Course 2S620, Lecture 3
Sherrington-Kirkpatrick model

C. Giardinà *

June 15, 2007

1 Definition of the model

The *mean-field* theory of spin-glasses is a solvable model which shows the existence of a new kind of phase transition. The reference model is the Sherrington-Kirkpatrick model, introduced in 1975. After some false starts, a meaningfull solution was proposed by G. Parisi in the eighties using a very clever ansatz for the symmetry breaking. The results was obtained by using the non-rigorous method of *replicas*, which we are going to review later. In 2005 the joint effort of F. Guerra and M. Talagrand allowed to establish the correctness of Parisi result as a theorem. While this has been an impressive achievement (it took 30 years!), it must also be noted that some relevant properties of the quenched equilibrium state, like ultrametricity, are still the subject of intensive investigations.

Similarly to the Curie-Weiss model, the Sherrington-Kirkpatrick model is obtained by considering an interaction between all possible couples of $N \in \mathbb{N}$ spins. With reference to the setting of the previous lecture, we define the volume $\Lambda = \{1, 2, \ldots, N\}$, with cardinality $|\Lambda| = N$ and the sample space $\Omega_N = \{-1, +1\}^N$ with cardinality $|\Omega_N| = 2^N$. Let $\{J_{ij}\}_{i,j\in\Lambda\times\Lambda}$ be a family of i.i.d. random variables. While the model can be defined for an arbitrary choice of the distribution of the random variable we will stick to the choice of a standard normal distribution, that is for any $(i, j) \in \Lambda \times \Lambda$ the J_{ij} are i.i.d. Gaussian random variable with

$$\mathbb{E}(J_{ij}) = 0 \qquad\qquad \mathbb{E}(J_{ij}^2) = 1 \tag{1.1}$$

The Hamiltonian then reads:

$$H_N(\sigma) = -\frac{1}{\sqrt{2N}} \sum_{i,j=1}^{N} J_{ij}\sigma_i\sigma_j \tag{1.2}$$

*Department of Mathematics and Computer Science, Eindhoven University, P.O. Box 513 - 5600 MB Eindhoven, The Netherlands, *c.giardina@tue.nl*

Note that the normalization $1/\sqrt{N}$ has been added in such a way that the thermodynamic quantities (free energy, internal energy, ..) are extensive, i.e. order N.

Contrary to the Curie-Weiss model, where the Hamiltonian is a function that associates a real number to each configuration $\sigma \in \Omega_N$, now the Hamiltonian is a random variable which for each $\sigma \in \Omega_N$ takes real values with an assigned probability. Being the Hamiltonian a linear combination of independent Gaussian random variables, it is also a gaussian random variable itself. However it will have correlations. The covariance structure is immediately calculated as

$$\mathbb{E}(H_N(\sigma)H_N(\tau)) = \frac{N}{2}q_N^2(\sigma,\tau) \tag{1.3}$$

where we have defined the site overlap between two spin configuration σ and τ as

$$q_N(\sigma,\tau) = \frac{1}{N}\sum_{i=1}^{N}\sigma_i\tau_i \tag{1.4}$$

The family $\{H_N(\sigma)\}_{\sigma\in\Omega_N}$ is thus a centered Gaussian family with covariance matrix $C = C_N(\sigma,\tau)$ specified by eq. (1.3) and (1.4).

As usual, the main thermodynamic quantity of interest will be the pressure per particle. Due to the randomness of $H_N(\sigma)$, we will have a random finite volume pressure

$$p_N(\beta,h,J) = \frac{1}{N}\log\left(\sum_{\sigma\in\Omega_N} e^{-\beta[H_N(\sigma)-h\sum_{i=1}^{N}\sigma_i]}\right) \tag{1.5}$$

The finite volume quenched pressure is defined as

$$p_N(\beta,h) = \mathbb{E}(p_N(\beta,h,J))$$

and the limiting quenched pressure will be

$$p(\beta,h) = \lim_{N\to\infty} p_N(\beta,h)$$

These definitions rase immediately some questions:

- does the limit $p(\beta,h)$ exists?
- what is the typical value of the random pressure $p_N(\beta,h,J)$ for large N?

We will answer this question in section 2 and 3, respectively.

1.1 Gaussian processes

We recall that the vector $X = (X_1, X_2, \ldots, X_n)$ has a centered gaussian distribution if its joint probability density function is given by

$$f_X(x) = \frac{1}{\sqrt{(2\pi)^n \det(C)}}\exp[-\frac{1}{2}\sum_{i,j=1}^{n} C_{i,j}^{-1}x_ix_j] \tag{1.6}$$

where C is a positive definite symmetric matrix, which has the meaning of covariance matrix, i.e. $C_{i,j} = \mathbb{E}(X_i X_j)$.

The following result, known as integration by parts formula or generalized Wick law, will be very useful.

Lemma 1.7. *For any differentiable function $g : \mathbb{R}^n \to \mathbb{R}$ with at most polynomial growth and for any $i = 1, \ldots n$ we have*

$$\mathbb{E}(X_i g(X)) = \sum_{j=1}^{n} C_{i,j} \mathbb{E}\left(\frac{\partial}{\partial X_i} g(X)\right) \tag{1.8}$$

2 Thermodynamic limit of the pressure

In analogy with the case of the Curie-Weiss model the existence of the infinite volume quenched pressure can be proved using a superadditive argument. The idea for the proof of superadditivity is based again on interpolation. However a rigorous proof was obtained only quite recently by F. Guerra and F.L.Toninelli. It turns out that the convenient setting is the one of comparison of Gaussian processes. The necessary tool is provided by the following result, due to Slepian and Kahane.

Lemma 2.1. *Let X and Y be two independent n-dimensional Gaussian vectors with covariances C^X and C^Y respectively. Assume that*

$$\begin{array}{rcll} C^X_{i,j} & \leq & C^Y_{i,j} & \text{if} \quad i \neq j \\ C^X_{i,i} & = & C^Y_{i,i} & \end{array} \tag{2.2}$$

Let $F : \mathbb{R}^n \to \mathbb{R}$ be a function such that its second derivatives satisfy

$$\frac{\partial^2 F}{\partial x_i \partial x_j}(x) \leq 0 \qquad \text{if} \quad i \neq j \tag{2.3}$$

Then

$$\mathbb{E}(F(X)) \geq \mathbb{E}(F(Y)) \tag{2.4}$$

Proof. For an interpolating parameter $t \in [0,1]$ consider the new Gaussian random vectors $V(t)$ and $W(t)$ defined by

$$V(t) = \sqrt{t}X + \sqrt{1-t}Y \tag{2.5}$$

$$W(t) = \frac{\partial}{\partial t} V(t) = \frac{1}{2\sqrt{t}} X - \frac{1}{2\sqrt{1-t}} Y \tag{2.6}$$

Define the function

$$\varphi(t) = \mathbb{E}[F(V(t))] \tag{2.7}$$

and observe that it satisfies the following boundary conditions:

$$\varphi(1) = \mathbb{E}(F(X)) \tag{2.8}$$

$$\varphi(0) = \mathbb{E}(F(Y)) \tag{2.9}$$

The function $\varphi(t)$ is non-decreasing. Indeed computing its derivative we find

$$\begin{aligned}\frac{\partial}{\partial t}\varphi(t) &= \mathbb{E}\left(\sum_{i=1}^{n} \frac{\partial}{\partial V_i(t)} F(V(t)) \cdot \frac{\partial}{\partial t} V_i(t)\right) \\ &= \mathbb{E}\left(\sum_{i=1}^{n} \frac{\partial}{\partial V_i(t)} F(V(t)) \cdot W_i(t)\right)\end{aligned}$$

Then, using Wick's rule, we have

$$\frac{\partial}{\partial t}\varphi(t) = \mathbb{E}\left(\sum_{i,j=1}^{n} \frac{\partial^2}{\partial V_j(t)\partial V_i(t)} F(V(t)) \cdot \mathbb{E}[W_i(t)V_j(t)]\right)$$

From eq. (2.5) and (2.6) we deduce

$$\mathbb{E}[W_i(t)V_j(t)] = \frac{1}{2}(C^X_{ij} - C^Y_{ij})$$

From the hypothesis (2.2) and (2.3) we then deduce that

$$\frac{\partial}{\partial t}\varphi(t) \geq 0$$

which, by the fundamental theorem of calculus, gives

$$\varphi(1) = \mathbb{E}(F(X)) \geq \varphi(0) = \mathbb{E}(F(Y))$$

□

Using the previous lemma one can establish the following theorem.

Theorem 2.10. *The thermodynamic limit of the pressure*

$$p(\beta, h) = \lim_{N\to\infty} p_N(\beta, h)$$

exists and is finite for all β and h.

Proof. The idea is to show that the sequence $P_N(\beta, h) = N p_N(\beta, h)$ is a superadditive sequence; this immediately entails that the limit $\lim_{N\to\infty} \frac{P_N(\beta,h)}{N}$ exists. To prove superadditivity means that we must show that for any $N_1, N_2, N \in \mathbb{N}$ with $N_1 + N_2 = N$

$$N p_N(\beta, h) \geq N_1 p_{N_1}(\beta, h) + N_2 p_{N_2}(\beta, h) \tag{2.11}$$

In order to see this we consider the function $F : \mathbb{R}^{2^N} \to \mathbb{R}$ which, for a 2^N-dimensional vector $x = \{x(\sigma)\}_{\sigma\in\Omega_N}$, is defined as

$$F(x) = \log\left(\sum_{\sigma\in\Omega_N} \exp[-\beta x(\sigma)] \exp[\beta h \sum_{i=1}^{N} \sigma_i]\right)$$

It is easily checked that the off-diagonal components of the second derivative matrix are all non positive, i.e.

$$\frac{\partial^2 F}{\partial x(\sigma)\partial x(\tau)}(x) \leq 0 \qquad \text{if} \quad \sigma \neq \tau$$

We now consider a splitting of an SK-system of size N into two subsystem of sizes N_1, N_2 with $N_1 + N_2 = N$. For a given spin configuration $\sigma = (\sigma_1, \ldots, \sigma_N)$ let us denote in the most natural way by $\sigma' = (\sigma_1, \ldots, \sigma_{N_1})$ the spin variable for the subsystem of size N_1 and by $\sigma'' = (\sigma_{N_1+1}, \ldots, \sigma_N)$ the N_2 spin variables for the second subsystem. Obviously we have $\sigma = \sigma' \cup \sigma''$. Define the two gaussian processes

$$X(\sigma) = H_N(\sigma)$$

$$Y(\sigma) = H_{N_1}(\sigma') + H_{N_2}(\sigma'')$$

where $H_N(\sigma)$, $H_{N_1}(\sigma')$ and $H_{N_2}(\sigma'')$ are three independent Gaussian processes with covariances respectively

$$C_N(\sigma,\tau) = \frac{N}{2} q_N^2(\sigma,\tau)$$

$$C_{N_1}(\sigma',\tau') = \frac{N_1}{2} q_{N_1}^2(\sigma',\tau')$$

$$C_{N_2}(\sigma'',\tau'') = \frac{N_2}{2} q_{N_2}^2(\sigma'',\tau'')$$

where

$$q_N(\sigma,\tau) = \frac{1}{N}\sum_{i=1}^{N} \sigma_i \tau_i$$

$$q_{N_1}(\sigma',\tau') = \frac{1}{N_1}\sum_{i=1}^{N_1} \sigma_i \tau_i$$

$$q_{N_2}(\sigma'',\tau'') = \frac{1}{N_2}\sum_{i=N_1+1}^{N} \sigma_i \tau_i$$

It is immediate to check that

$$\begin{array}{lcl} C^X_{\sigma,\tau} & \leq & C^Y_{\sigma,\tau} \qquad \text{if} \quad \sigma \neq \tau \\ C^X_{\sigma,\sigma} & = & C^Y_{\sigma,\sigma} \end{array} \tag{2.12}$$

since, by convexity,

$$q_N(\sigma,\tau)^2 \leq \frac{N_1}{N} q_{N_1}(\sigma',\tau')^2 + \frac{N_2}{N} q_{N_2}(\sigma'',\tau'')^2 \tag{2.13}$$

With these definitions, the superadditivity condition (2.11) is equivalent to show that

$$\mathbb{E}(F(X)) \geq \mathbb{E}(F(Y))$$

But this relation is immediately verified by applying the previous theorem, since all hypothesis are verified. This complete the proof of superadditivity. To show that the limit pressure $p(\beta, h)$ is finite it is enough to prove that the sequence $p_N(\beta, h)$ is bounded from above. This follows from an application of Jensen's inequality, which gives

$$\frac{1}{N}\mathbb{E}\left(\log \sum_{\sigma\in\Omega_N} e^{\beta H_N(\sigma)+\beta h \sum_{i=1}^N \sigma_i}\right) \leq \frac{\beta^2}{4} + \log(2\cosh(\beta h))$$

□

3 Self-averaging of the pressure

The next problem we are going to tackle is to establish that the random pressure coincides with the quenched pressure almost surely in the thermodynamic limit. This will be a consequence of the following concentration of measures result by M. Talagrand.

Theorem 3.1. *For a given $M \in \mathbb{N}$ consider a Lipschitz function $F : \mathbb{R}^M \to \mathbb{R}$ of Lipschitz constant A, that is*

$$|F(x) - F(y)| \leq A||x - y|| \tag{3.2}$$

where $||\cdot||$ denotes the Euclidean norm. If $\{X_i^1\}_{i=1,\dots,M}$ is a family of i.i.d. standard gaussian random variables, then the following holds for any $t > 0$

$$\mathbb{P}\left(|F(X^1) - \mathbb{E}(F(X^1))| \geq t\right) \leq 2\exp\left(-\frac{t^2}{4A^2}\right) \tag{3.3}$$

Proof. The proof is again a consequence of interpolation. We first observe that if we are able to prove that

$$\mathbb{E}\{\exp(s[F(X^1) - \mathbb{E}(F(X^1))])\} \leq \exp(s^2A^2) \tag{3.4}$$

then (3.3) follows. Indeed we have:

$$\mathbb{P}\left(|F(X^1) - \mathbb{E}(F(X^1))| \geq t\right) = \mathbb{P}\left(F(X^1) - \mathbb{E}(F(X^1)) \geq t\right) + \mathbb{P}\left(F(X^1) - \mathbb{E}(F(X^1)) \leq -t\right)$$

Considering the first probability on the right hand side of the previous equation. We have, by Markov inequality,

$$\begin{aligned}
\mathbb{P}\left(F(X^1) - \mathbb{E}(F(X^1)) \geq t\right) &= \mathbb{P}\left(\exp(s[F(X^1) - \mathbb{E}(F(X^1))]) \geq \exp(st)\right) \\
&\leq \frac{\mathbb{E}\{\exp(s[F(X^1) - \mathbb{E}(F(X^1))])\}}{\exp(st)} \\
&\leq \exp(s^2A^2 - st) \qquad (3.5)
\end{aligned}$$

where in the last line we used (3.4). Optimizing in s, we obtain the optimal value $s = t/2A^2$ and this gives

$$\mathbb{P}\left(F(X^1) - \mathbb{E}(F(X^1)) \geq t\right) = \exp\left(-\frac{t^2}{4A^2}\right) \tag{3.6}$$

One can repeat the same computation as in (3.5) for F replaced by $-F$, which has the same Lipschitz constant. This gives the bound

$$\mathbb{P}\left(F(X^1) - \mathbb{E}(F(X^1)) \leq -t\right) = \exp\left(-\frac{t^2}{4A^2}\right) \tag{3.7}$$

Adding up together (3.6) and (3.7), we obtain (3.3).

Thus to prove the theorem it suffices to prove eq. (3.4). To achieve this we introduce two $2M$-dimensional Guassian families X and Y as follows:

$$X = (X^1, X^2) \tag{3.8}$$

$$Y = (Y^1, Y^1) \tag{3.9}$$

where $\{X_i^1\}$, $\{X_i^2\}$, $\{Y_i^1\}$ for $i = 1, \ldots M$ are i.i.d. standard Gaussian random variables. For an arbitrary $2M$-dimensional vector $Z = (Z^1, Z^2)$, we then consider the function $G : \mathbb{R}^{2M} \to \mathbb{R}_+$ defined by

$$G(Z) = G(Z^1, Z^2) = \exp(s[F(Z^1) - F(Z^2)]) \tag{3.10}$$

For an interpolating parameter $t \in [0, 1]$ consider the new $2M$-dimensional Gaussian random vectors $V(t)$ and $W(t)$ defined by

$$V(t) = \sqrt{t}X + \sqrt{1-t}Y \tag{3.11}$$

$$W(t) = \frac{\partial}{\partial t}V(t) = \frac{1}{2\sqrt{t}}X - \frac{1}{2\sqrt{1-t}}Y \tag{3.12}$$

More explicitly:

$$V(t) = (V^1(t), V^2(t)) = \left(\sqrt{t}X^1 + \sqrt{1-t}Y^1, \sqrt{t}X^2 + \sqrt{1-t}Y^1\right) \tag{3.13}$$

$$W(t) = (W^1(t), W^2(t)) = \left(\frac{1}{2\sqrt{t}}X^1 - \frac{1}{2\sqrt{1-t}}Y^1, \frac{1}{2\sqrt{t}}X^2 - \frac{1}{2\sqrt{1-t}}Y^1\right) \tag{3.14}$$

Consider then the function

$$\varphi(t) = \mathbb{E}[G(V(t))]$$

which satisfies the following boundary conditions:

$$\begin{aligned} \varphi(1) &= \mathbb{E}(G(X^1, X^2)) \\ \varphi(0) &= \mathbb{E}(G(Y^1, Y^1)) = 1 \end{aligned} \tag{3.15}$$

The derivative of the function $\varphi(t)$ with respect to t reads

$$\begin{aligned} \frac{\partial}{\partial t}\varphi(t) &= \mathbb{E}\left(\sum_{i=1}^{2M} \frac{\partial}{\partial V_i(t)}G(V(t)) \cdot \frac{\partial}{\partial t}V_i(t)\right) \\ &= \mathbb{E}\left(\sum_{i=1}^{2M} \frac{\partial}{\partial V_i(t)}G(V(t)) \cdot W_i(t)\right) \end{aligned}$$

Then, using Wick's rule, we have

$$\frac{\partial}{\partial t}\varphi(t) = \mathbb{E}\left(\sum_{i,j=1}^{2M} \frac{\partial^2}{\partial V_j(t)\partial V_i(t)} G(V(t)) \cdot \mathbb{E}[W_i(t)V_j(t)]\right)$$

From eq. (3.13) and (3.14) we deduce

$$\mathbb{E}[W_i(t)V_j(t)] = -\frac{1}{2}(\delta_{i,j-M} - \delta_{i+M,j})$$

so that we arrive at

$$\frac{\partial}{\partial t}\varphi(t) = -\mathbb{E}\left(\sum_{i=1}^{M} \frac{\partial^2}{\partial V_i(t)\partial V_{i+M}(t)} G(V(t))\right) \tag{3.16}$$

By an explicit computation

$$\frac{\partial^2}{\partial V_i(t)\partial V_{i+M}(t)} G(V(t)) = -s^2 G(V(t)) \frac{\partial}{\partial V_i(t)} F(V^1(t)) \cdot \frac{\partial}{\partial V_{i+M}(t)} F(V^2(t))$$

We now remind a very remarkable property of Lipschitz function:

$$||\nabla F(x)||^2 \leq A^2 \tag{3.17}$$

Indeed we can write

$$A||x-y|| \geq |F(x) - F(y)| = |\langle \nabla F(x), y-x \rangle| + o(||y-x||^2)$$

which implies

$$A||x-y|| \geq ||\nabla F(x)|| \cdot ||y-x|| + o(||y-x||^2)$$

Choosing $y - x = \lambda \nabla F(x)$, for a constant λ, this implies (3.17). The property (3.17) can be used, together with Cauchy-Schwartz inequality, in eq. (3.16), which then gives

$$\frac{\partial}{\partial t}\varphi(t) \leq s^2 A^2 \varphi(t)$$

Integrating the previous differential inequality in the interval $[0,1]$ we find

$$\varphi(1) \leq \varphi(0) e^{s^2 A^2}$$

Remembering the boundary conditions (3.15) this means that

$$\mathbb{E}\left(\exp(s[F(X^1) - F(X^2)])\right) \leq e^{s^2 A^2}$$

Using Jensen's inequality with respect to the random vector X^2, which is independnet of X^1, one obtains

$$\mathbb{E}\left(\exp(s[F(X^1) - \mathbb{E}(F(X^2))])\right) \leq e^{s^2 A^2}$$

But of course $\mathbb{E}(F(X^2)) = \mathbb{E}(F(X^1))$. So this proves inequality (3.4). □

Having established the previous theorem, the following result gives the almost sure convergence of the random pressure to the quenched pressure.

Theorem 3.18. *For any β and h, and any $t \geq 0$,*

$$\mathbb{P}\left(|p_N(\beta,h,J) - p_N(\beta,h)| \geq t\right) \leq 2\exp\left(-\frac{Nt^2}{2\beta^2}\right) \tag{3.19}$$

In particular this implies

$$\mathbb{E}\left([p_N(\beta,h,J) - p_N(\beta,h)]^2\right) \leq \frac{4\beta^2}{N} \tag{3.20}$$

Proof. Consider the pressure $p_N(\beta,h,J)$ as a function of the N^2 i.i.d. standard Gaussian random variable $J = \{J_{i,j}\}_{i,j=1,\dots,N}$. This function is Lipschitz. Indeed, considering two independent N^2-dimensional i.i.d. standard Gaussian families J and J', we may write

$$p_N(\beta,h,J) - p_N(\beta,h,J') = \int_0^1 \frac{d}{dt} p_N(\beta,h,tJ + (1-t)J')dt$$

On the other hand

$$\frac{d}{dt} p_N(\beta,h,tJ + (1-t)J') = \frac{1}{N}\frac{\beta}{\sqrt{2N}}\sum_{i,j=1}^{N} \omega_t(\sigma_i\sigma_j)(J_{ij} - J'_{ij})$$

Using the fact that $|\sigma_i\sigma_j| \leq 1$ and Cauchy-Schwarz inequality we find that

$$|p_N(\beta,h,J) - p_N(\beta,h,J')| \leq \frac{\beta}{\sqrt{2N}}||J - J'||$$

which proves that the function $p_N(\beta,h,J)$ is Lipschitz, with Lipschitz constant $\frac{\beta}{\sqrt{2N}}$. The claim (3.19) follows from the previous theorem. The second claim (3.20) is then immediately proved using the identity

$$\mathbb{E}\left([p_N(\beta,h,J) - p_N(\beta,h)]^2\right) = 2\int_0^\infty t\mathbb{P}\left(|p_N(\beta,h,J) - p_N(\beta,h)| \geq t\right)dt$$

□

References

[1] F. Guerra, *Spin glasses*, (2005) http://arxiv.org/abs/cond-mat/0507581

[2] M. Talagrand, *Spin glasses: a challenge to mathematicians*, Spingler (2000)