N^{-1/2} \sum_{i<j \leqslant N} J_{ij}\sigma_i\sigma_j$ has variance $(N-1)/2$, and $(N+M)^{-\frac{1}{2}} \sum_{i<j \leqslant N} J_{ij}\sigma_i\sigma_j$ has variance $(N+M)^{-1} N(N-1)/2 \simeq (N-1)/2 - M/2$. We can therefore (approximately) represent the former by the latter plus an independent Gaussian $\sqrt{M/2}\kappa_\sigma$, where $(\kappa_\sigma)_{\sigma \in \Sigma_N}$ is a field with the covariances given by (31). If we set

$$
\eta_\sigma \overset{\text{def}}{=} \exp\left[\frac{\beta}{\sqrt{N+M}} \sum_{i<j \leqslant N} J_{ij}\sigma_i\sigma_j + h \sum_{i=1}^{N} \sigma_i \right],
$$

we see that

$$(34) \qquad \frac{Z_{N+M}}{Z_N} \simeq \frac{\sum_{\sigma \in \Sigma_N, \, \tau \in \Sigma_M} \eta_\sigma \exp\left[\sum_{i=1}^M (\beta y_{\sigma,i} + h)\, \tau_i\right]}{\sum_{\sigma \in \Sigma_N} \eta_\sigma \exp\left[\beta \sqrt{M/2}\kappa_\alpha\right]}.$$

Here we have used the ROSt from the N-spin SK model (with Gibbs weights coming from a slightly changed temperature parameter). Aizenman, Sims and Starr had the idea to investigate the above object when the N system is replaced by an *arbitrary* ROSt $\mathcal{R}$, and they considered the "relative finite M free energy" in the following way

$$(35) \qquad G_M(\beta, h, \mathcal{R}) \overset{\text{def}}{=} \frac{1}{M}\mathbb{E}\left(\log \frac{\sum_{\alpha, \tau \in \Sigma_M} \eta_\alpha \exp\left[\sum_{j=1}^M (\beta y_{\alpha,j} + h)\, \tau_j\right]}{\sum_\alpha \eta_\alpha \exp\left[\beta \sqrt{M/2}\kappa_\alpha\right]}\right),$$

where the $\mathbb{E}$ expectation is taken with respect both to the law of the random overlap structure and the cavity variables $y_{\alpha,i}$ and κ_α. The variant of Guerra's theorem in the generalized version of Aizenman, Sims and Starr is the following remarkable inequality, which holds for *any* M, and *any* random overlap structure.

THEOREM 5.4. — *For any M, and any random overlap structure $\mathcal{R}$ one has*

$$(36) \qquad f_M(\beta, h) \leqslant G_M(\beta, h, \mathcal{R}).$$

Remark 5.5. — Aizenman, Sims and Starr actually show that

$$f(\beta, h) = \inf_{\mathcal{R}} \lim_{M \to \infty} G_M(\beta, h, \mathcal{R}),$$

which follows by plugging in the SK-random overlap structure itself, using (34), and applying the subadditivity result of Guerra-Toninelli (3).

I am not going to prove the theorem. To a large extent it is a rerun of the computation done above in Section 3. Here is an outline. One uses the following interpolation:

$$H_M(\tau, \alpha, t) \overset{\text{def}}{=} \frac{\sqrt{1-t}}{\sqrt{M}} \sum_{1 \leqslant i < j \leqslant M} J_{ij}\tau_i\tau_j + \sqrt{\frac{M(1-t)}{2}}\kappa_\alpha + \sqrt{t}\sum_{i=1}^M y_{\alpha,i}\tau_i$$

and defines

$$\widehat{G}_M(\beta, h, t, \mathcal{R}) \overset{\text{def}}{=} \frac{1}{M}\mathbb{E}\left(\log \frac{\sum_{\alpha \in A, \, \tau \in \Sigma_M} \eta_\alpha \exp\left[\beta H_M(\tau, \alpha, t) + h\sum_{j=1}^M \tau_j\right]}{\sum_{\alpha \in A} \eta_\alpha \exp\left[\beta \sqrt{M/2}\kappa_\alpha\right]}\right),$$

where $\mathbb{E}$ is taken with respect to the overlap structure *and* the J's (which are supposed to be independent). For $t = 0$, the κ-part cancels, and one just gets $f_M(\beta, h)$. For $t = 1$, one gets $G_M(\beta, h, \mathcal{R})$. By a computation similar to the one in Section 3, one gets

$$(37) \qquad \frac{d\widehat{G}_M(\beta, h, , t, \mathcal{R})}{dt} = \frac{1}{2}\nu_t^{(2)}\left(\left(R_M(\tau, \tau') - q_{\alpha, \alpha'}\right)^2\right) \geqslant 0,$$

which immediately implies the theorem. $\nu_t^{(2)}$ is to be understood in the following way: For fixed environment (from J and the ROSt), one takes two independent copies of the (τ, α) distributed according to

$$p_t(\tau, \alpha) \stackrel{\text{def}}{=} \frac{\eta_\alpha \exp\left[\beta H_M(\tau, \alpha, t) + h \sum_i \tau_i\right]}{\text{Normalization}},$$

and afterwards, one takes the environment expectation.

In principle, one should of course take $\mathcal{R}_N^{\text{SK}}$ as the random overlap structure, in which case one gets $f(\beta, h)$ by Guerra-Toninelli in the $M \to \infty$ limit, as remarked above, but one has obviously not gained much. The striking fact however is that the inequality is true for *any* ROSt, and one can try to obtain good upper bounds by choosing random overlap structures for which one can compute G_M. It turns out that the good choice is Ruelle's $\mathcal{R}_K^{\text{Ruelle}}$ from Example 5.3:

LEMMA 5.6

$$(38) \qquad G_M\left(\beta, h, \mathcal{R}_K^{\text{Ruelle}}\right) = G_1\left(\beta, h, \mathcal{R}_K^{\text{Ruelle}}\right) = \mathcal{P}_K(m, q; \beta, h).$$

Proof. — We give a sketch of the computation. We can handle numerator and denominator in (35) separately. The denominator is simpler, so I will only discuss the numerator. We take $M = 1$. It will be clear from the computation that for general M the outcome is the same. With the representation of the $y_{\mathbf{i}}$ by (33), we get

(39)

$$\frac{1}{2} \sum_{\mathbf{i}, \tau \in \Sigma_1} \eta_{\mathbf{i}} \exp\left[(\beta y_{\mathbf{i}} + h)\tau\right] = \sum_{\mathbf{i}} \eta_{\mathbf{i}} \cosh(\beta y_{\mathbf{i}} + h)$$

$$= \sum_{(i_1, \ldots, i_K)} \eta_{i_1}^1 \eta_{i_1 i_2}^2 \cdots \cdots \eta_{i_1 i_2 \ldots i_K}^K \cosh\left(\beta \sum_{n=0}^{K} g_{i_1, \ldots, i_n}^{(n)} + h\right).$$

We condition on $\eta_{i_1}^1, \eta_{i_1 i_2}^2, \ldots, \eta_{i_1 i_2 \ldots i_{K-1}}^{K-1}$ and $g_{i_1}^{(1)}, g_{i_1 i_2}^{(2)}, \ldots, g_{i_1 i_2 \ldots i_{K-1}}^{(K-1)}$. Then $(\eta_{i_1 i_2 \ldots i_K}^K)_{i_K \in \mathbb{N}}$ is a PPP $(t \to m_K t^{-m_K - 1})$ whose points are multiplied by the independent random variables $\left(\cosh\left(\beta \sum_{n=0}^{K} g_{i_1, \ldots, i_n}^{(n)} + h\right)\right)_{i_K}$. From Proposition 4.1, we know that the conditional law (conditioned on anything up to level $K-1$) of the point process

$$\left(\eta_{i_1 i_2 \ldots i_K}^K \cosh\left(\beta \sum_{n=0}^{K} g_{i_1, \ldots, i_n}^{(n)} + h\right)\right)_{i_K}$$

is the same as that of

$$\left(C_K\left(\beta \sum_{n=0}^{K-1} g_{i_1, \ldots, i_n}^{(n)}\right) \eta_{i_1 i_2 \ldots i_K}^K\right)_{i_K},$$

where

$$C_K(\xi) \stackrel{\text{def}}{=} \left[E_Z \cosh^{m_K}\left(\xi + h + \beta\sqrt{q_{K+1} - q_K} Z\right)\right]^{1/m_K}, \ \xi \in \mathbb{R},$$

Z being a standard Gaussian random variable, and E_Z the expectation with respect to Z. C_K is a random variable which still depends on the $g^{(n)}$, $n \leqslant K - 1$. We proceed in the same way, replacing $\left(C_K \left(\beta \sum_{n=0}^{K-1} g_{i_1,\ldots,i_n}^{(n)} \right) \eta_{i_1 i_2 \ldots i_{K-1}}^{K-1} \right)_{i_{K-1} \in \mathbb{N}}$ by $\left(C_{K-1} \left(\beta \sum_{n=0}^{K-2} g_{i_1,\ldots,i_n}^{(n)} \right) \eta_{i_1 i_2 \ldots i_{K-1}}^{K-1} \right)_{i_{K-1} \in \mathbb{N}}$, where

$$C_{K-1}\left(\xi \right) \stackrel{\text{def}}{=} \left[E_Z C_K^{m_{K-1}} \left(\xi + h + \beta \sqrt{q_{K-1} - q_{K-2}} Z \right) \right]^{1/m_{K-1}},$$

and so on. We finally see that multiplying the points $\eta_{\mathbf{i}}$ by $\cosh\left(\beta y_{\mathbf{i}} + h \right)$ amounts (for the corresponding point process) in multiplying the points by the constant $E \log Y_1$ from (13). The denominator in (35) is simpler because there one has in every step just an integration of a Gaussian in the exponent. We therefore see that multiplying the points $\eta_{\mathbf{i}}$ by $\exp\left[\left(\beta/\sqrt{2} \right) \kappa_{\mathbf{i}} \right]$ simply leads to a multiplication of the point process by $\exp\left[\left(\beta^2/4 \right) \sum_{i=0}^{K} m_i \left(q_{i+1}^2 - q_i^2 \right) \right]$. In the definition of G_1 (35), we would now like to argue that $\sum_{\mathbf{i}} \eta_{\mathbf{i}}$ cancels out. There is the slight difficulty that this sum diverges almost surely, because of $m_K = 1$, but we can choose m_K slightly less than 1, in which case the sum is finite, and so cancels, and then we can let $m_K \to 1$ in the end. The upshot of this computation is that

$$
\begin{aligned}
G_1\left(\beta, h, \mathcal{R} \right) &= E \log Y_1 - \frac{\beta^2}{4} \sum_{i=0}^{K} m_i \left(q_{i+1}^2 - q_i^2 \right) + \log 2 \\
&= \mathcal{P}_K \left(m, q; \beta, h \right),
\end{aligned}
$$

the $\log 2$ is coming from dividing by 2 in (39). It is evident from this computation that we get the same for arbitrary M. (One is just having M factors of $\cosh\left(\cdot \right)$ with independent contents, so in every step of the above argument, the factoring remains.) □

Combining this result with Theorem 5.4, one gets

$$f_M\left(\beta, h \right) \leqslant \mathcal{P}_K\left(m, q; \beta, h \right)$$

for any K, and any sequence m and q. Therefore

$$f_M\left(\beta, h \right) \leqslant \inf_{K,m,q} \mathcal{P}_K\left(m, q; \beta, h \right).$$

This is Guerra's upper bound.

6. ON TALAGRAND'S PROOF OF THE PARISI FORMULA

From (37) it is evident that one could prove the Parisi formula, provided one gets the right hand side of this equality under control and shows that it approaches 0 by correctly choosing a sequence of random overlap structures. A difficulty arises from the fact that one does not know from the outset the correct random overlap structures.

In his remarkable paper [20], Talagrand had been able to finish this approach and to prove the Parisi formula. This is too complicated to be explained in any details, but I will sketch the strategy in the simplest situation where there is no replica symmetry breaking, i.e. in the region where the replica symmetric formula is correct. In order to prove that, Talagrand considers a replicated system with fixed overlap of the replicas. As one evidently wishes to apply this to the replicated interpolated system of Section 3, we work with measures of the type (23) with the additional parameter γ. Let therefore

$$Z_N^{(2)}(\beta, \gamma, h, u) \stackrel{\text{def}}{=} \sum_{\sigma, \tau \in \Sigma_N: \, R_n(\sigma, \tau) = u} \exp\left[\beta\left(H_N(\sigma) + H_N(\tau)\right) - \sum_i (h + \gamma g_i)(\sigma_i + \tau_i)\right],$$

and

$$(40) \qquad\qquad r_N(\beta, \gamma, h, u) \stackrel{\text{def}}{=} \frac{1}{N} \mathbb{E} \log Z_N^{(2)}(\beta, \gamma, h, u).$$

(One restricts u to the possible overlaps only, but this is not causing any difficulties).

The strategy is to prove first an upper bound for the free energy r_N of the replicated system with restricted overlaps again through an interpolation procedure similar to the proof of Guerra's bound, obtaining a Parisi formula for the replicated system with overlap restrictions. The expression is more complicated than (13). To give the reader an impression of the complexity the procedure gains quickly, I describe the modifications which are needed. We choose K and sequences m, q as in (10), (11), but we need another sequence $(c_j)_{0 \leqslant j \leqslant K}$, $c_j \in [-1, 1]$, describing correlations. Instead of the sequence (g_i) used in (13), we now take a sequence of independent two-dimensional Gaussian vectors (g_i^1, g_i^2), which have the same variances as before, $\mathrm{var}\left(g_i^j\right) = \beta^2 (q_{i+1} - q_i)$, but with correlations between g_i^1 and g_i^2 given by the c_i. Then one defines the sequence $(Y_j)_{0 \leqslant j \leqslant K+1}$ by downwards recursion similar to that in (12), with additional parameter $x, x_1, x_2 \in \mathbb{R}$, starting with

$$Y_{K+1}(x, x_1, x_2) = \cosh(x) \cosh\left(x_1 + \sum_{i=0}^{K} g_i^1\right) \cosh\left(x_2 + \sum_{i=0}^{K} g_i^2\right)$$

$$+ \sinh(x) \sinh\left(x_1 + \sum_{i=0}^{K} g_i^1\right) \sinh\left(x_2 + \sum_{i=0}^{K} g_i^2\right),$$

and finally one sets

$$\mathcal{P}_K(m, q, c, x; \beta, h, \gamma, u) = 2 \log 2 + E Y_0(x, h + \gamma g, h + \gamma g) - ux$$

$$- \frac{\beta^2}{2} \sum_{i=1}^{K} m_i (q_{i+1}^2 - q_i^2) - \frac{\beta^2}{2} \sum_{i=1}^{K} m_i (d_{i+1}^2 - d_i^2) + \frac{\beta^2}{2} (u - d_{k+1})^2,$$

where g is an additional standard Gaussian, and $d_k \stackrel{\text{def}}{=} \sum_{i=0}^{k-1} c_i (q_{i+1} - q_i)$. Then with

$$(41) \qquad\qquad U(\beta, \gamma, h, u) \stackrel{\text{def}}{=} \inf_{K, q, m, c, x} \mathcal{P}_K(m, q, c, x; \beta, h, \gamma, u)$$

one proves

$$(42) \qquad\qquad r_N\left(\beta,\gamma,h,u\right) \leqslant U\left(\beta,\gamma,h,u\right)$$

in a way parallel to the one for Guerra's replica symmetry breaking bound.

Fix now some parameter (β, h). Talagrand's strategy to show $f(\beta, h) = RS(\beta, h)$ (which is true only for small enough β) is to prove that for the parameter from the interpolation in Section 3, i.e. $\beta_t = \sqrt{t}\beta$, $\gamma_t = \sqrt{1-t}\sqrt{q}$, q being the "correct one", i.e. the solution of equation (8), one has

$$\limsup_{N\to\infty} r_N\left(\beta_t,\gamma_t,h,u\right) < 2f\left(\beta_t,\gamma_t,h\right),$$

for $u \neq q$, where $f(\beta,\gamma,h)$ is the free energy of the slightly extended SK-model defined in (23). By (42), this follows from

$$(43) \qquad\qquad U\left(\beta_t,\gamma_t,h,u\right) < 2f\left(\beta_t,\gamma_t,h\right).$$

This would lead to the conclusion that the probability for the replicated system (at interpolation parameter t) not to nearly match the correct q with the overlap, is negligible. By (21), this then implies that $f(\beta, h) = RS(\beta, h)$ (essentially, of course, as there are many details to fit in). At first sight, this seems to be a hopeless enterprise, as on the right hand side of the desired inequality, one has the free energy one is shooting for. Talagrand however remarked that it essentially suffices to prove an inequality where the right hand side in (43) is replaced by $2RS(\beta_t,\gamma_t,h)$.

The main line of argument is running as follows: Let $\psi(\beta, h)$ be the Parisi solution,

$$\psi\left(\beta, h\right) = \inf_{K,q,m} \mathcal{P}_K\left(m, q; \beta, h\right),$$

$\mathcal{P}_K$ from (13), and let $\psi(\beta,\gamma,h)$ be the Parisi solution for the Hamiltonian with the additional parameter γ. Let also $RS(\beta,\gamma,h)$ be the replica symmetric solution for the Hamiltonian with the additional parameter γ. Let β, h be such that $\psi(\beta, h) = RS(\beta, h)$, and furthermore, that this relation is correct on our path joining the coupled system with the uncoupled one, i.e. $\psi(\beta_t,\gamma_t,h) = RS(\beta_t,\gamma_t,h)$, with $\beta_t = \sqrt{t}\beta$, $\gamma_t = \sqrt{(1-t)q}$.

Assume we can prove that

$$(44) \qquad\qquad \sup_{u:|u-q|\geqslant y} U\left(\beta_t,\gamma_t,h,u\right) \leqslant 2RS\left(\beta_t,\gamma_t,h\right) - \varepsilon y^2$$

for some constant ε (possibly depending on β, h, but not on t), then we get

$$\nu_{N,t}^{(2)}\left\{\left(R_N\left(\sigma,\tau\right) - q\right)^2 \geqslant y\right\} \leqslant \mathrm{e}^{o(N)} \exp\left[2N\left(RS\left(\beta_t,\gamma_t,h\right) - \varepsilon y - f_N\left(\beta_t,\gamma_t,h\right)\right)\right],$$

and plugging in $y = C\left(RS\left(\beta_t,\gamma_t,h\right) - f_N\left(\beta_t,\gamma_t,h\right)\right) + \varepsilon_1$, with $C = 1/\varepsilon$, $\varepsilon_1 > 0$ arbitrary, we get that

$$\nu_{N,t}^{(2)}\left\{\left(R_N\left(\sigma,\tau\right) - q\right)^2 \geqslant C\left(RS\left(\beta_t,\gamma_t,h\right) - f_N\left(\beta_t,\gamma_t,h\right)\right) + \varepsilon_1\right\}$$

is exponentially decaying in N, uniformly in t, with a rate depending on ε_1. Therefore, we can choose N large enough (depending on β, h, ε_1) such that the above probability is small. Using that, and (22), it is then possible to prove

$$\lim_{N \to \infty} f_N(\beta, h) = RS(\beta, h).$$

(44) needs considerable work, but given Parisi-like formula (41), and the Guerra type bound (42), it suffices to (very) cleverly plug in parameters K, m, q, c, x on the right hand side to get the desired inequality.

One would hope that the regions where this is correct coincide with the one bounded by the AT-line, i.e. when (9) is satisfied. This however has not yet been proved. Anyway, Talagrand gives a not totally explicit characterization of the region in the (β, h)-plane where this line of arguments works (see [19], p. 153), and where the replica symmetric solution is correct.

In a remarkable tour de force, Talagrand has been able to extend this type of arguments to cover the whole temperature regime, and to verify the Parisi formula in the whole temperature region. An apparent difficulty is coming from the fact that in the low temperature region, one needs to let $K \to \infty$ in the infimum (14). A very pertinent observation of Talagrand is that along the interpolation, for any fixed $t_0 < 1$, one can work with finitely many symmetry breakings on $[0, t_0]$, where the number of levels of symmetry breaking has to increase to ∞ only when $t_0 \uparrow 1$. Therefore, along the interpolation path, one can work with a finite K, and instead of minimizing over one parameter q as in the high temperature region, one minimizes over $2K-1$ parameters $q_1, \ldots, q_K, m_1, \ldots, m_{K-1}$, and tries to use estimates like (44). It goes without saying that the whole argument becomes more complicated than the replica symmetric case by an order of magnitude.

7. SOME OPEN PROBLEMS

Despite the success of the approach by Guerra and Talagrand, the whole field is still very much in its beginning, and there remain many open problems. Much of the recent progress depends on the possibility to derive inequalities of the type explained in Section 3 and Section 5. It is completely open if there exist similar inequalities, perhaps only in an asymptotic sense, for other models like the perceptron, or the Hopfield net. Talagrand had used a different method to obtain results on these models in the high temperature region which is based on the "cavity method". There is not space enough here to give any explanations of this procedure (see [19], Chap. 3-5). The low temperature region for these models seems to be completely out of reach, presently.

Another open question concerns ultrametricity, even for the SK-model, and for instance to prove (17). A natural approach would be to try to somehow extend the approach used by Talagrand for a tripled system with three replicas $\sigma^1, \sigma^2, \sigma^3$, define a free energy like (40) with the overlaps $R_N\left(\sigma^1, \sigma^2\right), R_N\left(\sigma^1, \sigma^3\right), R_N\left(\sigma^1, \sigma^3\right)$ fixed, and derive a "Guerra type" bound, like (42) with a Parisi-formula like (41), and then to exclude overlaps which do not satisfy the ultrametricity condition. It seems that this natural approach leads to tremendous difficulties, and it has not been possible to make progress along this line, at least to my knowledge.

Another approach would be more in the line of Theorem 5.4 and Remark 5.5, and try to prove that the only random overlap structures which minimize the right hand side of (36) are ultrametric ones. This could follow from a characterization of the Ruelle point process through some invariance properties, like the one given in [17], but including correlations. This had been proposed by Michael Aizenman, but first of all, the necessary extension of [17] seems to be rather delicate, and secondly, it is not clear to me if this really would lead to a proof of ultrametricity in the form (17).

On a much more modest level, there has recently been some progress, namely a proof that some natural "non-hierarchical" generalizations of the GREM are in the limit hierarchical organized (see [4]), but it is very questionable if this sheds much light on similar problems for instance for the SK model.

REFERENCES

[1] M. AIZENMAN, J. LEBOWITZ & D. RUELLE – Some rigorous results on the Sherrington-Kirkpatrick model, *Comm. Math. Phys.* **112** (1987), p. 3–20.

[2] M. AIZENMAN, R. SIMS & S.L. STARR – An extended variational principle for the SK spin-glass model, *Phys. Rev. B* **68** (2003), p. 214–403.

[3] J.R.L. DE ALMEIDA & D.J. THOULESS – Stability of the Sherrington-Kirkpatrick solution of spin glasses, *J. Phys.* A **11** (1978), p. 983.

[4] E. BOLTHAUSEN & N. KISTLER – On a non-hierarchical version of the Generalized Random Energy Model, *Ann. Appl. Prob.* **16** (2006), p. 1–14.

[5] E. BOLTHAUSEN & A.-S. SZNITMAN – On Ruelle's probability cascades and an abstract cavity method, *Comm. Math. Phys.* **197** (1998), p. 247–276.

[6] A. BOVIER & I. KURKOVA – Derrida's Generalized Random Energy Models I & II, *Annales de l'Institut Henri Poincaré* **40** (2004), p. 439–495.

[7] B. DERRIDA – Random energy model: An exactly solvable model of disordered systems, *Phys. Rev. B* **24** (1981), p. 2613–2626.

[8] ———, A generalization of the random energy model that includes correlations between the energies, *J. Physique. Lett* **46** (1986), p. 401–407.

[9] F. GUERRA – Replica broken bounds in the mean field spin glass model, *Comm. Math. Phys.* **233** (2003), p. 1–12.

[10] F. GUERRA & F. L. TONINELLI – Quadratic replica coupling in the Sherrington-Kirkpatrick mean field spin glass model, *J. Math. Phys.* **43** (2002), p. 3704–3716.

[11] ⸺ , The thermodynamic limit in mean field spin glass models, *Comm. Math. Phys.* **230** (2002), p. 71–79.

[12] M. LEDOUX – The Concentration of Measure Phenomenon, vol. 89, AMS, 2002.

[13] M. MÉZARD, G. PARISI & M.A. VIRASORO – *Spin Glass Theory and Beyond*, World Scientific, 1987.

[14] H. NISHIMORI – *Statistical Physics of Spin Glasses and Information Processing*, Oxford Science Publications, 1999.

[15] G. PARISI – A sequence of approximate solutions to the S-K model for spin glasses, *J. Phys.* A **13 L-115** (1980).

[16] D. RUELLE – A mathematical reformulation of Derrida's REM and GREM, *Comm. Math. Phys.* **108** (1987), p. 225–239.

[17] A. RUZMAIKINA & M. AIZENMAN – Characterization of invariant measures at the leading edge for competing particle systems, *Ann. Prob.* **33** (2005), p. 83–113.

[18] D. SHERRINGTON & S. KIRKPATRICK – Solvable model of a spin glass, *Phys. Rev. Lett.* **35** (1972), p. 1792–1796.

[19] M. TALAGRAND – *Spin Glasses: A Challenge for Mathematicians*, Springer, Heidelberg, 2003.

[20] ⸺ , The Parisi formula, *Ann. Math.* **163** (2006), p. 221–263.

Erwin BOLTHAUSEN

Institute of Mathematics
University of Zurich
Winterthurerstrasse 190
CH-8057 ZURICH
(Suisse)
E-mail : eb@math.unizh.ch