A short course in mean field spin glasses

Anton Bovier
Weierstraß-Institut für Angewandte Analysis und Stochastik
Mohrenstraße 39
10117 Berlin, Germany

Institut für Mathematik
Technische Universität Berlin
Straße des 17. Juni 136
10623 Berlin, Germany

Contents

1

Introduction

The common feature of mean-field models is that the spatial structure of the lattice $\mathbb{Z}^d$ is abandoned in favour of a simpler setting, where sites are indexed by the natural numbers and all spins are supposed to interact with each other, irrespective of their distance.

1.1 SK-models

The naive analog of the Curie–Weiss Hamiltonian with random couplings would be

$$H_N[\omega](\sigma) = -\frac{1}{2N} \sum_{1\le i,j\le N} J_{ij}[\omega]\sigma_i\sigma_j \tag{1.1}$$

for, say, J_{ij} some family of i.i.d. random variables. The main interest in this model concerns the case when the random couplings have mean zero. In this case, we will see shortly that the normalization factor, N^{-1}, is actually inappropriate and must be replaced by $N^{-1/2}$, to obtain an interesting model. Namely, we certainly want the free energy to be an extensive quantity, i.e. to be of order N. This means that, for typical realizations of the disorder, there must be at least some spin configurations, σ, for which $H_N(\sigma) \sim CN$, for some $C > 0$. Thus, we must estimate $\mathbb{P}[\max_\sigma H_N(\sigma) \ge CN]$. But,

$$\begin{aligned}
\mathbb{P}[\max_\sigma H_N(\sigma) \ge CN] &\le \sum_{\sigma\in\mathcal{S}_N} \mathbb{P}[H_N(\sigma) \ge CN] \qquad (1.2)\\
&= \sum_{\sigma\in\mathcal{S}_N} \inf_{t\ge0} e^{-tCN} \mathbb{E} e^{t\frac{1}{2N}\sum_{i,j\in\Lambda_N\times\Lambda_N} J_{ij}[\omega]\sigma_i\sigma_j}\\
&= \sum_{\sigma\in\mathcal{S}_N} \inf_{t\ge0} e^{-tCN} \prod_{i,j\in\Lambda_N\times\Lambda_N} \mathbb{E} e^{t\frac{1}{2N} J_{ij}[\omega]\sigma_i\sigma_j}
\end{aligned}$$

where we assumed that the exponential moments of J_{ij} exist. A standard estimate then shows that, for some constant c, $\mathbb{E}e^{t\frac{1}{2N}J_{ij}[\omega]\sigma_i\sigma_j} \leq e^{c\frac{t^2}{2N^2}}$, and so

$$\mathbb{P}[\max_{\sigma} H_N(\sigma) \geq CN] \leq 2^N \inf_{t\geq 0} e^{-tCN} e^{ct^2/2} \leq 2^N e^{-\frac{C^2N^2}{2c}} \tag{1.3}$$

which tends to zero with N. Thus, our Hamiltonian is never of order N, but at best of order $\sqrt{N}$. The proper Hamiltonian for what is called the *Sherrington–Kirkpatrick model* (or short SK-model), is thus

$$H_N^{SK} \equiv -\frac{1}{\sqrt{2N}} \sum_{i,j\in\Lambda_N\times\Lambda_N} J_{ij}\sigma_i\sigma_j \tag{1.4}$$

where the random variables $J_{ij} = J_{ji}$ are i.i.d. for $i \leq j$ with mean zero (or at most $J_0 N^{-1/2}$) and variance normalized to one for $i \neq j$ and to two for $i = j$ In its original, and mostly considered, form, the distribution is moreover taken to be Gaussian. Note that $\sum_{ij} |N^{-1/2}J_{ij}\sigma_i\sigma_j| \sim N^{3/2}$, and that competing signs play a major role.

This model was introduced by Sherrington and Kirkpatrick in 1976 [22] as an attempt to furnish a simple, solvable mean-field model for the then newly discovered class of materials called *spin-glasses*. However, it turned out that the innocent looking modifications made to create a spin-glass model that looks similar to the Curie–Weiss model had thoroughly destroyed the simplifying properties that made the latter so easily solvable, and that a model with an enormously complex structure had been invented. Using highly innovative ideas based on ad hoc mathematical structures, Parisi (see [17]) produced in the mid-eighties a heuristic framework that explained the properties of the model. Only very recently, these predictions have to some extent been rigorously justified through work of F. Guerra [12] and M. Talagrand [26], which we will explain in these lectures.

1.2 Gaussian process

It will be useful to introduce a different point of view on the SK-model, which allows us to put it in a wider context. This point of view consists of regarding the Hamiltonian (1.4) as a *Gaussian random process*[1] indexed

[1] The choice of Gaussian couplings and hence Gaussian processes may appear too restrictive, and, from a physical point of view, poorly motivated. It turns out, however, that the Gaussian nature of the processes considered is not really important, and that a large class of models have the same asymptotics as the corresponding Gaussian ones (at least on the level of the free energy) [6]. It is however a good idea to start with the simplest situation.

by the set $\mathcal{S}_N$, i.e. by the N-dimensional hypercube. We will restrict our attention to the case when the J_{ij} are centred Gaussian random variables. In this case, $H_N(\sigma)$ is in fact a centred Gaussian random process which is fully characterized by its covariance function

$$\begin{aligned} \operatorname{cov}(H_N(\sigma), H_N(\sigma')) &= \frac{1}{2N} \sum_{1\leq i,j,l,k\leq N} \mathbb{E} J_{ij} J_{kl} \sigma_i \sigma_j \sigma'_k \sigma'_l \\ &= \frac{1}{N} \sum_{1\leq i,j\leq N} \sigma_i \sigma'_i \sigma_j \sigma'_j = N R_N(\sigma,\sigma')^2 \end{aligned} \tag{1.5}$$

where $R_N(\sigma,\sigma') \equiv N^{-1} \sum_{i=1}^N \sigma_i \sigma'_i$ is usually called the *overlap* between the two configurations σ and σ'. It is useful to recall that the overlap is closely related to the *Hamming distance* $d_{HAM}(\sigma,\sigma') \equiv \#(i \leq N : \sigma_i \neq \sigma'_i)$, namely $R_N(\sigma,\sigma') = (1 - 2N^{-1} d_{HAM}(\sigma,\sigma'))$.

Seen this way, the SK- model is a particular example of a class of models whose Hamiltonian are centred Gaussian random process on the hypercube with covariance depending only on $R_N(\sigma,\sigma')$,

$$\operatorname{cov}(H_N(\sigma), H_N(\sigma')) = N\xi(R_N(\sigma,\sigma')) \tag{1.6}$$

normalized such that $\xi(1) = 1$. A class of examples considered in the literature are the so-called p-spin SK-models, which are obtained by choosing $\xi(x) = |x|^p$. They enjoy the property that they may be represented in a form similar to the SK-Hamiltonian, except that the two-spin interaction must be replaced by a p-spin one:

$$H_N^{p-SK}(\sigma) = \frac{-1}{\sqrt{N^{p-1}}} \sum_{1\leq i_1,\dots,i_p\leq N} J_{i_1\dots i_p} \sigma_{i_1} \dots \sigma_{i_p} \tag{1.7}$$

with $J_{i_1,\dots,i_p}$ i.i.d. standard normal random variables[1]. As we will see later, the difficulties in studying the statistical mechanics of these models is closely linked to the understanding of the extremal properties of the corresponding random processes. While Gaussian processes have been heavily analyzed in the mathematical literature (see e.g. [16, 1]), the known results were not enough to recover the heuristic results obtained in the physics literature. This is one reason why this particular field of mean-field spin-glass models has considerable intrinsic interest for mathematics.

[1] Sometimes the terms where some indices in the sum (1.7) coincide are omitted. This can be included in our framework by making ξ explicitly N-dependent in a suitable way. Although this has an effect for instance on the fluctuations of the free energy (see [5]), for our present purposes this is not relevant and we choose the form with the simplest expression for the covariance.

1.3 The generalized random energy models

The class of models we have just introduced depends on the particular choice of the Hamming distance as the metric on the hypercube. It is only natural to think of further classes of models that can be obtained by other choices of the metric. A particularly important alternative choice is the *lexicographic distance*: Given two sequences σ and τ, we look at the first value of the index i for which the sequences differ, i.e. $\sigma_i \neq \tau_i$. Naturally, if this value is N, then $\sigma = \tau$ and thus their distance is zero, while, if $i = 1$, then we consider them maximally apart. The quantity

$$d_N(\sigma, \tau) \equiv N^{-1} (\min(i : \sigma_i \neq \tau_i) - 1) \tag{1.8}$$

is thus analogous to the overlap $R_N(\sigma, \tau)$. The corresponding Gaussian processes are then characterised by covariances given by

$$\mathrm{cov}(H_N(\sigma), H_N(\tau)) = NA(d_N(\sigma, \tau)) \tag{1.9}$$

where A can be chosen to be any non-decreasing function on $[0, 1]$, and can be thought of as a probability distribution function. The choice of the lexicographic distance entails some peculiar features. First, this distance is an *ultrametric*, i.e. for any three configurations σ, τ, ρ,

$$d_N(\sigma, \tau) = \min (d_N(\sigma, \rho), d_N(\tau, \rho)) \tag{1.10}$$

This fact will be seen to have remarkable consequences, that make the Gibbs measures of these models fully analyzable, even though they will show as much complexity as those of the SK-models. Moreover, a clever comparison between the two types of processes is instrumental in the analysis of the SK-models themselves, as will be explained later. I will therefore devote a considerable amount of attention to the analysis of these models before returning to the study of the SK-models.

1.4 Gibbs measures, partition functions, free energies

The principle object that statistical mechanics aims to study in such models are the *Gibbs measures*, $\mu_{b,N}$, that the random Hamiltonians induce on the state spaces $\mathcal{S}_N$, and their asymptotics as N tends to infinity. One defines

$$\mu_{\beta,N}(\sigma) \equiv \frac{2^{-N} \exp(-\beta H_N(\sigma))}{Z_{\beta,N}} \tag{1.11}$$

where the normalizing factor,

$$Z_{\beta,N} \equiv \mathbb{E}_\sigma \exp(-\beta H_N(\sigma)) = \sum -\sigma \in \mathcal{S}_N 2^{-N} \exp(-\beta H_N(\sigma)) \tag{1.12}$$

is called the partition function. Frequently one introduces an extra parameter, h, called the *magnetic field*, which consists in tilting the a-priori distribution of the sin variables σ_i by an exponential factor $e^{\beta h \sigma_i}$. We will come back to this only much later. The first quantity one would like to compute is the exponential asymptotics of the normalising partition function, which gives some hint as to how much the interaction is pushing the measure away form the uniform a-priory measure. The corresponding rate,

$$f_{\beta,N} \equiv -\frac{1}{\beta N} \ln Z_{\beta,N} \tag{1.13}$$

is called the *free energy*. The first task is to compute the limit

$$f_\beta \equiv \lim_{N\uparrow\infty} f_{\beta,N}$$

if it exists. We now know that this limit exists almost surely, and is a constant which can be expressed through a complicated variational principle. Since the minus-sign and the factor β^{-1} are sometimes annoying, we will often consider the nameless quantity

$$\Phi_{\beta,N} \equiv \frac{1}{N} \ln Z_{\beta,N} \tag{1.14}$$

instead of the free energy.

1.5 Gibbs states

In principle we would like to know much more about the nature of the Gibbs measures of these models then just the free energy. The real question concerns the geometry of the mass distribution on the hypercube. This is, however, a rather complicated issue that we will possibly touch towards the end of these notes in general, although we will illustrate this in a very simple case soon. Basic question to formulate are: Is the support of μ_β connected or disconnected? If not, what is the distribution of mass on the different components? What is the structure of the set of connected components?

2

The simples example: the REM

The most familiar processes for a probabilist are surely independent random variables. The corresponding model is known as the *random energy model* or *REM*.

The REM was introduced by Derrida [9, 10] in 1980. One might think that it would be an extremely misleading model, and that its study teaches us nothing about other models such as the SK model. This turn out not to be the case. We set

$$H_N(\sigma) = -\sqrt{N} X_\sigma \tag{2.1}$$

where X_σ, $\sigma \in \mathcal{S}_N$, are 2^N i.i.d. standard normal random variables.

2.1 Ground-state energy and free energy

The first and as we will see crucial information we require concerns the value of the maximum of the variables X_σ, i.e. the ground-state energy. For i.i.d. random variables, this is of course not very hard.

Lemma 2.1.1 *The family of random variables introduced above satisfies*

$$\lim_{N\uparrow\infty} \max_{\sigma\in\mathcal{S}_N} N^{-1/2} X_\sigma = \sqrt{2\ln 2} \tag{2.2}$$

both almost surely and in mean.

Proof Since everything is independent,

$$\mathbb{P}\left[\max_{\sigma\in\mathcal{S}_N} X_\sigma \leq u\right] = \left(1 - \frac{1}{\sqrt{2\pi}} \int_u^\infty e^{-x^2/2} dx\right)^{2^N} \tag{2.3}$$

and we just need to know how to estimate the integral appearing here. This is something we should get used to quickly, as it will occur all over the place. It will always be done using the fact that, for $u > 0$,

$$\frac{1}{u} e^{-u^2/2} \left(1 - 2u^{-2}\right) \leq \int_u^\infty e^{-x^2/2} dx \leq \frac{1}{u} e^{-u^2/2} \tag{2.4}$$

We see that for our probability to converge neither to zero nor to one, u must be chosen in such a way that the integral is of order 2^{-N}. With the help of the bounds (2.4), one can show with a little computation that, if we define $u_N(x)$ by

$$\frac{2^N}{\sqrt{2\pi}} \int_{u_N(x)}^\infty e^{-z^2/2} dx = e^{-x} \tag{2.5}$$

then (for $x > -\ln N / \ln 2$)

$$u_N(x) = \sqrt{2N \ln 2} + \frac{x}{\sqrt{2N \ln 2}} - \frac{\ln(N \ln 2) + \ln 4\pi}{2\sqrt{2N \ln 2}} + o(1/\sqrt{N}) \tag{2.6}$$

Thus

$$\mathbb{P}\left[\max_{\sigma \in \mathcal{S}_N} X_\sigma \leq u_N(x)\right] = \left(1 - 2^{-N} e^{-x}\right)^{2^N} \to e^{-e^{-x}} \tag{2.7}$$

In other terms, the random variable $u_N^{-1}\left(\max_{\sigma \in \mathcal{S}_N} X_\sigma\right)$ converges in distribution to a random variable with double-exponential distribution (this is the most classic result of *extreme value statistics*, see [15]). The assertion of the lemma is now a simple corollary of this fact. □

Next we turn to the analysis of the partition function. In this model, the partition function is just the sum of i.i.d. random variables, i.e.

$$Z_{\beta,N} \equiv 2^{-N} \sum_{\sigma \in \mathcal{S}_N} e^{\beta\sqrt{N} X_\sigma} \tag{2.8}$$

A first guess would be that a *law of large numbers* might hold, implying that $Z_{\beta,N} \sim \mathbb{E} Z_{\beta,N}$, and hence

$$\lim_{N \uparrow \infty} \Phi_{\beta,N} = \lim_{N \uparrow \infty} \frac{1}{N} \ln \mathbb{E} Z_{\beta,N} = \frac{\beta^2}{2}, \text{a.s.} \tag{2.9}$$

It turns out that this is indeed true, but only for small enough values of β, and that there is a critical value β_c associated with a breakdown of the law of large numbers. The analysis of this problem will allow us to compute the free energy exactly.

Theorem 2.1.2 *In the REM,*

$$\lim_{N \uparrow \infty} \mathbb{E} \Phi_{\beta,N} = \begin{cases} \frac{\beta^2}{2}, & \text{for } \beta \leq \beta_c \\ \frac{\beta_c^2}{2} + (\beta - \beta_c)\beta_c, & \text{for } \beta \geq \beta_c \end{cases} \tag{2.10}$$

where $\beta_c = \sqrt{2 \ln 2}$.

Proof We use the method of truncated second moments, which was introduced in the context of spin-glasses by M. Talagrand [23, 24, 25].

We will first derive an upper bound for $\mathbb{E}\Phi_{\beta,N}$. Note first that by Jensen's inequality, $\mathbb{E} \ln Z \leq \ln \mathbb{E} Z$, and thus

$$\mathbb{E}\Phi_{\beta,N} \leq \frac{\beta^2}{2} \tag{2.11}$$

On the other hand we have that

$$\begin{aligned} \mathbb{E}\frac{d}{d\beta}\Phi_{\beta,N} &= N^{-1/2}\mathbb{E}\frac{\mathbb{E}_\sigma X_\sigma e^{\beta\sqrt{N}X_\sigma}}{Z_{\beta,N}} \\ &\leq N^{-1/2}\mathbb{E}\max_{\sigma\in\mathcal{S}_N} X_\sigma \leq \beta\sqrt{2\ln 2}(1+C/N) \end{aligned} \tag{2.12}$$

for some constant C. Combining (2.11) and (2.12), we deduce that

$$\mathbb{E}\Phi_{\beta,N} \leq \inf_{\beta_0 \geq 0} \begin{cases} \frac{\beta^2}{2}, & \text{for } \beta \leq \beta_0 \\ \frac{\beta_0^2}{2} + (\beta - \beta_0)\sqrt{2\ln 2}(1 + C/N), & \text{for } \beta \geq \beta_0 \end{cases} \tag{2.13}$$

It is easy to see that the infimum is realized (ignore the C/N correction) for $\beta_0 = \sqrt{2\ln 2}$. This shows that the right-hand side of (2.10) is an upper bound.

It remains to show the corresponding lower bound. Note that, since $\frac{d^2}{d\beta^2}\Phi_{\beta,N} \geq 0$, the slope of $\Phi_{\beta,N}$ is non-decreasing, so that the theorem will be proven if we can show that $\Phi_{\beta,N} \to \beta^2/2$ for all $\beta < \sqrt{2\ln 2}$, i.e. that the law of large numbers holds up to this value of β. A natural idea to prove this is to estimate the variance of the partition function[1]. Naively, one would compute

$$\begin{aligned} \mathbb{E}Z_{\beta,N}^2 &= \mathbb{E}_\sigma \mathbb{E}_{\sigma'} \mathbb{E} e^{\beta\sqrt{N}(X_\sigma + X_{\sigma'})} \\ &= 2^{-2N}\left(\sum_{\sigma\neq\sigma'} e^{N\beta^2} + \sum_\sigma e^{2N\beta^2}\right) \\ &= e^{N\beta^2}\left[(1-2^{-N}) + 2^{-N}e^{N\beta^2}\right] \end{aligned} \tag{2.14}$$

where all we used is that for $\sigma \neq \sigma'$ X_σ and $X_{\sigma'}$ are independent. The second term in the square brackets is exponentially small if and only if $\beta^2 < \ln 2$. For such values of β we have that

[1] This idea can be traced to Aizenman, Lebowitz, and Ruelle [2], and, later, Comets and Neveu [7], who used it in the proofs of a central limit theorem for the free energy in the SK-model.

$$\begin{aligned}\mathbb{P}\left[\left|\ln\frac{Z_{\beta,N}}{\mathbb{E}Z_{\beta,N}}\right|>\epsilon N\right]&=\mathbb{P}\left[\frac{Z_{\beta,N}}{\mathbb{E}Z_{\beta,N}}<e^{-\epsilon N}\text{ or }\frac{Z_{\beta,N}}{\mathbb{E}Z_{\beta,N}}>e^{\epsilon N}\right]\\&\leq\mathbb{P}\left[\left(\frac{Z_{\beta,N}}{\mathbb{E}Z_{\beta,N}}-1\right)^2>\left(1-e^{-\epsilon N}\right)^2\right]\\&\leq\frac{\mathbb{E}Z^2_{\beta,N}/(\mathbb{E}Z_{\beta,N})^2-1}{(1-e^{-\epsilon N})^2}\\&\leq\frac{2^{-N}+2^{-N}e^{N\beta^2}}{(1-e^{-\epsilon N})^2}\end{aligned}\qquad(2.15)$$

which is more than enough to get (2.9). But of course this does not correspond to the critical value of β claimed in the proposition! Some reflection shows that the point here is that when computing $\mathbb{E}e^{\beta\sqrt{N}2X_\sigma}$, the dominant contribution comes from the part of the distribution of X_σ where $X_\sigma\sim 2\beta\sqrt{N}$, whereas in the evaluation of $\mathbb{E}Z_{\beta,N}$ the values of X_σ where $X_\sigma\sim\beta\sqrt{N}$ give the dominant contribution. Thus one is led to realize that instead of the second moment of Z one should compute a truncated version of it, namely, for $c\geq 0$,

$$\widetilde{Z}_{\beta,N}(c)\equiv\mathbb{E}_\sigma e^{\beta\sqrt{N}X_\sigma}\mathbb{1}_{X_\sigma<c\sqrt{N}}\qquad(2.16)$$

An elementary computation using (2.4) shows that, if $c>\beta$, then

$$\mathbb{E}\widetilde{Z}_{\beta,N}(c)=e^{\frac{\beta^2N}{2}}\left(1-\frac{e^{-N\beta^2/2}}{\sqrt{2\pi N}(c-\beta)}(1+O(1/N)\right)\qquad(2.17)$$

so that such a truncation essentially does not influence the mean partition function. Now compute the mean of the square of the truncated partition function (neglecting irrelevant $O(1/N)$ errors):

$$\mathbb{E}\widetilde{Z}^2_{\beta,N}(c)=(1-2^{-N})[\mathbb{E}\widetilde{Z}_{\beta,N}(c)]^2+2^{-N}\mathbb{E}e^{\beta\sqrt{N}2X_\sigma}\mathbb{1}_{X_\sigma<c\sqrt{N}})\qquad(2.18)$$

where

$$\mathbb{E}\,e^{2\beta\sqrt{N}X_\sigma}\mathbb{1}_{X_\sigma<c\sqrt{N}}=\begin{cases}e^{2\beta^2N}, & \text{if } 2\beta<c\\ 2^{-N}\frac{e^{2c\beta N-\frac{c^2N}{2}}}{(2\beta-c)\sqrt{2\pi N}}, & \text{otherwise,}\end{cases}\qquad(2.19)$$

Combined with (2.17) this implies that, for $c/2<\beta<c$,

$$\frac{2^{-N}\mathbb{E}\,e^{2\beta\sqrt{N}X_\sigma}\mathbb{1}_{X_\sigma<c\sqrt{N}}}{\left(\mathbb{E}\,\widetilde{Z}_{N,\beta}\right)^2}=\frac{e^{-N(c-\beta)^2-N(2\ln 2-c^2)/2}}{(2\beta-c)\sqrt{N}}\qquad(2.20)$$

Therefore, for all $c<\sqrt{2\ln 2}$, and all $\beta<c$,

$$\mathbb{E}\left[\frac{\widetilde{Z}_{\beta,N}(c) - \mathbb{E}\widetilde{Z}_{\beta,N}(c)}{\mathbb{E}\widetilde{Z}_{\beta,N}(c)}\right]^2 \leq e^{-Ng(c,\beta)} \tag{2.21}$$

with $g(c, \beta) > 0$. Thus Chebyshev's inequality implies that

$$\mathbb{P}\left[|\widetilde{Z}_{\beta,N}(c) - \mathbb{E}\widetilde{Z}_{\beta,N}(c)| > \delta\mathbb{E}\widetilde{Z}_{\beta,N}(c)\right] \leq \delta^{-2}e^{-Ng(c,\beta)} \tag{2.22}$$

and so, in particular,

$$\lim_{N\uparrow\infty} \frac{1}{N}\mathbb{E}\ln \widetilde{Z}_{\beta,N}(c) = \lim_{N\uparrow\infty} \frac{1}{N} \ln \mathbb{E}\widetilde{Z}_{\beta,N}(c) \tag{2.23}$$

for all $\beta < c < \sqrt{2\ln 2} = \beta_c$. But this implies that for all $\beta < \beta_c$, we can chose c such that

$$\lim_{N\uparrow\infty} \frac{1}{N} \ln \mathbb{E}Z_{\beta,N} \geq \lim_{N\uparrow\infty} \frac{1}{N} \ln \mathbb{E}\widetilde{Z}_{\beta,N}(c) = \frac{\beta^2}{2} \tag{2.24}$$

This proves the theorem. □

2.2 Fluctuations and limit theorems

In the previous section we went to some length to compute the limit of the free energy. However, computing the free energy is not quite enough to get a full understanding of a model, and in particular the Gibbs states. The limit of the free energy has been seen to be a non-random quantity. A question of central importance is to understand how and on what level the randomness shows up in the corrections to the limiting behaviour. This question has been fully analysed in [5], here I will only discuss the low-temperature regime $\beta > \sqrt{2\ln 2}$.

Theorem 2.2.1 *Let $\mathcal{P}$ denotes the Poisson point process on $\mathbb{R}$ with intensity measure $e^{-x}dx$. Then, in the REM, with $\alpha = \beta/\sqrt{2\ln 2}$, if $\beta > \sqrt{2\ln 2}$,*

$$e^{-N[\beta\sqrt{2\ln 2}-\ln 2]+\frac{\alpha}{2}[\ln(N\ln 2)+\ln 4\pi]} Z_{\beta,N} \xrightarrow{\mathcal{D}} \int_{-\infty}^{\infty} e^{\alpha z}\mathcal{P}(dz) \tag{2.25}$$

and

$$N\left(\Phi_{\beta,N} - \mathbb{E}\Phi_{\beta,N}\right) \xrightarrow{\mathcal{D}} \ln \int_{-\infty}^{\infty} e^{\alpha z}\mathcal{P}(dz) - \mathbb{E}\ln \int_{-\infty}^{\infty} e^{\alpha z}\mathcal{P}(dz). \tag{2.26}$$

Proof Basically, the idea is very simple. We expect that for β large, the partition function will be dominated by the configurations σ corresponding to the largest values of X_σ. Thus we split $Z_{\beta,N}$ carefully into

$$Z^x_{N,\beta} \equiv \mathbb{E}_\sigma e^{\beta\sqrt{N}X_\sigma} \mathbb{1}_{\{X_\sigma \leq u_N(x)\}} \tag{2.27}$$

and $Z_{\beta,N} - Z^x_{\beta,N}$. Let us first consider the last summand. We introduce the random variable

$$\mathcal{W}_N(x) = Z_{\beta,N} - Z^x_{\beta,N} = 2^{-N} \sum_{\sigma\in\mathcal{S}_N} e^{\beta\sqrt{N}X_\sigma} \mathbb{1}_{\{X_\sigma > u_N(x)\}} \tag{2.28}$$

It is convenient to rewrite this as (we ignore the sub-leading corrections to $u_N(x)$ and only keep the explicit part of (2.6))

$$\begin{aligned}\mathcal{W}_N(x) &= 2^{-N} \sum_{\sigma\in\mathcal{S}_N} e^{\beta\sqrt{N}u_N(u_N^{-1}(X_\sigma))} \mathbb{1}_{\{u_N^{-1}(X_\sigma)>x\}} \\ &= e^{N(\beta\sqrt{2\ln 2}-\ln 2)-\frac{\alpha}{2}[\ln(N\ln 2)+\ln 4\pi]} && (2.29)\\ &\quad\times \sum_{\sigma\in\mathcal{S}_N} e^{\alpha u_N^{-1}(X_\sigma)} \mathbb{1}_{\{u_N^{-1}(X_\sigma)>x\}} && (2.30)\\ &\equiv \frac{1}{C(\beta,N)} \sum_{\sigma\in\mathcal{S}_N} e^{\alpha u_N^{-1}(X_\sigma)} \mathbb{1}_{\{u_N^{-1}(X_\sigma)>x\}} && (2.31)\end{aligned}$$

where

$$\alpha \equiv \beta/\sqrt{2\ln 2} \tag{2.32}$$

and $C(b,N)$ is defined through the last identity. The key to most of what follows relies on the famous result on the convergence of the extreme value process to a Poisson point process (for a proof see, e.g., [15]):

Theorem 2.2.2 *Let $\mathcal{P}_N$ be point process on $\mathbb{R}$ given by*

$$\mathcal{P}_N \equiv \sum_{\sigma\in\mathcal{S}_N} \delta_{u_N^{-1}(X_\sigma)} \tag{2.33}$$

Then $\mathcal{P}_N$ converges weakly to a Poisson point process on $\mathbb{R}$ with intensity measure $e^{-x}dx$.

Clearly, the weak convergence of $\mathcal{P}_N$ to $\mathcal{P}$ implies convergence in law of the right-hand side of (2.29), provided that $e^{\alpha x}$ is integrable on $[x,\infty)$ w.r.t. the Poisson point process with intensity e^{-x}. This is, in fact, never a problem: the Poisson point process has almost surely support on a finite set, and therefore $e^{\alpha x}$ is always a.s. integrable. Note, however, that for $\beta \geq \sqrt{2\ln 2}$ the mean of the integral is infinite, indicating the passage to the low-temperature regime.

Lemma 2.2.3 *Let $\mathcal{W}_N(x), \alpha$ be defined as above, and let $\mathcal{P}$ be the Poisson point process with intensity measure $e^{-z}dz$. Then*

$$C(\beta,N)\mathcal{W}_N(x) \xrightarrow{\mathcal{D}} \int_x^\infty e^{\alpha z}\mathcal{P}(dz) \tag{2.34}$$

Next we show that the contribution of the truncated part of the partition function is negligible compared to this contribution. For this it is enough to compute the mean values

$$\begin{aligned}
\mathbb{E}Z^x_{\beta,N} &\sim e^{N\beta^2/2} \int_{-\infty}^{u_N(x)-1\beta\sqrt{N}} \frac{dz}{\sqrt{2}\pi} e^{-\frac{z^2}{2}} \\
&\sim e^{N\beta^2/2} \frac{e^{-(u_N(x)-\beta\sqrt{N})^2/2}}{\sqrt{2\pi}(\beta\sqrt{N}-u_N(x))} \\
&\sim \frac{2^{-N}e^{x(\alpha-1)}}{\alpha-1} e^{N(\beta\sqrt{2\ln 2}-\ln 2)-\frac{\alpha}{2}[\ln(N\ln 2)+\ln 4\pi]} \\
&= \frac{e^{x(\alpha-1)}}{\alpha-1}\frac{1}{C(\beta,N)}
\end{aligned} \tag{2.35}$$

so that

$$C(\beta,N)\mathbb{E}Z^x_{\beta,N} \sim \frac{e^{x(\alpha-1)}}{\alpha-1}$$

which tends to zero as $x \downarrow -\infty$, and so $C(\beta,N)\mathbb{E}Z^x_{\beta,N}$ converges to zero in probability. The assertions of Theorem 2.2.1 follow. □

2.3 The Gibbs measure

With our preparation on the fluctuations of the free energy, we have accumulated enough understanding about the partition function that we can deal with the Gibbs measures. Clearly, there are a number of ways of trying to describe the asymptotics of the Gibbs measures. Recalling the general discussion on random Gibbs measures from Part II, it should be clear that we are seeking a result on the convergence in distribution of random measures. To be able to state such a result, we have to introduce a topology on the spin configuration space that makes it uniformly compact. The first natural candidate would seem to be the product topology. However, given what we already know about the partition function, this topology does not appear ideally adapted to give adequate information. Recall that at low temperatures, the partition function was dominated by a 'few' spin configurations with exceptionally large energy. This is a feature that should remain visible in a limit theorem. A nice way to do this consists in mapping the hypercube to the interval $(0, 1]$ via

$$\mathcal{S}_N \ni \sigma \to r_N(\sigma) \equiv 1 - \sum_{i=1}^{N} (1 - \sigma_i) 2^{-i-1} \in (0, 1] \tag{2.36}$$

Define the pure point measure $\tilde{\mu}_{\beta,N}$ on $(0, 1]$ by

$$\tilde{\mu}_{\beta,N} \equiv \sum_{\sigma \in \mathcal{S}_N} \delta_{r_N(\sigma)} \mu_{\beta,N}(\sigma) \tag{2.37}$$

Our results will be expressed in terms of the convergence of these measures. It will be understood in the sequel that the space of measures on $(0, 1]$ is equipped with the topology of weak convergence, and all convergence results hold with respect to this topology.

The behaviour of the measure at low temperatures is much more interesting. Let us introduce the Poisson point process $\mathcal{R}$ on the strip $(0, 1] \times \mathbb{R}$ with intensity measure $\frac{1}{2} dy \times e^{-x} dx$. If (Y_k, X_k) denote the atoms of this process, define a new point process $\mathcal{M}_\alpha$ on $(0, 1] \times (0, 1]$ whose atoms are (Y_k, w_k), where

$$w_k \equiv \frac{e^{\alpha X_k}}{\int \mathcal{R}(dy, dx) e^{\alpha x}} \tag{2.38}$$

for $\alpha > 1$. With this notation we have that:

Theorem 2.3.1 *If $\beta > \sqrt{2 \ln 2}$, with $\alpha = \beta / \sqrt{2 \ln 2}$, then*

$$\tilde{\mu}_{\beta,N} \xrightarrow{\mathcal{D}} \tilde{\mu}_\beta \equiv \int_{(0,1] \times (0,1]} \mathcal{M}_\alpha(dy, dw) \delta_y w \tag{2.39}$$

Proof With $u_N(x)$ defined in (2.6), we define the point process $\mathcal{R}_N$ on $(0, 1] \times \mathbb{R}$ by

$$\mathcal{R}_N \equiv \sum_{\sigma \in \mathcal{S}_N} \delta_{(r_N(\sigma), u_N^{-1}(X_\sigma))} \tag{2.40}$$

A standard result of extreme value theory (see [15], Theorem 5.7.2) is easily adapted to yield that

$$\mathcal{R}_N \xrightarrow{\mathcal{D}} \mathcal{R}, \text{ as } N \uparrow \infty \tag{2.41}$$

where the convergence is in the sense of weak convergence on the space of sigma-finite measures endowed with the (metrisable) topology of vague convergence. Note that

$$\mu_{\beta,N}(\sigma) = \frac{e^{\alpha u_N^{-1}(X_\sigma)}}{\sum_\sigma e^{\alpha u_N^{-1}(X_\sigma)}} = \frac{e^{\alpha u_N^{-1}(X_\sigma)}}{\int \mathcal{R}_N(dy,dx)e^{\alpha x}} \tag{2.42}$$

Since $\int \mathcal{R}_N(dy,dx)e^{\alpha x} < \infty$ a.s., we can define the point process

$$\mathcal{M}_{\alpha,N} \equiv \sum_{\sigma\in\mathcal{S}_N} \delta_{\left(r_N(\sigma), \frac{\exp(\alpha u_N^{-1}(X_\sigma))}{\int \mathcal{R}_N(dy,dx)\exp(\alpha x)}\right)} \tag{2.43}$$

on $(0,1]\times(0,1]$. Then

$$\tilde{\mu}_{\beta,N} = \int \mathcal{M}_{\alpha,N}(dy,dw)\delta_y w \tag{2.44}$$

The only non-trivial point in the convergence proof is to show that the contribution to the partition functions in the denominator from atoms with $u_N(X_\sigma) < x$ vanishes as $x \downarrow -\infty$. But this is precisely what we have shown to be the case in the proof of part (v) of Theorem 2.2.1. Standard arguments then imply that first $\mathcal{M}_{\alpha,N} \stackrel{\mathcal{D}}{\to} \mathcal{M}_\alpha$, and consequently, (2.39). □

Remark 2.3.1 In [20], Ruelle introduced a process $\mathcal{W}_\alpha$ that is nothing but the marginal of $\mathcal{M}_\alpha$ on the 'masses', i.e. on the second variable, as an asymptotic description of the distribution of the masses of the Gibbs measure of the REM in the infinite-volume limit. Our result implies in particular that indeed $\sum_{\sigma\in\mathcal{S}_N} \delta_{\mu_{\beta,N}(\sigma)} \stackrel{\mathcal{D}}{\to} \mathcal{W}_\alpha$ if $\alpha > 1$. Neveu in [18] gave a sketch of the proof of this fact. Note that Theorem 2.3.1 contains in particular the convergence of the Gibbs measure in the product topology on $\mathcal{S}_N$, since cylinders correspond to certain subintervals of $(0,1]$. The formulation of Theorem 2.3.1 is very much in the spirit of the metastate approach to random Gibbs measures. The limiting measure is a measure on a continuous space, and each point measure on this set may appear as a 'pure state'. The 'metastate', i.e. the law of the random measure $\tilde{\mu}_\beta$, is a probability distribution, concentrated on the countable convex combinations of pure states, randomly chosen by a Poisson point process from an uncountable collection, while the coefficients of the convex combination are again random variables and are selected via another point process. The only aspect of metastates that is missing here is that we have not 'conditioned on the disorder'. The point is, however, that there is no natural filtration of the disorder space compatible with, say, the product topology, and thus in this model we have no natural urge to 'fix the disorder locally'; note, however, that it is possible to represent the i.i.d. family X_σ as a sum of 'local' couplings, i.e. let J_Δ, for any $\Delta \subset \mathbb{N}$ be i.i.d. standard normal variables. Then we can represent $X_\sigma = 2^{-N/2} \sum_{\Delta \subset \{1,\dots,N\}} \sigma_\Delta J_\Delta$; obviously these variables become independent of any of the J_Δ, with Δ fixed, so that conditioning on them would not change the metastate.

Let us discuss the properties of the limiting process $\tilde{\mu}_\beta$. It is easy to see that, with probability one, the support of $\tilde{\mu}_\beta$ is the entire interval $(0,1]$. But its mass is concentrated on a countable set, i.e. the measure is pure point. To see this, consider the rectangle $A_\epsilon \equiv (\ln \epsilon, \infty) \times (0,1]$. The process $\mathcal{R}$ restricted to this set has finite total intensity given by ϵ^{-1}, i.e. the number of atoms in that set is a Poissonian random variable with parameter ϵ^{-1}. If we remove the projection of these finitely many random points from $(0,1]$, the remaining mass is given by

$$\int_{(0,1]\times(-\infty,\ln\epsilon)} \mathcal{R}(dy,dx) \frac{e^{\alpha x}}{\int \mathcal{P}(dx')e^{\alpha x'}} = \int_{-\infty}^{\ln \epsilon} \mathcal{P}(dx) \frac{e^{\alpha x}}{\int \mathcal{P}(dx')e^{\alpha x'}} \tag{2.45}$$

We want to get a lower bound in probability on the denominator. The simplest possible bound is obtained by estimating the probability of the integral by the contribution of the largest atom, which of course follows the double-exponential distribution. Thus

$$\mathbb{P}\left[\int \mathcal{P}(dx)e^{\alpha x} \leq Z\right] \leq e^{-e^{-\ln Z/\alpha}} = e^{-Z^{-\frac{1}{\alpha}}} \tag{2.46}$$

Setting $\Omega_Z \equiv \{\mathcal{P} : \int \mathcal{P}(dx)e^{\alpha x} \leq Z\}$, we conclude that, for $\alpha > 1$,

$$\begin{aligned}
&\mathbb{P}\left[\int_{-\infty}^{\ln\epsilon} \mathcal{P}(dx) \frac{e^{\alpha x}}{\int \mathcal{P}(dx')e^{\alpha x'}} > \gamma\right] \qquad (2.47)\\
&\leq \mathbb{P}\left[\int_{-\infty}^{\ln\epsilon} \mathcal{P}(dx) \frac{e^{\alpha x}}{\int \mathcal{P}(dx')e^{\alpha x'}} > \gamma,\, \Omega_Z^c\right] + \mathbb{P}[\Omega_Z]\\
&\leq \mathbb{P}\left[\int_{-\infty}^{\ln\epsilon} \mathcal{P}(dx)e^{\alpha x} > \gamma Z,\, \Omega_Z^c\right] + \mathbb{P}[\Omega_Z]\\
&\leq \mathbb{P}\left[\int_{-\infty}^{\ln\epsilon} \mathcal{P}(dx)e^{\alpha x} > \gamma Z\right] + \mathbb{P}[\Omega_Z]\\
&\leq \frac{\mathbb{E}\int_{-\infty}^{\ln\epsilon} \mathcal{P}(dx)e^{\alpha x}}{\gamma} + \mathbb{P}[\Omega_Z] \leq \frac{\epsilon^{\alpha-1}}{(\alpha-1)\gamma Z} + e^{-Z^{-\frac{1}{\alpha}}}
\end{aligned}$$

Obviously, for any positive γ it is possible to choose Z as a function of ϵ in such a way that the right-hand side tends to zero. But this implies that, with probability one, all of the mass of the measure $\tilde{\mu}_\beta$ is carried by a countable set, implying that $\tilde{\mu}_\beta$ is pure point. A more refined analysis is given in [14].

So we see that the phase transition in the REM expresses itself via a change of the properties of the infinite-volume Gibbs measure mapped to the interval from the Lebesgue measure at high temperatures to a random dense pure point measure at low temperatures.

2.4 The replica overlap

While the random measure description of the phase transition in the REM yields an elegant description of the thermodynamic limit, the projection to the unit interval loses the geometric structure of the space $\mathcal{S}_N$. The description of this geometry is a central issue that we will continue to pursue. We would like to describe 'where' in $\mathcal{S}_N$ the mass of the Gibbs measure is located. In a situation where no particular reference configuration exists, a natural possibility is to compare two independent copies of spin configurations drawn from the same Gibbs distribution to each other. To make this precise, recall the function $r_N : \mathcal{S}_N \times \mathcal{S}_N \to (0,1]$ defined in (2.36). We are interested in the probability distribution of $R_N(\sigma, \sigma')$ under the product measure $\mu_{\beta,N} \otimes \mu_{\beta,N}$, i.e. define a probability measure, $\psi_{\beta,N}$, on $[-1,1]$ by

$$\psi_{\beta,N}[\omega](dz) \equiv \mu_{\beta,N}[\omega] \otimes \mu_{\beta,N}[\omega]\left(R_N(\sigma,\sigma') \in dz\right) \tag{2.48}$$

As we will see later, the analysis of the replica overlap is a crucial tool for studying the Gibbs measures of more complicated models. The following exposition is intended to give a first introduction to this approach:

Theorem 2.4.1 *(i) For all $\beta < \sqrt{2\ln 2}$*

$$\lim_{N\uparrow\infty} \psi_{\beta,N} = \delta_0, \; a.s. \tag{2.49}$$

(ii) For all $\beta > \sqrt{2\ln 2}$

$$\psi_{\beta,N} \xrightarrow{\mathcal{D}} \delta_0 \left(1 - \int \mathcal{W}_\alpha(dw) w^2\right) + \delta_1 \int \mathcal{W}_\alpha(dw) w^2 \tag{2.50}$$

Proof We write for any $\Delta \subset [-1,1]$

$$\psi_{\beta,N}(\Delta) = Z_{\beta,N}^{-2} \mathbb{E}_\sigma \mathbb{E}_{\sigma'} \sum_{\substack{t\in\Delta \\ R_N(\sigma,\sigma')=t}} e^{\beta\sqrt{N}(X_\sigma + X_{\sigma'})} \tag{2.51}$$

First, the denominator is bounded from below by $[\widetilde{Z}_{\beta,N}(c)]^2$, and by (2.22), with probability of order $\delta^{-2}\exp(-Ng(c,\beta))$, this in turn is larger than $(1-\delta)^2[\mathbb{E}\widetilde{Z}_{\beta,N}(c)]^2$. Let first $\beta < \sqrt{2\ln 2}$. Assume initially that $\Delta \subset (0,1) \cup [-1,0)$. We conclude that

$$\mathbb{E}\psi_{\beta,N}(\Delta) \le \frac{1}{(1-\delta)^2}\mathbb{E}_\sigma\mathbb{E}_{\sigma'} \sum_{\substack{t\in\Delta \\ R_N(\sigma,\sigma')=t}} 1 + \delta^{-2}e^{-g(c,\beta)N} \tag{2.52}$$
$$\sim \frac{1}{\sqrt{2\pi N}}\frac{1}{(1-\delta)^2}\sum_{t\in\Delta}\frac{2e^{-NI(t)}}{\sqrt{1-t^2}} + \delta^{-2}e^{-g(c,\beta)N}$$

for any $\beta < c < \sqrt{2\ln 2}$, where $I : [-1,1] \to \mathbb{R}$ denotes the Cramèr entropy function defined in (**??**). Here we used that, if $(1-t)N = 2\ell$, $\ell = 0,\dots,N$, then

$$\mathbb{E}_\sigma\mathbb{E}_{\sigma'}\mathbb{1}_{R_N(\sigma,\sigma')=t} = 2^{-N}\binom{N}{\ell} \tag{2.53}$$

and the approximation of the binomial coefficient given in (**??**) and (**??**)). Under our assumptions on Δ, we see immediately from this representation that the right-hand side of (2.52) is clearly exponentially small in N. It remains to consider the mass at the point 1, i.e.

$$\psi_{\beta,N}(1) = Z_{\beta,N}^{-2}\mathbb{E}_\sigma 2^{-N}e^{2\beta\sqrt{N}X_\sigma} \tag{2.54}$$

But we can split

$$\mathbb{E}_\sigma e^{2\beta\sqrt{N}X_\sigma} = Z_{2\beta,N}^x + (Z_{\beta,N} - Z_{2\beta,N} - Z_{2\beta,N}^x) \tag{2.55}$$

For the first, we use that

$$\mathbb{E}Z_{2\beta,N}^x \le 2^{-N}e^{2\beta N\sqrt{2\ln 2}} \tag{2.56}$$

and for the second we use that it is

$$e^{N2\beta\sqrt{2\ln 2}e^{\alpha}[\ln(N\ln 2)+4\pi]}\sum_\sigma e^{2\alpha u_N^{-1}(x_\sigma)} \tag{2.57}$$

Both terms are exponentially smaller than $2^N e^{\beta^2 N}$, and thus the mass of $\psi_{\beta,N}$ at 1 also vanishes. This proves (2.49).

Let now $\beta > \sqrt{2\ln 2}$. We use the truncation introduced in Section 2.2. Note first that, for any interval Δ,

$$\left|\psi_{\beta,N}(\Delta) - Z_{\beta,N}^{-2}\mathbb{E}_\sigma\mathbb{E}_{\sigma'}\sum_{\substack{t\in\Delta \\ R_N(\sigma,\sigma')=t}}\mathbb{1}_{X_\sigma,X_{\sigma'}\ge u_N(x)}e^{\beta\sqrt{N}(X_\sigma+X_{\sigma'})}\right| \le \frac{2Z_{\beta,N}^x}{Z_{\beta,N}} \tag{2.58}$$

We have already seen in the proof of Theorem (2.2.1) (see (**??**)), that the right-hand side of (2.58) tends to zero in probability, as first $N \uparrow \infty$ and then $x \downarrow -\infty$. On the other hand, for $t \ne 1$,

$$\mathbb{P}\left[\exists_{\sigma,\sigma,:R_N(\sigma,\sigma')=t} : X_\sigma > u_N(x) \wedge X_{\sigma'} > u_N(x)\right] \tag{2.59}$$
$$\leq \mathbb{E}_{\sigma\sigma'} \mathbb{I}_{R_N(\sigma,\sigma')=t} \, 2^{2N} \mathbb{P}\left[X_\sigma > u_N(x)\right]^2 = \frac{2e^{-I(t)N}e^{-2x}}{\sqrt{2\pi N}\sqrt{1-t^2}}$$

by the definition of $u_N(x)$ (see (2.5)). This implies again that any interval $\Delta \subset [-1,1) \cup [-1,0)$ has zero mass. To conclude the proof it is enough to compute $\psi_{\beta,N}(1)$. Clearly

$$\psi_{\beta,N}(1) = \frac{2^{-N}\mathbb{Z}_{2\beta,N}}{Z^2_{\beta,N}} \tag{2.60}$$

By (v) of Theorem 2.2.1, one sees easily that

$$\psi_{\beta,N}(1) \xrightarrow{\mathcal{D}} \frac{\int e^{2\alpha z}\mathcal{P}(dz)}{\left(\int e^{\alpha z}\mathcal{P}(dz)\right)^2} \tag{2.61}$$

Expressing the left-hand side of (2.61) in terms of the point process $\mathcal{W}_\alpha$, defined in (2.38), yields the expression for the mass of the atom at 1; since the only other atom is at zero, the assertion (ii) follows from the fact that $\psi_{\beta,N}$ is a probability measure. This concludes the proof. □

3

Gaussian comparison and applications

Comparison of Gaussian processes has a long-standing tradition in the analysis of extremal and regularity properties of Gaussian processes. It should not have come as a surprise that it provides a key tool in the analysis of spin-glasses.

3.1 A theorem of Slepian-Kahane

Lemma 3.1.2 *Let X and Y be two independent n-dimensional Gaussian vectors. Let D_1 and D_2 be subsets of $\{1,\dots,n\}\times\{1,\dots,n\}$. Assume that*

$$\begin{aligned} \mathbb{E}X_iX_j \leq \mathbb{E}Y_iY_j, &\quad \text{if} \quad (i,j)\in D_1 \\ \mathbb{E}X_iX_j \geq \mathbb{E}Y_iY_j, &\quad \text{if} \quad (i,j)\in D_2 \\ \mathbb{E}X_iX_j = \mathbb{E}Y_iY_j, &\quad \text{if} \quad (i,j)\notin D_1\cup D_2 \end{aligned} \tag{3.1}$$

Let f be a function on $\mathbb{R}^n$, such that its second derivatives satisfy

$$\begin{aligned} \frac{\partial^2}{\partial x_i \partial x_j} f(x) \geq 0, &\quad \text{if} \quad (i,j)\in D_1 \\ \frac{\partial^2}{\partial x_i \partial x_j} f(x) \leq 0, &\quad \text{if} \quad (i,j)\in D_2 \end{aligned} \tag{3.2}$$

Then

$$\mathbb{E}f(X) \leq \mathbb{E}f(Y) \tag{3.3}$$

Proof The first step of the proof consists of writing

$$f(X) - f(Y) = \int_0^1 dt \frac{d}{dt} f(X^t) \tag{3.4}$$

where we define the interpolating process

$$X^t \equiv \sqrt{t}\,X + \sqrt{1-t}\,Y \tag{3.5}$$

Next observe that

$$\frac{d}{dt}f(X^t) = \frac{1}{2}\sum_{i=1}^{n}\frac{\partial}{\partial x_i}f(X^t)\left(t^{-1/2}X_i - (1-t)^{-1/2}Y_i\right) \tag{3.6}$$

Finally, we use the generalization of the Gaussian integration by parts formula (**??**) to the multivariate setting:

Lemma 3.1.3 *Let X_i, $i \in \{1,\dots,n\}$ be a multivariate Gaussian process, and let $g : \mathbb{R}^n \to \mathbb{R}$ be a differentiable function of at most polynomial growth. Then*

$$\mathbb{E}g(X)X_i = \sum_{j=1}^{n}\mathbb{E}(X_iX_j)\mathbb{E}\frac{\partial}{\partial x_j}g(X) \tag{3.7}$$

Applied to the mean of the left-hand side of (3.6) this yields

$$\mathbb{E}f(X) - \mathbb{E}f(Y) = \frac{1}{2}\sum_{i,j}\int_{0,1}dt\,(\mathbb{E}X_iX_j - \mathbb{E}Y_iY_j)\,\mathbb{E}\frac{\partial^2}{\partial x_j\partial x_i}f(X^t) \tag{3.8}$$

from which the assertion of the theorem can be read off. □

Note that Equation (3.8) has the flavour of the fundamental theorem of calculus on the space of Gaussian processes.

All classical comparison results in the theory of extremes of Gaussian processes can be seen as special (or limiting) cases of this result. In all applications, one chooses a process Y with well understood properties to obtain estimates for a process X under investigations. For instance, choosing for Y just iid rv's allows us to bound the free energy of any of our models through that of the REM. Unfortunately, these bounds will as such be too crude to be interesting. A main obstacle to progress along these lines was the fact that there appear no immediate candidates for comparison processes about which we can say something that would allow to improve these bounds.

3.2 The thermodynamic limit through comparison

A key observation that has led to the recent breakthroughs in spin glass theory, made by F. Guerra and F.-L. Toninelli, was that Theorem 3.1.2 can be used to prove the existence of the limit of the free energy. This problem had remained open for decades, as none of the usual sub-additivity arguments appeared to work.

Theorem 3.2.4 *[13] Assume that X_σ is a normalized Gaussian process on $\mathcal{S}_N$ with covariance*

$$\mathbb{E}X_\sigma X_\tau = \xi(R_N(\sigma,\tau)) \tag{3.9}$$

where $\xi : [-1,1] \to [0,1]$ is convex and even. Then

$$\lim_{N\uparrow\infty} \frac{-1}{\beta N}\mathbb{E}\ln \mathbb{E}_\sigma e^{\beta\sqrt{N}X_\sigma} \equiv f_\beta \tag{3.10}$$

exists.

Proof The proof of this fact is frightfully easy, once you think about using Theorem 3.1.2. Choose any $1 < M < N$. Let $\sigma = (\hat\sigma, \check\sigma)$ where $\hat\sigma = (\sigma_1, \sigma_2, \dots, \sigma_M)$, and $\check\sigma = (\sigma_{M+1}, \dots, \sigma_N)$. Define independent Gaussian processes $\hat X$ and $\check X$ on $\mathcal{S}_M$ and $\mathcal{S}_{N-M}$, respectively, such that

$$\mathbb{E}\hat X_{\hat\sigma}\hat X_{\hat\tau} = \xi(R_M(\hat\sigma,\hat\tau)) \tag{3.11}$$

and

$$\mathbb{E}\check X_{\check\sigma}\check X_{\check\tau} = \xi(R_{N-M}(\check\sigma,\check\tau) \tag{3.12}$$

Set

$$Y_\sigma \equiv \sqrt{\tfrac{M}{N}}\hat X_{\hat\sigma} + \sqrt{\tfrac{N-M}{N}}\check X_{\check\sigma} \tag{3.13}$$

Clearly,

$$\begin{aligned}\mathbb{E}Y_\sigma Y_\tau &= \tfrac{M}{N}\xi(R_M(\hat\sigma,\hat\tau)) + \tfrac{N-M}{N}\xi(R_{N-M}(\check\sigma,\check\tau)) \qquad (3.14)\\ &\geq \xi\left(\tfrac{M}{N}R_M(\hat\sigma,\hat\tau) + \tfrac{N-M}{N}(R_{N-M}(\check\sigma,\check\tau))\right) = \xi(R_N(\sigma,\tau))\end{aligned}$$

Define real-valued functions $F_N(x) \equiv \ln \mathbb{E}_\sigma e^{\beta\sqrt{N}x_\sigma}$ on $\mathbb{R}^{2^N}$. It is straightforward that

$$\mathbb{E}F_N(Y) = \mathbb{E}F_M(X) + \mathbb{E}F_{N-M}(X) \tag{3.15}$$

A simple computation shows that, for $\sigma \neq \tau$,

$$\frac{\partial^2}{\partial x_\sigma \partial x_\tau}F_N(x) = -\frac{2^{-2N}\beta^2 N e^{\beta\sqrt{N}(x_\sigma + x_\tau)}}{Z^2_{\beta,N}} \leq 0 \tag{3.16}$$

Thus, Theorem 3.1.2 tells us that

$$\mathbb{E}F_N(X) \geq \mathbb{E}F_N(Y) = \mathbb{E}F_M(X) + \mathbb{E}F_{N-M}(X) \tag{3.17}$$

This implies that the sequence $-\mathbb{E}F_N(X)$ is subadditive, and this in turn implies (see Section **??**) that the free energy exists , provided it is bounded, which is easy to verify (see e.g. the discussion on the correct normalisation in the SK model). □

The same ideas can be used for other types of Gaussian processes, e.g. the GREM-type models discussed above [8].

Convergence of the free energy in mean implies readily almost sure convergence. This follows from a general *concentration of measure* principle for functions of Gaussian random variables, analogous to Theorem **??**. The following result can be found e.g. in [16], page 23:

Theorem 3.2.5 *Let $X_1, \dots, X_M$ be independent standard normal random variables, and let $f : R^M \to R$ be Lipschitz continuous with Lipschitz constant $\|f\|_{Lip}$. Set $g \equiv f(X_1, \dots, X_M)$. Then*

$$\mathbb{P}\left[|g - \mathbb{E}g| > x\right] \leq 2 \exp\left(-\frac{x^2}{2\|f\|_{Lip}^2}\right) \tag{3.18}$$

Corollary 3.2.6 *Assume that the function ξ is analytic with positive Taylor coefficients. Then*

$$\mathbb{P}\left[|f_{\beta,N} - \mathbb{E}f_{\beta,N}| > x\right] \leq 2 \exp\left(-\frac{Nx^2}{2\beta^2}\right) \tag{3.19}$$

In particular, $\lim_{N\uparrow\infty} f_{\beta,N} = f_\beta$*, almost surely.*

Proof If $\xi(x) = \sum_{p=1}^{\infty} a_p^2 x^p$, we can construct X_σ as

$$X_\sigma = \sum_{p=1}^{\infty} a_p N^{-p/2} \sum_{1 \leq i_1, \dots, i_p \leq N} J^{(p)}_{i_1 \dots, i_p} \sigma_{i_1} \dots \sigma_{i_p} \tag{3.20}$$

with standard i.i.d. Gaussians $J^{(p)}_{i_1 \dots i_p}$. Check that, as a function of these variables, the free energy is Lipschitz with Lipschitz constant $\beta N^{-1/2}$. □

3.3 The Parisi solution and Guerra's bounds

The original approach to the computation of the free energy in the SK models in the theoretical physics literature is quite remarkable. Morally, it constitutes an attempt to compute n moments, where n seems strangely to go to infinity and to zero simultaneously. This idea is the basis of the so-called *replica method*, or *replica trick* , that is a widely used tool in the heuristic analysis of disordered systems. While we must leave a detailed exposition of the heuristic approaches to the standard textbooks [17, 19, 11], it maybe worthwhile to discuss this approach briefly. The basic idea is the observation that $\lim_{n\downarrow 0}(x^n - 1) = \ln x$, and that for integer n, $\mathbb{E}Z_{\beta,N}$ can be computed, at least in the sense that it is possible to perform the average over the disorder and to obtain a deterministic expression (as we have seen in the case $n = 2$). The obvious problem is that the computation of integer moments of $Z_{\beta,N}$ does not immediately allow us to infer information on the limit $n \downarrow 0$, where n is to be considered real valued.

Having seen that the comparison theorem yields existence of the free energy almost for free, it is a bit more surprising that it allows to give a variational principle that allows to compare the free energy to Parisi's solution. As we do not yet know what Parisi's solution is, we will learn about it in the process.

3.3.1 An extended caparison principle

As I have mentioned, comparison of the free energy of SK models to simpler models do not immediately seem to work. The idea is to use comparison on a much richer class of processes. Basically, rather than comparing one process to another, we construct an extended process on a product space and use comparison on this richer space. Let us first explain this in an abstract setting. We have a process X on a space $\mathcal{S}$ equipped with a probability measure $\mathbb{E}_\sigma$. We want to compute as usual the average of the logarithm of the partition function $F(X) = \ln \mathbb{E}_\sigma e^{\beta X_\sigma}$. Now consider a second space $\mathcal{T}$ equipped with a probability law $\mathbb{E}_\alpha$. Choose a Gaussian process, Y, independent of X, on this space, and define a further independent process, Z, on the product space $\mathcal{S} \times \mathcal{T}$. Define real valued functions, G, H, on the space of real valued functions on $\mathcal{T}$ and $\mathcal{S} \times \mathcal{T}$, respectively, via $G(y) \equiv \ln \mathbb{E}_\alpha e^{\beta y_\alpha}$ and $H(z) = \ln \mathbb{E}_\sigma \mathbb{E}_\alpha e^{\beta z_{\sigma,\alpha}}$. Note that $H(X+Y) = F(X) + G(Y)$. Assume that the covariances are chosen such that

$$\mathrm{cov}(X_\sigma, X_{\sigma'}) + \mathrm{cov}(Y_\alpha, Y_{\alpha'}) \geq \mathrm{cov}(Z_{\sigma,\alpha}, Z_{\sigma',\alpha'}) \tag{3.21}$$

Since we know that the second derivatives of H are negative, we get from Theorem 3.1.2 that

$$\mathbb{E}F(X) + \mathbb{E}G(Y) = \mathbb{E}H(X+Y) \leq \mathbb{E}H(Z) \tag{3.22}$$

This is a useful relation if we know how to compute $\mathbb{E}G(Y)$ and $\mathbb{E}H(Z)$. This idea may look a bit crazy at first sight, but we must remember that we have a lot of freedom in choosing the auxiliary spaces and processes to our convenience. Before turning to the issue whether we can find useful computable processes Y and Z, let us see why we could hope to find in this way *sharp* bounds.

3.4 The extended variational principle and thermodynamic equilibrium

To do so, we will show that, in principle, we can represent the free energy in the thermodynamic limit in the form $\mathbb{E}H(Z) - \mathbb{E}G(Y)$. To this end let $\mathcal{S} = \mathcal{S}_M$ and $\mathcal{T} = \mathcal{S}_N$, both equipped with their natural probability measure $\mathbb{E}_\sigma$. We will think of $N \gg M$, and both tending to infinity eventually. We write again $\mathcal{S} \times \mathcal{T} \ni \sigma = (\hat\sigma, \check\sigma)$. Consider the process X_σ on $\mathcal{S}_{N+M}$ with covariance $\xi(R_{N+M}(\sigma, \sigma'))$. We would like to write this as

$$X_\sigma = \hat X_{\hat\sigma} + \check X_{\check\sigma} + Z_\sigma \tag{3.23}$$

where all three processes are independent. Note that here and in the sequel equalities between random variables are understood to hold in distribution. Moreover, we demand that

$$\operatorname{cov}(\hat X_{\hat\sigma}, \hat X_{\hat\sigma'}) = \xi(\tfrac{M}{N+M} R_M(\hat\sigma, \hat\sigma')) \tag{3.24}$$

and

$$\operatorname{cov}(\check X_{\check\sigma}, \check X_{\check\sigma'}) = \xi(\tfrac{N}{N+M} R_N(\check\sigma, \check\sigma')) \tag{3.25}$$

Obviously, this implies that

$$\begin{aligned} \operatorname{cov}(Z_\sigma, Z_{\sigma'}) = \xi\left(\tfrac{M}{N+M} R_M(\hat\sigma, \hat\sigma') + \tfrac{N}{N+M} R_N(\check\sigma, \check\sigma')\right) \\ - \xi\left(\tfrac{M}{N+M} R_M(\hat\sigma, \hat\sigma')\right) - \xi\left(\tfrac{N}{N+M} R_N(\check\sigma, \check\sigma')\right) \end{aligned} \tag{3.26}$$

(we will not worry about the existence of such a decomposition; if $\xi(x) = x^p$, we can use the explicit representation in terms of p-spin interactions to construct them). Now we first note that, by super-additivity [3]

$$\lim_{M\uparrow\infty} \frac{1}{\beta M} \liminf_{N\uparrow\infty} \mathbb{E} \log \frac{Z_{\beta,N+M}}{Z_{\beta,N}} = -f_\beta \tag{3.27}$$

Thus we need a suitable representation for $\frac{Z_{\beta,N+M}}{Z_{\beta,N}}$. But

$$\frac{Z_{\beta,N+M}}{Z_{\beta,N}} = \frac{\mathbb{E}_\sigma e^{\beta\sqrt{N+M}\left(\check X_{\check\sigma} + Z_\sigma + \hat X_{\hat\sigma}\right)}}{\mathbb{E}_{\check\sigma} e^{\beta\sqrt{N+M}\left(\sqrt{(1-M/(N+M))} X_{\check\sigma}\right)}} \tag{3.28}$$

Now we want to express the random variables in the denominator in the form

$$\sqrt{(1 - M/(N+M))} X_{\check\sigma} = \check X_{\check\sigma} + Y_{\check\sigma} \tag{3.29}$$

where Y is independent of $\check{X}$. Comparing covariances, this implies that

$$\operatorname{cov}(Y_{\check\sigma}, Y_{\check\sigma'}) = (1 - M/(N+M))\xi(R_N(\check\sigma, \check\sigma')) - \xi\left(\tfrac{N}{N+M} R_N(\check\sigma, \check\sigma')\right) \tag{3.30}$$

As we will be interested in taking the limit $N \uparrow \infty$ before $M \uparrow \infty$, we may expand in $M/(N+M)$ to see that to leading order in $M/(N+M)$,

$$\operatorname{cov}(Y_{\check\sigma}, Y_{\check\sigma'}) \sim \tfrac{M}{N+M} R_N(\check\sigma, \check\sigma')\xi'\left(\tfrac{N}{N+M} R_N(\check\sigma, \check\sigma')\right) - \tfrac{M}{N+M}\xi\left(\tfrac{N}{N+M} R_N(\check\sigma, \check\sigma')\right) \tag{3.31}$$

Finally, we note that the random variables $\hat{X}_{\hat\sigma}$ are negligible in the limit $N \uparrow \infty$, since their variance is smaller than $\xi(M/(N+M))$ and hence their maximum is bounded by $\sqrt{\xi(M/(N+M))M \ln 2}$, which even after multiplication with $\sqrt{N+M}$ gives no contribution in the limit if ξ tends to zero faster than linearly at zero, which we can safely assume. Thus we see that we can indeed express the free energy as

$$f_\beta = -\lim_{M\uparrow\infty} \liminf_{N\uparrow\infty} \frac{1}{\beta M} \mathbb{E} \ln \frac{\mathbb{E}_{\hat\sigma} \widetilde{\mathbb{E}}_{\check\sigma} e^{\beta\sqrt{N+M} Z_{\hat\sigma,\check\sigma}}}{\widetilde{\mathbb{E}}_{\check\sigma} e^{\beta\sqrt{N+M} Y_{\check\sigma}}} \tag{3.32}$$

where the measure $\widetilde{\mathbb{E}}_{\check\sigma}$ can be chosen as a probability measure defined by $\widetilde{\mathbb{E}}_{\check\sigma}(\cdot) = \mathbb{E}_{\check\sigma} e^{\beta\sqrt{N+M}\check{X}_{\check\sigma}}(\cdot)/\check{Z}_{\beta,N,M}$ where $\check{Z}_{\beta,N,M} \equiv \mathbb{E}_{\check\sigma} e^{\beta\sqrt{N+M}\check{X}_{\check\sigma}}$. Of course this representation is quite pointless, because it is certainly uncomputable, since $\widetilde{\mathbb{E}}$ is effectively the limiting Gibbs measure that we are looking for. However, at this point there occurs a certain miracle: the (asymptotic) covariances of the processes X, Y, Z satisfy

$$\xi(x) + y\xi'(y) - \xi(y) \geq x\xi'(y) \tag{3.33}$$

for all $x, y \in [-1, 1]$, if ξ is convex and even. This comes as a surprise, since we did not do anything to impose such a relation! But it has the remarkable consequence that asymptotically, by virtue of Lemma 3.1.2 it implies the bound

$$\mathbb{E} \ln \mathbb{E}_{\hat\sigma} e^{\beta\sqrt{M} X_{\hat\sigma}} \leq \mathbb{E} \ln \frac{\mathbb{E}_{\hat\sigma} \widetilde{\mathbb{E}}_{\check\sigma} e^{\beta\sqrt{N+M} Z_{\hat\sigma,\check\sigma}}}{\widetilde{\mathbb{E}}_{\check\sigma} e^{\beta\sqrt{N+M} Y_{\check\sigma}}} \tag{3.34}$$

(if the processes are taken to have the asymptotic form of the covariances). Moreover, this bound will hold *even* if we replace the measure $\widetilde{\mathbb{E}}$ by some other probability measure, and even if we replace the overlap R_N on the space $\mathcal{S}_N$ by some other function, e.g. the ultrametric d_N. Seen the other way around, we can conclude that a lower bound of the form (3.22) can actually be made as good as we want, provided we choose the right measure $\widetilde{\mathbb{E}}$. This observation is due to Aizenman, Sims, and Starr [3]. They call the auxiliary structure made from a space $\mathcal{T}$, a probability measure $\mathbb{E}_\alpha$ on $\mathcal{T}$, a normalized distance q on $\mathcal{T}$, and the corresponding processes, Y and Z, a *random overlap structure*

$$\mathrm{cov}(Y_\alpha, Y_{\alpha'}) = q(\alpha,\alpha')\xi'(q(\alpha,\alpha')) - \xi(q(\alpha,\alpha')) \tag{3.35}$$

and the process $Z_{\sigma,\alpha}$ on $\mathcal{S}_N \times [0,1]$ with covariance

$$\mathrm{cov}(Z_{\sigma,\alpha}, Z_{\sigma',\alpha'}) \equiv R_N(\sigma,\sigma')\xi'(q(\alpha,\alpha')) \tag{3.36}$$

With these choices, and naturally X_σ our original process with covariance $\xi(R_N)$, the equation (3.21) is satisfied, and hence the inequality (3.22) holds, no matter what choice of q and $\mathbb{E}_\alpha$ we make. Restricting these choices to the random genealogies obtained from Neveu's process by a time change with some probability distribution function m, and $\mathbb{E}_\alpha$ the Lebesgue measure on $[0,1]$, gives the bound we want.

This bound would be quite useless if we could not compute the right-hand side. Fortunately, one can get rather explicit expressions. We need to compute two objects:

$$\mathbb{E}_\alpha \mathbb{E}_\sigma e^{\beta\sqrt{N} Z_{\sigma,\alpha}} \tag{3.37}$$

and

$$\mathbb{E}_\alpha e^{\beta\sqrt{N} Y_\alpha} \tag{3.38}$$

In the former we use that Z has the representation

$$Z_{\sigma,\alpha} = N^{-1/2} \sum_{i=1}^N \sigma_i z_{\alpha,i} \tag{3.39}$$

where the processes $z_{\alpha,i}$ are independent for different i and have covariance

$$\mathrm{cov}(z_{\alpha,i}, z_{\alpha',i}) = \xi'(q(\alpha,\alpha')) \tag{3.40}$$

Thus at least the σ- average is trivial:

$$\mathbb{E}_\alpha \mathbb{E}_\sigma e^{\beta\sqrt{N} Z_{\sigma,\alpha}} = \mathbb{E}_\alpha \prod_{i=1}^N e^{\ln\cosh(\beta z_{\alpha,i})} \tag{3.41}$$

Thus we see that, in any case, we obtain bounds that only involve objects that we introduced ourselves and that thus can be manipulated to be computable. In fact, such computations have been done in the context of the Parisi solution [17]. A useful mathematical reference is [4].

This is the form derived in Aizenman, Sims, and Starr [3].

3.5 Parisi auxiliary systems

After this digression, that only served to show that *in principle* we can find a perfect auxiliary system, we now turn to the search for *useful* auxiliary systems. A sufficiently rich class was proposed (in disguise) by Parisi. There are several ways to describe this, and we will discuss at least two ways here, and, time permitting, give a motivation in the last part of the talk.

Ruelle's probability cascades. In this paragraph we describe the auxiliary random structure introduced by Parisi in terms of Ruelle's probability cascades without very much motivation. They can be shown to arise naturally in the analysis of the GREMs that we introduced in the introduction. I will explain this relation briefly in Chapter 4.

The space $\mathcal{T}$ in this case is chosen as a n-level infinite tree indexed by multi-indices $\underline{i} = (i_1, \dots, i_n,\ i_k \in \mathbb{N}$. To each $\underline{i}$, we associate an energy

$$x_{\underline{i}} = \sum_{\ell=1}^{k} \gamma_\ell x^\ell_{i_1,\dots,i_\ell}$$

where $x^\ell_{i_1,\dots,i_\ell}$ is the i_ℓ-th atom of the Poisson point process $\mathcal{P}^\ell_{x_{i_1},\dots,x_{i_{\ell-1}}}$, and all Poisson point processes $\mathcal{P}^\ell_{x_{i_1},\dots,x_{i_{\ell-1}}}$ are independent for different sub- and superscripts. The collection of all these Poisson point processes is called a Poisson cascade. The numbers γ_ℓ must form a decreasing sequence, and $\gamma_n > 1$.

Given such a Poisson cascade, we define the probability measure ξ on $\mathcal{T}$ through its atoms

$$w(\underline{i}) \equiv \frac{\exp\left(x_{\underline{i}}\right)}{\sum_{\underline{j}\in\mathcal{T}}\left(\alpha x_{\underline{j}}\right)} \tag{3.42}$$

Note that the Poisson processes $\sum_i \delta\{e^{\gamma_\ell x_{i_1},\dots,i_{\ell-1},i_\ell}\}$ are iid Poisson point processes with intensity measure $\gamma_\ell^{-1} z^{-1-1/\gamma_\ell} dz$, and denoting the corresponding atoms by $\xi^\ell_{i_1,\dots,i_\ell}$, we can write w also as

$$w(\underline{i}) = \frac{\prod_{\ell=1}^n \xi^\ell_{i_1,\dots,i_\ell}}{\sum_{\underline{j}\in\mathcal{T}} \prod_{\ell=1}^n \xi^\ell_{j_1,\dots,j_\ell}}$$

The random measure ξ on $\mathcal{T}$ will play the rôle of $\mathbb{E}_\alpha$.

$\mathcal{T}$ is naturally endowed with its tree overlap, $D_n(\underline{i},\underline{j}) \equiv n^{-1}(\min\{\ell : i_\ell \neq j_\ell\} - 1)$. This distance will play the rôle of the distance q on $\mathcal{T}$. Finally, we define the processes $Y_{\underline{i}}$ and $Z_{\underline{i},\sigma}$ with covariances

$$\mathrm{cov}(Y_{\underline{i}}, Y_{\underline{j}}) = D_n(\underline{i}, \underline{j})\xi'(D_n(\underline{i}, \underline{j})) - \xi(d_n(\underline{i}, \underline{j})) \equiv h(D_n(\underline{i}, \underline{j})) \tag{3.43}$$

and the process $Z_{\sigma,\underline{i}}$ on $\mathcal{S}_N \times \mathcal{T}$ with covariance

$$\mathrm{cov}(Z_{\sigma,\underline{j}}, Z_{\sigma',\underline{j}}) \equiv R_N(\sigma, \sigma')\xi'(D_n(\underline{i}, \underline{j})) \tag{3.44}$$

It is easy to see that such processes can be constructed as long as h, ξ' are increasing functions. E.g.

$$Y_{\underline{i}} = \sum_{\ell=1}^{n} \sqrt{h(\ell/n) - h((\ell-1)/n)}\, Y^{(\ell)}_{\underline{i}_1 \cdots \underline{i}_\ell} \tag{3.45}$$

where $Y^{(\ell)}_{\underline{i}_1 \cdots \underline{i}_\ell}$ are independent standard normal random variables. In this way, we have constructed an explicit random overlap structure, which corresponds indead to the one generating the Parisi solution.

3.6 Computing with Poisson cascades

At first glance it still seems impossible to do any computation with this complicated structure. Mirculously, this is not true. The key point it contained in the following lemma, which reflects on a very amazing invariance property of Poisson point processes.

Lemma 3.6.1 *Assume that $\mathcal{P}$ be a Poisson process with intensity measure $e^{-x}dx$. and let $Y_{i,j}$, $i \in \mathbb{N}$, $j \in \mathbb{N}$, be iid standard normal random variables. Let Y be a random variable that has the same distribution as $Y_{\sigma,1}$. Let $g_i : \mathbb{R} \to \mathbb{R}$ be smooth functions, such that, for all $|m| \leq 2$, there exist $C < \infty$, independent of N, such that*

$$\mathbb{E}_Y e^{m g_i(Y)} \equiv e^{L_i(m)} < C \tag{3.46}$$

Let x_i be the atoms of the Poisson process $\mathcal{P}$.

$$\mathbb{E} \ln \frac{\sum_{i=1}^{\infty} e^{\alpha x_i + \sum_{j=1}^{M} g_i(Y_{i,j})}}{\sum_{i=1}^{\infty} e^{\alpha x_i}} = \sum_{i=1}^{M} \frac{L_i(m)}{m} \tag{3.47}$$

where $m = 1/\alpha$.

Proof Let for simplicity $M = 1$. The numerator on the left in (3.47) can be written as

$$\int e^{\alpha z} \widetilde{\mathcal{P}}(dz)$$

where $\widetilde{\mathcal{P}}$ is the point process

$$\widetilde{\mathcal{P}} \equiv \sum_j \delta_{z_j + \alpha^{-1} g(Y_j)}$$

This follows from a general fact about Poisson point processes: if $\mathcal{N} \equiv \sum_i \delta_{x_i}$ is a Poisson point process with intensity measure λ on E, and Y_i are iid random variables with distribution ρ, then

$$\widetilde{\mathcal{N}} \sum_i \delta_{x_i + Y_i}$$

is a Poisson process with intensity measure $\lambda * \rho$ on the set $E + \text{supp}\rho$. This follows from the representation of $\mathcal{N}$ as

$$\mathcal{N} = \sum_{i=1}^{N_\lambda} \delta_{X_i}$$

where N_λ is Poisson with parameter $\int_E \lambda(dx) \equiv |\lambda|$ (if this is finite), and X_i iid random variables with distribution $\lambda/|\lambda|$. Clearly

$$\widetilde{\mathcal{N}} = \sum_i \delta_{x_i + Y_i} = \sum_{i=1}^{N_\lambda} \delta_{X_i + Y_i}$$

is again of the form of a PPP, and the distriburion of $X_i + Y_i$ is $\lambda * \rho / |\lambda|$. Since the total intensity of the process is the parameter of N_λ, $\lambda|$, it follows that the intensity measure of this process is the one we claimed.

Thus, in our case, $\widetilde{\mathcal{P}}$ is a PPP with intensity measure the convolution of the measure $e^{-z}dz$ and the distribution of the random variable $\alpha^{-1}g(Y)$. A simple computation shows that this is $\mathbb{E}_Y e^{g(Y)/\alpha} e^{-z} dz$, i.e. a multiple of the original intensity measure! Thus the Poisson point process $\sum_j \delta_{e^{\alpha z_j + g(Y_j)}}$ has intensity measure $\mathbb{E}_Y e^{g(Y)/\alpha} \alpha^{-1} x^{-1/\alpha - 1} dx$. Finally, one makes the elementary but surprising and remarkable observation that the Poisson point process $\sum_j \delta_{e^{\alpha z} [\mathbb{E}_Y e^{g(Y)/\alpha}]^\alpha}$ has the same intensity measure, and therefore, $\sum_j e^{\alpha z_j + g(Y_j)}$ has the same law as $\sum_j e^{\alpha z_j} [\mathbb{E}_Y e^{g(Y)/\alpha}]^\alpha$: multiplying each atom with an iid random variable leads to the same process as multiplying each atom by a suitable constant! The assertion of the Lemma follows immediately.

□

Remark 3.6.1 Noody seems to know who made this discovery. Michael Aizenman told me about it and attributed to David Ruelle, but one cannot find it in his paper. A slightly different proof from the one above can be found in [21].

Let us look first at (3.38). We can then write

$$\sum_{\underline{i}} e^{x_{\underline{i}}+\beta\sqrt{M}Y_{\underline{i}}} = \sum_{\underline{i}} e^{\beta\sqrt{N}X_{\tilde{\sigma}}+\beta\sqrt{M}Y_{\tilde{\sigma}_1\dots\tilde{\sigma}_{n-1}}+\sqrt{h(x_n)-h(x_{n-1})}Y^{(n)}_{\tilde{\sigma}_1\dots\tilde{\sigma}_n}}$$

$$= \sum_{\underline{i}_1\cdots\underline{i}_{n-1}} e^{\sum_{\ell=1}^{n-1}\gamma_\ell x_{\underline{i}_1\dots\tilde{\sigma}_{\ell-1}}+\beta\sqrt{M}Y_{\underline{i}_1\dots\underline{i}_{n-1}}}$$

$$\times\sum_{\underline{i}_n} e^{g_n x_{\underline{i}_1\cdots\underline{i}_n}+\beta\sqrt{M}\sqrt{h(1)-h(1-1/n)}Y^{(n)}_{\underline{i}_1\cdots\underline{i}_n}} \tag{3.48}$$

Using Lemma 3.6.1, the last factor can be replaced by

$$\mathbb{E}_{\underline{i}_n} e^{\gamma_n x_{\underline{i}_1\cdots\underline{i}_n}+\beta\sqrt{M}\sqrt{h(1)-h(1-1/n)}Y^{(n)}_{\underline{i}_1\cdots\underline{i}_n}} \tag{3.49}$$

$$\to \left[\int \frac{dz}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} e^{zm_n\beta\sqrt{M}\sqrt{h(1)-h(1-1/n)}}\right]^{1/m_n} \sum_{i_n} e^{\gamma_n x_{\underline{i}}} \tag{3.50}$$

$$= e^{\frac{\beta^2 M}{2}m_n(h(1)-h(1-1/n))}\sum_{i_n} e^{\gamma_n x_{\underline{i}}} \tag{3.51}$$

(we use througout $m_n = 1/\gamma_n$). Note that the last factor is independent of the random variables $x_{i_1,\dots,i_\ell}$ with $\ell < n$. Thus

$$\mathbb{E}\ln\sum_{\underline{i}} e^{\alpha x_{\underline{i}}+\beta\sqrt{M}Y_{\underline{i}}} = \mathbb{E}\ln\sum_{i_1,\dots,i_{n-1}} e^{\sum_{\ell=1}^{n-1}\gamma_\ell x_{i_1,\dots,i_{n-1}}+\beta\sqrt{M}Y_{i_1,\dots,i_{n-1}}}$$

$$+\frac{\beta^2 M}{2}m_n(h(1)-h(1-1/n)) + \mathbb{E}\ln\sum_{i_n} e^{\gamma_n x_{\underline{i}}} \tag{3.52}$$

The first term now has the same form as the original one with n replaced by $n-1$, and thus the procedure can obviously iterated. As the final result, we get that a consequence, we get that

$$M^{-1}\mathbb{E}\ln\frac{\sum_{\underline{i}} e^{x_{\underline{i}}+\beta\sqrt{M}Y_{\underline{i}}}}{\sum_{\underline{i}} e^{x_{\underline{i}}}}$$

$$= \sum_{\ell=1}^{n}\frac{\beta^2}{2}m_\ell(h(1-\ell/n)-h(1-(\ell-1)/n))$$

$$= \frac{\beta^2}{2}\int_0^1 m(x)x\xi''(x)dx \tag{3.53}$$

The computation of the expression (3.37) is now very similar, but gives a more complicated result since the analogs of the expressions (3.49) cannot be computed explicitly. Thus, after the k-th step, we end up with a new function of the remaining random variables $Y_{i_1\dots i_{n-k}}$. The result can be expressed in the form

$$\frac{1}{M}\mathbb{E}\ln \xi_{\underline{i}}\mathbb{E}_\sigma e^{\beta\sqrt{M}Z_{\sigma,\underline{i}}} = \zeta(0,h,m,\beta) \tag{3.54}$$

(here h is the magnetic field (which we have so far hidden in the notation) that can be taken as a parameter of the a priori distribution on the σ such that $\mathbb{E}_{\sigma_i}(\cdot) \equiv \frac{1}{2\cosh(\beta h)}\sum_{\sigma_i=\pm 1} e^{\beta h\sigma_i}(\cdot))$ where $\zeta(1,h) = \ln\cosh(\beta h)$, and

$$\zeta(x_{a-1},h) = \frac{1}{m_a}\ln\int\frac{dz}{\sqrt{2\pi}}e^{-z^2/2}e^{m_a\zeta(x_a,h+z\sqrt{\xi'(x_a)-\xi'(x)})} \tag{3.55}$$

(we put $x_a = a/n$).

In all of the preceeding discussion, the choice of the parameter n and of the numbers $m_i = 1/\gamma_1$ is still free. From

We can now announce Guerra's bound in the following form:

Theorem 3.6.2 *[12] Let $\zeta(t,h,m,b)$ be the function defined in terms of the recursion (3.55). Then*

$$\lim_{N\uparrow\infty} N^{-1}\mathbb{E}\ln Z_{\beta,h,N} \leq \inf_m \zeta(0,h,m,\beta) - \frac{\beta^2}{2}\int_0^1 m(x)x\xi''(x)dx \tag{3.56}$$

where the infimum is over all probability distribution functions m on the unit interval.

Remark 3.6.2 It is also interesting to see that the recursive form of the function ζ above can also be represented in a closed form as the solution of a partial differential equation. Consider the case $\xi(x) = x^2/2$. Then ζ is the solution of the differential equation

$$\frac{\partial}{\partial t}\zeta(t,h) + \frac{1}{2}\left(\frac{\partial^2}{\partial h^2}\zeta(t,h) + m(t)\left(\frac{\partial}{\partial h}\zeta(t,h)\right)^2\right) = 0 \tag{3.57}$$

with final condition

$$\zeta(1,h) = \ln\cosh(\beta h) \tag{3.58}$$

If m is a step function, it is easy to see that a solution is obtained by setting, for $x \in [x_{a-1}, x_a)$,

$$\zeta(x,h) = \frac{1}{m_a}\ln\mathbb{E}_z e^{m_a\zeta(x_a,h+z\sqrt{x_a-x})} \tag{3.59}$$

For general convex ξ, analogous expressions can be obtained through changes of variables [12].

3.6.1 Talagrand's theorem

In both approaches, it pays to write down the expression of the difference between the free energy and the lower bound, since this takes a very suggestive form.

To do this, we just have to use formula (3.8) with

$$X^t_{\sigma,\alpha} \equiv \sqrt{t}(X_\sigma + Y_\alpha) + \sqrt{1-t} Z_{\sigma,\alpha} \tag{3.60}$$

and $f(X^t)$ replaced by $H(X^t) = \ln \mathbb{E}_\sigma \mathbb{E}_\alpha e^{\beta\sqrt{N} Z^t_{\sigma,\alpha}}$. This gives the equality

$$\begin{aligned} H(X+Y) - H(Z) = \frac{1}{2}\mathbb{E}\int_0^1 dt \widetilde{\mu}^{\otimes 2}_{\beta,t,N}(d\sigma, d\alpha)\Big(&\xi(R_N(\sigma,\sigma')) \\ &+ q(\alpha,\alpha')\xi'(q(\alpha,\alpha')) \\ &- \xi(q(\alpha,\alpha')) - R_N(\sigma,\sigma')\xi'(q(\alpha,\alpha'))\Big) \end{aligned} \tag{3.61}$$

where the measure $\widetilde{\mu}_{\beta,t,N}$ is defined as

$$\widetilde{\mu}_{\beta,t,N}(\cdot) \equiv \frac{\mathbb{E}_\sigma \mathbb{E}_\alpha e^{\beta\sqrt{N} X^t_{\sigma,\alpha}}(\cdot)}{\mathbb{E}_\sigma \mathbb{E}_\alpha e^{\beta\sqrt{N} X^t_{\sigma,\alpha}}} \tag{3.62}$$

where we interpret the measure $\widetilde{\mu}_{\beta,t,N}$ as a joint distribution on $\mathcal{S}_N \times [0,1]$. Note that for convex and even ξ, the function $\xi(R_N(\sigma,\sigma')) + q(\alpha,\alpha')\xi'(q(\alpha,\alpha')) - \xi(q(\alpha,\alpha'))$ vanishes if and only if $R_N(\sigma,\sigma') = q(\alpha,\alpha')$. Thus for the left hand side of (3.61) to vanish, the replicated interpolating product measure should (for almost all t), concentrate on configurations where the overlaps in the σ-variables coincide with the genealogical distances of the α-variables. Thus we see that the inequality in Theorem 3.6.2 will turn into an equality if it is possible to choose the parameters of the reservoir system in such a way that the the overlap distribution on $\mathcal{S}_N$ aligns with the genealogical distance distribution in the reservoir once the systems are coupled by the interpolation.

This latter fact was proven very recently, and not long after the discovery of Guerra's bound, by M. Talagrand [26].

Theorem 3.6.3 *[26] Let $\zeta(t,h,m,b)$ be the function defined in terms of (3.57) and (3.58). Then*

$$\lim_{N\uparrow\infty} N^{-1}\mathbb{E}\ln Z_{\beta,h,N} = \inf_m \left(\zeta(0,h,m,.\beta) - \frac{\beta^2}{2}\int_0^1 m(x)x\xi''(x)dx\right) \tag{3.63}$$

where the infimum is over all probability distribution functions m on the unit interval.

I will not give the complex proof which the interested reader should study in the original paper [26], but I will make some comments on the key ideas. First, Talagrand proves more than the assertion 3.63. What he actually proves is the following. For any $\epsilon > 0$, there exists a positive integer $n(\epsilon) < \infty$, and a probability distribution function m_n that is a step function with n steps, such that for all $t > \epsilon$,

$$\lim_{N\uparrow\infty} \mathbb{E}\widetilde{\mu}^{\otimes 2}_{\beta,t,N}(d\sigma, d\alpha)\Big(\xi(R_N(\sigma,\sigma')) + q(\alpha,\alpha')\xi'(q(\alpha,\alpha')) \\ -\xi(q(\alpha,\alpha')) - R_N(\sigma,\sigma')\xi'(q(\alpha,\alpha'))\Big) = 0 \qquad (3.64)$$

if the measure $\widetilde{\mu}_{b,t,N}$ corresponds to the genealogical distance obtained from this function m. That is to say, if the coupling parameter t is large enough, the SK model can be aligned to a GREM with any desired number of hierarchies.

Second, the proof naturally proceeds by showing that the measure $\mathbb{E}\widetilde{\mu}^{\otimes 2}_{\beta,t,N}$ seen as a distribution of the "overlaps" concentrates on the set where the R_N and q's are the same. Such a fact is usually proven by looking at a suitable Laplace transform, whose calculation amounts again to the estimate of a free energy, this time in a replicated, coupled system. Since the main effort goes into an upper bound, Guerra's techniques can again be used to provide help, even though the details of the computations now get very involved.

Bibliography

[1] R.J. Adler. On excursion sets, tube formulas and maxima of random fields. *Ann. Appl. Probab.*, 10(1):1–74, 2000.

[2] M. Aizenman, J.L. Lebowitz, and D. Ruelle. Some rigorous results on the Sherrington-Kirkpatrick spin glass model. *Comm. Math. Phys.*, 112(1):3–20, 1987.

[3] M. Aizenman, R. Sims, and S.L. Starr. An extended variational principle for the SK spin-glass model. *Phys. Rev. B*, 6821(21):4403–4403, 2003.

[4] E. Bolthausen and A.-S. Sznitman. On Ruelle's probability cascades and an abstract cavity method. *Comm. Math. Phys.*, 197(2):247–276, 1998.

[5] A. Bovier, I. Kurkova, and M. Löwe. Fluctuations of the free energy in the REM and the p-spin SK models. *Ann. Probab.*, 30(2):605–651, 2002.

[6] P. Carmona and Y. Hu. Universality in Sherrington-Kirkpatrick's Spin Glass Model. preprint arXiv:math.PR/0403359 v2, May 2004.

[7] F. Comets and J. Neveu. The Sherrington-Kirkpatrick model of spin glasses and stochastic calculus: the high temperature case. *Comm. Math. Phys.*, 166(3):549–564, 1995.

[8] P. Contucci, M. Degli Esposti, C. Giardinà, and S. Graffi. Thermodynamical limit for correlated Gaussian random energy models. *Comm. Math. Phys.*, 236(1):55–63, 2003.

[9] B. Derrida. Random-energy model: limit of a family of disordered models. *Phys. Rev. Lett.*, 45(2):79–82, 1980.

[10] B. Derrida. Random-energy model: an exactly solvable model of disordered systems. *Phys. Rev. B (3)*, 24(5):2613–2626, 1981.

[11] V. Dotsenko. *Introduction to the replica theory of disordered statistical systems*. Collection Aléa-Saclay: Monographs and Texts in Statistical Physics. Cambridge University Press, Cambridge, 2001.

[12] F. Guerra. Broken replica symmetry bounds in the mean field spin glass model. *Comm. Math. Phys.*, 233(1):1–12, 2003.

[13] F. Guerra and F. L. Toninelli. The thermodynamic limit in mean field spin glass models. *Comm. Math. Phys.*, 230(1):71–79, 2002.

[14] M.F. Kratz and P. Picco. A representation of Gibbs measure for the random energy model. *Ann. Appl. Probab.*, 14(2):651–677, 2004.

[15] M.R. Leadbetter, G. Lindgren, and H. Rootzén. *Extremes and related properties of random sequences and processes*. Springer Series in Statistics. Springer-Verlag, New York, 1983.

[16] M. Ledoux and M. Talagrand. *Probability in Banach spaces*, volume 23 of *Ergebnisse der Mathematik und ihrer Grenzgebiete (3) [Results in Mathematics and Related Areas (3)]*. Springer-Verlag, Berlin, 1991.

[17] M. Mézard, G. Parisi, and M.A. Virasoro. *Spin glass theory and beyond*, volume 9 of *World Scientific Lecture Notes in Physics*. World Scientific Publishing Co. Inc., Teaneck, NJ, 1987.

[18] J. Neveu. A continuous state branching process in relation with the GREM model of spin glass theory. rapport interne 267, Ecole Polytechnique Paris, 1992.

[19] H. Nishimori. *Statistical physics of spin glasses and information processing*. International Series of Monographs on Physics 111. Oxford University Press, Oxford, 2001.

[20] D. Ruelle. A mathematical reformulation of Derrida's REM and GREM. *Comm. Math. Phys.*, 108(2):225–239, 1987.

[21] A. Ruzmaikina and M. Aizenman. Characterization of invariant measures at the leading edge for competing particle systems. *Ann. Probab.*, 33(1):82–113, 2005.

[22] D. Sherrington and S. Kirkpatrick. Solvable model of a spin glas. *Phys. Rev. Letts.*, 35:1792–1796, 1972.

[23] M. Talagrand. Concentration of measure and isoperimetric inequalities in product spaces. *Inst. Hautes Études Sci. Publ. Math.*, 81:73–205, 1995.

[24] M. Talagrand. Rigorous results for the Hopfield model with many patterns. *Probab. Theory Related Fields*, 110(2):177–276, 1998.

[25] M. Talagrand. Rigorous low-temperature results for the mean field p-spins interaction model. *Probab. Theory Related Fields*, 117(3):303–360, 2000.

[26] M. Talagrand. The Parisi solution. *Ann. Math.*, to appear, 2005.

Index