DOI: 10.1214/ECP.v10-1145

Elect. Comm. in Probab. **10** (2005), 155–166

ELECTRONIC COMMUNICATIONS in PROBABILITY

A QUESTION ABOUT THE PARISI FUNCTIONAL

DMITRY PANCHENKO[1]
Department of Mathematics
Massachusetts Institute of Technology
77 Massachusetts Ave, 2-181
Cambridge, MA, 02139
email: panchenk@math.mit.edu

Submitted 30 December 2004, *accepted in final form* 19 July 2005

AMS 2000 Subject classification: 60K35, 82B44
Keywords: Spin glasses, the Parisi formula.

Abstract

We conjecture that the Parisi functional in the Sherrington-Kirkpatrick model is convex in the functional order parameter. We prove a partial result that shows the convexity along "one-sided" directions. An interesting consequence of this result is the log-convexity of L_m norm for a class of random variables.

1 A problem and some results.

Let $\mathcal{M}$ be a set of all nondecreasing and right-continuous functions $m:[0,1]\to[0,1]$. Let us consider two convex smooth functions Φ and $\xi:\mathbb{R}\to\mathbb{R}$ both symmetric, $\Phi(-x)=\Phi(x)$ and $\xi(-x)=\xi(x)$, and $\Phi(0)=\xi(0)=0$. We will also assume that Φ is of moderate growth so that all integrals below are well defined.

Given $m\in\mathcal{M}$, consider a function $\Phi(q,x)$ for $q\in[0,1], x\in\mathbb{R}$ such that $\Phi(1,x)=\Phi(x)$ and

$$\frac{\partial\Phi}{\partial q}=-\frac{1}{2}\xi''(q)\Big(\frac{\partial^2\Phi}{\partial x^2}+m(q)\Big(\frac{\partial\Phi}{\partial x}\Big)^2\Big). \tag{1.1}$$

Let us consider a functional $\mathcal{P}:\mathcal{M}\to\mathbb{R}$ defined by $\mathcal{P}(m)=\Phi(0,h)$ for some $h\in\mathbb{R}$.

Main question: Is $\mathcal{P}$ a convex functional on $\mathcal{M}$?

The same question was asked in [7]. Unfortunately, despite considerable effort, we were not able to give complete answer to this question. In this note we will present a partial result that shows convexity along the directions $\lambda m+(1-\lambda)n$ when $m(q)\geq n(q)$ for all $q\in[0,1]$. It is possible that the answer to this question lies in some general principle that we are not aware of. A good starting point would be to find an alternative proof of the simplest case of constant m given in Corollary 1 below.

[1]RESEARCH PARTIALLY SUPPORTED BY NSF

The functional $\mathcal{P}$ arises in the Sherrington-Kirkpatrick mean field model where with the choice of $\Phi(x) = \log \mathrm{ch} x$, the following *Parisi formula*

$$\inf_{m\in\mathcal{M}} \Big(\log 2 + \mathcal{P}(m) - \frac{1}{2}\int_0^1 m(q) q \xi''(q) dq\Big) \tag{1.2}$$

gives the free energy of the model. A rigorous proof of this result was given by Michel Talagrand in [5]. Since the last term is a linear functional of m, convexity of $\mathcal{P}(m)$ would imply the uniqueness of *the functional order parameter* $m(q)$ that minimizes (1.2). A particular case of $\xi(x) = \beta^2 x^2/2$ for $\beta > 0$ would also be of interest since it corresponds to the original SK model [2].

In the case when m is a step function, the solution of (1.1) can be written explicitly, since for a constant m the function $g(q,x) = \exp m\Phi(q,x)$ satisfies the heat equation

$$\frac{\partial g}{\partial q} = -\frac{1}{2}\xi''(q)\frac{\partial^2 g}{\partial x^2}.$$

Given $k \geq 1$, let us consider a sequence

$$0 = m_0 \leq m_1 \leq \ldots \leq m_k = 1$$

and a sequence

$$q_0 = 0 \leq q_1 \leq \ldots \leq q_k \leq q_{k+1} = 1.$$

We will denote $\boldsymbol{m} = (m_0, \ldots, m_k)$ and $\boldsymbol{q} = (q_0, \ldots, q_{k+1})$. Let us define a function $m \in \mathcal{M}$ by

$$m(q) = m_l \text{ for } q_l \leq q < q_{l+1}. \tag{1.3}$$

For this step function $\mathcal{P}(m)$ can be defined as follows. Let us consider a sequence of independent Gaussian random variables $(z_l)_{0\leq l\leq k}$ such that

$$\mathbb{E} z_l^2 = \xi'(q_{l+1}) - \xi'(q_l).$$

Define $\Phi_{k+1}(x) = \Phi(x)$ and recursively over $l \geq 0$ define

$$\Phi_l(x) = \frac{1}{m_l}\log \mathbb{E}_l \exp m_l \Phi_{l+1}(x + z_l) \tag{1.4}$$

where $\mathbb{E}_l$ denotes the expectation in $(z_i)_{i\geq l}$ and in the case of $m_l = 0$ this means $\Phi_l(x) = \mathbb{E}_l \Phi_{l+1}(x + z_l)$. Then $\mathcal{P}(m)$ for m in (1.3) is be given by

$$\mathcal{P}_k = \mathcal{P}_k(\boldsymbol{m}, \boldsymbol{q}) = \Phi_0(h). \tag{1.5}$$

For simplicity of notations, we will sometimes omit the dependence of $\mathcal{P}_k$ on $\boldsymbol{q}$ and simply write $\mathcal{P}_k(\boldsymbol{m})$. Let us consider another sequence $\boldsymbol{n} = (n_0, \ldots, n_k)$ such that

$$0 = n_0 \leq n_1 \leq \ldots \leq n_k = 1.$$

The following is our main result.

Theorem 1 *If $n_j \leq m_j$ for all j or $n_j \geq m_j$ for all j then*

$$\mathcal{P}_k(\boldsymbol{n}) - \mathcal{P}_k(\boldsymbol{m}) \geq \nabla \mathcal{P}_k(\boldsymbol{m}) \cdot (\boldsymbol{n} - \boldsymbol{m}) = \sum_{0\leq j\leq k} \frac{\partial \mathcal{P}_k}{\partial m_j}(\boldsymbol{m})(n_j - m_j). \tag{1.6}$$

Remark. In Theorem 1 one does not have to assume that the coordinates of vectors $\boldsymbol{m}$ and $\boldsymbol{n}$ are bounded by 1 or arranged in an increasing order. The proof requires only slight modifications which for simplicity will be omitted.

Since the functional $\mathcal{P}$ is uniformly continuous on $\mathcal{M}$ with respect to L_1 norm (see [1] or [7]), approximating any function by the step functions implies that $\mathcal{P}$ is continuous along the directions $\lambda m + (1-\lambda)n$ when $m(q) \geq n(q)$ for all $q \in [0,1]$.

Of course, (1.6) implies that $\mathcal{P}_k(\boldsymbol{m})$ is convex in each coordinate. This yields an interesting consequence for the simplest case of a constant function $m(q) = m$, which formally corresponds to the case of $k = 2$,

$$0 = m_0 \leq m \leq m_2 = 1 \text{ and } 0 = q_0 = q_1 \leq q_2 = q_3 = 1.$$

In this case,

$$\mathcal{P}_k = f(m) = \frac{1}{m} \log \mathbb{E} \exp m\Phi(h + \sigma z). \tag{1.7}$$

Here $\sigma^2 = \xi'(1)$ can be made arbitrary by the choice of ξ. (1.6) implies the following.

Corollary 1 *If $\Phi(x)$ is convex and symmetric them $f(m)$ defined in (1.7) is convex.*

Corollary 1 implies that the L_m norm of $\exp \Phi(h + \sigma z)$ is log-convex in m. This is a stronger statement than the well-known consequence of Hölder's inequality that the L_m norm is always log-convex in $1/m$. At this point it does not seem obvious how to give an easier proof even in the simplest case of Corollary 1 than the one we give below. For example, it is not clear how to show directly that

$$f''(m) = m^{-3}(\mathbb{E}V \log^2 V - (\mathbb{E}V \log V)^2 - 2\mathbb{E}V \log V) \geq 0,$$

where $V = \exp m(\Phi(h + \sigma z) - f(m))$.

Finally, let us note some interesting consequences of the convexity of $f(m)$. First, $f''(0) \geq 0$ implies that the third cumulant of $\eta = \Phi(h + \sigma z)$ is nonnegative,

$$\mathbb{E}\eta^3 - 3\mathbb{E}\eta^2\mathbb{E}\eta + 2(\mathbb{E}\eta)^3 \geq 0. \tag{1.8}$$

Another interesting consequence of Corollary 1 is the following. If we define by continuity $f(0) = \mathbb{E}\eta = \mathbb{E}\Phi(h + \sigma z)$ and write $\lambda = \lambda \cdot 1 + (1-\lambda) \cdot 0$ then convexity of $f(m)$ implies

$$\mathbb{E}\exp(\lambda\eta) \leq (\mathbb{E}\exp\eta)^{\lambda^2} \exp(\lambda(1-\lambda)\mathbb{E}\eta). \tag{1.9}$$

If $A = \log \mathbb{E}\exp(\eta - \mathbb{E}\eta) < \infty$ then Chebyshev's inequality and (1.9) imply that

$$\mathbb{P}(\eta \geq \mathbb{E}\eta + t) \leq \mathbb{E}\exp(\lambda\eta - \lambda\mathbb{E}\eta - \lambda t) \leq \exp(\lambda^2 A - \lambda t)$$

and minimizing over $\lambda \in [0,1]$ we get,

$$\mathbb{P}(\eta \geq \mathbb{E}\eta + t) \leq \begin{cases} \exp(-t^2/4A), & t \leq 2A \\ \exp(A - t), & t \geq 2A. \end{cases} \tag{1.10}$$

This result can be slightly generalized.

Corollary 2 *If $\eta = \Phi(|\boldsymbol{h} + \boldsymbol{z}|)$ for some $\boldsymbol{h} \in \mathbb{R}^n$ and standard Gaussian $\boldsymbol{z} \in \mathbb{R}^n$ then the function $m^{-1} \log \mathbb{E}\exp m\eta$ is convex in m and, thus, (1.9) and (1.10) hold.*

The proof follows along the lines of the proof of Corollary 1 (or Theorem 1 in the simplest case of Corollary 1) and will be omitted.

2 Proof of Theorem 1.

The proof of Theorem 1 will be based on the following observations. First of all, we will compute the derivative of $\mathcal{P}_k$ with respect to q_l. We will need the following notations. For $0 \le l \le k$ we define

$$V_l = V_l(x, z_l) = \exp m_l(\Phi_{l+1}(x + z_l) - \Phi_l(x)). \tag{2.1}$$

Let $Z = h + z_0 + \ldots + z_k$ and $Z_l = h + z_0 + \ldots + z_{l-1}$ and define

$$X_l = \Phi_l(Z_l) \text{ and } W_l = V_l(Z_l, z_l) = \exp m_l(X_{l+1} - X_l).$$

Then the following holds.

Lemma 1 *For $1 \le l \le k$, we have,*

$$\frac{\partial \mathcal{P}_k}{\partial q_l} = -\frac{1}{2}(m_l - m_{l-1})\xi''(q_l)U_l \tag{2.2}$$

where

$$U_l = U_l(\boldsymbol{m}, \boldsymbol{q}) = \mathbb{E} W_1 \ldots W_{l-1}\Big(\mathbb{E}_l W_l \ldots W_k \Phi'(Z)\Big)^2. \tag{2.3}$$

Proof. The proof can be found in Lemma 3.6 in [7] (with slightly different notations).

□

It turns out that the function U_l is nondecreasing in each m_j which is the main ingredient in the proof of Theorem 1.

Theorem 2 *For any $1 \le l \le k$ the function U_l defined in (2.3) is nondecreasing in each m_j for $1 \le j \le k$.*

First, let us show how Lemma 1 and Theorem 2 imply Theorem 1.

Proof of Theorem 1. Let us assume that $n_j \le m_j$ for all $j \le k$. The opposite case can be handled similarly. If we define

$$\boldsymbol{m}^l = (n_0, \ldots, n_l, m_{l+1}, \ldots, m_k)$$

then

$$\mathcal{P}_k(\boldsymbol{n}) - \mathcal{P}_k(\boldsymbol{m}) = \sum_{0 \le l \le k} \big(\mathcal{P}_k(\boldsymbol{m}^l) - \mathcal{P}_k(\boldsymbol{m}^{l-1})\big).$$

We will prove that

$$\mathcal{P}_k(\boldsymbol{m}^l) - \mathcal{P}_k(\boldsymbol{m}^{l-1}) \ge \frac{\partial \mathcal{P}_k(\boldsymbol{m})}{\partial m_l}(n_l - m_l) \tag{2.4}$$

which, obviously, will prove Theorem 1. Let us consider vectors

$$\boldsymbol{m}_+^l = (n_0, \ldots, n_l, m_l, m_{l+1}, \ldots, m_k)$$

and

$$\boldsymbol{q}^l(t) = (q_0, \ldots, q_l, q_{l+1}(t), q_{l+1}, q_{l+2}, \ldots, q_k),$$

where $q_{l+1}(t) = q_l + t(q_{l+1} - q_l)$. Notice that we inserted one coordinate in vectors $\boldsymbol{m}^l$ and $\boldsymbol{q}$. For $0 \le t \le 1$, we consider

$$\varphi(t) = \mathcal{P}_{k+1}(\boldsymbol{m}_+^l, \boldsymbol{q}^l(t)).$$

It is easy to see that $\varphi(t)$ interpolates between $\varphi(1) = \mathcal{P}_k(\boldsymbol{m}^l)$ and $\varphi(0) = \mathcal{P}_k(\boldsymbol{m}^{l-1})$. By Lemma 1,

$$\varphi'(t) = -\frac{1}{2}(m_l - n_l)\xi''(q_{l+1}(t))U_{l+1}$$

where U_{l+1} is defined in terms of $\boldsymbol{m}^l_+$ and $\boldsymbol{q}^l(t)$. Next, let us consider

$$\boldsymbol{m}^l_\varepsilon = (m_0, \ldots, m_{l-1}, m_l - \varepsilon(m_l - n_l), m_l, m_{l+1}, \ldots, m_k)$$

and define

$$\varphi_\varepsilon(t) = \mathcal{P}_{k+1}(\boldsymbol{m}^l_\varepsilon, \boldsymbol{q}^l(t)).$$

First of all, we have $\varphi_\varepsilon(0) = \mathcal{P}_k(\boldsymbol{m})$ and $\varphi_\varepsilon(1) = \mathcal{P}_k(\boldsymbol{m}_\varepsilon)$, where

$$\boldsymbol{m}_\varepsilon = (m_0, \ldots, m_{l-1}, m_l - \varepsilon(m_l - n_l), m_{l+1}, \ldots, m_k).$$

Again, by Lemma 1,

$$\varphi'_\varepsilon(t) = -\frac{1}{2}\varepsilon(m_l - n_l)\xi''(q_{l+1}(t))U^\varepsilon_{l+1}$$

where U^ε_{l+1} is defined in terms of $\boldsymbol{m}^l_\varepsilon$ and $\boldsymbol{q}^l(t)$. It is obvious that for $\varepsilon \in [0,1]$ each coordinate of $\boldsymbol{m}^l_\varepsilon$ is not smaller than the corresponding coordinate of $\boldsymbol{m}^l$ and, therefore, Theorem 2 implies that $U_{l+1} \le U^\varepsilon_{l+1}$. This implies

$$\frac{1}{\varepsilon}\varphi'_\varepsilon(t) \le \varphi'(t)$$

and, therefore,

$$\frac{1}{\varepsilon}(\varphi_\varepsilon(1) - \varphi_\varepsilon(0)) \le \varphi(1) - \varphi(0)$$

which is the same as

$$\frac{1}{\varepsilon}(\mathcal{P}_k(\boldsymbol{m}_\varepsilon) - \mathcal{P}_k(\boldsymbol{m})) \le \mathcal{P}_k(\boldsymbol{m}^l) - \mathcal{P}_k(\boldsymbol{m}^{l-1}).$$

Letting $\varepsilon \to 0$ implies (2.4) and this finishes the proof of Theorem 1.

□

3 Proof of Theorem 2.

Let us start by proving some preliminary results. Consider two classes of (smooth enough) functions

$$\mathcal{C} = \{f : \mathbb{R} \to [0,\infty) : f(-x) = f(x), f'(x) \ge 0 \text{ for } x \ge 0\} \tag{3.1}$$

and

$$\mathcal{C}' = \{f : \mathbb{R} \to [0,\infty) : f(-x) = -f(x), f'(x) \ge 0 \text{ for } x \ge 0\}. \tag{3.2}$$

The next Lemma describes several facts that will be useful in the proof of Theorem 2.

Lemma 2 *For all $1 \leq l \leq k$ and $V_l = V_l(x, z_l)$ defined in (2.1) we have,*
(a) $\Phi_l(x)$ is convex, $\Phi_l(x) \in \mathcal{C}$ and

$$\Phi_l'(x) = \mathbb{E}_l V_l \ldots V_k \Phi'(x + z_l + \ldots + z_k) \in \mathcal{C}'.$$

(b) If $f_1 \in \mathcal{C}$ and $f_2 \in \mathcal{C}'$ then for $x \geq 0$

$$\mathbb{E}_l V_l f_1(x + z_l) f_2(x + z_l) \geq \mathbb{E}_l V_l f_1(x + z_l) \mathbb{E}_l V_l f_2(x + z_l).$$

(c) If $f(-x) = -f(x)$ and $f(x) \geq 0$ for $x \geq 0$ then $g(x) = \mathbb{E}_l V_l f(x + z_l)$ is such that

$$g(-x) = -g(x) \text{ and } g(x) \geq 0 \text{ if } x \geq 0.$$

(d) If $f \in \mathcal{C}$ then $\mathbb{E}_l V_l f(x + z_l) \in \mathcal{C}$. *(e) If $f \in \mathcal{C}'$ then $\mathbb{E}_l V_l f(x + z_l) \in \mathcal{C}'$.*
(f) $f(x) = \mathbb{E}_l V_l \log V_l \in \mathcal{C}$.

Proof. (a) Since Φ_{k+1} is convex, symmetric and nonnegative then $\Phi_l(x)$ is convex, symmetric and nonnegative by induction on l in (1.4). Convexity is the consequence of Hölder's inequality and the symmetry follows from the symmetry of Φ_{l+1} and the symmetry of the Gaussian distribution. Obviously, this implies that $\Phi_l'(x) \in \mathcal{C}'$.
(b) Let z_l' be an independent copy of z_l and, for simplicity of notations, let $\sigma^2 = \mathbb{E} z_l^2$. Since $\mathbb{E}_l V_l = 1$ (i.e. we can think of V_l as the change of density), we can write,

$$\begin{aligned} & \mathbb{E}_l V_l f_1(x + z_l) f_2(x + z_l) - \mathbb{E}_l V_l f_1(x + z_l) \mathbb{E}_l V_l f_2(x + z_l) = \\ = \; & \mathbb{E}_l V_l(x, z_l) V_l(x, z_l') \Big(f_1(x + z_l) - f_1(x + z_l') \Big) \Big(f_2(x + z_l) - f_2(x + z_l') \Big) I(z_l \geq z_l') \end{aligned} \tag{3.3}$$

Since $V_l(x, z_l) V_l(x, z_l') = \exp m_l(\Phi_l(x + z_l) + \Phi_l(x + z_l') - 2\Phi_l(x))$, if we make the change of variables $s = x + z_l$ and $t = x + z_l'$ then the right hand side of (3.3) can be written as

$$\frac{1}{2\pi\sigma^2} \exp(-2 m_l \Phi_l(x)) \int\limits_{\{s \geq t\}} K(s, t) \exp\Big(-\frac{1}{2\sigma^2} ((s - x)^2 + (t - x)^2) \Big) ds dt, \tag{3.4}$$

where

$$K(s, t) = \exp m_l(\Phi_l(s) + \Phi_l(t)) \Big(f_1(s) - f_1(t) \Big) \Big(f_2(s) - f_2(t) \Big).$$

We will split the region of integration $\{s \geq t\} = \Omega_1 \cup \Omega_2$ in the last integral into two disjoint sets

$$\Omega_1 = \{(s, t) : s \geq t, |s| \geq |t|\}, \;\; \Omega_2 = \{(s, t) : s \geq t, |s| < |t|\}.$$

In the integral over Ω_2 we will make the change of variables $s = -v, t = -u$ so that for $(s, t) \in \Omega_2$ we have $(u, v) \in \Omega_1$ and $ds dt = du dv$. Also,

$$K(s, t) = K(-v, -u) = -K(u, v)$$

since Φ_l is symmetric by (a), $f_1 \in \mathcal{C}, f_2 \in \mathcal{C}'$ and, therefore,

$$\Big(f_1(-v) - f_1(-u) \Big) \Big(f_2(-v) - f_2(-u) \Big) = -\Big(f_1(u) - f_1(v) \Big) \Big(f_2(u) - f_2(v) \Big).$$

Therefore,

$$\int_{\Omega_2} K(s,t)\exp\Big(-\frac{1}{2\sigma^2}((s-x)^2+(t-x)^2)\Big)dsdt = -\int_{\Omega_1} K(u,v)\exp\Big(-\frac{1}{2\sigma^2}((u+x)^2+(v+x)^2)\Big)dudv$$

and (3.4) can be rewritten as

$$\frac{1}{2\pi\sigma^2}\exp(-2m_l\Phi_l(x))\int_{\Omega_1} K(s,t)L(s,t,x)dsdt \tag{3.5}$$

where

$$L(s,t,x) = \exp\Big(-\frac{1}{2\sigma^2}((s-x)^2+(t-x)^2)\Big) - \exp\Big(-\frac{1}{2\sigma^2}((s+x)^2+(t+x)^2)\Big).$$

Since $f_1 \in \mathcal{C}$, for $(s,t)\in\Omega_1$ we have $f_1(s)-f_1(t) = f_1(|s|)-f_1(|t|)\geq 0$. Moreover, since for $(s,t)\in\Omega_1$ we have $t\leq s$, the fact that $f_2\in\mathcal{C}'$ implies that $f_2(s)-f_2(t)\geq 0$. Combining these two observations we get that $K(s,t)\geq 0$ on Ω_1. Finally, for $(s,t)\in\Omega_1$ we have $L(s,t,x)\geq 0$ because

$$(s-x)^2+(t-x)^2 \leq (s+x)^2+(t+x)^2 \Longleftrightarrow x(s+t)\geq 0,$$

and the latter holds because $x\geq 0$ and $s+t\geq 0$ on Ω_1. This proves that (3.5), (3.4) and, therefore, the right hand side of (3.3) are nonnegative.

(c) Let $g(x) = \mathbb{E}_l V_l(x,z_l) f(x+z_l)$. Then

$$g(-x) = \mathbb{E}_l V_l(-x,z_l)f(-x+z_l) = \mathbb{E}_l V(-x,-z_l)f(-x-z_l) = -\mathbb{E}_l V_l(x,z_l)f(x+z_l) = -g(x).$$

Next, if $x\geq 0$ and $\sigma^2 = \mathbb{E}_l z_l^2$ then

$$\begin{aligned} g(x) &= \exp(-m_l\Phi_l'(x))\mathbb{E}_l\exp(m_l\Phi_{l+1}(x+z_l))f(x+z_l) = \exp(-m_l\Phi_l'(x))\frac{1}{\sqrt{2\pi}\sigma}\times \\ &\quad\times\int_{s\geq 0}\exp(m_l\Phi_{l+1}(s))f(s)\Big(\exp\big(-\frac{1}{2\sigma^2}(x-s)^2\big)-\exp\big(-\frac{1}{2\sigma^2}(x+s)^2\big)\Big)ds\geq 0 \end{aligned}$$

because $(x-s)^2\leq (x+s)^2$ for $x,s\geq 0$ and $f(s)\geq 0$ for $s\geq 0$.

(d) Take $f\in\mathcal{C}$. Positivity of $\mathbb{E}_l V_l f(x+z_l)$ is obvious and symmetry follows from

$$\mathbb{E}_l V_l(-x,z_l)f(-x+z_l) = \mathbb{E}_l V_l(-x,-z_l)f(-x-z_l) = \mathbb{E}_l V_l(x,z_l)f(x+z_l). \tag{3.6}$$

Let $x\geq 0$. Recalling the definition (2.1), the derivative

$$\frac{\partial}{\partial x}\mathbb{E}_l V_l(x,z_l)f(x+z_l) = \mathrm{I} + m_l\mathrm{II}$$

where $\mathrm{I} = \mathbb{E}_l V_l(x,z_l)f'(x+z_l)$ and

$$\begin{aligned} \mathrm{II} &= \mathbb{E}_l V_l(x,z_l)f(x+z_l)(\Phi_{l+1}'(x+z_l)-\Phi_l'(x)) \\ &= \mathbb{E}_l V_l(x,z_l)f(x+z_l)\Phi_{l+1}'(x+z_l) - \mathbb{E}_l V_l(x,z_l)f(x+z_l)\mathbb{E}_l V_l\Phi_{l+1}'(x+z_l), \end{aligned}$$

since (1.4) yields that $\Phi_l'(x) = \mathbb{E}_l V_l(x,z_l)\Phi_{l+1}'(x+z_l)$. By (a), $\Phi_{l+1}'\in\mathcal{C}'$, and since $f\in\mathcal{C}$, (b) implies that $\mathrm{II}\geq 0$. The fact that $\mathrm{I}\geq 0$ for $x\geq 0$ follows from (c) because $f'(-x) = -f'(x)$ and $f'(x)\geq 0$ for $x\geq 0$.

(e) Take $f \in \mathcal{C}'$. Antisymmetry of $\mathbb{E}_l V_l f(x+z_l)$ follows from

$$\mathbb{E}_l V_l(-x,z_l) f(-x+z_l) = \mathbb{E}_l V_l(-x,-z_l) f(-x-z_l) = -\mathbb{E}_l V_l(x,z_l) f(x+z_l).$$

As in (d), the derivative can be written as

$$\frac{\partial}{\partial x}\mathbb{E}_l V_l(x,z_l) f(x+z_l) = \mathrm{I} + m_l \mathrm{II}$$

where $\mathrm{I} = \mathbb{E}_l V_l(x,z_l) f'(x+z_l)$ and

$$\mathrm{II} = \mathbb{E}_l V_l(x,z_l) f(x+z_l)\Phi'_{l+1}(x+z_l) - \mathbb{E}_l V_l(x,z_l) f(x+z_l)\mathbb{E}_l V_l \Phi'_{l+1}(x+z_l).$$

First of all, $\mathrm{I} \geq 0$ because $f' \geq 0$ for $f \in \mathcal{C}'$. As in (3.3) we can write

$$\mathrm{II} = \mathbb{E}_l V_l(x,z_l) V_l(x,z'_l)\big(f(x+z_l) - f(x+z'_l)\big)\big(\Phi'_{l+1}(x+z_l) - \Phi'_{l+1}(x+z'_l)\big) I(z_l \geq z'_l).$$

But both f and Φ'_{l+1} are in the class $\mathcal{C}'$ and, therefore, both nondecreasing which, obviously, implies that they are similarly ordered, i.e. for all $a,b \in \mathbb{R}$,

$$(f(a) - f(b))(\Phi'_{l+1}(a) - \Phi'_{l+1}(b)) \geq 0 \tag{3.7}$$

and as a result $\mathrm{II} \geq 0$.
(f) Symmetry of $g(x) = \mathbb{E}_l V_l \log V_l$ follows as above and positivity follows from Jensen's inequality, convexity of $x \log x$ and the fact that $\mathbb{E}_l V_l = 1$. Next, using that $\Phi'_l(x) = \mathbb{E}_l V_l \Phi'_{l+1}(x+z_l)$ we can write

$$\begin{aligned}
g'(x) &= m_l \mathbb{E}_l (1+\log V_l) V_l (\Phi'_{l+1}(x+z_l) - \Phi'_l(x)) \\
&= m_l^2 \mathbb{E}_l V_l (\Phi_{l+1}(x+z_l) - \Phi_l(x))(\Phi'_{l+1}(x+z_l) - \Phi'_l(x)) \\
&= m_l^2 \mathbb{E}_l V_l \Phi_{l+1}(x+z_l)(\Phi'_{l+1}(x+z_l) - \Phi'_l(x)) \\
&= m_l^2 \Big(\mathbb{E}_l V_l \Phi_{l+1}(x+z_l)\Phi'_{l+1}(x+z_l) - \mathbb{E}_l V_l \Phi_{l+1}(x+z_l)\mathbb{E}_l V_l \Phi'_{l+1}(x+z_l)\Big).
\end{aligned}$$

Since $\Phi_{l+1} \in \mathcal{C}$ and $\Phi'_{l+1} \in \mathcal{C}'$, (b) implies that for $x \geq 0$, $g'(x) \geq 0$ and, therefore, $g \in \mathcal{C}$.

□

Proof of Theorem 2.
We will consider two separate cases.
Case 1. $j \leq l-1$. First of all, using Lemma 2 (a) we can rewrite U_l as

$$U_l = \mathbb{E} W_1 \ldots W_{l-1} f_l(Z_l)$$

where

$$f_l(x) = (\Phi'_l(x))^2 \in \mathcal{C} \text{ since } \Phi'_l(x) \in \mathcal{C}'. \tag{3.8}$$

Using that

$$X_j = \frac{1}{m_j}\log \mathbb{E}_j \exp m_j X_{j+1}$$

we get

$$\frac{\partial X_j}{\partial m_j} = \frac{1}{m_j}\mathbb{E}_j W_j X_{j+1} - \frac{1}{m_j^2}\log \mathbb{E}_j \exp m_j X_{j+1} = \frac{1}{m_j}\mathbb{E}_j W_j (X_{j+1} - X_j).$$

For $p \le j$, we get

$$\frac{\partial X_p}{\partial m_j} = \frac{1}{m_j}\mathbb{E}_p W_p \ldots W_j (X_{j+1} - X_j),$$

and for $p > j$, X_p does not depend on m_j. Therefore,

$$\frac{\partial}{\partial m_j} W_1 \ldots W_{l-1} = \frac{\partial}{\partial m_j} \exp\Bigl(\sum_{p \le l-1} m_p (X_{p+1} - X_p)\Bigr)$$
$$= W_1 \ldots W_{l-1}\Bigl((X_{j+1} - X_j) - \frac{1}{m_j}\sum_{p \le j}(m_p - m_{p-1})\mathbb{E}_p W_p \ldots W_j (X_{j+1} - X_j)\Bigr).$$

Hence,

$$\begin{aligned} m_j \frac{\partial U_l}{\partial m_j} &= m_j \mathbb{E} W_1 \ldots W_{l-1} f_l(Z_l)(X_{j+1} - X_j) \\ &\quad - \sum_{p \le j}(m_p - m_{p-1})\mathbb{E} W_1 \ldots W_{l-1} f_l(Z_l)\mathbb{E}_p W_p \ldots W_j (X_{j+1} - X_j). \end{aligned}$$

If we denote $f_j(Z_{j+1}) = \mathbb{E}_{j+1} W_{j+1} \ldots W_{l-1} f_l(Z_l)$ then we can rewrite

$$\begin{aligned} m_j \frac{\partial U_l}{\partial m_j} &= m_j \mathbb{E} W_1 \ldots W_j f_j(Z_{j+1})(X_{j+1} - X_j) \\ &\quad - \sum_{p \le j}(m_p - m_{p-1})\mathbb{E} W_1 \ldots W_{p-1}\mathbb{E}_p W_p \ldots W_j f_j(Z_{j+1})\mathbb{E}_p W_p \ldots W_j (X_{j+1} - X_j). \end{aligned} \tag{3.9}$$

First of all, let us show that

$$\mathbb{E}_j W_j f_j(Z_{j+1})(X_{j+1} - X_j) \ge \mathbb{E}_j W_j f_j(Z_{j+1})\mathbb{E}_j W_j (X_{j+1} - X_j). \tag{3.10}$$

Since X_j does not depend on z_j and $\mathbb{E}_j W_j = 1$, this is equivalent to

$$\mathbb{E}_j W_j f_j(Z_{j+1}) X_{j+1} \ge \mathbb{E}_j W_j f_j(Z_{j+1})\mathbb{E}_j W_j X_{j+1}. \tag{3.11}$$

Here f_j and X_{j+1} are both functions of $Z_{j+1} = Z_j + z_j$. Since by (3.8), $f_l(Z_l)$ seen as a function of Z_l is in $\mathcal{C}$, applying Lemma 2 (d) inductively we get that $f_j(Z_{j+1})$ seen as a function of Z_{j+1} is also in $\mathcal{C}$. By Lemma 2 (a), X_{j+1} seen as a function of Z_{j+1} is also in $\mathcal{C}$. Therefore, f_j and X_{j+1} are similarly ordered i.e.

$$\bigl(f_j(Z_{j+1}) - f_j(Z'_{j+1})\bigr)\bigl(X_{j+1}(Z_{j+1}) - X_{j+1}(Z'_{j+1})\bigr) \ge 0$$

and, therefore, using the same trick as in (3.3) we get (3.11) and, hence, (3.10). By Lemma 2 (d), $\mathbb{E}_j W_j f_j(Z_{j+1})$ seen as a function of Z_j is in $\mathcal{C}$ and by Lemma 2 (f), $\mathbb{E}_j W_j (X_{j+1} - X_j) = m_j^{-1}\mathbb{E}_j W_j \log W_j$ seen as a function of Z_j is also in $\mathcal{C}$. Therefore, they are similarly ordered and again

$$\mathbb{E}_p W_p \ldots W_{j-1}\mathbb{E}_j W_j f_j(Z_{j+1})\mathbb{E}_j W_j (X_{j+1} - X_j) \ge \mathbb{E}_p W_p \ldots W_j f_j(Z_{j+1})\mathbb{E}_p W_p \ldots W_j (X_{j+1} - X_j).$$

Combining this with (3.10) implies that

$$\mathbb{E} W_1 \ldots W_j f_j(Z_{j+1})(X_{j+1} - X_j) \ge \mathbb{E} W_1 \ldots W_{p-1}\mathbb{E}_p W_p \ldots W_j f_j(Z_{j+1})\mathbb{E}_p W_p \ldots W_j (X_{j+1} - X_j).$$

Since $m_j = \sum_{p\le j}(m_p - m_{p-1})$, this and (3.9) imply that $\partial U_l/\partial m_j \ge 0$ which completes the proof of Case 1.

□

Case 2. $j \ge l$. If we denote

$$g_l = g_l(Z_l) = \mathbb{E}_l W_l \dots W_l \Phi'(Z),\;\; f_l = f_l(Z_l) = g_l^2$$

then a straightforward calculation similar to the one leading to (3.9) gives

$$\begin{aligned} m_j\frac{\partial U_l}{\partial m_j} = & - \sum_{p\le l-1}(m_p - m_{p-1})\mathbb{E}W_1\dots W_{l-1} f_l \mathbb{E}_p W_p \dots W_j (X_{j+1} - X_j) \\ & - (2m_l - m_{l-1})\mathbb{E}W_1 \dots W_{l-1} f_l \mathbb{E}_l W_l \dots W_j (X_{j+1} - X_j) \\ & - \sum_{l+1\le p\le j} 2(m_p - m_{p-1})\mathbb{E}W_1\dots W_{l-1} g_l \mathbb{E}_l W_l \dots W_k \Phi'(Z)\mathbb{E}_p W_p \dots W_j (X_{j+1} - X_j) \\ & + 2m_j \mathbb{E}W_1\dots W_{l-1} g_l \mathbb{E}_l W_l \dots W_k \Phi'(Z)(X_{j+1} - X_j). \end{aligned} \tag{3.12}$$

To show that this is positive we notice that

$$2m_j = \sum_{p\le l-1}(m_p - m_{p-1}) + (2m_l - m_{l-1}) + \sum_{l+1\le p\le j} 2(m_p - m_{p-1})$$

and we will show that the last term with factor $2m_j$ is bigger than all other terms with negative factors. If we denote

$$h(Z_{j+1}) = \mathbb{E}_{j+1} W_{j+1}\dots W_k \Phi'(Z)$$

then since $\Phi' \in \mathcal{C}'$, using Lemma 2 (e) inductively, we get that $h(Z_{j+1})$ seen as a function of Z_{j+1} is in $\mathcal{C}'$. Each term in the third line of (3.12) (without the factor $2(m_p - m_{p-1})$) can be rewritten as

$$\mathbb{E}W_1\dots W_{l-1} g_l \mathbb{E}_l W_l \dots W_{p-1}\mathbb{E}_p W_p \dots W_j h(Z_{j+1})\mathbb{E}_p W_p \dots W_j (X_{j+1} - X_j), \tag{3.13}$$

the term in the second line of (3.12) (without the factor $2m_l - m_{l-1}$)) is equal to (3.13) for $p = l$, and the term in the fourth line (without $2m_j$) can be written as

$$\mathbb{E}W_1\dots W_{l-1} g_l \mathbb{E}_l W_l \dots W_j h(Z_{j+1})(X_{j+1} - X_j). \tag{3.14}$$

We will show that (3.14) is bigger than (3.13) for $l \le p \le j$. This is rather straightforward using Lemma 2. Notice that $g_l = g_l(Z_l)$ seen as a function of Z_l is in $\mathcal{C}'$ by Lemma 2 (a). If we define for $l \le p \le j$,

$$r_p(Z_l) = \mathbb{E}_l W_l \dots W_{p-1}\mathbb{E}_p W_p \dots W_j h(Z_{j+1})\mathbb{E}_p W_p \dots W_j (X_{j+1} - X_j)$$

and

$$r(Z_l) = \mathbb{E}_l W_l \dots W_j h(Z_{j+1})(X_{j+1} - X_j)$$

then the difference of (3.14) and (3.13) is

$$\mathbb{E}W_1\dots W_{l-1} g_l(Z_l)(r(Z_l) - r_p(Z_l)). \tag{3.15}$$

Using the argument similar to (3.6) (and several other places above), it should be obvious that $r_p(-Z_l) = -r_p(Z_l)$ since X_i's are symmetric and h is antisymmetric. Similarly, $r(-Z_l) = -r(Z_l)$. Therefore, if we can show that

$$r(Z_l) - r_p(Z_l) \ge 0 \text{ for } Z_l \ge 0 \tag{3.16}$$

then, since $g_l \in \mathcal{C}'$, we would get that

$$g_l(Z_l)(r(Z_l) - r_p(Z_l)) \geq 0 \text{ for all } Z_l$$

and this would prove that (3.15) is nonnegative. Let us first show that (3.16) holds for $p = j$. In this case, since X_j does not depend on z_j and, therefore, $\mathbb{E}_j W_j X_j = X_j$, (3.16) is equivalent to

$$\mathbb{E}_l W_l \ldots W_{j-1} \mathbb{E}_j W_j h(Z_{j+1}) X_{j+1} \geq \mathbb{E}_l W_l \ldots W_{j-1} \mathbb{E}_j W_j h(Z_{j+1}) \mathbb{E}_j W_j X_{j+1}, \tag{3.17}$$

for $Z_l \geq 0$. Let us define

$$\Delta_j(Z_j) = \mathbb{E}_j W_j h(Z_{j+1}) X_{j+1} - \mathbb{E}_j W_j h(Z_{j+1}) \mathbb{E}_j W_j X_{j+1}.$$

As above, $\Delta_j(-Z_j) = -\Delta_j(Z_j)$ and by Lemma 2 (b), $\Delta_j(Z_j) \geq 0$ for $Z_j \geq 0$, since $h \in \mathcal{C}'$ and $X_{j+1} \in \mathcal{C}$. Therefore, by Lemma 2 (c),

$$\Delta_{j-1}(Z_{j-1}) := \mathbb{E}_{j-1} W_{j-1} \Delta_j(Z_{j-1} + z_j) \geq 0 \text{ if } Z_{j-1} \geq 0$$

and, easily, $\Delta_{j-1}(-Z_{j-1}) = -\Delta_{j-1}(Z_{j-1})$. Therefore, if for $i \geq l$ we define

$$\Delta_i(Z_i) = \mathbb{E}_i W_i \Delta_{i+1}(Z_i + z_i)$$

we can proceed by induction to show that $\Delta_i(-Z_i) = -\Delta_i(Z_i)$ and $\Delta_i(Z_i) \geq 0$ for $Z_i \geq 0$. For $i = l$ this proves (3.17) and, therefore, (3.16) for $p = j$. Next, we will show that

$$r_{p+1}(Z_l) - r_p(Z_l) \geq 0 \text{ for } Z_l \geq 0 \tag{3.18}$$

for all $l \leq p < j$, and this, of course, will prove (3.16). If we define

$$f_1(Z_{p+1}) = \mathbb{E}_{p+1} W_{p+1} \ldots W_j h(Z_{j+1}) \text{ and } f_2(Z_{p+1}) = \mathbb{E}_{p+1} W_{p+1} \ldots W_j (X_{j+1} - X_j)$$

then (3.18) can be rewritten as

$$\mathbb{E}_l W_l \ldots W_{p-1} \mathbb{E}_p W_p f_1(Z_{p+1}) f_2(Z_{p+1}) \geq \mathbb{E}_l W_l \ldots W_{p-1} \mathbb{E}_p W_p f_1(Z_{p+1}) \mathbb{E}_p f_2(Z_{p+1}) \text{ for } Z_l \geq 0.$$

Since $h(Z_{j+1}) \in \mathcal{C}'$, recursive application of Lemma 2 (e) implies that $f_1(Z_{p+1}) \in \mathcal{C}'$. Since $\mathbb{E}_j W_j (X_{j+1} - X_j) = m_j^{-1} \mathbb{E}_j W_j \log W_j$ seen as a function of Z_j is in $\mathcal{C}$ by Lemma 2 (f), recursive application of Lemma 2 (d) implies that $f_2(Z_{p+1}) \in \mathcal{C}$. If we now define

$$\Delta_p(Z_p) = \mathbb{E}_p W_p f_1(Z_{p+1}) f_2(Z_{p+1}) - \mathbb{E}_p W_p f_1(Z_{p+1}) \mathbb{E}_p W_p f_2(Z_{p+1}),$$

then, as above, $\Delta_p(-Z_p) = -\Delta_p(Z_p)$ and by Lemma 2 (b), $\Delta_p(Z_p) \geq 0$ for $Z_p \geq 0$, since $f_1 \in \mathcal{C}'$ and $f_2 \in \mathcal{C}$. Therefore, by Lemma 2 (c),

$$\Delta_{p-1}(Z_{p-1}) := \mathbb{E}_{p-1} W_{p-1} \Delta_p(Z_{p-1} + p_j) \geq 0 \text{ if } Z_{p-1} \geq 0$$

and, easily, $\Delta_{p-1}(-Z_{p-1}) = -\Delta_{p-1}(Z_{p-1})$. Therefore, if for $i \geq l$ we define

$$\Delta_i(Z_i) = \mathbb{E}_i W_i \Delta_{i+1}(Z_i + z_i)$$

we can proceed by induction to show that $\Delta_i(-Z_i) = -\Delta_i(Z_i)$ and $\Delta_i(Z_i) \geq 0$ for $Z_i \geq 0$. For $i = l$ this proves (3.18). Thus, we finally proved that (3.14) is bigger than (3.13) for $p \geq l$. To prove that (3.12) is nonnegative it remains to show that each term in the first line of (3.12) (without the factor $-(m_p - m_{p-1})$) is smaller than (3.14). Clearly, it is enough to show that

$$\mathbb{E}W_1 \dots W_{l-1} f_l \mathbb{E}_p W_p \dots W_j (X_{j+1} - X_j) \le \mathbb{E}W_1 \dots W_{l-1} f_l \mathbb{E}_l W_l \dots W_j (X_{j+1} - X_j) \quad (3.19)$$

since the right hand side of (3.19) is equal to (3.13) for $p = l$ which was already shown to be smaller than (3.14). The proof of (3.19) can be carried out using the same argument as in the proof of (3.10) in Case 1 and this finishes the proof of Case 2.

□

References

[1] F. Guerra. Broken replica symmetry bounds in the mean field spin glass model. *Comm. Math. Phys.* **233** no. 1 (2003), 1-12.

[2] D. Sherrington, S. Kirkpatrick. Solvable model of a spin glass. *Phys. Rev. Lett.* **35** (1972), 1792-1796.

[3] M. Talagrand. *Spin Glasses: a Challenge for Mathematicians.* (2003) Springer-Verlag.

[4] M. Talagrand. The generalized Parisi formula. *C. R. Math. Acad. Sci. Paris* **337** no. 2 (2003), 111-114.

[5] M. Talagrand. Parisi formula. To appear in *Ann. Math.* (2003).

[6] M. Talagrand. On the meaning of Parisi's functional order parameter. *C. R. Math. Acad. Sci. Paris* **337** no. 9 (2003), 625-628.

[7] M. Talagrand. Parisi measures. Preprint (2004).