On differentiability of the Parisi formula.

Dmitry Panchenko*
Department of Mathematics
Texas A&M University

Abstract

It was proved by Michel Talagrand in [10] that the Parisi formula for the free energy in the Sherrington-Kirkpatrick model is differentiable with respect to inverse temperature parameter. We present a simpler proof of this result by using approximate solutions in the Parisi formula and give one example of application of the differentiability to prove non self-averaging of the overlap outside of the replica symmetric region.

Key words: Sherrington-Kirkpatrick model, Parisi formula.
Mathematics Subject Classification: 60K35, 82B44

1.1 Introduction and main results.

Let us consider a p-spin Sherrington-Kirkpatrick Hamiltonian

$$H_{N,p}(\boldsymbol{\sigma}) = \frac{1}{N^{(p-1)/2}} \sum_{1\leq i_1,\ldots,i_p\leq N} g_{i_1,\ldots,i_p}\sigma_{i_1}\ldots\sigma_{i_p}$$

indexed by spin configurations $\boldsymbol{\sigma} \in \Sigma_N = \{-1,+1\}^N$ where $(g_{i_1,\ldots,i_p})$ are i.i.d. standard Gaussian random variables. A mixed p-spin Hamiltonian is defined as the sum

$$H_N(\boldsymbol{\sigma}) = \sum_{p\geq 1} \beta_p\, H_{N,p}(\boldsymbol{\sigma}) \tag{1.1}$$

over a finite set of indices $p \geq 1$. The covariance of H_N can be easily computed

$$\mathbb{E} H_N(\boldsymbol{\sigma}^1) H_N(\boldsymbol{\sigma}^2) = N\xi(R_{1,2}), \tag{1.2}$$

where

$$R_{1,2} = \frac{1}{N}\sum_{i\leq N} \sigma_i^1\sigma_i^2 \quad \text{and} \quad \xi(x) = \sum_{p\geq 1}\beta_p^2 x^p.$$

A quantity $R_{1,2}$ is called the overlap of configurations $\boldsymbol{\sigma}^1, \boldsymbol{\sigma}^2$. To avoid the trivial case when all the spins decouple we assume that $\beta_p \neq 0$ for at least one $p \geq 2$ so that $\xi''(x) > 0$ for $x > 0$. Given an external field parameter $h \in \mathbb{R}$, the free energy is defined by

$$F_N(\boldsymbol{\beta}) = \frac{1}{N}\,\mathbb{E}\log\sum_{\boldsymbol{\sigma}}\exp\bigl(H_N(\boldsymbol{\sigma}) + h\sum_{i\leq N}\sigma_i\bigr). \tag{1.3}$$

The problem of computing the thermodynamic limit of the free energy $\lim_{N\to\infty} F_N$ is one of the central questions in the analysis of the SK model and the value of this limit was predicted by Giorgio Parisi in [5]

*Department of Mathematics, Texas A&M University, email: panchenk@math.tamu.edu. This work is partially supported by NSF grant.

as a part of his celebrated theory that goes far beyond the computation of the free energy. The prediction of Parisi was confirmed with mathematical rigor by Michel Talagrand in [11] following a breakthrough of Francesco Guerra in [2] where a replica symmetry breaking interpolation was introduced. Validity of the Parisi formula provides a lot of information about the model and, in particular, about the distribution of the overlap under the Gibbs measure corresponding to the Hamiltonian $H_N(\boldsymbol{\sigma})$. In the next section we will show one important application of the Parisi formula which is based on its differentiability with respect to inverse temperature parameters. Namely, we will prove a stronger version of the result of Pastur and Shcherbina in [6] about the non self-averaging of the overlap at low temperature.

In the remainder of this section we present a simplified version of the argument of Talagrand in [10] and prove the differentiability of the Parisi formula. Let us start by recalling the definition of the Parisi formula. Let $\mathcal{M}$ be the set of cumulative distribution functions on $[0,1]$. We will identify a c.d.f. m with a distribution it defines and simply call m itself a distribution on $[0,1]$. A distribution with at most k atoms is defined by

$$m(q) = \sum_{0\leq l\leq k} m_l I(q_l \leq q < q_{l+1}) \tag{1.4}$$

for some sequences

$$0 = m_0 \leq m_1 \leq \ldots \leq m_{k-1} \leq m_k = 1,$$
$$0 = q_0 \leq q_1 \leq \ldots \leq q_k \leq q_{k+1} = 1.$$

Consider independent Gaussian r.v. $(z_l)_{0\leq l\leq k}$ such that $\mathbb{E}z_l^2 = \xi'(q_{l+1}) - \xi'(q_l)$. Let

$$X_k = \log \mathrm{ch}\Big(\sum_{0\leq l\leq k} z_l + h\Big)$$

and recursively for $1 \leq l \leq k$ define

$$X_{l-1} = \frac{1}{m_l}\log \mathbb{E}_l \exp m_l X_l \tag{1.5}$$

where $\mathbb{E}_l$ denotes the expectation in (z_p) for $l \leq p \leq k$. Define

$$\mathcal{P}(m,\boldsymbol{\beta}) = \mathbb{E}X_0 - \frac{1}{2}\sum_{1\leq l\leq k} m_l(\theta(q_{l+1}) - \theta(q_l)). \tag{1.6}$$

where $\theta(x) = x\xi'(x) - \xi(x)$. On the set of discrete $m \in \mathcal{M}$ as in (1.4) the functional $\mathcal{P}(m,\boldsymbol{\beta})$ is Lipschitz in m with respect to L_1 norm (see [2], [10]). Therefore, it can be extended by continuity to a Lipschitz functional on the entire space $\mathcal{M}$. The Parisi formula is then defined by

$$\mathcal{P}(\boldsymbol{\beta}) = \inf_{m\in\mathcal{M}} \mathcal{P}(m,\boldsymbol{\beta}). \tag{1.7}$$

This infimum is obviously achieved by continuity and compactness. Any $m \in \mathcal{M}$ that achieves the infimum is called a Parisi measure. It is conjectured ([4]) that $\mathcal{P}(m,\boldsymbol{\beta})$ is convex in m in which case the Parisi measure would be unique.

By Hölder's inequality, $F_N(\boldsymbol{\beta})$ is convex in $\boldsymbol{\beta}$ and, thus, its limit $\mathcal{P}(\boldsymbol{\beta})$ is also convex. Convexity implies that $\mathcal{P}(\boldsymbol{\beta})$ is differentiable in each parameter β_p almost everywhere and it was proved in [10] that $\mathcal{P}(\boldsymbol{\beta})$ is in fact differentiable for all values of β_p. The proof was based on a careful analysis of the functional $\mathcal{P}(m,\boldsymbol{\beta})$ in the neighborhood of a Parisi measure and parts of the proof were rather technical due to the fact that a Parisi measure is not necessarily discrete. We will prove a slightly weaker analogue of Theorem 1.2 in [10] but we will bypass these difficulties by working with approximations of a Parisi measure by discrete measures of the type (1.4). The main difference is that we express the derivative in (1.8) below in terms of some Parisi measure instead of any Parisi measure as in [10].

Theorem 1 *The derivative of the Parisi formula $\mathcal{P}(\boldsymbol{\beta})$ with respect to any β_p exists and*

$$\frac{\partial \mathcal{P}(\boldsymbol{\beta})}{\partial \beta_p} = \beta_p\Big(1 - \int q^p dm_{\boldsymbol{\beta}}(q)\Big) \quad \textit{for all} \quad p \geq 1 \tag{1.8}$$

for some Parisi measure $m_{\boldsymbol{\beta}}$.

To prove Theorem 1 we will first obtain a similar statement for discrete approximations of a Parisi measure; this result corresponds to Proposition 3.2 in [10].

Lemma 1 *Given $k \geq 1$, suppose that $m \in \mathcal{M}$ achieves the minimum of $\mathcal{P}(m, \boldsymbol{\beta})$ over all distributions with at most k atoms as in (1.4). Then*

$$\frac{\partial \mathcal{P}}{\partial \beta_p}(m, \boldsymbol{\beta}) = \beta_p \Big(1 - \int q^p dm(q)\Big).$$

Proof. Suppose that m has k' atoms in $(0,1)$ for some $k' \leq k$. For simplicity of notations, let us assume that $k' = k$. Let us start by noting that $\mathbb{E}X_0$ depends on $\boldsymbol{\beta}$ only through $\xi'(1)$ and $\xi'(q_l)$ for $1 \leq l \leq k$. Let us make the dependence on $\xi'(1)$ explicit. Since

$$X_{k-1} = \log \mathrm{ch}\Big(\sum_{0 \leq l \leq k-1} z_l + h\Big) + \frac{1}{2}(\xi'(1) - \xi'(q_k))$$

we can continue recursive construction (1.5) to show that

$$\mathbb{E}X_0 = \frac{1}{2}\,\xi'(1) + \frac{1}{2}\,f(\xi'(q_1), \ldots, \xi'(q_k))$$

for some smooth function $f(x_1, \ldots, x_k) : \mathbb{R}^k \to \mathbb{R}$. Then, rearranging the terms in (1.6)

$$\mathcal{P}(m, \boldsymbol{\beta}) = \frac{1}{2}\,\xi(1) + \frac{1}{2}\,f(\xi'(q_1), \ldots, \xi'(q_k)) + \frac{1}{2}\sum_{1 \leq l \leq k}(m_l - m_{l-1})\theta(q_l). \tag{1.9}$$

Since m achieves the minimum, for $1 \leq l \leq k$

$$2\frac{\partial \mathcal{P}}{\partial q_l} = \frac{\partial f}{\partial x_l}\,\xi''(q_l) + (m_l - m_{l-1})q_l \xi''(q_l) = 0$$

and since $\xi''(q) > 0$ for $q > 0$ this implies that

$$\frac{\partial f}{\partial x_l} = -(m_l - m_{l-1})q_l. \tag{1.10}$$

Since

$$\xi(q) = \sum_{p \geq 1} \beta_p^2\, q^p, \;\; \xi'(q) = \sum_{p \geq 1} p\, \beta_p^2\, q^{p-1} \;\text{ and }\; \theta(q) = \sum_{p \geq 1}(p-1)\beta_p^2\, q^p,$$

using (1.9) and (1.10) we compute

$$\begin{aligned} \frac{\partial \mathcal{P}}{\partial \beta_p} &= \beta_p + \sum_{1 \leq l \leq k} \frac{\partial f}{\partial x_l}\, p\, \beta_p\, q_l^{p-1} + \sum_{1 \leq l \leq k}(m_l - m_{l-1})(p-1)\beta_p\, q_l^p \\ &= \beta_p - \beta_p \sum_{1 \leq l \leq k}(m_l - m_{l-1})q_l^p = \beta_p\Big(1 - \int q^p dm(q)\Big) \end{aligned}$$

and this finishes the proof.

□

Proof of Theorem 1. First of all, let us fix all but one parameter in $\boldsymbol{\beta}$ and think of all the functions that depend on $\boldsymbol{\beta}$ as functions of one variable $\beta = \beta_p$. Let m^k be a distribution from Lemma 1. By definition of Parisi formula and Lipschitz property of $\mathcal{P}(m, \beta)$ we have $\mathcal{P}(m^k, \beta) \downarrow \mathcal{P}(\beta)$ as $k \to \infty$ or, in other words,

$$0 \leq \mathcal{P}(m^k, \beta) - \mathcal{P}(\beta) \leq \varepsilon_k \tag{1.11}$$

for some sequence $\varepsilon_k \downarrow 0$. To prove that a convex function $\mathcal{P}(\beta)$ is differentiable we need to show that its subdifferential $\partial\mathcal{P}(\beta)$ contains a unique point. Let $a \in \partial\mathcal{P}(\beta)$. Then by convexity of $\mathcal{P}$, (1.11) and the fact that $\mathcal{P}(\beta') \leq \mathcal{P}(m^k, \beta')$ for all β',

$$a \leq \frac{\mathcal{P}(\beta + y) - \mathcal{P}(\beta)}{y} \leq \frac{\mathcal{P}(m^k, \beta + y) - \mathcal{P}(m^k, \beta) + \varepsilon_k}{y}$$

and

$$a \geq \frac{\mathcal{P}(\beta) - \mathcal{P}(\beta - y)}{y} \geq \frac{\mathcal{P}(m^k, \beta) - \mathcal{P}(m^k, \beta - y) - \varepsilon_k}{y}$$

for $y > 0$. It is a simple exercise to check that for any discrete $m \in \mathcal{M}$ the second derivative $\partial^2\mathcal{P}(m, \beta)/\partial\beta^2$ stays bounded if β stays bounded and the bound is uniform in m (see [11] or [10]). Therefore, using Taylor's expansion around $y = 0$ on the right hand side of the above inequalities gives

$$\frac{\partial\mathcal{P}}{\partial\beta}(m^k, \beta) - Ly - \frac{\varepsilon_k}{y} \leq a \leq \frac{\partial\mathcal{P}}{\partial\beta}(m^k, \beta) + Ly + \frac{\varepsilon_k}{y}.$$

Taking $y = \sqrt{\varepsilon_k}$ we obtain

$$a = \frac{\partial\mathcal{P}}{\partial\beta}(m^k, \beta) + \mathcal{O}(\sqrt{\varepsilon_k}) = \beta\Big(1 - \int q^p dm^k(q)\Big) + \mathcal{O}(\sqrt{\varepsilon_k})$$

by Lemma 1. Finally, taking a subsequence of (m^k) that converges in L_1 norm to some Parisi measure $m_{\boldsymbol{\beta}}$ proves that

$$a = \beta\Big(1 - \int q^p dm_{\boldsymbol{\beta}}(q)\Big).$$

This uniquely determines a and, thus, $a = \mathcal{P}'(\beta)$.

□

1.2 Non self-averaging of the overlap.

In this section we make an assumption that all indices in (1.1) are even numbers with one possible exception of $p = 1$, i.e. besides a trivial linear term we consider only even spin interaction terms. The reason for this is because the validity of the Parisi formula was proved in [11] under certain conditions on the function ξ which essentially correspond to the choice of only even spin interaction terms. Under this assumption, by [11],

$$\lim_{N\to\infty} F_N(\boldsymbol{\beta}) = \mathcal{P}(\boldsymbol{\beta})$$

and since both $F_N(\boldsymbol{\beta})$ and $\mathcal{P}(\boldsymbol{\beta})$ are convex functions and, by Theorem 1, $\mathcal{P}(\boldsymbol{\beta})$ is differentiable in β_p, we get

$$\lim_{N\to\infty} \frac{\partial F_N}{\partial\beta_p} = \frac{\partial\mathcal{P}}{\partial\beta_p} = \beta_p\Big(1 - \int q^p dm_{\boldsymbol{\beta}}(q)\Big).$$

By Gaussian integration by parts one can easily see that,

$$\frac{\partial F_N}{\partial\beta_p} = \beta_p\Big(1 - \mathbb{E}\big\langle R_{1,2}^p\big\rangle\Big)$$

where $\langle\cdot\rangle$ is the Gibbs average with respect to the Hamiltonian $H_N(\boldsymbol{\sigma})$ and, therefore, for any $p \geq 1$ such that $\beta_p > 0$ we get

$$\lim_{N\to\infty} \mathbb{E}\langle R_{1,2}^p\rangle = \int q^p dm_{\boldsymbol{\beta}}(q). \tag{1.12}$$

Thus, from Theorem 1 one obtains information about moments of the overlap, in particular, about the existence of their thermodynamic limit. (This result is not new, it appears in [9] and [10].) If Hamiltonian $H_N(\boldsymbol{\sigma})$ contains all even p-spin interaction terms then (1.12) holds for all even $p \geq 2$ and, thus, the distribution of

$|R_{1,2}|$ is approximated by the Parisi measure $m_{\boldsymbol{\beta}}$. It is predicted by the Parisi theory that this is also true when only a finite number of even p-spin interaction terms are present; however, this is an open problem. (1.12) provides information only about the moments of the overlap corresponding to the terms present in the Hamiltonian.

We will now use this information to give two examples of non self-averaging of the overlap. To put these examples in perspective, let us first recall several well-known results about the classical 2-spin SK model, $H_N = \beta H_{N,2}$, without external field, $h = 0$. Let us recall that inverse temperature parameter β is said to belong to replica symmetric region if the infimum in the Parisi formula (1.7) is achieved on Dirac measure δ_0 concentrated at zero. In this simplest case the Parisi formula $\mathcal{P}(\beta)$ is called a replica symmetric solution. It was proved by Aizenman, Lebowitz and Ruelle in [1] that replica symmetric solution holds for $\beta^2 \leq 2$ and it was proved by Toninelli in [12] that it does not hold for $\beta^2 > 2$ (the result in [12] is more general, it also covers the case with external field). In other words, the set of $\beta^2 \leq 2$ is the replica symmetric region. Note that the reason we have $\beta^2 \leq 2$ instead of a more familiar $\beta^2 \leq 1$ is because for simplicity we defined the Hamiltonian $H_{N,2}$ as the sum over all indices i_1 and i_2 rather than $i_1 < i_2$. A well-known result of Pastur and Shcherbina in [6] states that if

$$\lim_{N\to\infty} \mathbb{E}(\langle R_{1,2}\rangle - \mathbb{E}\langle R_{1,2}\rangle)^2 = 0 \tag{1.13}$$

then replica symmetric solution holds. Therefore, for $\beta^2 > 2$ (1.13) can not hold and this implies that $\limsup_{N\to\infty} \mathbb{E}\langle R_{1,2}^2\rangle > 0$. Differentiability of the Parisi formula implies that the limit $\lim_{N\to\infty} \mathbb{E}\langle R_{1,2}^2\rangle$ in (1.12) exists and, consequently, the result of Pastur and Shcherbina can be used to deduce that this limit is strictly positive when $\beta^2 > 2$. However, one can give a more direct proof of a more general result without invoking [6].

Example 1 ($h = 0, \beta_1 = 0$). This case is similar to the classical SK model without external field, only now p-spin interactions for even $p > 2$ are also allowed. A replica symmetric region is again defined as the set of parameters $\boldsymbol{\beta}$ such that the infimum in (1.7) is achieved on Dirac measure δ_0 concentrated at zero, but the description of this region is slightly more complicated (see Theorem 2.11.16 in [8]). Using the continuity of the functional $m \to \mathcal{P}(m, \boldsymbol{\beta})$ with respect to the L_1 norm (see [2], [10]), outside of the replica symmetric region any Parisi measure $m_{\boldsymbol{\beta}}$ must satisfy $m_{\boldsymbol{\beta}}(\{q > 0\}) > 0$. Therefore, by (1.12), for any even $p \geq 2$ such that $\beta_p > 0$ we have

$$\lim_{N\to\infty} \mathbb{E}\langle R_{1,2}^p\rangle > 0. \tag{1.14}$$

Since by symmetry, $\langle R_{1,2}\rangle = 0$, this proves non self-averaging of the overlap outside of the replica symmetric region.

Example 2 ($h \neq 0$, $\beta_{p_1}, \beta_{p_2} \neq 0$ for some $p_1 < p_2$). A similar argument can be used in the presence of external field if at least two different even p-spin interaction terms are present. In this case, due to the absence of symmetry, a replica symmetric region is defined as the set of parameters $\boldsymbol{\beta}$ such that the infimum in (1.7) is achieved on Dirac measure δ_x concentrated at any point $x \in [0,1]$ rather than zero. Again, by continuity of $m \to \mathcal{P}(m, \boldsymbol{\beta})$, on the complement of the replica symmetric region any Parisi measure $m_{\boldsymbol{\beta}}$ must satisfy

$$\int |q - x| dm_{\boldsymbol{\beta}}(q) \geq \varepsilon$$

for all $x \in [0,1]$ and some $\varepsilon > 0$. This means that $m_{\boldsymbol{\beta}}$ is not concentrated near any one point $x \in [0,1]$ and, therefore,

$$\Bigl(\int q^{p_1} dm_{\boldsymbol{\beta}}(q)\Bigr)^{1/p_1} \leq \Bigl(\int q^{p_2} dm_{\boldsymbol{\beta}}(q)\Bigr)^{1/p_2} - \delta$$

for some $\delta > 0$. By (1.12), for large enough N,

$$\bigl(\mathbb{E}\langle R_{1,2}^{p_1}\rangle\bigr)^{1/p_1} \leq \bigl(\mathbb{E}\langle R_{1,2}^{p_2}\rangle\bigr)^{1/p_2} - \frac{\delta}{2}$$

which means that the Gibbs measure can not concentrate near one point and, therefore,

$$\mathbb{E}\bigl\langle (R_{1,2} - \mathbb{E}\langle R_{1,2}\rangle)^2\bigr\rangle \geq \delta' > 0. \tag{1.15}$$

□

Even though these examples strengthen and generalize the result of Pastur and Shcherbina in [6], unfortunately, the argument used above does not apply to the most interesting case of the classical 2-spin model with external field, $\beta_2 \neq 0, h \neq 0$, and it is not clear how to prove (1.15) in that case.

Acknowledgments. The author would like to thank the referees for many helpful comments and suggestions that lead to the improvement of the paper.

References

[1] Aizenman, M., Lebowitz, J. L., Ruelle, D. (1987) Some rigorous results on the Sherrington-Kirkpatrick spin glass model. *Comm. Math. Phys.* **112**, no. 1, 3-20.

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

[3] Machta, J., Newman, C.M., Stein D.L. (2007) Percolation in the Sherrington-Kirk- patrick spin glass. Preprint, arXiv:0710.1399.

[4] Panchenko, D. (2005) A question about the Parisi functional. *Electron. Comm. Probab.* **10**, 155-166.

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

[6] Pastur, L. A., Shcherbina, M. V. (1991) Absence of self-averaging of the order parameter in the Sherrington-Kirkpatrick model. *J. Statist. Phys.* **62**, no. 1-2, 1-19.

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

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

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

[10] Talagrand, M. (2006) Parisi measures. *J. Funct. Anal.* **231**, no. 2, 269-286.

[11] Talagrand, M. (2006) Parisi formula. *Ann. of Math. (2)* **163**, no. 1, 221-263.

[12] Toninelli, F.L. (2002) About the Almeida-Thouless transition line in the Sherrington-Kirkpatrick mean-field spin glass model, *Europhys. Lett.* **60**, 764-767.