



ELSEVIER

Available online at www.sciencedirect.com



C. R. Acad. Sci. Paris, Ser. I 337 (2003) 477–480



Probability Theory/Mathematical Physics

On Guerra's broken replica-symmetry bound

Michel Talagrand

Équipe d'analyse de l'institut mathématique, 4, place Jussieu, 75230 Paris cedex 05, France

Received 21 August 2003; accepted 1 September 2003

Presented by Yves Meyer

Abstract

Consider a random Hamiltonian $H_N(\vec{\sigma})$ for $\vec{\sigma} \in \Sigma_N = \{0, 1\}^N$. We assume that the family $(H_N(\vec{\sigma}))$ is jointly Gaussian centered and that for $\vec{\sigma}^1, \vec{\sigma}^2 \in \Sigma_N$, $N^{-1} E H_N(\vec{\sigma}^1) H_N(\vec{\sigma}^2) = \xi(N^{-1} \sum_{i \leq N} \sigma_i^1 \sigma_i^2)$ for a certain function ξ on \mathbb{R} . F. Guerra proved the remarkable fact that the free energy of the system with Hamiltonian $H_N(\vec{\sigma}) + h \sum_{i \leq N} \sigma_i$ is bounded below by the free energy of the Parisi solution provided that ξ is *convex* on \mathbb{R} . We prove that this fact remains (asymptotically) true when the function ξ is only assumed to be convex on \mathbb{R}^+ . This covers in particular the case of the p -spin interaction model for any p . **To cite this article:** *M. Talagrand, C. R. Acad. Sci. Paris, Ser. I 337 (2003)*.

© 2003 Académie des sciences. Published by Éditions scientifiques et médicales Elsevier SAS. All rights reserved.

Résumé

Sur une majoration célèbre de F. Guerra. Considérons un hamiltonien aléatoire $H_N(\vec{\sigma})$ où $\vec{\sigma} \in \Sigma_N = \{0, 1\}^N$. Nous supposons la famille $(H_N(\vec{\sigma}))$ gaussienne centrée et que pour tous $\vec{\sigma}^1, \vec{\sigma}^2 \in \Sigma_N$, on ait $N^{-1} E H_N(\vec{\sigma}^1) H_N(\vec{\sigma}^2) = \xi(N^{-1} \sum_{i \leq N} \sigma_i^1 \sigma_i^2)$ pour une certaine fonction ξ sur \mathbb{R} . F. Guerra a prouvé récemment le fait remarquable que l'énergie libre du système d'hamiltonien $H_N(\vec{\sigma}) + h \sum_{i \leq N} \sigma_i$ est bornée inférieurement par l'énergie libre de la solution de Parisi lorsque ξ est *convexe* sur \mathbb{R} . Nous montrons que ceci reste asymptotiquement vrai si l'on suppose seulement que la fonction ξ est convexe sur \mathbb{R}^+ . Ce résultat s'applique en particulier au cas du modèle d'interaction à p -spin pour tout p . **Pour citer cet article:** *M. Talagrand, C. R. Acad. Sci. Paris, Ser. I 337 (2003)*.

© 2003 Académie des sciences. Published by Éditions scientifiques et médicales Elsevier SAS. All rights reserved.

1. Statement of the result

We consider for each N a Gaussian Hamiltonian H_N on Σ_N , that is a centered jointly Gaussian family of r.v. indexed by Σ_N . We assume that for a certain sequence $c(N) \rightarrow 0$ and a certain function $\xi : \mathbb{R} \rightarrow \mathbb{R}$, we have

$$\forall \vec{\sigma}^1, \vec{\sigma}^2 \in \Sigma_N, \quad \left| \frac{1}{N} E H_N(\vec{\sigma}^1) H_N(\vec{\sigma}^2) - \xi(R_{1,2}) \right| \leq c(N), \quad (1)$$

where $R_{1,2} = R_{1,2}(\vec{\sigma}^1, \vec{\sigma}^2) = N^{-1} \sum_{i \leq N} \sigma_i^1 \sigma_i^2$.

We fix once and for all a number h (that represents the strength of an “external field”). Consider an integer k and numbers $0 = m_0 \leq m_1 \leq \dots \leq m_{k-1} \leq m_k = 1$ and $0 = q_0 \leq q_1 \leq \dots \leq q_{k+1} = 1$. To lighten notation, we write

$$\mathbf{m} = (m_0, \dots, m_{k-1}, m_k); \quad \mathbf{q} = (q_0, \dots, q_k, q_{k+1}). \quad (2)$$

1631-073X/\$ – see front matter © 2003 Académie des sciences. Published by Éditions scientifiques et médicales Elsevier SAS. All rights reserved.

doi:10.1016/j.crma.2003.09.001

Consider independent standard Gaussian r.v. $(z_p)_{0 \leq p \leq k}$ and define $a_p = \sqrt{\xi'(q_{p+1}) - \xi'(q_p)}$. We define the r.v. $X_{k+1} = \log \text{ch}(h + \sum_{0 \leq p \leq k} a_p z_p)$ and recursively, for $\ell \geq 0$ we define $X_\ell = m_\ell^{-1} \log E_\ell \exp m_\ell X_{\ell+1}$, where E_ℓ denotes expectation in the r.v. $z_p, p \geq \ell$. When $m_\ell = 0$ this means $X_\ell = E_\ell X_{\ell+1}$. Thus X_0 is a number. We set $\theta(q) = q\xi'(q) - \xi(q)$ and

$$\mathcal{P}(\mathbf{m}, \mathbf{q}) = \log 2 + X_0 - \frac{1}{2} \sum_{1 \leq \ell \leq k} m_\ell (\theta(q_{\ell+1}) - \theta(q_\ell)). \tag{3}$$

To lighten the exposition, we do not follow the convention of Physics to put a minus sign in front of the Hamiltonian.

Theorem 1.1 (Guerra’s bound [2]). *If ξ is convex, we have*

$$\frac{1}{N} E \log \sum_{\vec{\sigma}} \exp \left(H_N(\vec{\sigma}) + h \sum_{i \leq N} \sigma_i \right) \leq \mathcal{P} := \inf \mathcal{P}_k(\mathbf{m}, \mathbf{q}) + c(N), \tag{4}$$

where the infimum is computed over the all values of k and the parameters \mathbf{m}, \mathbf{q} .

Strictly speaking, Guerra proves this result only for $\xi(x) = ax^p, p$ even, $a > 0$ but almost no changes are required to his proof to obtain the above statement. Our main result is the following:

Theorem 1.2. *If ξ is convex on \mathbb{R}^+ , we have*

$$\limsup_{N \rightarrow \infty} \frac{1}{N} E \log \sum_{\vec{\sigma}} \exp \left(H_N(\vec{\sigma}) + h \sum_{i \leq N} \sigma_i \right) \leq \mathcal{P} := \inf \mathcal{P}_k(\mathbf{m}, \mathbf{q}), \tag{5}$$

where the infimum is computed over all the values of the parameters.

This result applies in particular to the case of the p -spin interaction model for all values of p while Theorem 1.1 applies only to the case where p is even.

It is proved in [4] that when ξ is convex, $\xi(x) = \xi(-x), \xi(0) = \xi'(0) = 0$ and $\xi'(x) > 0$ for $x > 0$ there is equality in (5) (and the limsup is a limit). It is natural to conjecture that this remains the case under the conditions on Theorem 1.2 but this question remains open.

2. Elements of proof

The central idea is Guerra’s interpolation scheme. Given m_1, \dots, m_k and q_1, q_2, \dots, q_k as above, consider for $i \leq N$ and $0 \leq \ell \leq k$ independent standard Gaussian r.v. $z_{i,\ell}$, independent of the randomness of H_N , and for $0 \leq t \leq 1$ consider

$$H_t(\vec{\sigma}) = \sqrt{t} H_N(\vec{\sigma}) + \sqrt{1-t} \sum_{i \leq N} \sigma_i \sum_{0 \leq \ell \leq k} a_\ell z_{i,\ell} + h \sum_{i \leq N} \sigma_i.$$

Set $F_{k+1,t} = \log \sum_{\vec{\sigma}} \exp(H_t(\vec{\sigma}))$ and define recursively $F_{\ell,t} = m_\ell^{-1} \log E_\ell \exp m_\ell F_{\ell+1,t}$ for $\ell \geq 1$, where E_ℓ denote expectation in the r.v. $z_{i,p}, p \geq \ell$. Set $\phi(t) = N^{-1} E F_{1,t}$, where expectation is now in the randomness of H_N and the r.v. $z_{i,0}$. For $1 \leq \ell \leq k$, define $W_\ell = \exp m_\ell (F_{\ell+1,t} - F_{\ell,t})$ and for a function f on Σ_N , let $\gamma_\ell(f) = E_\ell (W_\ell \cdots W_k \langle f \rangle_t)$, where $\langle \cdot \rangle_t$ denote averaging for the Gibbs measure with Hamiltonian H_t . It is a probability on Σ_N ; we denote by $\gamma_\ell^{\otimes 2}$ its square on Σ_N^2 , and we denote by μ_ℓ the probability on Σ_N^2 given by $\mu_\ell(f) = E(W_1 \cdots W_{\ell-1} \gamma_\ell^{\otimes 2}(f))$ for a function f on Σ_N^2 . When $\xi'(0) = 0$, Guerra [2] proves the identity

$$\phi'(t) = -\frac{1}{2} \sum_{1 \leq \ell \leq k} m_\ell (\theta(q_{\ell+1}) - \theta(q_\ell)) - \frac{1}{2} \sum_{1 \leq \ell \leq k} (m_\ell - m_{\ell-1}) \mu_\ell (\xi(R_{1,2}) - R_{1,2} \xi'(q_\ell) + \theta(q_\ell)) + \mathcal{R}, \tag{6}$$

where $|\mathcal{R}| \leq c(N)$. When one assumes that ξ is convex, the term $\xi(R_{1,2}) - R_{1,2}\xi'(q_\ell) + \theta(q_\ell)$ is non-negative, so that (6) implies that $\phi(1) \leq \phi(0) - (1/2) \sum_{1 \leq \ell \leq k} m_\ell(\theta(q_{\ell+1}) - \theta(q_\ell)) + c(N)$. It is then easy to compute $\phi(0)$ since for the corresponding Hamiltonian there is not interaction between the different sites, and this proves Theorem 1.1.

The natural approach to prove Theorem 1.2 would then be to follow the same proof, and to show that, in some sense, as seen for the point of view of the functional μ_ℓ , the quantity $R_{1,2}$ is essentially non-negative, in which case it suffices to know that ξ is convex on \mathbb{R}^+ to assert that the term $\xi(R_{1,2}) - R_{1,2}\xi'(q_\ell) + \theta(q_\ell)$ is non-negative.

Having in mind purposes somewhat similar to the present ones, the author could establish ([3], Section 6.6) that, under very general conditions, the quantity $R_{1,2}$ is essentially non-negative when seen from the point of view of a random Gibbs measure G , in the sense that for each $\varepsilon > 0$, as $N \rightarrow \infty$, we have $E \langle 1_{\{R_{1,2} \leq \varepsilon\}} \rangle \rightarrow 0$, where E denotes expectation in the disorder and $\langle \cdot \rangle$ denotes average for the Gibbs measure $G^{\otimes 2}$. The condition under which this result is true is (roughly speaking) the validity of the Ghirlanda–Guerra identities. These authors discovered how these identities become rather miraculously true when one adds to the random Hamiltonian of the Gibbs’ measure a suitable “lower order” term that does not change the limit in (5). It is then natural to expect that one could adapt this approach to prove that the quantity $R_{1,2}$ is essentially positive from the point of view of the functional μ_ℓ . Unfortunately, when one tries to extend the Ghirlanda–Guerra identities to this setting of the functionals μ_ℓ , one seems to run into intractable computational difficulties, a fact that puzzled the author for a long time. The way out of this morass was (unintentionally) provided by the paper [1]. The goal of these authors is to provide a more general bound than that of (4). In view of the result of [4] it seems unlikely that this goal will be achieved, but another benefit of their approach is that they propose a setting in which the arguments of [3], Section 6.6 can be extended.

The basic idea of [1] is to consider quantities of the type $E \log(\sum_{\alpha, \vec{\sigma}} w_\alpha \exp H(\vec{\sigma}, \alpha))$. When the family of weights w_α is suitably chosen, using Derrida–Ruelle cascades, this allows one to recover quantities such as in the right-hand side of (3). Considering an integer k , we take as indexes α the $k + 1$ -tuples (n_0, \dots, n_k) where $n_0, \dots, n_k \geq 1$. For $p \geq 1$, $i \geq 1$ and integers n_0, \dots, n_p we consider independent standard normal r.v. z_{i,n_0, \dots, n_p} and y_{n_0, \dots, n_p} , independent of the Hamiltonian. Consider q as in (2) and for $0 \leq p \leq k$ define $a_p = \sqrt{\xi'(q_{p+1}) - \xi'(q_p)}$ and $b_p = \sqrt{\theta(q_{p+1}) - \theta(q_p)}$ and consider

$$H_{N,t}(\vec{\sigma}, \alpha) = \sqrt{t} \left(H_N(\vec{\sigma}) + \sqrt{N} \sum_{0 \leq p \leq k} b_p y_{n_0, \dots, n_p} \right) + \sqrt{1-t} \left(\sum_{i \leq N} \sigma_i \sum_{0 \leq p \leq k} a_p z_{i,n_0, \dots, n_p} \right) + h \sum_{i \leq N} \sigma_i.$$

Consider the quantities $Z_{N,t} = \sum_{\alpha, \vec{\sigma}} w_\alpha \exp H_{N,t}(\vec{\sigma}, \alpha)$ and $\psi_N(t) = N^{-1} E \log Z_{N,t}$. The computation corresponding to the relation (6) is the relation (that holds whenever $\xi'(0) = 0$ and the weights w_α are independent of all the other r.v. already introduced)

$$\psi'_N(t) = -\frac{1}{2} E \left\{ \xi(R_{1,2}) - R_{1,2} \xi'(R_{\alpha,\beta}) + \theta(R_{\alpha,\beta}) \right\}_t + \mathcal{R}, \tag{7}$$

where $|\mathcal{R}| \leq c(N)$ and where, for $\alpha = (n_0, \dots, n_k)$ and $\beta = (n'_0, \dots, n'_k)$ we have $R_{\alpha,\beta} = q_\ell$, where ℓ is the largest such that $n_0 = n'_0, \dots, n_{\ell-1} = n'_{\ell-1}$, and, where, for a function $f(\vec{\sigma}^1, \vec{\sigma}^2, \alpha, \beta)$ we have

$$\langle f \rangle_t = Z_{N,t}^{-2} \sum w_\alpha w_\beta f(\vec{\sigma}^1, \vec{\sigma}^2, \alpha, \beta) \exp(H_{N,t}(\vec{\sigma}^1, \alpha) + H_{N,t}(\vec{\sigma}^2, \beta)), \tag{8}$$

where the summation is over all values of $\alpha, \beta, \vec{\sigma}^1, \vec{\sigma}^2$. Thus, when ξ is convex, (7) yields $\psi(1) \leq \psi(0) + c(N)$. We then choose the weights w_α as follows (probability cascades). Assuming, without loss of generality that $0 < m_0$ and $m_1 < 1$ (rather than $m_0 = 0$ and $m_1 = 1$), we consider a non-increasing enumeration $(x_{n_0, \dots, n_{p-1}, \ell})_{\ell \geq 1}$ of the points generated by a Poisson point process on \mathbb{R} of intensity measure $\exp(-m_p y) dy$, these processes being independent of each other as the indices n_0, \dots, n_{p-1} vary, and we take $w_\alpha = U^{-1} \exp(\sum_{0 \leq p \leq k} x_{n_0, \dots, n_{p-1}, n_p})$,

where U is the normalization factor that ensures that the sum of the weights is 1. It can be shown with this choice that the relation $\psi(1) \leq \psi(0) + c(N)$ yields (4). To prove Theorem 1.2 let us introduce the “perturbation term”

$$H'_N(\vec{\sigma}, \vec{\beta}) = a(N) \sum_{\ell \geq 1} \frac{\beta_\ell}{N^{(\ell-1)/2}} \sum_{1 \leq i_1 < \dots < i_\ell \leq N} g_{i_1 \dots i_\ell} \sigma_{i_1} \dots \sigma_{i_\ell},$$

where the r.v. $g_{i_1 \dots i_\ell}$ are i.i.d. standard normal, independent of all the r.v. previously considered, where $\vec{\beta} = (\beta_\ell)_{\ell \geq 1}$, $|\beta_\ell| \leq 1$, and where $a(N) \rightarrow 0$ will be specified later. The purpose of the perturbation term is to give rise to the Ghirlanda–Guerra equalities, as in [3], Section 6.4. We then consider

$$\psi_N(t) = \frac{1}{N} \int E \log \sum_{\alpha, \vec{\sigma}} w_\alpha \exp(H_{N,t}(\vec{\sigma}, \alpha) + H'_N(\vec{\sigma}, \vec{\beta})) d\vec{\beta},$$

where (as everywhere below) the integral is over the domain $|\beta_\ell| \leq 1$ for each ℓ . As in (7) we now have that

$$\psi'_N(t) = -\frac{1}{2} \int E \langle \xi(R_{1,2}) - R_{1,2} \xi(R_{\alpha,\beta}) + \theta(R_{\alpha,\beta}) \rangle_{t, \vec{\beta}} d\vec{\beta} + \mathcal{R}, \quad (9)$$

where $\langle \cdot \rangle_{t, \vec{\beta}}$ has the obvious definition similar to (8). To prove Theorem 1.2 it suffices to show that if the sequence $a(N)$ goes to 0 slowly enough, for any $\varepsilon > 0$ we have (uniformly in t) that

$$\lim_{N \rightarrow \infty} \int E \langle 1_{\{R_{1,2} \leq -\varepsilon\}} \rangle_{t, \vec{\beta}} d\vec{\beta} = 0.$$

This is done by a rather straightforward adaptation of the arguments of [3], Sections 6.5 and 6.7. The one point that is not immediate is that to make the arguments work one needs to know that if we set $A(\alpha) = \sum_{\vec{\sigma}} \exp(H_{N,t}(\vec{\sigma}, \alpha) + H'_N(\vec{\sigma}, \vec{\beta}))$, the function $N^{-1} E \log \sum_{\alpha} w_\alpha A(\alpha)$ has fluctuations of order $\ll 1$, and because of the weights w_α we cannot deduce this immediately from Gaussian concentration of measure. But we simply observe that a few of the indices α carry most of the weight, and that by Gaussian concentration, each quantity $\log A(\alpha)$ has fluctuations of order $N^{1/2}$, and its distribution is independent of α .

References

- [1] M. Aizenman, R. Sims, S. Starr, An extended variational principle for the SK spin-glass model, cond-mat/0306386.
- [2] F. Guerra, Replica broken bounds in the mean field spin glass model, Comm. Math. Phys. 233 (2003) 1–12.
- [3] M. Talagrand, Spin Glasses, in: A Challenge to Mathematicians, Springer-Verlag, 2003.
- [4] M. Talagrand, The generalized Parisi formula, C. R. Acad. Sci. Paris, Ser. I 337 (2003) 111–114.