Short Course on the Sherrington Kirkpatrick Model
Lecture 1: The Replica Method

Shannon Starr

2 August 2004

Material which is not otherwise attributed in these lectures is motivated by joint work with Michael Aizenman and Robert Sims, or, in other cases, is material communicated to me in private conversations by Michael Aizenman.

1 Introduction to the Sherrington-Kirkpatrick Model

Suppose $N \in \mathbb{N}_+ = \{1, 2, \dots, \}$ and let $J = (J_{ij} : i, j \in \{1, \dots, N\})$ be a family of iid normal, Gaussian random variables, $\mathcal{N}(0, 1)$. We denote $\mathbb{Z}_2 = \{+1, -1\}$, which is the state space for a single Ising spin. Then the Sherrington-Kirkpatrick Hamiltonian is a random function $H_N(\cdot, J) : \mathbb{Z}_2^N \to \mathbb{R}$ wherein, for $\sigma = (\sigma_1, \dots, \sigma_N) \in \mathbb{Z}_2^N$, we have

$$H_N(\sigma, J) \, = \, -\sum_{i=1}^N \sum_{j=1}^N \frac{J_{ij}}{\sqrt{2N}} \sigma_i \sigma_j \, . \tag{1}$$

It is common to add to this basic Hamiltonian an external magnetic field XSterm given by $\sum_{i=1}^N h\sigma_i$. We define

$$\widetilde{H}_N(\sigma, J, h) \, = \, H_N(\sigma, J) \, - \, h \sum_{i=1}^N \sigma_i \, . \tag{2}$$

The relevant quantities to calculate are the random partition function

$$Z_N(\beta, J, h) \, = \, \sum_{\mathbb{Z}_2^N} 2^{-N} \exp\Big(-\beta \widetilde{H}_N(\sigma, J, h)\Big) \, , \tag{3}$$

and the random finite-volume approximation to the pressure

$$p_N(\beta, J, h) \, = \, N^{-1} \log\left(Z_N(\beta, J, h)\right) \, , \tag{4}$$

and the so-called "quenched pressure", which is the expectation of the random pressure with respect to the random variables J:

$$p_N(\beta, h) \, = \, \mathbb{E}\{p_N(\beta, h)\} \, = \, \int_{\mathbb{R}^{N^2}} \prod_{i=1}^N \prod_{j=1}^N \left[\frac{e^{-J_{ij}^2/2} dJ_{ij}}{\sqrt{2\pi}}\right] p_N(\beta, J, h) \, . \tag{5}$$

In fact it is the last one which primarily interests us. The main problem is to calculate the quenched pressure. But, following the physicists (nonrigorous) arguments, we will begin by calculating the positive integer moments of the random variable $Z_N(\beta, J, h)$.

2 The Positive Integer Moments of Partition Function

We will now lay the framework for Parisi's replica symmetry breaking ansatz. The analysis we present follows the paper of Jan van Hemmen and R G Palmer [3]. That paper is a critical review of the approach of Sherrington and Kirkpatrick in [5].

One begins by considering the moments (more specifically, the positive integer moments) of the random variable $Z_N(\beta, J, h)$. That is, we want to consider $\mathbb{E}\{Z_N(\beta, J, h)^n\}$, for $n \in \mathbb{N}_+$. Specifically, we define, for $n \in \mathbb{N}_+$,

$$\phi_N(\beta, h, n) = \frac{1}{Nn} \log\left(\mathbb{E}\left\{Z_N(\beta, J, h)^n\right\}\right) . \tag{6}$$

One starts by observing that we can rewrite

$$\begin{aligned} Z_N(\beta, J, h)^n &= \left(\sum_{\sigma \in \mathbb{Z}_2^N} 2^{-N} \exp\left(-\beta \widetilde{H}_N(\sigma, J, h)\right)\right)^n \\ &= \sum_{\sigma^{(1)} \in \mathbb{Z}_2^N} \cdots \sum_{\sigma^{(n)} \in \mathbb{Z}_2^N} 2^{-Nn} \exp\left(-\beta \sum_{a=1}^n \widetilde{H}_N(\sigma^{(a)}, J, h)\right) . \end{aligned} \tag{7}$$

Note that in order to calculate the nth power of $Z_N(\beta, J, h)$, we have introduced n "replicas". We like to rewrite this as

$$Z_N(\beta, J, h)^n = \sum_{\sigma^{(1)} \in \mathbb{Z}_2^N} \cdots \sum_{\sigma^{(n)} \in \mathbb{Z}_2^N} 2^{-Nn} \exp\left(\beta h \sum_{a=1}^n \sum_{i=1}^N \sigma_i^{(a)}\right) \exp\left(-\beta \sum_{a=1}^n H_N(\sigma^{(a)}, J)\right) . \tag{8}$$

The only random variable above is $\sum_{a=1}^n H_N(\sigma^{(a)}, J)$. Note that because this is a sum of Gaussian centered random variables, it is itself a Gaussian random variable. Moreover, we can calculate the variance

$$\begin{aligned} \mathbb{E}\left\{\left(\sum_{a=1}^n H_N(\sigma^{(a)}, J)\right)^2\right\} &= \sum_{a=1}^n \sum_{b=1}^n \sum_{i=1}^N \sum_{j=1}^N \sum_{k=1}^N \sum_{\ell=1}^N \frac{\mathbb{E}\{J_{ij} J_{k\ell}\}}{2N} \sigma_i^{(a)} \sigma_j^{(a)} \sigma_k^{(b)} \sigma_\ell^{(b)} \\ &= \sum_{a=1}^n \sum_{b=1}^n \sum_{i=1}^N \sum_{j=1}^N \frac{1}{2N} \sigma_i^{(a)} \sigma_j^{(a)} \sigma_i^{(b)} \sigma_j^{(b)} \end{aligned} \tag{9}$$

We have used the fact that $\mathbb{E}\{J_{ij} J_{k\ell}\} = \delta_{ik}\delta_{j\ell}$. Defining

$$\mathcal{R}_N(\sigma, \sigma') = \frac{1}{N} \sum_{i=1}^N \sigma_i \sigma'_i . \tag{10}$$

we therefore have

$$\mathbb{E}\left\{\left(\sum_{a=1}^{n} H_N(\sigma^{(a)}, J)\right)^2\right\} = \sum_{a=1}^{n}\sum_{b=1}^{n} \frac{N}{2}\mathcal{R}_N(\sigma^{(a)}, \sigma^{(b)})^2 \,. \tag{11}$$

Now suppose that z is a Gaussian centered random variable with variance σ^2. Then observe that

$$\begin{aligned} \mathbb{E}\{e^{\lambda z}\} &= \int_{-\infty}^{\infty} \frac{e^{-z^2/2\sigma^2}dz}{\sqrt{2\pi\sigma^2}}\, e^{\lambda z} \\ &= \int_{-\infty}^{\infty} \frac{e^{-(z-\sigma^2\lambda)^2/2\sigma^2}dz}{\sqrt{2\pi\sigma^2}}\, e^{\lambda^2\sigma^2/2} \\ &= e^{\lambda^2\sigma^2/2}\,. \end{aligned} \tag{12}$$

Therefore, we can calculate the expectation

$$\begin{aligned} &\mathbb{E}\{Z_N(\beta, J, h)^n\} \\ &= \sum_{\sigma^{(1)}\in\mathbb{Z}_2^N} \cdots \sum_{\sigma^{(n)}\in\mathbb{Z}_2^N} 2^{-Nn} \exp\left(\beta h \sum_{a=1}^{n}\sum_{i=1}^{N} \sigma_i^{(a)}\right) \mathbb{E}\left\{\exp\left(-\beta \sum_{a=1}^{n} H_N(\sigma^{(a)}, J)\right)\right\} \\ &= \sum_{\sigma^{(1)}\in\mathbb{Z}_2^N} \cdots \sum_{\sigma^{(n)}\in\mathbb{Z}_2^N} 2^{-Nn} \exp\left(\beta h \sum_{a=1}^{n}\sum_{i=1}^{N} \sigma_i^{(a)}\right) \exp\left(\frac{\beta^2 N}{4} \sum_{a=1}^{n}\sum_{b=1}^{n} \mathcal{R}_N(\sigma^{(a)}, \sigma^{(b)})^2\right) . \end{aligned} \tag{13}$$

At this point, we have taken the expectation of the random variable, and we are left with a purely deterministic number. However, it is a number which we cannot easily evaluate. Our primary goal is to calculate the limit, inside $\frac{1}{Nn}\log$, as $N \to \infty$, with n fixed. It turns out that to effectively calculate this limit, it is useful to reintroduce randomness, or rather to introduce new randomness.

3 The van Hemmen and den Ouden Transformation

The formula which we wrote above, that

$$\mathbb{E}\{e^{\lambda z}\} = e^{\lambda^2\sigma^2/2}\,, \tag{14}$$

can be turned around to "linearize" quadratic exponentials, at the cost of introducing new Gaussian random variables. In that context, it is often called the Hubbard-Stratonovich transformation. We would like to perform such a transformation to "linearize" the terms $\mathcal{R}_N(\sigma^{(a)}, \sigma^{(b)})^2$ which appear in the exponential. Since there are n^2 such terms, we must introduce n^2 random variables. Let us call them Q_{ab} for $a = 1, \dots, n$ and $b = 1, \dots, n$.

To be more specific, we want to realize

$$\exp\left(\frac{\beta^2 N}{4}\mathcal{R}_N(\sigma^{(a)}, \sigma^{(b)})^2\right) \tag{15}$$

as $e^{\lambda^2\sigma^2/2}$. We choose $\sigma^{-2} = \frac{\beta^2 N}{2}$, so that $\lambda = \frac{\beta^2 N}{2}\mathcal{R}_N(\sigma^{(a)},\sigma^{(b)})$. One can then check

$$\frac{1}{2}\lambda^2\sigma^2 \;=\; \frac{1}{2}\cdot\frac{2}{\beta^2 N}\cdot\frac{\beta^4 N^2}{4}\mathcal{R}_N(\sigma^{(a)},\sigma^{(b)})^2 \;=\; \frac{\beta^2 N}{4}\mathcal{R}_N(\sigma^{(a)},\sigma^{(b)})^2\,, \tag{16}$$

as it should be.

Then we can write

$$\exp\left(\frac{\beta^2 N}{4}\mathcal{R}_N(\sigma^{(a)},\sigma^{(b)})^2\right) \;=\; \int_{-\infty}^{\infty}\frac{dQ_{ab}}{\sqrt{4\pi/N}}\exp\left(-\frac{\beta^2 N}{4}Q_{ab}^2 + \frac{\beta^2 N}{2}\mathcal{R}_N(\sigma^{(a)},\sigma^{(b)})Q_{ab}\right)\,. \tag{17}$$

Putting all n^2 such terms together, we have

$$\begin{aligned}&\exp\left(\frac{\beta^2 N}{4}\sum_{a=1}^{n}\sum_{b=1}^{n}\mathcal{R}_N(\sigma^{(a)},\sigma^{(b)})^2\right)\\ &= \int_{\mathbb{R}^{n^2}}\prod_{a=1}^{n}\prod_{b=1}^{n}\left[\frac{N^{1/2}dQ_{ab}}{(4\pi)^{1/2}}\right]\exp\left(-\frac{\beta^2 N}{4}\sum_{a=1}^{n}\sum_{b=1}^{n}Q_{ab}^2 + \frac{\beta^2 N}{2}\sum_{a=1}^{n}\sum_{b=1}^{n}Q_{ab}\mathcal{R}_N(\sigma^{(a)},\sigma^{(b)})\right)\,.\end{aligned} \tag{18}$$

Therefore, substituting this formula into our (previously deterministic) expression for $\mathbb{E}\{Z_N(\beta,J,h)^n\}$, and commuting the sums and the integrals, we obtain

$$\begin{aligned}\mathbb{E}\{Z_N(\beta,J,h)^n\} \;=\; &\int_{\mathbb{R}^{n^2}}\prod_{a=1}^{n}\prod_{b=1}^{n}\left[\frac{N^{1/2}dQ_{ab}}{(4\pi)^{1/2}}\right]\\ &\sum_{\substack{\sigma_i^{(a)}\in\mathbb{Z}_2\\ a=1,\dots,n\\ i=1,\dots,N}} 2^{-nN}\exp\left(-\frac{\beta^2 N}{4}\sum_{a,b=1}^{n}Q_{ab}^2 + \beta h\sum_{a=1}^{n}\sum_{i=1}^{N}\sigma_i^{(a)} + \frac{\beta^2}{2}\sum_{a,b=1}^{n}\sum_{i=1}^{N}Q_{ab}\sigma_i^{(a)}\sigma_i^{(b)}\right)\,.\end{aligned} \tag{19}$$

We observe that we can split the exponential as

$$\begin{aligned}&\exp\left(-\frac{\beta^2 N}{4}\sum_{a,b=1}^{n}Q_{ab}^2 + \beta h\sum_{a=1}^{n}\sum_{i=1}^{N}\sigma_i^{(a)} + \frac{\beta^2}{2}\sum_{a,b=1}^{n}\sum_{i=1}^{N}\sigma_i^{(a)}\sigma_i^{(b)}\right)\\ &\qquad = \prod_{i=1}^{N}\exp\left(-\frac{\beta^2}{4}\sum_{a,b=1}^{n}Q_{ab}^2 + \beta h\sum_{a=1}^{n}\sigma_i^{(a)} + \frac{\beta^2}{2}\sum_{a,b=1}^{n}Q_{ab}\sigma_i^{(a)}\sigma_i^{(b)}\right)\,.\end{aligned} \tag{20}$$

Therefore, since we have a sum over all spins, we can rewrite it as

$$
\begin{aligned}
&\sum_{\substack{\sigma_i^{(a)}\in\mathbb{Z}_2\\ a=1,\dots,n\\ i=1,\dots,N}} 2^{-nN}\exp\left(-\frac{\beta^2 N}{4}\sum_{a,b=1}^{n}Q_{ab}^2+\beta h\sum_{a=1}^{n}\sum_{i=1}^{N}\sigma_i^{(a)}+\frac{\beta^2}{2}\sum_{a,b=1}^{n}\sum_{i=1}^{N}Q_{ab}\sigma_i^{(a)}\sigma_i^{(b)}\right)\\
&=\prod_{i=1}^{N}\left[\sum_{\sigma_i^{(1)},\dots,\sigma_i^{(n)}\in\mathbb{Z}_2}2^{-n}\exp\left(-\frac{\beta^2}{4}\sum_{a,b=1}^{n}Q_{ab}^2+\beta h\sum_{a=1}^{n}\sigma_i^{(a)}+\frac{\beta^2}{2}\sum_{a,b=1}^{n}Q_{ab}\sigma_i^{(a)}\sigma_i^{(b)}\right)\right]\\
&=\left[\sum_{\sigma^{(1)},\dots,\sigma^{(n)}\in\mathbb{Z}_2}2^{-n}\exp\left(-\frac{\beta^2}{4}\sum_{a,b=1}^{n}Q_{ab}^2+\beta h\sum_{a=1}^{n}\sigma^{(a)}+\frac{\beta^2}{2}\sum_{a,b=1}^{n}Q_{ab}\sigma^{(a)}\sigma^{(b)}\right)\right]
\end{aligned}
\tag{21}
$$

We prefer to relabel the remaining n spins as $S_1,\dots,S_n\in\mathbb{Z}_2$ instead of $\sigma^{(1)},\dots,\sigma^{(n)}$.

Therefore, at last, we arrive at the formula

$$
\begin{aligned}
\mathbb{E}\{Z_N(\beta,J,h)^n\} \;=\; &\int_{\mathbb{R}^{n^2}}\prod_{a=1}^{n}\prod_{b=1}^{n}\left[\frac{N^{1/2}dQ_{ab}}{(4\pi)^{1/2}}\right]\\
&\left[\sum_{S_1,\dots,S_n\in\mathbb{Z}_2}2^{-n}\exp\left(-\frac{\beta^2}{4}\sum_{a,b=1}^{n}Q_{ab}^2+\beta h\sum_{a=1}^{n}S_a+\frac{\beta^2}{2}\sum_{a,b=1}^{n}Q_{ab}S_aS_b\right)\right]^N .
\end{aligned}
\tag{22}
$$

Finally, in calculating

$$
\begin{aligned}
\phi_N(\beta,h,n) \;=\; &\frac{1}{Nn}\log\left(\mathbb{E}\left\{Z_N(\beta,J,h)^n\right\}\right)\\
=\; &\frac{1}{Nn}\log\int_{\mathbb{R}^{n^2}}\prod_{a=1}^{n}\prod_{b=1}^{n}\left[\frac{N^{1/2}dQ_{ab}}{(4\pi)^{1/2}}\right]\\
&\left[\sum_{S_1,\dots,S_n\in\mathbb{Z}_2}2^{-n}\exp\left(-\frac{\beta^2}{4}\sum_{a,b=1}^{n}Q_{ab}^2+\beta h\sum_{a=1}^{n}S_a+\frac{\beta^2}{2}\sum_{a,b=1}^{n}Q_{ab}S_aS_b\right)\right]^N ,
\end{aligned}
\tag{23}
$$

we use the standard method of steepest descents (also called the "saddle point method" or "method of stationary phase") to evaluate the limit. (See, for example, [4], Chapter 7, for some words about the rigorous use of the saddle point method. In our case, we can fairly easily obtain rigorous bounds to justify the saddle point method.)

We obtain for $\phi(\beta,h,n):=\lim_{N\to\infty}\phi_N(\beta,h,n)$, that

$$
\phi(\beta,h,n) \;=\; \max_{(Q_{ab}\in\mathbb{R}\,:\,a,b=1,\dots,n)}\left[-\frac{\beta^2}{4n}\sum_{a,b=1}^{n}Q_{ab}^2+\frac{1}{n}\log\left(\sum_{S_1,\dots,S_n\in\mathbb{Z}_2}2^{-n}e^{\mathcal{H}(S,\beta,h,Q)}\right)\right] , \tag{24}
$$

where

$$
\mathcal{H}(S,\beta,h,Q) \;=\; \beta h\sum_{a=1}^{n}S_a+\frac{\beta^2}{2}\sum_{a,b=1}^{n}Q_{ab}S_aS_b\,. \tag{25}
$$

This is a rigorous result, in the spirit of a large deviation estimate. It is called the van Hemmen and den Ouden transformation in the paper by van Hemmen and Palmer.

Note that, in calculating the nth moment, and taking the limit $N \to \infty$, we have arrived at a problem of optimizing the pressure, minus a cost term, for an n-spin Hamiltonian, with respect to the couplings Q.

4 Ultrametric choices of Q

Let us summarize what we have done so far. We began with the SK Hamiltonian. We wanted to calculate $\lim_{N\to\infty}(Nn)^{-1}\log(\mathbb{E}\{Z_N(\beta,J,h)^n\})$. We first used the Hubbard-Stratonovich formula to take the expectation with respect to all the Gaussian couplings $J = (J_{ij} : i,j = 1,\dots,N)$. We then used the Hubbard-Stratonovich formula to reintroduce Gaussians $Q = (Q_{ab} : a,b = 1,\dots,n)$ in order to "linearize" a quadratic term in the exponential. Then, by a large deviation estimate, we were able to write the limit as $N \to \infty$ as an optimization over the n^2 real numbers $Q = (Q_{ab} : a,b = 1,\dots,n)$.

Whenever we have had a quadratic term in the exponential, we used the Hubbard-Stratonovich formula to "linearize" it. Well, in the current optimization problem, one has yet more quadratic terms in the exponential. For, in the formula

$$\phi(\beta,h,n) = \max_{(Q_{ab}\in\mathbb{R}\,:\,a,b=1,\dots,n)} \Phi_n(Q,\beta,h) \tag{26}$$

$$\Phi_n(Q,\beta,h) = -\frac{\beta^2}{4n}\sum_{a,b=1}^{n} Q_{ab}^2 + \frac{1}{n}\log\left(\sum_{S_1,\dots,S_n\in\mathbb{Z}_2} 2^{-n}e^{\mathcal{H}(S,\beta,h,Q)}\right) \tag{27}$$

$$\mathcal{H}(S,\beta,h,Q) = \beta h\sum_{a=1}^{n} S_a + \frac{\beta^2}{2}\sum_{a,b=1}^{n} Q_{ab}S_aS_b\,, \tag{28}$$

there are quadratic terms $\sum_{a,b=1}^{n} Q_{ab}S_aS_b$ in the definition of the Hamiltonian. Only, we would need to diagonalize the matrix Q_{ab}, and hope that all eigenvalues are positive, to really "linearize" this term.

Parisi realized that there is a nice class of matrices where one can actually do the Hubbard-Stratonovich transformation more directly. These are the ultrametric matrices, which we will describe next. However, since the observation was originally made by Sherrington and Kirkpatrick, although only for "replica symmetric" ultrametric matrices, we will start with that.

4.1 Replica Symmetric Choice

Suppose that Q has the special form

$$Q_{ab} = \begin{cases} q_1 & \text{if } a = b, \\ q_0 & \text{if } a \neq b\,. \end{cases} \tag{29}$$

Suppose, furthermore, that $0 \leq q_0 \leq q_1$. Then we can rewrite

$$\sum_{a,b=1}^{n} Q_{ab} S_a S_b \;=\; \sum_{a,b=1}^{n} q_0 S_a S_b \;+\; \sum_{a=1}^{n} (q_1 - q_0) S_a^2 \;=\; q_0 \left(\sum_{a=1}^{n} S_a \right)^2 + \sum_{a=1}^{n} (q_1 - q_0) S_a^2 \,. \tag{30}$$

We should note now, that we have neglected some times before to make the formal simplification that $\sigma_i^2 = 1$. Likewise, at this point, we will also ignore the fact that $S_a^2 = 1$. It will come out of a later calculation anyway, but we prefer to preserve a certain feeling of symmetry at this moment.

Therefore, we introduce normal, Gaussian centered random variables z_0 and $(z_1(a) \,:\, a = 1, \ldots, n)$, which are all supposed to be independent. With our very specific (and so far unjustified) assumption on the form of Q, we can then rewrite

$$\begin{aligned} e^{\mathcal{H}(S,\beta,h,Q)} &= \exp\left(\beta h \sum_{a=1}^{n} S_a + \frac{\beta^2}{2} \sum_{a,b=1}^{n} Q_{ab} S_a S_b \right) \\ &= \mathbb{E}\left\{ \exp\left(\beta \sum_{a=1}^{n} [h + \sqrt{q_0}\, z_0 + \sqrt{q_1 - q_0}\, z_1(a)] S_a \right) \right\} . \end{aligned} \tag{31}$$

This is, of course, once more using the Hubbard-Stratonovich formula.

Commuting the sum and expectation, and substituting the formula above, we obtain

$$\begin{aligned} \sum_{S_1,\ldots,S_n \in \mathbb{Z}_2} 2^{-n} e^{\mathcal{H}(S,\beta,h,Q)} &= \mathbb{E}\left\{ \sum_{S_1,\ldots,S_n \in \mathbb{Z}_2} 2^{-n} \exp\left(\beta \sum_{a=1}^{n} [h + \sqrt{q_0}\, z_0 + \sqrt{q_1 - q_0}\, z_1(a)] S_a \right) \right\} \\ &= \mathbb{E}\left\{ \prod_{a=1}^{n} \cosh\left(\beta [h + \sqrt{q_0}\, z_0 + \sqrt{q_1 - q_0}\, z_1(a)] \right) \right\} . \end{aligned} \tag{32}$$

Using the fact that all the $(z_1(a) \,:\, a = 1, \ldots, n)$ are i.i.d., and using the conditional expectation, we can rewrite this as

$$\begin{aligned} \sum_{S_1,\ldots,S_n \in \mathbb{Z}_2} 2^{-n} e^{\mathcal{H}(S,\beta,h,Q)} &= \mathbb{E}\left\{ \mathbb{E}\left\{ \prod_{a=1}^{n} \cosh\left(\beta [h + \sqrt{q_0}\, z_0 + \sqrt{q_1 - q_0}\, z_1(a)] \right) \,\middle|\, z_0 \right\} \right\} \\ &= \mathbb{E}\left\{ \prod_{a=1}^{n} \mathbb{E}\left\{ \cosh\left(\beta [h + \sqrt{q_0}\, z_0 + \sqrt{q_1 - q_0}\, z_1(a)] \right) \,\middle|\, z_0 \right\} \right\} . \end{aligned} \tag{33}$$

If we define a new $\mathcal{N}(0,1)$ random variable, independent of everything else, called z_1, then we can simplify this further as

$$\sum_{S_1,\ldots,S_n \in \mathbb{Z}_2} 2^{-n} e^{\mathcal{H}(S,\beta,h,Q)} = \mathbb{E}\left\{ \mathbb{E}\left\{ \cosh\left(\beta [h + \sqrt{q_0}\, z_0 + \sqrt{q_1 - q_0}\, z_1] \right) \,\middle|\, z_0 \right\}^n \right\} \tag{34}$$

Therefore, for such (very special) Q, as written above, we have

$$\begin{aligned}\Phi_n(Q,\beta,h) \;=\; & -\frac{\beta^2 n}{4}q_0^2 - \frac{\beta^2}{4}(q_1^2-q_0^2) \\ & + \frac{1}{n}\log\left(\mathbb{E}\left\{\mathbb{E}\left\{\cosh\left(\beta[h+\sqrt{q_0}\,z_0+\sqrt{q_1-q_0}\,z_1]\right)\,\Big|\,z_0\right\}^n\right\}\right)\,. \end{aligned} \tag{35}$$

This is actually Sherrington and Kirkpatrick's solution, except that they assumed that $q_1 = 1$. Then one obtains

$$\begin{aligned}\Phi_n(Q,\beta,h) \;=\; & -\frac{\beta^2 n}{4}q_0^2 - \frac{\beta^2}{4}(1-q_0^2) \\ & + \frac{1}{n}\log\left(\mathbb{E}\left\{\mathbb{E}\left\{\cosh\left(\beta[h+\sqrt{q_0}\,z_0+\sqrt{1-q_0}\,z_1]\right)\,\Big|\,z_0\right\}^n\right\}\right)\,. \end{aligned} \tag{36}$$

We find it convenient, for later comparisons, to define $x_0 = n$ and $x_1 = 1$, and to define the piecewise-constant (nonincreasing) function $X : (0, 1]$ such that

$$X(q) \;=\; \begin{cases} x_0 & 0 < q \le q_0\,, \\ x_1 & q_0 < q \le 1\,. \end{cases} \tag{37}$$

Note that (unlike the situation which will arise later) $x_0 > x_1$ and both are positive integers. We define

$$\mathcal{Z}_1(\beta,h) \;=\; \cosh\left(\beta\left[h+\sqrt{q_0}z_0+\sqrt{q_1-q_0}\,z_1\right]\right)\,. \tag{38}$$

This is a random variable measurable with respect to $\mathcal{F}_1 := \sigma(z_0, z_1)$, the σ-algebra generated by z_0 and z_1. If we define $\mathcal{F}_0 := \sigma(z_0)$, the σ-algebra generated by z_0, then we may define

$$\mathcal{Z}_0(\beta,h) \;=\; \mathbb{E}\left\{(\mathcal{Z}_1)^{x_1}\,\Big|\,\mathcal{F}_0\right\}^{1/x_1}\,. \tag{39}$$

In this case, we can rewrite

$$\Phi_n(Q,\beta,h) \;=\; \frac{1}{n}\log\left(\mathbb{E}\{\mathcal{Z}_0(\beta,h)^n\}\right) \;-\; \frac{\beta^2}{4}\int_0^1 q\,X(q)\,dq \tag{40}$$

4.2 K-level RSB choice

We now consider the K-level "replica symmetry breaking" choice. To motivate the name "replica symmetry breaking" note that one may ask why we might hope that Q would have such a simple form as in the last problem. The naive response is that since the functional $\Phi_n(Q,\beta,h)$ is invariant with respect to permutations of the indices $a = 1,\dots,n$, one might hope that Q has the same structure. But then it is trivial to see, that this is the only possible form for a matrix $Q = (Q_{ab}\ :\ a, b = 1,\dots,n)$, which is invariant under all permutations, i.e., such that $Q_{a,b} = Q_{\pi(a),\pi(b)}$ for all $\pi \in S_n$.

We now consider a more general form for Q. This form has a connection to other representations of the symmetric group. But, for us, the most important immediate fact is that it allows a simple calculation. (It would perhaps be worthwhile for someone more knowledgeable than us to examine whether the connection to nontrivial representations of S_n is important.)

Define $x_0 = n$, and suppose that there are positive integers $x_1, \dots, x_K$, such that

$$x_K | x_{K-1} | \dots | x_1 | x_0 \,, \tag{41}$$

where $x|y$ means that "y divides x". I.e., x/y is an integer. We define $x_{K+1} = 1$. Suppose that there are also numbers $0 \leq q_0 \leq q_1 \leq \cdots \leq q_K \leq q_{K+1} = 1$. Then one considers a matrix Q such that

$$Q_{ab} \,=\, \begin{cases} 1 & \text{if } a = b\,, \\ q_K & \text{if } a \neq b\,, \text{ but } a \equiv b \,(\mathrm{mod}\, x_0/x_K)\,, \\ q_{K-1} & \text{if } a \not\equiv b \,(\mathrm{mod}\, x_0/x_K)\,, \text{ but } a \equiv b \,(\mathrm{mod}\, x_0/x_{K-1})\,, \\ \dots & \\ q_1 & \text{if } a \not\equiv b \,(\mathrm{mod}\, x_0/x_2)\,, \text{ but } a \equiv b \,(\mathrm{mod}\, x_0/x_1)\,, \\ q_0 & \text{if } a \not\equiv b \,(\mathrm{mod}\, x_0/x_1)\,. \end{cases} \tag{42}$$

In other words, we can say that $Q_{ab} = q_{K+1}$ if $a \equiv b \,(\mathrm{mod}\, x_0/x_{K+1})$, and that for $k \leq K$, $Q_{ab} = q_k$ iff $a \not\equiv b \,(\mathrm{mod}\, x_0/x_{k+1})$, but $a \equiv b \,(\mathrm{mod}\, x_0/x_k)$.

Then, going through the analysis as above we can determine the exact identity for Q of this form. Define $z_0, \dots, z_{K+1}$ to be i.i.d. normal, Gaussian random variables. Define $\mathcal{F}_k = \sigma(z_0, \dots, z_k)$, the σ-algebra generated by $z_0, \dots, z_k$ for $k = 0, \dots, K+1$. Define $\mathcal{Z}_{K+1}(\beta, h)$ as the $\mathcal{F}_{K+1}$-measurable random variable

$$\mathcal{Z}_{K+1}(\beta, h) \,=\, \cosh\left(\beta\left[h + \sqrt{q_0} z_0 + \sum_{k=1}^{K+1} \sqrt{q_k - q_{k-1}}\, z_k\right]\right)\,. \tag{43}$$

For $k = 0, \dots, K$, make the inductive definition of $\mathcal{Z}_k(\beta, h)$ which is $\mathcal{F}_k$-measurable,

$$\mathcal{Z}_k(\beta, h) \,=\, \mathbb{E}\left\{(\mathcal{Z}_{k+1})^{x_{k+1}} \,\middle|\, \mathcal{F}_k\right\}^{1/x_{k+1}}\,. \tag{44}$$

Then

$$\Phi_n(Q, \beta, h) \,=\, -\frac{\beta^2}{4n} \sum_{a,b=1}^{n} Q_{ab}^2 \,+\, \frac{1}{n} \log \mathbb{E}\{\mathcal{Z}_0(\beta, h)^n\}\,. \tag{45}$$

We observe that $Q_{ab} = q_k$ iff $a \not\equiv b \,(\mathrm{mod}\, x_0/x_{k+1})$, but $a \equiv b \,(\mathrm{mod}\, x_0/x_k)$. The number of pairs with $a \equiv b \,(\mathrm{mod}\, x_0/x_k)$ is equal to

$$\begin{aligned} \#\{(a,b) \in [1,n]^2 \,:\, a \equiv b \,(\mathrm{mod}\, x_0/x_k)\} \,&=\, \sum_{j=1}^{x_0/x_k} \#\{(a,b) \in [1,n]^2 \,:\, a \equiv b \equiv j \,(\mathrm{mod}\, x_0/x_k)\} \\ &=\, \sum_{j=1}^{x_0/x_k} \left(\#\{a \in [1,n] \,:\, a \equiv j \,(\mathrm{mod}\, x_0/x_k)\}\right)^2 \\ &=\, \frac{x_0}{x_k} \cdot x_k^2 \\ &=\, n x_k\,. \end{aligned} \tag{46}$$

Therefore,

$$-\frac{\beta^2}{4n}\sum_{a,b=1}^{n} Q_{ab}^2 \;=\; -\frac{\beta^2}{4}\sum_{k=0}^{K}(x_k - x_{k+1})q_k^2 \;-\; \frac{\beta^2}{4}q_{K+1}^2$$
$$\;=\; -\frac{\beta^2}{4}\sum_{k=1}^{K+1}(q_k^2 - q_{k-1}^2)x_k\,. \tag{47}$$

Therefore, defining

$$X(q) \;=\; \begin{cases} n = x_0 & 0 < q \leq q_0\,, \\ x_1 & q_0 < q \leq q_1\,, \\ \ldots & \\ x_K & q_{K-1} < q \leq q_K\,, \\ 1 = x_{K+1} & q_K \leq q \leq q_{K+1} = 1\,, \end{cases} \tag{48}$$

we have

$$\Phi_n(Q,\beta,h) \;=\; \frac{1}{n}\log\left(\mathbb{E}\{\mathcal{Z}_0(\beta,h)^n\}\right) \;-\; \frac{\beta^2}{4}\int_0^1 q\,X(q)\,dq \tag{49}$$

5 The van Hemmen, Lieb, Palmer theorem

So far, what was done in the last section, and indeed in all sections leading up to this one has been rigorous. No one has proved that the optimal choice of Q is ultrametric, in general. But using a reflection-positivity argument, van Hemmen, Palmer and Elliot Lieb proved the following theorem. The proof is contained in an appendix to [3]

Theorem 5.1 *If $h = 0$ and $n \in \mathbb{N}_+$, then*

$$\phi(\beta,h,n) \;=\; \max_{(Q_{ab}\in\mathbb{R}\,:\,a,b=1,\ldots,n)} \Phi_n(Q,\beta,h) \tag{50}$$

$$\Phi_n(Q,\beta,h) \;=\; -\frac{\beta^2}{4n}\sum_{a,b=1}^{n} Q_{ab}^2 + \frac{1}{n}\log\left(\sum_{S_1,\ldots,S_n\in\mathbb{Z}_2} 2^{-n}e^{\mathcal{H}(S,\beta,h,Q)}\right) \tag{51}$$

$$\mathcal{H}(S,\beta,h,Q) \;=\; \beta h\sum_{a=1}^{n} S_a + \frac{\beta^2}{2}\sum_{a,b=1}^{n} Q_{ab}S_aS_b\,, \tag{52}$$

is optimized by one Q of the form

$$Q_{ab} \;=\; \begin{cases} 1 & a = b\,, \\ q_0 & a \neq b\,, \end{cases} \tag{53}$$

for some $0 \leq q_0 \leq 1$. When $h = 0$, there are generally 2^n other solutions obtained by choosing $S \in \mathbb{Z}_2^n$ and taking $Q_{ab} \to S(a)Q_{ab}S(b)$.

It seems to be generally accepted that the following is true, even though it has still not been proved

Conjecture 5.2 *If $n \in \mathbb{N}_+$, and $h \neq 0$, then there is a unique optimizer Q, which has the same form as in the theorem.*

In principle, one believes that this conjecture should be provable in the following form, using Talagrand's techniques: *The optimal Q has the ultrametric form.*

6 Parisi's Ansatz

Let us define, for arbitrary real $r \neq 0$, the number

$$\phi_N(\beta, h, r) \;=\; \frac{1}{Nr} \log\left(\mathbb{E}\{Z_N(\beta, J, h)^r\}\right) . \tag{54}$$

This exists, because $Z_N(\beta, J, h)^r$ is a positive, measurable function for every r. One can observe that, by l'Hospital's theorem

$$\begin{aligned} \lim_{r\to 0} \phi_N(\beta, h, r) \;&=\; \lim_{r\to 0} \frac{1}{Nr} \log\left(\mathbb{E}\{Z_N(\beta, J, h)^r\}\right) \\ &=\; \lim_{r\to 0} \frac{(d/dr)\log\left(\mathbb{E}\{Z_N(\beta, J, h)^r\}\right)}{(d/dr)Nr} \\ &=\; \lim_{r\to 0} \frac{1}{N} \cdot \frac{\mathbb{E}\{\log(Z_N(\beta, J, h))Z_N(\beta, J, h)^r\}}{\mathbb{E}\{Z_N(\beta, J, h)^r\}} \\ &=\; p_N(\beta, h) . \end{aligned} \tag{55}$$

Moreover, one may hope that the following limits exist

$$\phi(\beta, h, r) \;=\; \lim_{N\to\infty} \phi_N(\beta, h, r) . \tag{56}$$

Indeed, it can be shown to be true that the limit exists, using the techniques of Guerra and Toninelli. The proof is just like the proof of the existence of the pressure in the $N \to \infty$ limit, [2]. It is worthwhile to point out that there is also an extended variational principle for $\phi(\beta, h, r)$ for all r, which follows by arguments just like that in [1]. One may hope that $p(\beta, h) := \lim_{N\to\infty} p_N(\beta, h)$ satisfies

$$p(\beta, h) \;=\; \lim_{r\to 0} \phi(\beta, h, r) . \tag{57}$$

Parisi's ansatz is motivated by these considerations.

Parisi's ansatz consists in two parts:

1. The observation that the function $r \mapsto \phi(\beta, h, r)$ changes from being convex in r for $1 \leq r < \infty$, to being concave in r for $0 < r \leq 1$. It is also convex for $-\infty < r < 1$. Parisi argues that this means that for $0 < r < 1$, instead of maximizing a functional $\Phi_r(Q, \beta, h)$ one should minimize it. Unfortunately this functional is not well defined.

2. Parisi says that one can naturally extend $\Phi_n(Q, \beta, h)$ to $\Phi_r(Q, \beta, h)$ by restricting to ultrametric choices of Q, and formally extending the formula one obtains, which includes n as a parameter. He claims that $\phi(\beta, h, r)$ for $0 < r < 1$ can be calculated by solving the resulting minimization problem, and that $p(\beta, h)$ can be inferred from the limit.

Without addressing the reasonability of his assumptions, let us give the precise formulation of his conjecture for the case of the pressure. Suppose $K \in \mathbb{N}$. Choose numbers $0 \leq q_0 \leq q_1 \leq \cdots \leq q_K \leq q_{K+1} = 1$. Also choose numbers $0 = x_0 < x_1 < \cdots < x_K < x_{K+1} = 1$. Define the function

$$X(q) \;=\; \begin{cases} 0 = x_0 & 0 < q \leq q_0\,, \\ x_1 & q_0 < q \leq q_1\,, \\ \dots & \\ x_K & q_{K-1} < q \leq q_K\,, \\ 1 = x_{K+1} & q_K \leq q \leq q_{K+1} = 1\,, \end{cases} \tag{58}$$

which is now a nondecreasing function from $[0,1]$ to $[0,1]$. Let $z_0,\dots,z_{K+1}$ be i.i.d., normal Gaussian random variables. Define $\mathcal{F}_k = \sigma(z_0,\dots,z_k)$, the σ-algebra generated by $z_0,\dots,z_k$ for $k = 0,\dots,K+1$. Define $\mathcal{Z}_{K+1}(\beta,h)$ as the $\mathcal{F}_{K+1}$-measurable random variable

$$\mathcal{Z}_{K+1}(\beta,h) \;=\; \cosh\left(\beta\left[h + \sqrt{q_0}\,z_0 + \sum_{k=1}^{K+1}\sqrt{q_k - q_{k-1}}\,z_k\right]\right)\,. \tag{59}$$

For $k = 0,\dots,K$, make the inductive definition of $\mathcal{Z}_k(\beta,h)$ which is $\mathcal{F}_k$-measurable,

$$\mathcal{Z}_k(\beta,h) \;=\; \mathbb{E}\left\{(\mathcal{Z}_{k+1})^{x_{k+1}}\,\middle|\,\mathcal{F}_k\right\}^{1/x_{k+1}}\,. \tag{60}$$

Then

$$\Phi_0^{(K)}(\beta;h;q_0,\dots,q_K;x_1,\dots,x_K) \;=\; \mathbb{E}\{\log(\mathcal{Z}_0(\beta,h))\} \;-\; \frac{\beta^2}{4}\int_0^1 q\,X(q)\,dq\,. \tag{61}$$

Conjecture 6.1 *The pressure for the Sherrington-Kirkpatrick model can be calculated as*

$$p(\beta,h) \;=\; \lim_{K\to\infty}\;\min_{\substack{0\leq q_0\leq\cdots\leq q_K\leq 1\\ 0<x_1<\cdots<x_K<1}}\;\Phi_0^{(K)}(\beta;h;q_0,\dots,q_K;x_1,\dots,x_K)\,. \tag{62}$$

7 Endnote

Though we will discuss Guerra and Toninelli's results [2] as well as those in [1] in detail later, we felt it necessary to mention here that much of what Parisi guessed can be proved correct. In particular, it is true that $\phi_r(\beta,h)$ exists for all $r \in \mathbb{R}\setminus\{0\}$. Also, by the extended variational principle, one can write $\phi_r(\beta,h)$ as a maximization problem for $1 \leq r < \infty$ or $-\infty < r < 0$, or as a minimization problem for $0 < r < 1$. As it turns out (and it makes sense when considering $\log(\cdot)$ as a derivative of $(\cdot)^r$ at $r = 0$) that it is also a minimization problem to calculate $p(\beta,h)$.

References

[1] M. Aizenman, R. Sims, and S. Starr. Extended variational principle for the Sherrington-Kirkpatrick spin-glass model. *Phys. Rev. B* **68** (2003) 214403. http://front.math.ucdavis.edu/cond-mat/0306386

[2] F. Guerra and F. Toninelli. The thermodynamic limit in mean field spin glass models. *Comm. Math. Phys.*, 230(1):71–79, 2002. http://front.math.ucdavis.edu/cond-mat/0204280

[3] J.L. van Hemmen and R.G. Palmer The replica method and a solvable spin glass model. *J. Phys. A* 12(4):563–580, 1979. http://www.iop.org/EJ/abstract/-search=6932651.1/0305-4470/12/4/016

[4] M.J. Hoffman and J.E. Marsden, Basic complex analysis. Second edition. W. H. Freeman and Company, New York, 1987.

[5] D. Sherrington and S. Kirkpatrick. Solvable model of a spin-glass. *Phys. Rev. Lett.*, **35**, 1792–1796, 1975. http://link.aps.org/abstract/PRL/v35/p1792

(Wherever possible we give the most easily accessible online location of the paper possible; i.e., one which costs no fee to obtain.)