6. The Diluted SK Model and the K-Sat Problem

6.1 Introduction

In the SK model, each individual (or spin) interacts with every other individual. For large N, this does not make physical sense. Rather, we would like that, as $N \to \infty$, a given individual typically interacts only with a bounded number of other individuals. This motivates the introduction of the *diluted SK model*. In this model, the Hamiltonian is given by

$$-H_N(\boldsymbol{\sigma}) = \beta \sum_{i<j} g_{ij}\gamma_{ij}\sigma_i\sigma_j \ . \tag{6.1}$$

As usual, $(g_{ij})_{i<j}$ are i.i.d. standard Gaussian r.v.s. The quantities $\gamma_{ij} \in \{0,1\}$ determine which of the interaction terms are actually present in the Hamiltonian. There is an interaction term between σ_i and σ_j only when $\gamma_{ij} = 1$. The natural choice for these quantities is to consider a parameter $\gamma > 0$ (that does not depend on N) indicating "how diluted is the interaction", and to decide that the quantities γ_{ij} are i.i.d. r.v.s with $\mathsf{P}(\gamma_{ij} = 1) = \gamma/N$, $\mathsf{P}(\gamma_{ij} = 0) = 1 - \gamma/N$, and are independent from the r.v.s g_{ij}. Thus, the expected number of terms in (6.1) is

$$\frac{\gamma}{N}\frac{N(N-1)}{2} = \frac{\gamma(N-1)}{2} \,,$$

and the expected number of terms that contain σ_i is about $\gamma/2$. That is, the average number of spins that interact with one given spin is about $\gamma/2$. One should observe that the usual normalizing factor $1/\sqrt{N}$ does not occur in (6.1).

If we draw an edge between i and j when $\gamma_{ij} = 1$, the resulting random graph is well understood [12]. When $\gamma < 1$, this graph has only small connected components, so there is no "global interaction" and the situation is not so interesting. In order to get a challenging model we must certainly allow the case where γ takes any positive value.

In an apparently unrelated direction, let us remind the reader that the motivation of Chapter 2 is the problem as to whether certain random subsets of $\{-1,1\}^N$ have a non-empty intersection. In Chapter 2, we considered "random half-spaces". These somehow "depend on all coordinates". What would

M. Talagrand, *Mean Field Models for Spin Glasses*, Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge / A Series of Modern Surveys in Mathematics 54, DOI 10.1007/978-3-642-15202-3_6, © Springer-Verlag Berlin Heidelberg 2011

happen if instead we considered sets depending only on a given number p of coordinates? For example sets of the type

$$\{\boldsymbol{\sigma}\,;\ (\sigma_{i_1},\ldots,\sigma_{i_p}) \neq (\eta_1,\ldots,\eta_p)\} \tag{6.2}$$

where $1 \le i_1 < i_2 < \ldots < i_p \le N$, and $\eta_1,\ldots,\eta_p = \pm 1$?

The question of knowing whether M random independent sets of the type (6.2) have a non-empty intersection is known in theoretical computer science as the random K-sat problem, and is of considerable interest. (There K is just another notation for what we call p. "Sat" stands for "satisfiability", as the problem is presented under the equivalent form of whether one can assign values to N Boolean variables in order to satisfy a collection of M random logical clauses of a certain type.) By a random subset of the type (6.2), we of course mean a subset that is chosen uniformly at random among all possible such subsets. This motivates the introduction of the Hamiltonian

$$-H_N(\boldsymbol{\sigma}) = -\beta \sum_{k \le M} W_k(\boldsymbol{\sigma}) \tag{6.3}$$

where $W_k(\boldsymbol{\sigma}) = 0$ if $(\sigma_{i(k,1)},\ldots,\sigma_{i(k,p)}) \neq (\eta_{k,1},\ldots,\eta_{k,p})$, and $W_k(\boldsymbol{\sigma}) = 1$ otherwise. The indices $1 \le i(k,1) < i(k,2) < \ldots < i(k,p) \le N$ and the numbers $\eta_{k,i} = \pm 1$ are chosen randomly uniformly over all possible choices. The interesting case is when M is proportional to N.

In a beautiful paper, S. Franz and S. Leone [60] observed that many technicalities disappear (and that one obtains a similar model) if rather than insisting that the Hamiltonian contains exactly a given number of terms, this number of terms is a Poisson r.v. M (independent of the other sources of randomness). Since we are interested in the case where M is proportional to N we will assume that $\mathsf{E}M$ is proportional to N, i.e. $\mathsf{E}M = \alpha N$, where of course α does not depend on N.

To cover simultaneously the cases of (6.1) and (6.3), we consider a random real-valued function θ on $\{-1,1\}^p$, i.i.d. copies $(\theta_k)_{k\ge 1}$ of θ, and the Hamiltonian

$$-H_N(\boldsymbol{\sigma}) = \sum_{k \le M} \theta_k(\sigma_{i(k,1)},\ldots,\sigma_{i(k,p)})\,. \tag{6.4}$$

Here, M is a Poisson r.v. of expectation αN, $1 \le i(k,1) < \ldots < i(k,p) \le N$, the sets $\{i(k,1),\ldots,i(k,p)\}$ for $k \ge 1$ are independent and uniformly distributed, and the three sources of randomness (these sets, M, and the θ_k) are independent of each other. There is no longer a coefficient β, since this coefficient can be thought of as a part of θ. For example, a situation very similar to (6.1) is obtained for $p = 2$ and $\theta(\sigma_1,\sigma_2) = \beta g \sigma_1 \sigma_2$ where g is standard Gaussian. It would require no extra work to allow an external field in the formula (6.4). We do not do this for simplicity, but we stress that our approach does not require any special symmetry property. (On the other hand, precise specific results such as those of [78] seem to rely on such properties.)

It turns out that the mean number of terms of the Hamiltonian that depend on a given spin is of particular relevance. This number is $\gamma = \alpha p$ (where α is such that $\mathsf{E} M = \alpha N$), and for simplicity of notation this will be our main parameter rather than α.

The purpose of this chapter is to describe the behavior of the system governed by the Hamiltonian (6.4) under a "high-temperature condition" asserting in some sense that this Hamiltonian is small enough. This condition will involve the r.v. S given by

$$S = \sup |\theta(\sigma_1, \ldots, \sigma_p)| , \tag{6.5}$$

where the supremum is of course over all values of $\sigma_1, \sigma_2, \ldots, \sigma_p = \pm 1$, and has the following property: if γ (and p) are given, then the high-temperature condition is satisfied when S is small enough.

Generally speaking, the determination of exactly under which conditions there is high-temperature behavior is a formidable problem. The best that our methods can possibly achieve is to reach qualitatively optimal conditions, that capture "a fixed proportion of the high-temperature region". This seems to be the case of the following condition:

$$16 p \gamma \mathsf{E}\, S \exp 4S \le 1 . \tag{6.6}$$

Since the mean number of spins interacting with a given spin remains bounded independently of N, the central limit theorem does not apply, and the ubiquitous Gaussian behavior of the previous chapters is now absent. Despite this fundamental difference, and even though this is hard to express explicitly, there are many striking similarities.

We now outline the organization of this chapter. A feature of our approach is that, in contrast with what happened for the previous models, we do not know how to gain control of the model "in one step". Rather, we will first prove in Section 6.2 that for large N a small collection of spins are approximately independent under a condition like (6.6). This is the main content of Theorem 6.2.2. The next main step takes place in Section 6.4, where in Theorem 6.4.1 we prove that under a condition like (6.6), a few quantities $\langle \sigma_1 \rangle, \ldots, \langle \sigma_k \rangle$ are approximately independent with law μ_γ where μ_γ is a probability measure on $[0, 1]$, that is described in Section 6.3 as the fixed point of a (complicated) operator. This result is then used in the last part of Section 6.4 to compute $\lim_{N \to \infty} p_N(\gamma)$, where $p_N(\gamma) = N^{-1} \mathsf{E} \log \sum \exp(-H_N(\boldsymbol{\sigma}))$, still under a "high-temperature" condition of the type (6.6). In Section 6.5 we prove under certain conditions an upper bound for $p_N(\gamma)$, that is true for *all values of* γ and that asymptotically coincides with the limit previously computed under a condition of the type (6.6). In Section 6.6 we investigate the case of continuous spins, and in Section 6.7 we demonstrate the very strong consequences of a suitable concavity hypothesis on the Hamiltonian, and we point out a number of rather interesting open problems.

6.2 Pure State

The purpose of this section is to show that under (6.6) "the system is in a pure state", that is, the spin correlations vanish. In fact we will prove that

$$\mathsf{E}\,|\langle\sigma_1\sigma_2\rangle - \langle\sigma_1\rangle\langle\sigma_2\rangle| \le \frac{K}{N} \tag{6.7}$$

where K depends only on p and γ. The proof, by induction over N, is similar in spirit to the argument occurring at the end of Section 1.3. In order to make the induction work, it is necessary to carry a suitable induction hypothesis, that will prove a stronger statement than (6.7). This stronger statement will be useful later in its own right.

Given $k \ge 1$ we say that two functions f, f' on Σ_N^n *depend on k coordinates* if we can find indices $1 \le i_1 < \ldots < i_k \le N$ and functions $\overline{f}, \overline{f}'$ from $\{-1,1\}^{kn}$ to $\mathbb{R}$ such that

$$f(\boldsymbol{\sigma}^1, \ldots, \boldsymbol{\sigma}^n) = \overline{f}(\sigma^1_{i_1}, \ldots, \sigma^1_{i_k}, \sigma^2_{i_1}, \ldots, \sigma^2_{i_k}, \ldots, \sigma^n_{i_1}, \ldots, \sigma^n_{i_k})$$

and similarly for f'. The reason we define this for two functions is to stress that both functions depend on the same set of k coordinates.

For $i \le N$, consider the transformation T_i of Σ_N^n that, for a point $(\boldsymbol{\sigma}^1, \ldots, \boldsymbol{\sigma}^n)$ of Σ_N^n, exchanges the i-th coordinates of $\boldsymbol{\sigma}^1$ and $\boldsymbol{\sigma}^2$, and leaves all the other coordinates unchanged.

The following technical condition should be interpreted as an "approximate independence condition".

Definition 6.2.1. *Given three numbers $\gamma_0 > 0$, $B > 0$ and $B^* > 0$, we say that Property $C(N, \gamma_0, B, B^*)$ holds if the following is true. Consider two functions f, f' on Σ_N^n, and assume that they depend on k coordinates. Assume that $f \ge 0$, that for a certain $i \le N$ we have*

$$f' \circ T_i = -f' \,, \tag{6.8}$$

and that for a certain number Q we have $|f'| \le Qf$ at each point of Σ_N^n. Then if $\gamma \le \gamma_0$ we have

$$\mathsf{E}\left|\frac{\langle f'\rangle}{\langle f\rangle}\right| \le \frac{(kB + B^*)Q}{N} \,. \tag{6.9}$$

Condition $C(N, \gamma_0, B, B^*)$ is not immediately intuitive. It is an "approximate independence condition" because if the spins were really independent, the condition $f' \circ T_i = -f'$ would imply that $\langle f'\rangle = \langle f' \circ T_i\rangle = \langle -f'\rangle$ so that $\langle f'\rangle = 0$.

To gain intuition, let us relate condition $C(N, \gamma_0, B, B^*)$ with (6.7). We take $n = 2$, $f = 1$,

$$f'(\boldsymbol{\sigma}^1, \boldsymbol{\sigma}^2) = \sigma^1_1(\sigma^1_2 - \sigma^2_2) \,,$$

so that (6.8) holds for $i = 2$, $k = 2$ and $|f'| \leq 2f$. Thus under condition $C(N, \gamma_0, B, B^*)$ we get by (6.9) that

$$\mathsf{E}\,|\langle \sigma_1^1(\sigma_2^1 - \sigma_2^2)\rangle| \leq \frac{2B + B^*}{N}$$

i.e.

$$\mathsf{E}\,|\langle \sigma_1 \sigma_2 \rangle - \langle \sigma_1 \rangle \langle \sigma_2 \rangle| \leq \frac{2B + B^*}{N}\,,$$

which is (6.7). More generally, basically the same argument shows that when condition $C(N, \gamma_0, B, B^*)$ holds (for each N and numbers B and B^* that do not depend on N), to compute Gibbs averages of functions that depend only on a number of spin that remains bounded independently of N, one can pretend that these spins are independent under Gibbs' measure. We will return to this important idea later.

Theorem 6.2.2. *There exists a number $K_0 = K_0(p, \gamma_0)$ such that if $\gamma \leq \gamma_0$ and*

$$16\gamma_0 p \mathsf{E}\, S \exp 4S \leq 1, \tag{6.10}$$

then Property $C(N, \gamma_0, K_0, K_0)$ holds for each N.

When property $C(N, \gamma_0, K_0, K_0)$ holds, for two functions f, f' on Σ_N^n, that depend on k coordinates, and with $f \geq 0$, $|f'| \leq Qf$, then under (6.8), and if $\gamma \leq \gamma_0$, we have

$$\mathsf{E}\left|\frac{\langle f' \rangle}{\langle f \rangle}\right| \leq \frac{(kK_0 + K_0)Q}{N} \leq \frac{2kK_0 Q}{N}\,. \tag{6.11}$$

The point of distinguishing in the definition of $C(N, \gamma_0, B, B^*)$ the values B and B^* will become apparent during the proofs.

To prove Theorem 6.2.2, we will proceed by induction over N. The smallest value of N for which the model is defined is $N = p$. We first observe that $|\langle f' \rangle| \leq Q\langle f \rangle$, so that $C(p, \gamma_0, K_1, K_1^*)$ is true if $K_1 \geq p$. We will show that if K_1 and K_1^* are suitably chosen, then under (6.10) we have

$$C(N-1, \gamma_0, K_1, K_1^*) \Rightarrow C(N, \gamma_0, K_1, K_1^*)\,. \tag{6.12}$$

This will prove Theorem 6.2.2.

The main idea to prove (6.12) is to relate the N-spin system with an $(N-1)$-spin system through the cavity method, and we first need to set up this method. We write $-H_N(\boldsymbol{\sigma}) = -H_{N-1}(\boldsymbol{\sigma}) - H(\boldsymbol{\sigma})$, where

$$-H_{N-1}(\boldsymbol{\sigma}) = \sum \theta_k(\sigma_{i(k,1)}, \ldots, \sigma_{i(k,p)})\,, \tag{6.13}$$

where the sum is over those $k \leq M$ for which $i(k,p) \leq N-1$, and where $H(\boldsymbol{\sigma})$ is the sum of the other terms of (6.4), those for which $i(k,p) = N$.

Since the set $\{i(k,1),\ldots,i(k,p)\}$ is uniformly distributed over the subsets of $\{1,\ldots,N\}$ of cardinality p, the probability that $i(k,p)=N$ is exactly p/N.

A remarkable property of Poisson r.v.s is as follows: when M is a Poisson r.v., if $(X_k)_{k\geq 1}$ are i.i.d. $\{0,1\}$-valued r.v.s then $\sum_{k\leq M} X_k$ and $\sum_{k\leq M}(1-X_k)$ are independent Poisson r.v.s with mean respectively $\mathsf{E}M\mathsf{E}X_k$ and $\mathsf{E}M\mathsf{E}(1-X_k)$. The simple proof is given in Lemma A.10.1. Using this for $X_k=1$ if $i(k,p)=N$ and $X_k=0$ otherwise implies that the numbers of terms in $H(\boldsymbol{\sigma})$ and $H_{N-1}(\boldsymbol{\sigma})$ are independent Poisson r.v.s of mean respectively $(p/N)\alpha N=\gamma$ and $\alpha N-\gamma$. Thus the pair $(-H_{N-1}(\boldsymbol{\sigma}),-H(\boldsymbol{\sigma}))$ has the same distribution as the pair

$$\left(\sum_{k\leq M'}\theta'_k(\sigma_{i'(k,1)},\ldots,\sigma_{i'(k,p)}),\sum_{j\leq r}\theta_j(\sigma_{i(j,1)},\ldots,\sigma_{i(j,p-1)},\sigma_N)\right). \tag{6.14}$$

Here M' and r are Poisson r.v.s of mean respectively $\alpha N-\gamma$ and γ; θ'_k and θ_j are independent copies of θ; $i'(k,1)<\ldots<i'(k,p)$ and the set $\{i'(k,1),\ldots,i'(k,p)\}$ is uniformly distributed over the subsets of $\{1,\ldots,N-1\}$ of cardinality p; $i(j,1)<\ldots<i(j,p-1)\leq N-1$ and the set $I_j=\{i(j,1),\ldots,i(j,p-1)\}$ is uniformly distributed over the subsets of $\{1,\ldots,N-1\}$ of cardinality $p-1$; all these random variables are globally independent.

The following exercise describes another way to think of the Hamiltonian H_N, which provides a different intuition for the fact that the pair $(-H_{N-1}(\boldsymbol{\sigma}),-H(\boldsymbol{\sigma}))$ has the same distribution as the pair (6.14).

Exercise 6.2.3. For each p-tuple $\mathbf{i}=(i_1,\ldots,i_p)$ with $1\leq i_1<\ldots<i_p\leq N$, and each $j\geq 1$ let us consider independent copies $\theta_{\mathbf{i},j}$ of θ, and define

$$-H_{\mathbf{i}}(\boldsymbol{\sigma})=\sum_{j\leq r_{\mathbf{i}}}\theta_{\mathbf{i},j}(\sigma_{i_1},\ldots,\sigma_{i_p})\,,$$

where $r_{\mathbf{i}}$ are i.i.d. Poisson r.v.s (independent of all other sources of randomness) with

$$\mathsf{E}r_{\mathbf{i}}=\frac{\alpha M}{\binom{M}{p}}\,.$$

Prove that the Hamiltonian H_N has the same distribution as the Hamiltonian $\sum_{\mathbf{i}} H_{\mathbf{i}}$.

Since the properties of the system governed by the Hamiltonian H_N depend only of the distribution of this Hamiltonian, from now on in this section we will assume that, using the same notation as in (6.14),

$$-H_N(\boldsymbol{\sigma})=-H_{N-1}(\boldsymbol{\sigma})-H(\boldsymbol{\sigma})\,, \tag{6.15}$$

where

$$-H_{N-1}(\boldsymbol{\sigma})=\sum_{k\leq M'}\theta'_k(\sigma_{i'(k,1)},\ldots,\sigma_{i'(k,p)})\,, \tag{6.16}$$

and

$$-H(\boldsymbol{\sigma}) = \sum_{j \le r} \theta_j(\sigma_{i(j,1)}, \ldots, \sigma_{i(j,p-1)}, \sigma_N) \,. \tag{6.17}$$

Let us stress that in this section and in the next, the letter r will stand for the number of terms in the summation (6.17), which is a Poisson r.v. of expectation γ.

We observe from (6.16) that if we write $\boldsymbol{\rho} = (\sigma_1, \ldots, \sigma_{N-1})$ when $\boldsymbol{\sigma} = (\sigma_1, \ldots, \sigma_N)$, $-H_{N-1}(\boldsymbol{\sigma}) = -H_{N-1}(\boldsymbol{\rho})$ is the Hamiltonian of a $(N-1)$-spin system, except that we have replaced γ by a different value γ_-. To compute γ_- we recall that the mean number of terms of the Hamiltonian H_{N-1} is $\alpha N - \gamma$, so that the mean number γ_- of terms that contain a given spin is

$$\gamma_- = \frac{p}{N-1}(\alpha N - \gamma) = \gamma \frac{N-p}{N-1} \,, \tag{6.18}$$

since $p\alpha = \gamma$. We note that $\gamma_- \le \gamma$, so that

$$\gamma < \gamma_0 \Rightarrow \gamma_- \le \gamma_0 \,, \tag{6.19}$$

a fact that will help the induction.

Given a function f on Σ_N^n, the algebraic identity

$$\langle f \rangle = \frac{\langle \mathrm{Av}\, f\mathcal{E} \rangle_-}{\langle \mathrm{Av}\, \mathcal{E} \rangle_-} \tag{6.20}$$

holds. Here,

$$\mathcal{E} = \mathcal{E}(\boldsymbol{\sigma}^1, \ldots, \boldsymbol{\sigma}^n) = \exp\left(\sum_{\ell \le n} -H(\boldsymbol{\sigma}^\ell) \right) \,, \tag{6.21}$$

and as usual Av means average over $\sigma_N^1, \ldots, \sigma_N^n = \pm 1$. Thus $\mathrm{Av}\, f\mathcal{E}$ is a function of $(\boldsymbol{\rho}^1, \ldots, \boldsymbol{\rho}^n)$ only, and $\langle \mathrm{Av}\, f\mathcal{E} \rangle_-$ means that it is then averaged for Gibbs' measure relative to the Hamiltonian (6.13).

In the right-hand side of (6.20), we have two distinct sources of randomness: the randomness in $\langle \cdot \rangle_-$ and the randomness in $\mathcal{E}$. It will be essential that these sources of randomness are probabilistically independent. In the previous chapters we were taking expectation given $\langle \cdot \rangle_-$. We find it more convenient to now take expectation given $\mathcal{E}$. This expectation is denoted by E_-, so that, according to (6.20) we have

$$\mathsf{E} \left| \frac{\langle f' \rangle}{\langle f \rangle} \right| = \mathsf{E} \left| \frac{\langle \mathrm{Av}\, f'\mathcal{E} \rangle_-}{\langle \mathrm{Av}\, f\mathcal{E} \rangle_-} \right| = \mathsf{E}\mathsf{E}_- \left| \frac{\langle \mathrm{Av}\, f'\mathcal{E} \rangle_-}{\langle \mathrm{Av}\, f\mathcal{E} \rangle_-} \right| \,. \tag{6.22}$$

After these preparations we describe the structure of the proof. Let us consider a pair (f', f) as in Definition 6.2.1. The plan is to write

$$\mathrm{Av}\, f'\mathcal{E} = \frac{1}{2} \sum_s f'_s$$

for some functions f'_s on Σ^n_{N-1}, such that the number of terms does not depend on N, and that all pairs $(f'_s, \mathrm{Av}\, f\mathcal{E})$ have the property of the pair (f', f), but in the $(N-1)$-spin system. Since

$$\frac{\mathrm{Av} f'\mathcal{E}}{\mathrm{Av}\, f\mathcal{E}} = \frac{1}{2}\sum_s \frac{f'_s}{\mathrm{Av}\, f\mathcal{E}} ,$$

we can now apply the induction hypothesis to each term to get a bound for the sum and hence for

$$\mathsf{E}_- \left| \frac{\langle \mathrm{Av}\, f'\mathcal{E}\rangle_-}{\langle \mathrm{Av}\, f\mathcal{E}\rangle_-} \right| ,$$

and finally (6.22) completes the induction step.

We now start the proof. We consider a pair (f', f) as in Definition 6.2.1, that is $|f'| \le Qf$, $f' \circ T_i = -f'$ for some $i \le N$, and f, f' depend on k coordinates. We want to bound $\mathsf{E}|\langle f'\rangle / \langle f\rangle|$, and for this we study the last term of (6.22). Without loss of generality, we assume that $i = N$ and that f and f' depend on the coordinates $1, \ldots, k-1, N$. First, we observe that, since we assume $|f'| \le Qf$, we have $|f'\mathcal{E}| \le Qf\mathcal{E}$, so that $|\mathrm{Av}\, f'\mathcal{E}| \le \mathrm{Av}\, |f'\mathcal{E}| \le Q\mathrm{Av}\, f\mathcal{E}$, and thus

$$\mathsf{E}_- \left| \frac{\langle \mathrm{Av}\, f'\mathcal{E}\rangle_-}{\langle \mathrm{Av}\, f\mathcal{E}\rangle_-} \right| \le Q . \tag{6.23}$$

We recall (6.21) and (6.17), and in particular that r is the number of terms in the summation (6.17) and is a Poisson r.v. of expectation γ. We want to apply the induction hypothesis to compute the left-hand side of (6.23). The expectation E_- is expectation given $\mathcal{E}$, and it helps to apply the induction hypothesis if the functions $\mathrm{Av} f'\mathcal{E}$ and $\mathrm{Av} f\mathcal{E}$ are not too complicated. To ensure this it will be desirable that all the points $i(j,\ell)$ for $j \le r$ and $\ell \le p-1$ are different and $\ge k$. In the rare event Ω (we recall that Ω denotes an event, and not the entire probability space) where this not the case, we will simply use the crude bound (6.23) rather than the induction hypothesis. Recalling that $i(j,1) < \ldots < i(j, p-1)$, to prove that Ω is a rare event we write $\Omega = \Omega_1 \cup \Omega_2$ where

$$\Omega_1 = \Big\{ \exists j \le r\,,\ i(j,1) \le k-1 \Big\}$$

$$\Omega_2 = \Big\{ \exists j, j' \le r\,,\ j \ne j'\,,\ \exists \ell, \ell' \le p-1\,,\ i(j,\ell) = i(j',\ell') \Big\} .$$

These two events depend only on the randomness of $\mathcal{E}$. Let us recall that for $j \le r$ the sets

$$I_j = \{i(j,1), \ldots, i(j,p-1)\} \tag{6.24}$$

are independent and uniformly distributed over the subsets of $\{1, \ldots, N-1\}$ of cardinality $p-1$. The probability that any given $i \le N-1$ belongs to I_j is therefore $(p-1)/(N-1)$. Thus the probability that $i(j,1) \le k-1$, i.e. the probability that there exists $\ell \le k-1$ that belongs to I_j is at most $(p-1)(k-1)/(N-1)$. Therefore

$$\mathsf{P}(\Omega_1) \le \frac{(p-1)(k-1)}{N-1}\mathsf{E}r \le \frac{kp\gamma}{N} \, .$$

Here and below, we do not try to get sharp bounds. There is no point in doing this, as anyway our methods cannot reach the best possible bounds. Rather, we aim at writing explicit bounds that are not too cumbersome. For $j < j' \le r$, the probability that a given point $i \le N-1$ belongs to both sets I_j and $I_{j'}$ is $((p-1)/(N-1))^2$. Thus the random number U of points $i \le N-1$ that belong to two different sets I_j for $j \le r$ satisfies

$$\mathsf{E}\,U = (N-1)\left(\frac{p-1}{N-1}\right)^2 \mathsf{E}\frac{r(r-1)}{2} \le \frac{p^2\gamma^2}{2N} \, ,$$

using that $\mathsf{E}r(r-1) = (\mathsf{E}r)^2$ since r is a Poisson r.v., see (A.64). Since U is integer valued, we have $\mathsf{P}(\{U \ne 0\}) \le \mathsf{E}U$ and since $\Omega_2 = \{U \ne 0\}$ we get

$$\mathsf{P}(\Omega_2) \le \frac{p^2\gamma^2}{2N} \, ,$$

so that finally, since $\Omega = \Omega_1 \cup \Omega_2$, we obtain

$$\mathsf{P}(\Omega) \le \frac{kp\gamma + p^2\gamma^2}{N} \, . \tag{6.25}$$

Using (6.22), (6.23) and (6.25), we have

$$\begin{aligned} \mathsf{E}\left|\frac{\langle f' \rangle}{\langle f \rangle}\right| &= \mathsf{E}\left(1_\Omega \mathsf{E}_-\left|\frac{\langle \mathrm{Av}\, f'\mathcal{E}\rangle_-}{\langle \mathrm{Av}\, f\mathcal{E}\rangle_-}\right|\right) + \mathsf{E}\left(1_{\Omega^c} \mathsf{E}_-\left|\frac{\langle \mathrm{Av}\, f'\mathcal{E}\rangle_-}{\langle \mathrm{Av}\, f\mathcal{E}\rangle_-}\right|\right) \\ &\le \frac{kp\gamma + p^2\gamma^2}{N} Q + \mathsf{E}\left(1_{\Omega^c} \mathsf{E}_-\left|\frac{\langle \mathrm{Av}\, f'\mathcal{E}\rangle_-}{\langle \mathrm{Av}\, f\mathcal{E}\rangle_-}\right|\right) \, . \end{aligned} \tag{6.26}$$

The next task is to use the induction hypothesis to study the last term above. When Ω does not occur (i.e. on Ω^c), all the points $i(j,\ell)$, $j \le r$, $\ell \le p-1$ are different and are $\ge k$. Recalling the notation (6.24) we have

$$J = \{i(j,\ell);\; j \le r,\; \ell \le p-1\} = \bigcup_{j \le r} I_j \, ,$$

so that $\operatorname{card} J = r(p-1)$ and

$$J \cap \{1, \dots, k-1, N\} = \emptyset \, . \tag{6.27}$$

For $i \le N-1$ let us denote by U_i the transformation of Σ^n_{N-1} that exchanges the coordinates σ_i^1 and σ_i^2 of a point $(\boldsymbol{\rho}^1, \dots, \boldsymbol{\rho}^n)$ of Σ^n_{N-1}, and that leaves all the other coordinates unchanged. That is, U_i is to $N-1$ what T_i is to N.

Lemma 6.2.4. *Assume that f' satisfies (6.8) for $i = N$, i.e. $f' \circ T_N = -f'$ and depends only on the coordinates in $\{1, \ldots, k-1, N\}$. Then when Ω does not occur (i.e. on Ω^c) we have*

$$(\mathrm{Av}\, f'\mathcal{E}) \circ \prod_{i \in J} U_i = -\mathrm{Av}\, f'\mathcal{E} \ . \tag{6.28}$$

Here $\prod_{i \in J} U_i$ denotes the composition of the transformations U_i for $i \in J$ (which does not depend on the order in which this composition is performed). This (crucial...) lemma means that something of the special symmetry of f' (as in (6.8)) is preserved when one replaces f' by $\mathrm{Av}\, f'\mathcal{E}$.

Proof. Let us write $T = \prod_{i \in J} T_i$. We observe first that

$$f' \circ T = f'$$

because f' depends only on the coordinates in $\{1, \ldots, k-1, N\}$, a set disjoint from J. Thus

$$f' \circ T \circ T_N = f' \circ T_N = -f' \tag{6.29}$$

since $f' \circ T_N = -f'$. We observe now that $T \circ T_N$ exchanges σ_i^1 and σ_i^2 for all $i \in J \cup \{N\}$. These values of i are precisely the coordinates of which $\mathcal{E}$ depends, so that

$$\mathcal{E} \circ T \circ T_N(\boldsymbol{\sigma}^1, \boldsymbol{\sigma}^2, \ldots, \boldsymbol{\sigma}^n) = \mathcal{E}(\boldsymbol{\sigma}^2, \boldsymbol{\sigma}^1, \ldots, \boldsymbol{\sigma}^n) = \mathcal{E}(\boldsymbol{\sigma}^1, \boldsymbol{\sigma}^2, \ldots, \boldsymbol{\sigma}^n) \ ,$$

and hence

$$\mathcal{E} \circ T \circ T_N = \mathcal{E} \ .$$

Combining with (6.29) we get

$$(f'\mathcal{E}) \circ T \circ T_N = (f' \circ T \circ T_N)(\mathcal{E} \circ T \circ T_N) = -f'\mathcal{E}$$

so that, since T_N^2 is the identity,

$$(f'\mathcal{E}) \circ T = -(f'\mathcal{E}) \circ T_N \ . \tag{6.30}$$

Now, for any function f we have $\mathrm{Av}(f \circ T_N) = \mathrm{Av} f$ and $\mathrm{Av}(f \circ T) = (\mathrm{Av} f) \circ \prod_{i \in J} U_i$. Therefore we obtain

$$\mathrm{Av}\,((f'\mathcal{E}) \circ T_N) = \mathrm{Av}\, f'\mathcal{E}$$

$$\mathrm{Av}((f'\mathcal{E}) \circ T) = (\mathrm{Av}\, f'\mathcal{E}) \circ \prod_{i \in J} U_i \ ,$$

so that applying Av to (6.30) proves (6.28). □

Let us set $k' = r(p-1) = \mathrm{card}\, J$, and let us enumerate as $i_1, \ldots, i_{k'}$ the points of J. Now (6.28) implies

$$\operatorname{Av} f'\mathcal{E} = \frac{1}{2}\Big(\operatorname{Av} f'\mathcal{E} - (\operatorname{Av} f'\mathcal{E}) \circ \prod_{s \le k'} U_{i_s}\Big) = \frac{1}{2} \sum_{1 \le s \le k'} f'_s \ , \tag{6.31}$$

where

$$f'_s = (\operatorname{Av} f'\mathcal{E}) \circ \prod_{u \le s-1} U_{i_u} - (\operatorname{Av} f'\mathcal{E}) \circ \prod_{u \le s} U_{i_u} \ . \tag{6.32}$$

Since U_i^2 is the identity, we have

$$f'_s \circ U_{i_s} = -f'_s \ . \tag{6.33}$$

In words, (6.31) decomposes $\operatorname{Av} f'\mathcal{E}$ as a sum of $k' = r(p-1)$ pieces that possess the symmetry property required to use the induction hypothesis. In order to apply this induction hypothesis, it remains to establish the property that will play for the pairs $(f'_s, \operatorname{Av} f\mathcal{E})$ the role the inequality $|f'| \le Qf$ plays for the pair (f', f). This is the purpose of the next lemma. For $j \le r$ we set

$$S_j = \sup |\theta_j(\varepsilon_1, \varepsilon_2, \dots, \varepsilon_p)| \ ,$$

where the supremum is over all values of $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_p = \pm 1$. We recall the notation (6.24).

Lemma 6.2.5. *Assume that Ω does not occur and that $i_s \in I_v$ for a certain (unique) $v \le r$. Then*

$$|f'_s| \le 4QS_v \exp\Big(4 \sum_{u \le r} S_u\Big) \operatorname{Av} f\mathcal{E} \ . \tag{6.34}$$

A crucial feature of this bound is that it does not depend on the number n of replicas.

Proof. Let us write

$$\mathcal{E}' = \exp\Big(\sum_{3 \le \ell \le n} -H(\boldsymbol{\sigma}^\ell)\Big) \ ; \ \mathcal{E}'' = \exp\Big(\sum_{\ell = 1,2} -H(\boldsymbol{\sigma}^\ell)\Big) \ ,$$

so that $\mathcal{E} = \mathcal{E}'\mathcal{E}''$. Since $|H(\boldsymbol{\sigma})| \le \sum_{j \le r} S_j$, we have

$$\mathcal{E}'' \ge \exp\Big(-2 \sum_{j \le r} S_j\Big) \ ,$$

and therefore

$$\mathcal{E} \ge \mathcal{E}' \exp\Big(-2 \sum_{j \le r} S_j\Big) \ . \tag{6.35}$$

This implies

$$\operatorname{Av} f\mathcal{E} \ge (\operatorname{Av} f\mathcal{E}') \exp\Big(-2 \sum_{j \le r} S_j\Big) \ . \tag{6.36}$$

Next,

$$\begin{aligned} f'_s &= (\mathrm{Av}\, f'\mathcal{E}) \circ \prod_{u\le s-1} U_{i_u} - (\mathrm{Av}\, f'\mathcal{E}) \circ \prod_{u\le s} U_{i_u} \\ &= \mathrm{Av}\bigg((f'\mathcal{E}) \circ \prod_{u\le s-1} T_{i_u} - (f'\mathcal{E}) \circ \prod_{u\le s} T_{i_u} \bigg) \\ &= \mathrm{Av}\bigg(f'\bigg(\mathcal{E} \circ \prod_{u\le s-1} T_{i_u} - \mathcal{E} \circ \prod_{u\le s} T_{i_u} \bigg)\bigg) , \end{aligned} \tag{6.37}$$

using in the last line that $f' \circ T_{i_u} = f'$ for each u, since f' depends only on the coordinates $1, \ldots, k-1, N$. Recalling that $\mathcal{E} = \mathcal{E}''\mathcal{E}'$, and observing that for each i, we have $\mathcal{E}' \circ T_i = \mathcal{E}'$, we get

$$\mathcal{E} \circ \prod_{u\le s-1} T_{i_u} - \mathcal{E} \circ \prod_{u\le s} T_{i_u} = \mathcal{E}'\bigg(\mathcal{E}'' \circ \prod_{u\le s-1} T_{i_u} - \mathcal{E}'' \circ \prod_{u\le s} T_{i_u} \bigg) ,$$

and, if we set

$$\Delta = \sup\bigg| \mathcal{E}'' \circ \prod_{u\le s-1} T_{i_u} - \mathcal{E}'' \circ \prod_{u\le s} T_{i_u} \bigg| = \sup |\mathcal{E}'' - \mathcal{E}'' \circ T_{i_s}| ,$$

we get from (6.37) that, using that $|f'| \le Qf$ in the first inequality and (6.35) in the second one,

$$|f'_s| \le \Delta \mathrm{Av}\, (|f'|\mathcal{E}') \le Q\Delta \mathrm{Av}(f\mathcal{E}') \le Q\Delta \mathrm{Av}(f\mathcal{E}) \exp\bigg(2\sum_{j\le r} S_j \bigg) . \tag{6.38}$$

To bound Δ, we write $\mathcal{E}'' = \prod_{j\le r} \mathcal{E}_j$, where

$$\mathcal{E}_j = \exp \sum_{\ell=1,2} \theta_j(\sigma^\ell_{i(j,1)}, \ldots, \sigma^\ell_{i(j,p-1)}, \sigma^\ell_N) .$$

We note that $\mathcal{E}_j \circ T_{i_s} = \mathcal{E}_j$ if $j \ne v$, because then $\mathcal{E}_j$ depends only on the coordinates in I_j, and $i_s \notin I_j$ if $j \ne v$, since $i_s \in I_v$ and $I_j \cap I_v = \emptyset$. Thus

$$\mathcal{E}'' - \mathcal{E}'' \circ T_{i_s} = (\mathcal{E}_v - \mathcal{E}_v \circ T_{i_s}) \prod_{j\ne v} \mathcal{E}_j .$$

Now, using the inequality $|e^x - e^y| \le |x-y|e^a \le 2ae^a$ for $|x|, |y| \le a$ and $a = 2S_v$, we get

$$|\mathcal{E}_v - \mathcal{E}_v \circ T_{i_s}| \le 4S_v \exp 2S_v .$$

Since for all j we have $\mathcal{E}_j \le \exp 2S_j$, we get $\Delta \le 4S_v \exp 2\sum_{j\le r} S_j$. Combining with (6.38) completes the proof. □

Proposition 6.2.6. *Assume that $N \geq p+1$ and that condition $C(N-1, \gamma_0, B, B^*)$ holds. Consider f' and f as in Definition 6.2.1, and assume that $\gamma \leq \gamma_0$. Then*

$$\mathsf{E}\left|\frac{\langle f'\rangle}{\langle f\rangle}\right| \leq \frac{Qp}{N}\Big(k(\gamma+4BD\exp 4D)+4pBU\mathsf{E}r^2V^{r-1}+p\gamma^2+4B^*D\exp 4D\Big)\,, \tag{6.39}$$

where

$$D = \gamma \mathsf{E}\, S \exp 4S\,.$$

Proof. We keep the notation of Lemmas 6.2.4 and 6.2.5. Since $\gamma_- \leq \gamma$, we can use $C(N-1, \gamma_0, B, B^*)$ to conclude from (6.33) and (6.34) that, since f'_s and $\mathrm{Av}\,\mathcal{E} f$ depend on $k-1+r(p-1) \leq k+rp$ coordinates, and since $1/(N-1) \leq 2/N$ because $N \geq 2$, on Ω^c we have

$$\mathsf{E}_-\left|\frac{\langle f'_s\rangle_-}{\langle \mathrm{Av}\, f\mathcal{E}\rangle_-}\right| \leq \frac{8Q}{N}((k+rp)B+B^*)S_v\exp\Bigg(4\sum_{j\leq r}S_j\Bigg)\,.$$

Let us denote by E_θ expectation in the r.v.s $\theta_1, \ldots, \theta_r$ only. Then we get

$$\mathsf{E}_\theta\mathsf{E}_-\left|\frac{\langle f'_s\rangle_-}{\langle \mathrm{Av}\, f\mathcal{E}\rangle_-}\right| \leq \frac{8Q}{N}((k+rp)B+B^*)UV^{r-1}\,,$$

where

$$U = \mathsf{E}\, S\exp 4S\,; \qquad V = \mathsf{E}\exp 4S\,.$$

Combining with (6.31), and since there are $k' = r(p-1) \leq rp$ terms we get

$$\mathsf{E}_\theta\,\mathsf{E}_-\left|\frac{\langle \mathrm{Av}\, f'\mathcal{E}\rangle_-}{\langle \mathrm{Av}\, f\mathcal{E}\rangle_-}\right| \leq \frac{4Qp}{N}((kr+r^2p)B+rB^*)UV^{r-1}\,.$$

This bound assumes that Ω does not occur; but combining with (6.26) we obtain the bound

$$\mathsf{E}\left|\frac{\langle f'\rangle}{\langle f\rangle}\right| \leq \frac{Qp}{N}\Big(k\gamma+p\gamma^2+4B\big(kU\mathsf{E}rV^{r-1}+pU\mathsf{E}r^2V^{r-1}\big)+4B^*U\mathsf{E}rV^{r-1}\Big)\,.$$

Since r is a Poisson r.v. of expectation γ a straightforward calculation shows that $\mathsf{E}rV^{r-1} = \gamma\exp\gamma(V-1)$. Since $e^x \leq 1+xe^x$ for all $x \geq 0$ (as is trivial using power series expansion) we have $V \leq 1+4U$, so $\exp\gamma(V-1) \leq \exp 4\gamma U$ and $U\mathsf{E}rV^{r-1} \leq D\exp 4D$. The result follows. □

Proof of Theorem 6.2.2. If

$$D_0 = \gamma_0 \mathsf{E}\, S\exp 4S$$

is small enough that $16pD_0 \leq 1$ then

$$4pD_0\exp 4D_0 \leq 1/2\,, \tag{6.40}$$

and (6.39) implies

$$\mathsf{E}\left|\frac{\langle f'\rangle}{\langle f\rangle}\right| \le \frac{Q}{N}\left(k\left(p\gamma_0+\frac{B}{2}\right)+4p^2BU\mathsf{E}r^2V^{r-1}+p^2\gamma^2+\frac{B^*}{2}\right) .$$

Thus condition

$$C(N,\gamma_0,p\gamma_0+B/2,4p^2BU\mathsf{E}r^2V^{r-1}+p^2\gamma_0^2+B^*/2)$$

holds. That is, we have proved under (6.40) that

$$C(N-1,\gamma_0,B,B^*)\Rightarrow C(N,\gamma_0,p\gamma_0+B/2,4p^2BU\mathsf{E}r^2V^{r-1}+p^2\gamma_0^2+B^*/2)) . \tag{6.41}$$

Now, we observe that $U\mathsf{E}r^2V^{r-1}\le K^{\sim}$ and that if $K_1=2p\gamma_0$ and $K_1^*=8p^2K_1K^{\sim}+2p^2\gamma_0^2$, (6.41) shows that (6.12) holds, and we have completed the induction. □

Probably at this point it is good to stop for a while and to wonder what is the nature of the previous argument. In essence this is "contraction argument". The operation of "adding one spin" essentially acts as a type of contraction, as is witnessed by the factor 1/2 in front of B and B^* in the right-hand side of (6.41). As it turns out, almost every single argument used in this work to control a model under a "high-temperature condition" is of the same type, whether this is rather apparent, as in Section 1.6, or in a more disguised form as here. (The one exception being Latala's argument of Section 1.4.)

We explained at length in Section 1.4 that we expect that at high-temperature, as long as one considers a number of spins that remains bounded independently of N, Gibbs' measure is nearly a product measure. For the present model, this property follows from Theorem 6.2.2 and we now give quantitative estimates to that effect, in the setting we need for future uses.

Let us consider the product measure μ on Σ_{N-1} such that

$$\forall\, i\le N-1\,,\ \int\sigma_i\,\mathrm{d}\mu(\boldsymbol{\rho})=\langle\sigma_i\rangle_-\ ,$$

and let us denote by $\langle\cdot\rangle_\bullet$ an average with respect to μ. Equivalently, for a function f on Σ_{N-1}, we have

$$\langle f\rangle_\bullet=\langle f(\sigma_1^1,\dots,\sigma_{N-1}^{N-1})\rangle_-\ , \tag{6.42}$$

where σ_i^i is the i-th coordinate of the i-th replica $\boldsymbol{\rho}^i$. The following consequence of property $C(N,\gamma_0,K_0,K_0)$ will be used in Section 6.4. It expresses, in a form that is particularly adapted to the use of the cavity method the fact that under property $C(N,\gamma_0,K_0,K_0)$, a given number of spins (independent of N) become nearly independent for large N.

Proposition 6.2.7. *If property $C(N,\gamma_0,K_0,K_0)$ holds for each N, and if $\gamma \le \gamma_0$, the following occurs. Consider for $j \le r$ sets $I_j \subset \{1,\dots,N\}$ with* $\operatorname{card} I_j = p$, *$N \in I_j$, and such that $j \ne j' \Rightarrow I_j \cap I_{j'} = \{N\}$. For $j \le r$ consider functions W_j on Σ_N depending only on the coordinates in I_j and let $S_j = \sup |W_j(\boldsymbol{\sigma})|$. Let*

$$\mathcal{E} = \exp \sum_{j \le r} W_j(\boldsymbol{\sigma}) \ .$$

Then, recalling the definition (6.42), we have

$$\mathsf{E}_- \left| \frac{\langle \operatorname{Av} \sigma_N \mathcal{E} \rangle_-}{\langle \operatorname{Av} \mathcal{E} \rangle_-} - \frac{\langle \operatorname{Av} \sigma_N \mathcal{E} \rangle_\bullet}{\langle \operatorname{Av} \mathcal{E} \rangle_\bullet} \right| \le \frac{8r(p-1)^2 K_0}{N-1} \sum_{j \le r} \exp 2S_j \ . \tag{6.43}$$

This is a powerful principle, since it is very much easier to work with the averages $\langle\cdot\rangle_\bullet$ than with the Gibbs averages $\langle\cdot\rangle_-$. We will use this result when r is as usual the number of terms in (6.17) but since in (6.43) the expectation E_- is only in the randomness of $\langle\cdot\rangle_-$ we can, in the proof, think of the quantities r and W_j as being non-random.

Proof. Let $f' = \operatorname{Av} \sigma_N \mathcal{E}$ and $f = \operatorname{Av} \mathcal{E}$. For $0 \le i \le N-1$, let us define

$$f_i = f_i(\boldsymbol{\rho}^1, \dots, \boldsymbol{\rho}^{N-1}) = f(\sigma_1^1, \sigma_2^2, \dots, \sigma_i^i, \sigma_{i+1}^1, \dots, \sigma_{N-1}^1)$$

and f'_i similarly. The idea is simply that "we make the spins independent one at a time". Thus

$$\frac{\langle \operatorname{Av} \sigma_N \mathcal{E} \rangle_-}{\langle \operatorname{Av} \mathcal{E} \rangle_-} = \frac{\langle f'_1 \rangle_-}{\langle f_1 \rangle_-} \ ; \ \frac{\langle \operatorname{Av} \sigma_N \mathcal{E} \rangle_\bullet}{\langle \operatorname{Av} \mathcal{E} \rangle_\bullet} = \frac{\langle f'_{N-1} \rangle_-}{\langle f_{N-1} \rangle_-} \ , \tag{6.44}$$

and the left-hand side of (6.43) is bounded by

$$\sum_{2 \le i \le N-1} \mathsf{E}_- \left| \frac{\langle f'_{i-1} \rangle_-}{\langle f_{i-1} \rangle_-} - \frac{\langle f'_i \rangle_-}{\langle f_i \rangle_-} \right| \ .$$

The terms in the summation are zero unless i belongs to the union of the sets I_j, $j \le r$, for otherwise f' and f do not depend on the i-th coordinate and $f_i = f_{i-1}$, $f'_i = f'_{i-1}$. We then try to bound the terms in the summation when $i \in I_j$ for a certain $j \le r$. Since $|f'_i| \le f_i$ we have

$$\begin{aligned} \left| \frac{\langle f'_{i-1} \rangle_-}{\langle f_{i-1} \rangle_-} - \frac{\langle f'_i \rangle_-}{\langle f_i \rangle_-} \right| &\le \left| \frac{\langle f'_{i-1} - f'_i \rangle_-}{\langle f_{i-1} \rangle_-} \right| + \left| \frac{\langle f'_i \rangle_- \langle f_{i-1} - f_i \rangle_-}{\langle f_{i-1} \rangle_- \langle f_i \rangle_-} \right| \\ &\le \left| \frac{\langle f'_{i-1} - f'_i \rangle_-}{\langle f_{i-1} \rangle_-} \right| + \left| \frac{\langle f_{i-1} - f_i \rangle_-}{\langle f_{i-1} \rangle_-} \right| \end{aligned}$$

so that, taking expectation in the previous inequality we get

$$\mathsf{E}_- \left|\frac{\langle f'_{i-1}\rangle_-}{\langle f_{i-1}\rangle_-} - \frac{\langle f'_i\rangle_-}{\langle f_i\rangle_-}\right| \le \mathsf{E}_- \left|\frac{\langle f'_{i-1} - f'_i\rangle_-}{\langle f_{i-1}\rangle_-}\right| + \mathsf{E}_- \left|\frac{\langle f_{i-1} - f_i\rangle_-}{\langle f_{i-1}\rangle_-}\right| . \quad (6.45)$$

We will use $C(N-1,\gamma_0,K_0,K_0)$ to bound these terms. First, we observe that the function $f'_{i-1} - f'_i$ changes sign if we exchange σ_i^1 and σ_i^i. Next, we observe that since W_u does not depend on σ_i for $u \ne j$ (where j is defined by $i \in I_j$) we have

$$\begin{aligned} \mathcal{E}' &:= \exp \sum_{u \ne j} W_u(\sigma_1^1, \sigma_2^2, \ldots, \sigma_i^i, \sigma_{i+1}^1, \ldots, \sigma_N^1) \\ &= \exp \sum_{u \ne j} W_u(\sigma_1^1, \sigma_2^2, \ldots, \sigma_{i-1}^{i-1}, \sigma_i^1, \sigma_{i+1}^1, \ldots, \sigma_N^1) \ . \end{aligned}$$

Then

$$f_{i-1} = \mathrm{Av}\, \mathcal{E}(\sigma_1^1, \ldots, \sigma_{i-1}^{i-1}, \sigma_i^1, \ldots, \sigma_N^1) \ge \exp(-S_j) \mathrm{Av}\, \mathcal{E}' ,$$

where Av denotes average over $\sigma_N^1 = \pm 1$. In a similar fashion, we get $|f'_{i-1}| \le \exp S_j \mathrm{Av}\, \mathcal{E}'$, $|f'_i| \le \exp S_j \mathrm{Av}\, \mathcal{E}'$, and thus

$$|f'_{i-1} - f'_i| \le (2 \exp 2S_j) f_{i-1} \ ,$$

so that using (6.11) property $C(N-1,\gamma_0,K_0,K_0)$ implies

$$\mathsf{E}_- \left|\frac{\langle f'_{i-1} - f'_i\rangle_-}{\langle f_{i-1}\rangle_-}\right| \le \frac{4K_0}{N-1} r(p-1) \exp 2S_j \ , \quad (6.46)$$

because these functions depend on $r(p-1)$ coordinates. We proceed similarly to handle the last term on the right-hand side of (6.45). We then perform the summation over $i \le N-1$. A new factor $p-1$ occurs because each set I_j contains $p-1$ such values of i. □

6.3 The Functional Order Parameter

As happened in the previous models, we expect that if we fix a number n and take N very large, at a given disorder, n spins $(\sigma_1, \ldots, \sigma_n)$ will asymptotically be independent, and that the r.v.s $\langle\sigma_1\rangle, \ldots, \langle\sigma_n\rangle$ will asymptotically be independent. In the case of the SK model, the limiting law of $\langle\sigma_i\rangle$ was the law of $\mathrm{th}(\beta z\sqrt{q} + h)$ where z is a standard Gaussian r.v. and thus this law depended only on the single parameter q.

The most striking feature of the present model is that the limiting law is now a complicated object, that no longer depends simply on a few parameters. It is therefore reasonable to think of this limiting law μ as being itself a kind of parameter (the correct value of which has to be found). This is what the physicists mean when they say "that the order parameter of the model is a function" because they identify a probability distribution μ on $\mathbb{R}$ with the tail function $t \mapsto \mu([t,\infty))$.

The purpose of the present section is to find the correct value of this parameter. As is the case of the SK model this value will be given as the solution of a certain equation. The idea of the construction we will perform is very simple. While using the cavity method in the previous section, we have seen in (6.34) (used for $n = 1$ and $f(\boldsymbol{\sigma}) = \sigma_N$) that

$$\langle \sigma_N \rangle = \frac{\langle \mathrm{Av} \sigma_N \mathcal{E} \rangle_-}{\langle \mathrm{Av} \mathcal{E} \rangle_-} , \tag{6.47}$$

where

$$\mathcal{E} = \exp \sum_{j \le r} \theta_j(\sigma_{i(j,1)}, \ldots, \sigma_{i(j,p-1)}, \sigma_N) . \tag{6.48}$$

In the limit $N \to \infty$ the sets $I_j = \{i(j,1), \ldots, i(j,p-1)\}$ are disjoint. The quantity $\mathcal{E}$ depends on a number of spins that in essence does not depend on N. If we know the asymptotic behavior of any fixed number (i.e. of any number that does not depend on N) of the spins $(\sigma_i)_{i<N}$, we can then compute the behavior of the spin σ_N. This behavior has to be the same as the behavior of the spins σ_i for $i < N$, and this gives rise to a "self-consistency equation".

To define formally this equation, consider a Poisson r.v. r with $\mathsf{E} r = \gamma$, and independent of the r.v.s θ_j. For $\boldsymbol{\sigma} \in \{-1,1\}^{\mathbb{N}}$ and $\varepsilon \in \{-1,1\}$ we define

$$\mathcal{E}_r = \mathcal{E}_r(\boldsymbol{\sigma}, \varepsilon) = \exp \sum_{1 \le j \le r} \theta_j(\sigma_{(j-1)(p-1)+1}, \ldots, \sigma_{j(p-1)}, \varepsilon) . \tag{6.49}$$

This definition will be used **many times** in the sequel. We note that $\mathcal{E}_r$ depends on $\boldsymbol{\sigma}$ only through the coordinates of rank $\le r(p-1)$.

Given a sequence $\mathbf{x} = (x_i)_{i \ge 1}$ with $|x_i| \le 1$ we denote by $\lambda_{\mathbf{x}}$ the probability on $\{-1,1\}^{\mathbb{N}}$ that "has a density $\prod_i (1 + x_i \sigma_i)$ with respect to the uniform measure". More formally, $\lambda_{\mathbf{x}}$ is the product measure such that $\int \sigma_i \mathrm{d}\lambda_{\mathbf{x}}(\boldsymbol{\sigma}) = x_i$ for each i. We denote by $\langle \cdot \rangle_{\mathbf{x}}$ an average for $\lambda_{\mathbf{x}}$.

Similarly, if $\mathbf{x} = (x_i)_{i \le M}$ we also denote by $\lambda_{\mathbf{x}}$ the probability measure on $\Sigma_M = \{-1,1\}^M$ such that $\int \sigma_i \mathrm{d}\lambda_{\mathbf{x}}(\boldsymbol{\sigma}) = x_i$ and we denote by $\langle \cdot \rangle_{\mathbf{x}}$ an average for $\lambda_{\mathbf{x}}$, so that we have

$$\langle f \rangle_{\mathbf{x}} = \int \prod_{i \le M} (1 + x_i \sigma_i) f(\boldsymbol{\sigma}) \mathrm{d}\boldsymbol{\sigma} ,$$

where $\mathrm{d}\boldsymbol{\sigma}$ means average for the uniform measure on Σ_M.

These definitions are also of **central importance** in this chapter. The idea underlying these definitions has already been used implicitly in (6.42) since for a function f on Σ_{N-1} we have

$$\langle f \rangle_\bullet = \langle f \rangle_{\mathbf{Y}} , \tag{6.50}$$

where $\mathbf{Y} = (\langle\sigma_1\rangle_-, \ldots, \langle\sigma_{N-1}\rangle_-)$.

Consider a probability measure μ on $[-1,1]$, and an i.i.d. sequence $\mathbf{X} = (X_i)_{i\geq 1}$ such that X_i is of law μ. We define $T(\mu)$ as the law of the r.v.

$$\frac{\langle \mathrm{Av}\, \varepsilon \mathcal{E}_r \rangle_{\mathbf{X}}}{\langle \mathrm{Av}\, \mathcal{E}_r \rangle_{\mathbf{X}}} \,, \tag{6.51}$$

where Av denotes the average over $\varepsilon = \pm 1$. We note that $\mathcal{E}$ depends on $\boldsymbol{\sigma}$ and ε, so that $\mathrm{Av}\, \varepsilon \mathcal{E}_r$ and $\mathrm{Av}\, \mathcal{E}_r$ depend on $\boldsymbol{\sigma}$ only and (6.51) makes sense. The intuition is that if μ is the law of $\langle \sigma_i \rangle$ for $i < N$, then $T(\mu)$ is the law of $\langle \sigma_N \rangle$. This is simply because if the spins "decorrelate" as we expect, and if in the limit any fixed number of the averages $\langle \sigma_i \rangle_i$ are i.i.d. of law μ, then the right-hand side of (6.47) will in the limit have the same distribution as the quantity (6.51).

Theorem 6.3.1. *Assume that*

$$4\gamma p \mathsf{E}(S \exp 2S) \leq 1 \,. \tag{6.52}$$

Then there exists a unique probability measure μ on $[-1,1]$ such that

$$\mu = T(\mu) \,.$$

The proof will consist of showing that T is a contraction for the Monge-Kantorovich transportation-cost distance d defined in (A.66) on the set of probability measures on $[-1,1]$ provided with the usual distance. In the present case, this distance is simply given by the formula

$$d(\mu_1, \mu_2) = \inf \mathsf{E}|X - Y| \,,$$

where the infimum is taken over all pairs of r.v.s (X, Y) such that the law of X is μ_1 and the law of Y is μ_2. The very definition of d shows that to bound $d(\mu_1, \mu_2)$ there is no other method than to produce a pair (X, Y) as above such that $\mathsf{E}|X - Y|$ is appropriately small. Such a pair will informally be called a coupling of the r.v.s X and Y.

Lemma 6.3.2. *For a function f on $\{-1,1\}^{\mathbb{N}}$, we have*

$$\frac{\partial}{\partial x_i} \langle f \rangle_{\mathbf{x}} = \langle \Delta_i f \rangle_{\mathbf{x}} \tag{6.53}$$

where $\Delta_i f(\boldsymbol{\eta}) = (f(\boldsymbol{\eta}_i^+) - f(\boldsymbol{\eta}_i^-))/2$, and where $\boldsymbol{\eta}_i^+$ (resp. $\boldsymbol{\eta}_i^-$) is obtained by replacing the i-th coordinate of $\boldsymbol{\eta}$ by 1 (resp. -1).

Proof. The measure λ_x on $\{-1,1\}$ such that $\int \eta \, \mathrm{d}\lambda_x(\eta) = x$ gives mass $(1+x)/2$ to 1 and mass $(1-x)/2$ to -1, so that for a function f on $\{-1,1\}$ we have

$$\langle f \rangle_x = \int f(\eta)\, \mathrm{d}\lambda_x(\eta) = \frac{1}{2}(f(1) + f(-1)) + \frac{x}{2}(f(1) - f(-1)) \,.$$

Thus, using in the second inequality the trivial fact that $a = \langle a \rangle_x$ for any number a implies

$$\frac{\mathrm{d}}{\mathrm{d}x} \langle f \rangle_x = \frac{1}{2}(f(1) - f(-1)) = \left\langle \frac{1}{2}(f(1) - f(-1)) \right\rangle_x \,. \tag{6.54}$$

Since $\lambda_{\mathbf{x}}$ is a product measure, using (6.54) given all the coordinates different from i, and then Fubini's theorem, we obtain (6.53). □

Lemma 6.3.3. *If $\mathcal{E}_r$ is as in (6.49), if $1 \le j \le r$ and if $(j-1)(p-1) < i \le j(p-1)$, then*

$$\left| \frac{\partial}{\partial x_i} \frac{\langle \mathrm{Av}\, \varepsilon \mathcal{E}_r \rangle_{\mathbf{x}}}{\langle \mathrm{Av}\, \mathcal{E}_r \rangle_{\mathbf{x}}} \right| \le 2S_j \exp 2S_j$$

where $S_j = \sup |\theta_j|$. For the other values of i the left-hand side of the previous inequality is 0.

Proof. Lemma 6.3.2 implies:

$$\frac{\partial}{\partial x_i} \frac{\langle \mathrm{Av}\, \varepsilon \mathcal{E}_r \rangle_{\mathbf{x}}}{\langle \mathrm{Av}\, \mathcal{E}_r \rangle_{\mathbf{x}}} = \frac{\langle \Delta_i (\mathrm{Av}\, \varepsilon \mathcal{E}_r) \rangle_{\mathbf{x}}}{\langle \mathrm{Av}\, \mathcal{E}_r \rangle_{\mathbf{x}}} - \frac{\langle \mathrm{Av}\, \varepsilon \mathcal{E}_r \rangle_{\mathbf{x}} \langle \Delta_i \mathrm{Av}\, \mathcal{E}_r \rangle_{\mathbf{x}}}{\langle \mathrm{Av}\, \mathcal{E}_r \rangle_{\mathbf{x}}^2} \,. \tag{6.55}$$

Now

$$|\Delta_i (\mathrm{Av}\, \varepsilon \mathcal{E}_r)| = |\mathrm{Av}\, (\varepsilon \Delta_i \mathcal{E}_r)| \le \mathrm{Av}\, |\Delta_i \mathcal{E}_r| \,.$$

We write $\mathcal{E}_r = \mathcal{E}' \mathcal{E}''$, where $\mathcal{E}' = \exp \theta_j(\sigma_{(j-1)(p-1)+1}, \ldots, \sigma_{j(p-1)}, \varepsilon)$, and where $\mathcal{E}''$ does not depend on σ_i. Thus, using that $|e^x - e^y| \le |x-y| e^a \le 2ae^a$ for $|x|, |y| \le a$, we get (keeping in mind the factor 1/2 in the definition of Δ_i, that offsets the factor 2 above) that $\Delta_i \mathcal{E}' \le S_j \exp S_j$, and since $\mathcal{E}'' \le \mathcal{E}_r \exp S_j$ we get

$$|\Delta_i \mathcal{E}_r| = |\mathcal{E}'' \Delta_i \mathcal{E}'| \le (S_j \exp S_j) \mathcal{E}'' \le (S_j \exp 2S_j) \mathcal{E}_r$$

and thus

$$\left| \frac{\langle \Delta_i (\mathrm{Av}\, \varepsilon \mathcal{E}_r) \rangle_{\mathbf{x}}}{\langle \mathrm{Av}\, \mathcal{E}_r \rangle_{\mathbf{x}}} \right| \le S_j \exp 2S_j \,.$$

The last term of (6.55) is bounded similarly. □

Proof of Theorem 6.3.1. This is a fixed point argument. It suffices to prove that under (6.52), for any two probability measures μ_1 and μ_2 on $[-1, 1]$, we have

$$d(T(\mu_1), T(\mu_2)) \le \frac{1}{2} d(\mu_1, \mu_2) \,. \tag{6.56}$$

First, Lemma 6.3.3 yields that given $\mathbf{x}, \mathbf{y} \in [-1, 1]^{\mathbb{N}}$ it holds:

$$\left|\frac{\langle \mathrm{Av}\,\varepsilon\mathcal{E}_r\rangle_{\mathbf{x}}}{\langle \mathrm{Av}\,\mathcal{E}_r\rangle_{\mathbf{x}}} - \frac{\langle \mathrm{Av}\,\varepsilon\mathcal{E}_r\rangle_{\mathbf{y}}}{\langle \mathrm{Av}\,\mathcal{E}_r\rangle_{\mathbf{y}}}\right| \le 2\sum_{j\le r} S_j \exp 2S_j \sum_{(j-1)(p-1)<i\le j(p-1)} |x_i - y_i| \ . \tag{6.57}$$

Consider a pair (X, Y) of r.v.s and independent copies $(X_i, Y_i)_{i\ge1}$ of this pair. Let $\mathbf{X} = (X_i)_{i\ge1}$, $\mathbf{Y} = (Y_i)_{i\ge1}$, so that from (6.57) we have

$$\left|\frac{\langle \mathrm{Av}\,\varepsilon\mathcal{E}_r\rangle_{\mathbf{X}}}{\langle \mathrm{Av}\,\mathcal{E}_r\rangle_{\mathbf{X}}} - \frac{\langle \mathrm{Av}\,\varepsilon\mathcal{E}_r\rangle_{\mathbf{Y}}}{\langle \mathrm{Av}\,\mathcal{E}_r\rangle_{\mathbf{Y}}}\right| \le 2\sum_{j\le r} S_j \exp 2S_j \sum_{(j-1)(p-1)<i\le j(p-1)} |X_i - Y_i| \ . \tag{6.58}$$

Let us assume that the randomness of the pairs (X_i, Y_i) is independent of the other sources of randomness in (6.58). Taking expectations in (6.58) we get

$$\mathsf{E}\left|\frac{\langle \mathrm{Av}\,\varepsilon\mathcal{E}_r\rangle_{\mathbf{X}}}{\langle \mathrm{Av}\,\mathcal{E}_r\rangle_{\mathbf{X}}} - \frac{\langle \mathrm{Av}\,\varepsilon\mathcal{E}_r\rangle_{\mathbf{Y}}}{\langle \mathrm{Av}\,\mathcal{E}_r\rangle_{\mathbf{Y}}}\right| \le 2\gamma(p-1)(\mathsf{E}S\exp 2S)\mathsf{E}|X-Y| \ . \tag{6.59}$$

If X and Y have laws μ_1 and μ_2 respectively, then

$$\frac{\langle \mathrm{Av}\,\varepsilon\mathcal{E}_r\rangle_{\mathbf{X}}}{\langle \mathrm{Av}\,\mathcal{E}_r\rangle_{\mathbf{X}}} \quad \text{and} \quad \frac{\langle \mathrm{Av}\,\varepsilon\mathcal{E}_r\rangle_{\mathbf{Y}}}{\langle \mathrm{Av}\,\mathcal{E}_r\rangle_{\mathbf{Y}}}$$

have laws $T(\mu_1)$ and $T(\mu_2)$ respectively, so that (6.59) implies

$$d(T(\mu_1), T(\mu_2)) \le 2\gamma(p-1)(\mathsf{E}S\exp 2S)\mathsf{E}|X-Y| \ .$$

Taking the infimum over all possible choices of X and Y yields

$$d(T(\mu_1), T(\mu_2)) \le 2\gamma(p-1)d(\mu_1,\mu_2)\mathsf{E}S\exp 2S \ ,$$

so that (6.52) implies (6.56). □

Let us denote by T_γ the operator T when we want to insist on the dependence on γ. The unique solution of the equation $\mu = T_\gamma(\mu)$ depends on γ, and we denote it by μ_γ when we want to emphasize this dependence.

Lemma 6.3.4. *If γ and γ' satisfy (6.52) we have*

$$d(\mu_\gamma, \mu_{\gamma'}) \le 4|\gamma - \gamma'| \ .$$

Proof. Without loss of generality we can assume that $\gamma \le \gamma'$. Since $\mu_\gamma = T_\gamma(\mu_\gamma)$ and $\mu_{\gamma'} = T_{\gamma'}(\mu_{\gamma'})$, we have

$$\begin{aligned} d(\mu_\gamma, \mu_{\gamma'}) &\le d(T_\gamma(\mu_\gamma), T_\gamma(\mu_{\gamma'})) + d(T_\gamma(\mu_{\gamma'}), T_{\gamma'}(\mu_{\gamma'})) \\ &\le \frac{1}{2} d(\mu_\gamma, \mu_{\gamma'}) + d(T_\gamma(\mu_{\gamma'}), T_{\gamma'}(\mu_{\gamma'})) \ , \end{aligned} \tag{6.60}$$

using (6.56). To compare $T_\gamma(\mu)$ and $T_{\gamma'}(\mu)$ the basic idea is that there is natural coupling between a Poisson r.v. of expectation γ and another Poisson r.v. of expectation γ' (an idea that will be used again in the next section).

Namely if r'' is a Poisson r.v. with $\mathsf{E}r'' = \gamma'' := \gamma' - \gamma$, and r'' is independent of the Poisson r.v. r such that $\mathsf{E}r = \gamma$ then $r + r''$ is a Poisson r.v. of expectation γ'. Consider $\mathcal{E}_r$ as in (6.49) and, with the same notation,

$$\mathcal{E}' = \exp \sum_{r<j\le r+r''} \theta_j(\sigma_{(j-1)(p-1)+1}, \dots, \sigma_{j(p-1)}, \varepsilon) ,$$

so that $\mathcal{E}_r\mathcal{E}' = \mathcal{E}_{r+r''}$. Consider an i.i.d. sequence $\mathbf{X} = (X_i)_{i\ge 1}$ of common law μ. Then the r.v.s

$$\frac{\langle \mathrm{Av}\, \varepsilon \mathcal{E}_r \rangle_{\mathbf{X}}}{\langle \mathrm{Av}\, \mathcal{E}_r \rangle_{\mathbf{X}}} \quad \text{and} \quad \frac{\langle \mathrm{Av}\, \varepsilon \mathcal{E}_r \mathcal{E}' \rangle_{\mathbf{X}}}{\langle \mathrm{Av}\, \mathcal{E}_r \mathcal{E}' \rangle_{\mathbf{X}}}$$

have respectively laws $T_\gamma(\mu)$ and $T_{\gamma'}(\mu)$. Thus

$$\begin{aligned} d(T_\gamma(\mu), T_{\gamma'}(\mu)) &\le \mathsf{E}\left| \frac{\langle \mathrm{Av}\, \varepsilon \mathcal{E}_r \rangle_{\mathbf{X}}}{\langle \mathrm{Av}\, \mathcal{E}_r \rangle_{\mathbf{X}}} - \frac{\langle \mathrm{Av}\, \varepsilon \mathcal{E}_r \mathcal{E}' \rangle_{\mathbf{X}}}{\langle \mathrm{Av}\, \mathcal{E}_r \mathcal{E}' \rangle_{\mathbf{X}}} \right| \\ &\le 2\mathsf{P}(r'' \ne 0) = 2(1 - e^{-(\gamma'-\gamma)}) \le 2(\gamma' - \gamma) , \end{aligned} \tag{6.61}$$

so that (6.60) implies that $d(\mu_\gamma, \mu_{\gamma'}) \le d(\mu_\gamma, \mu_{\gamma'})/2 + 2(\gamma' - \gamma)$, hence the desired result. □

Exercise 6.3.5. Consider three functions U, V, W on Σ_N^n. Assume that $V \ge 0$, that for a certain number Q, we have $|U| \le QV$, and let $S^* = \sup_{\boldsymbol{\sigma}^1,\dots,\boldsymbol{\sigma}^n} |W|$. Prove that for any Gibbs measure $\langle\cdot\rangle$ we have

$$\left| \frac{\langle U \exp W \rangle}{\langle V \exp W \rangle} - \frac{\langle U \rangle}{\langle V \rangle} \right| \le 2QS^* \exp 2S^* .$$

Exercise 6.3.6. Use the idea of Exercise 6.3.5 to control the influence of $\mathcal{E}'$ in (6.61) and to show that if γ and γ' satisfy (6.52) then $d(\mu_\gamma, \mu_{\gamma'}) \le 4|\gamma - \gamma'|\mathsf{E}S \exp 2S$.

6.4 The Replica-Symmetric Solution

In this section we will first prove that asymptotically as $N \to \infty$ any fixed number of the quantities $\langle \sigma_i \rangle$ are i.i.d. of law μ_γ, where μ_γ was defined in the last section. We will then compute the quantity $\lim_{N\to\infty} p_N(\gamma) = \lim_{N\to\infty} N^{-1}\mathsf{E} \log Z_N(\gamma)$.

Theorem 6.4.1. *Assume that*

$$16p\gamma_0 \mathsf{E}S \exp 4S \le 1 . \tag{6.62}$$

Then there exists a number $K_2(p, \gamma_0)$ such that if we define for $n \ge 0$ the numbers $A(n)$ as follows:

$$A(0) = K_2(p, \gamma_0) \mathsf{E} \exp 2S \;, \tag{6.63}$$

$$A(n+1) = A(0) + \big(40p^3(\gamma_0 + \gamma_0^3) \mathsf{E} S \exp 2S\big) A(n) \;, \tag{6.64}$$

then the following holds. If $\gamma \le \gamma_0$, given any integers $k \le N$ and n we can find i.i.d. r.v.s $z_1, \ldots, z_k$ of law μ_γ such that

$$\mathsf{E} \sum_{i \le k} |\langle \sigma_i \rangle - z_i| \le 2^{1-n} k + \frac{k^3 A(n)}{N} \;. \tag{6.65}$$

In particular when

$$80p^3(\gamma_0 + \gamma_0^3) \mathsf{E} S \exp 2S \le 1 \;, \tag{6.66}$$

we can replace (6.65) by

$$\mathsf{E} \sum_{i \le k} |\langle \sigma_i \rangle - z_i| \le \frac{2k^3 K_2(\gamma_0, p)}{N} \mathsf{E} \exp 2S \;. \tag{6.67}$$

The last statement of the Theorem simply follows from the fact that under (6.66) we have $A(n) \le 2A(0)$, so that we can take n very large in (6.90). When (6.66) need not hold, optimisation over n in (6.65) yields a bound $\le KkN^{-\alpha}$ for some $\alpha > 0$ depending only on γ_0, p and S.

The next problem need not be difficult. This issue came at the very time where the book was ready to be sent to the publisher, and it did not seem appropriate to either delay the publication or to try to make significant changes in a rush.

Research Problem 6.4.2. (level 1-) Is it true that (6.67) follows from (6.62)? More specifically, when $\gamma_0 \gg 1$, and when S is constant, does (6.67) follow from a condition of the type $K(p)\gamma_0 S \le 1$?

Probably the solution of this problem will not require essentially new ideas. Rather, it should require technical work and improvement of the estimates from Lemma 6.4.3 to Lemma 6.4.7, trying in particular to bring out more "small factors" such as $\mathsf{E} S \exp 2S$, in the spirit of Exercice 6.3.6. It seems however that it will also be necessary to proceed to a finer study of what happens on the set Ω defined page 349.

It follows from Theorem 6.2.2 that we can assume throughout the proof that property $C(\gamma_0, N, K_0, K_0)$ holds for every N. It will be useful to consider the metric space $[-1,1]^k$, provided with the distance d given by

$$d((x_i)_{i \le k}, (y_i)_{i \le k}) = \sum_{i \le k} |x_i - y_i| \;. \tag{6.68}$$

The Monge-Kantorovich transportation-cost distance on the space of probability measures on $[-1,1]^k$ that is induced by (6.68) will also be denoted by d. We define

$$D(N,k,\gamma_0)=\sup_{\gamma\le\gamma_0} d\big(\mathcal{L}(\langle\sigma_1\rangle,\ldots,\langle\sigma_k\rangle),\mu_\gamma^{\otimes k}\big) \tag{6.69}$$

where $\mathcal{L}(X_1,\ldots,X_k)$ denotes the law of the random vector $(X_1,\ldots,X_k)$.

By definition of the transportation-cost distance in the right-hand side of (6.69), the content of Theorem 6.4.1 is that if γ_0 satisfies (6.62) we have $D(N,k,\gamma_0)\le 2^{1-n}k+k^3A(n)/N$ for each $k\le N$ and each n. This inequality will be proved by obtaining a suitable induction relation between the quantities $D(N,k,\gamma_0)$. The overall idea of the proof is to use the cavity method to express $\langle\sigma_1\rangle,\ldots,\langle\sigma_k\rangle$ as functions of a smaller spin system, and to use Proposition 6.2.7 and the induction hypothesis to perform estimates on the smaller spin system.

We start by a simple observation. Since $\sum_{i\le k}|x_i-y_i|\le 2k$ for $x_i,y_i\in[-1,1]$, we have $D(N,k,\gamma_0)\le 2k$. Assuming, as we may, that $K_2(p,\gamma_0)\ge 4p$, we see that there is nothing to prove unless $N\ge 2pk^2$ so in particular $N\ge p+k$ and $N\ge 2k$. We will always assume below that this is the case. We also observe that, by symmetry,

$$\mathcal{L}(\langle\sigma_1\rangle,\ldots,\langle\sigma_k\rangle)=\mathcal{L}(\langle\sigma_{N-k+1}\rangle,\ldots,\langle\sigma_N\rangle)\ .$$

The starting point of the proof of Theorem 6.4.1 is a formula similar to (6.20), but where we remove the last k coordinates rather than the last one. Writing now $\boldsymbol{\rho}=(\sigma_1,\ldots,\sigma_{N-k})$, we consider the Hamiltonian

$$-H_{N-k}(\boldsymbol{\rho})=\sum_s\theta_s(\sigma_{i(s,1)},\ldots,\sigma_{i(s,p)})\ , \tag{6.70}$$

where the summation is restricted to those $s\le M$ for which $i(s,p)\le N-k$. This is the Hamiltonian of an $(N-k)$-spin system, except that we have replaced γ by a different value γ_-. To compute γ_- we observe that since the set $\{i(s,1),\ldots,i(s,p)\}$ is uniformly distributed among the subsets of $\{1,\ldots,N\}$ of cardinality p, the probability that $i(s,p)\le N-k$, i.e. the probability that this set is a subset of $\{1,\ldots,N-k\}$ is exactly

$$\tau=\frac{\binom{N-k}{p}}{\binom{N}{p}}\ ,$$

so that the mean number of terms of this Hamiltonian is $N\alpha\tau$, and

$$\gamma_-(N-k)=pN\alpha\tau=\gamma N\tau\ ,$$

and thus

$$\gamma_-=\gamma\frac{(N-k-1)\cdots(N-k-p+1)}{(N-1)\cdots(N-p+1)}\ . \tag{6.71}$$

In particular $\gamma_-\le\gamma_0$ whenever $\gamma\le\gamma_0$. Let us denote again by $\langle\cdot\rangle_-$ an average for the Gibbs measure with Hamiltonian (6.70). (The value of k will be clear from the context.) Given a function f on Σ_N, we then have

$$\langle f \rangle = \frac{\langle \mathrm{Av}\, f \mathcal{E} \rangle_-}{\langle \mathrm{Av}\, \mathcal{E} \rangle_-} \ , \tag{6.72}$$

where Av means average over $\sigma_{N-k+1}, \ldots, \sigma_N = \pm 1$, and where

$$\mathcal{E} = \exp \sum \theta_s(\sigma_{i(s,1)}, \ldots, \sigma_{i(s,p)}) \ ,$$

for a sum over those values of $s \le M$ for which $i(s,p) \ge N-k+1$. As before, in distribution,

$$\mathcal{E} = \exp \sum_{j \le r} \theta_j(\sigma_{i(j,1)}, \ldots, \sigma_{i(j,p)}) \ , \tag{6.73}$$

where now the sets $\{i(j,1), \ldots, i(j,p)\}$ are uniformly distributed over the subsets of $\{1, \ldots, N\}$ of cardinality p that intersect $\{N-k+1, \ldots, N\}$, and where r is a Poisson r.v. The expected value of r is the mean number of terms in the Hamiltonian $-H_N$ that are not included in the summation (6.70), so that

$$\mathsf{E}r = \alpha N \left(1 - \frac{\binom{N-k}{p}}{\binom{N}{p}} \right) = \frac{\gamma N}{p} \left(1 - \frac{(N-k) \cdots (N-k-p+1)}{N \cdots (N-p+1)} \right) \ . \tag{6.74}$$

The quantity r will keep this meaning until the end of the proof of Theorem 6.4.1, and the quantity $\mathcal{E}$ will keep the meaning of (6.73). It is good to note that, since $N \ge 2kp$, for $\ell \le p$ we have

$$\frac{N-k-\ell}{N-\ell} = 1 - \frac{k}{N-\ell} \ge 1 - \frac{2k}{N} \ .$$

Therefore

$$\frac{(N-k) \cdots (N-k-p-1)}{N \cdots (N-p+1)} \ge \left(1 - \frac{2k}{N} \right)^p \ge 1 - \frac{2kp}{N} \ , \tag{6.75}$$

and thus

$$\mathsf{E}r \le 2k\gamma \ . \tag{6.76}$$

We observe the identity

$$\mathcal{L}(\langle \sigma_{N-k+1} \rangle, \ldots, \langle \sigma_N \rangle) = \mathcal{L}\left(\frac{\langle \mathrm{Av}\, \sigma_{N-k+1} \mathcal{E} \rangle_-}{\langle \mathrm{Av}\, \mathcal{E} \rangle_-}, \ldots, \frac{\langle \mathrm{Av}\, \sigma_N \mathcal{E} \rangle_-}{\langle \mathrm{Av}\, \mathcal{E} \rangle_-} \right) \ . \tag{6.77}$$

The task is now to use the induction hypothesis to approximate the right-hand side of (6.77); this will yield the desired induction relation. There are three sources of randomness on the right-hand of (6.77). There is the randomness associated with the $(N-k)$-spin system of Hamiltonian (6.70); the randomness associated to r and the sets $\{i(j,1), \ldots, i(j,p)\}$; and the randomness associated to the functions θ_s, $s \le r$. These three sources of randomness are independent of each other.

To use the induction hypothesis, it will be desirable that for $j \le r$ the sets

$$I_j = \{i(j,1), \ldots, i(j,p-1)\} \tag{6.78}$$

are disjoint subsets of $\{1, \ldots, N-k\}$, so we first control the size of the rare event Ω where this is not the case. We have $\Omega = \Omega_1 \cup \Omega_2$, where

$$\Omega_1 = \Big\{\exists j \le r,\ i(j,p-2) \ge N-k+1\Big\}$$

$$\Omega_2 = \Big\{\exists\, j, j' \le r,\ j \ne j'\ ,\ \exists\, \ell, \ell' \le p-1\ , i(j,\ell) = i(j',\ell')\Big\}\ .$$

Proceeding as in the proof of (6.25) we easily reach the crude bound

$$\mathsf{P}(\Omega) \le \frac{4k^2}{N}(\gamma p + \gamma^2 p^2)\ . \tag{6.79}$$

We recall that, as defined page 341, given a sequence $\mathbf{x} = (x_1, \ldots, x_{N-k})$ with $|x_i| \le 1$ and a function f on Σ_{N-k}, we denote by $\langle f \rangle_{\mathbf{x}}$ the average of f with respect to the product measure $\lambda_{\mathbf{x}}$ on Σ_{N-k} such that $\int \sigma_i \,\mathrm{d}\lambda_{\mathbf{x}}(\boldsymbol{\rho}) = x_i$ for $1 \le i \le N-k$.

We now start a sequence of lemmas that aim at deducing from (6.77) the desired induction relations among the quantities $D(N,k,\gamma_0)$. There will be four steps in the proof. In the first step below, in each of the brackets in the right-hand side of (6.77) we replace the Gibbs measure $\langle\cdot\rangle_-$ by $\langle\cdot\rangle_{\mathbf{Y}}$ where $\mathbf{Y} = (\langle\sigma_1\rangle_-, \ldots, \langle\sigma_{N-k}\rangle_-)$. The basic reason why this creates only a small error is that $C(N,\gamma_0,K_0,K_0)$ holds true for each N, a property which is used as in Proposition 6.2.7.

Lemma 6.4.3. *Consider the sequence*

$$\mathbf{Y} = (\langle\sigma_1\rangle_-, \ldots, \langle\sigma_{N-k}\rangle_-)\ .$$

Set

$$u_\ell = \langle\sigma_{N-k+\ell}\rangle = \frac{\langle \mathrm{Av}\,\sigma_{N-k+\ell}\mathcal{E}\rangle_-}{\langle \mathrm{Av}\,\mathcal{E}\rangle_-}\ ;\ v_\ell = \frac{\langle \mathrm{Av}\,\sigma_{N-k+\ell}\mathcal{E}\rangle_{\mathbf{Y}}}{\langle \mathrm{Av}\,\mathcal{E}\rangle_{\mathbf{Y}}}\ .$$

Then we have

$$d(\mathcal{L}(u_1, \ldots, u_k), \mathcal{L}(v_1, \ldots, v_k)) \le \frac{k^3}{N} K(p,\gamma_0) \mathsf{E} \exp 2S\ . \tag{6.80}$$

Proof. From now on E_- denotes expectation in the randomness of the $N-k$ spin system only. When Ω does not occur, there is nothing to change to the proof of Proposition 6.2.7 to obtain that

$$\mathsf{E}_- |u_\ell - v_\ell| \le \frac{8r(p-1)^2 K_0}{N-k} \sum_{j \le r} \exp 2S_j\ ,$$

where we recall that r denotes the number of terms in the summation in (6.73), and is a Poisson r.v. which satisfies $\mathsf{E}r \le 2k\gamma$. We always have $\mathsf{E}_-|u_\ell - v_\ell| \le 2$, so that

$$\mathsf{E}_-|u_\ell - v_\ell| \le \frac{8r(p-1)^2 K_0}{N-k} \sum_{j\le r} \exp 2S_j + 2\mathbf{1}_\Omega \; . \tag{6.81}$$

Taking expectation in (6.81) then yields

$$\begin{aligned}\mathsf{E}|u_\ell - v_\ell| &\le \frac{8(p-1)^2 K_0}{N-k}\mathsf{E}\exp 2S\,\mathsf{E}r^2 + 2\mathsf{P}(\Omega)\\ &\le \frac{k^2 K(p,\gamma_0)}{N}\mathsf{E}\exp 2S \;,\end{aligned}$$

using (6.79), that $N-k \ge N/2$ and that $\mathsf{E}r^2 = \mathsf{E}r + (\mathsf{E}r)^2 \le 2\gamma k + 4\gamma^2 k^2$. Since the left-hand side of (6.80) is bounded by $\sum_{\ell\le k}\mathsf{E}|u_\ell - v_\ell|$, the result follows. □

In the second step, we replace the sequence $\mathbf{Y}$ by an appropriate i.i.d. sequence of law μ_{γ_-}. The basic reason this creates only a small error is the "induction hypothesis" i.e. the control of the quantities $D(N-k, m, \gamma_0)$.

Proposition 6.4.4. *Consider an independent sequence* $\mathbf{X} = (X_1, \dots, X_{N-k})$ *where each* X_i *has law* $\mu_- := \mu_{\gamma_-}$*. We set*

$$w_\ell = \frac{\langle \mathrm{Av}\,\sigma_{N-k+\ell}\mathcal{E}\rangle_{\mathbf{X}}}{\langle \mathrm{Av}\mathcal{E}\rangle_{\mathbf{X}}} \;, \tag{6.82}$$

and we recall the quantities v_ℓ *of the previous lemma. Then we have*

$$\begin{aligned} d(\mathcal{L}(v_1,\dots,v_k), \mathcal{L}(w_1,\dots,w_k)) &\le \frac{k^3}{N}K(p,\gamma_0) \\ &\quad + 4\mathsf{E}S\exp 2S\mathsf{E}D(N-k, r(p-1), \gamma_0) \;, \end{aligned} \tag{6.83}$$

where the last expectation is taken with respect to the Poisson r.v. r*.*

The proof will rely on the following lemma.

Lemma 6.4.5. *Assume that* Ω *does not occur. Consider* $\ell \le k$ *and*

$$\mathcal{E}_\ell = \exp \sum \theta_j(\sigma_{i(j,1)}, \dots, \sigma_{i(j,p-1)}, \sigma_{N-k+\ell}) \;, \tag{6.84}$$

where the summation is over those $j \le r$ *for which* $i(j,p) = N-k+\ell$*. Then for any sequence* $\mathbf{x}$ *we have*

$$\frac{\langle \mathrm{Av}\,\sigma_{N-k+\ell}\mathcal{E}\rangle_{\mathbf{x}}}{\langle \mathrm{Av}\,\mathcal{E}\rangle_{\mathbf{x}}} = \frac{\langle \mathrm{Av}\,\sigma_{N-k+\ell}\mathcal{E}_\ell\rangle_{\mathbf{x}}}{\langle \mathrm{Av}\,\mathcal{E}_\ell\rangle_{\mathbf{x}}} \;. \tag{6.85}$$

Consequently

$$\frac{\partial}{\partial x_i} \frac{\langle \mathrm{Av}\, \sigma_{N-k+\ell}\mathcal{E}\rangle_{\mathbf{x}}}{\langle \mathrm{Av}\, \mathcal{E}\rangle_{\mathbf{x}}} = 0 \tag{6.86}$$

unless $i \in I_j$ *for some* j *with* $i(j,p) = N-k+\ell$. *In that case we have moreover*

$$\left| \frac{\partial}{\partial x_i} \frac{\langle \mathrm{Av}\, \sigma_{N-k+\ell}\mathcal{E}\rangle_{\mathbf{x}}}{\langle \mathrm{Av}\, \mathcal{E}\rangle_{\mathbf{x}}} \right| \le 4S_j \exp 2S_j \,. \tag{6.87}$$

Proof. Define $\mathcal{E}'_\ell$ by $\mathcal{E} = \mathcal{E}_\ell \mathcal{E}'_\ell$. Since Ω does not occur, the quantities $\sigma_{N-k+\ell}\mathcal{E}_\ell$ and $\mathcal{E}_{\ell'}$ depend on disjoint sets of coordinates. Consequently

$$\mathrm{Av}\, \sigma_{N-k+\ell}\mathcal{E} = (\mathrm{Av}\, \sigma_{N-k+\ell}\mathcal{E}_\ell)(\mathrm{Av}\, \mathcal{E}'_\ell) \tag{6.88}$$

$$\mathrm{Av}\, \mathcal{E} = (\mathrm{Av}\, \mathcal{E}_\ell)(\mathrm{Av}\, \mathcal{E}'_\ell) \,. \tag{6.89}$$

In both (6.88) and (6.89) the two factors on the right depend on disjoint sets of coordinates. Since $\langle\cdot\rangle_{\mathbf{x}}$ is a product measure, we get

$$\langle \mathrm{Av}\, \sigma_{N-k+\ell}\mathcal{E}\rangle_{\mathbf{x}} = \langle \mathrm{Av}\, \sigma_{N-k+\ell}\mathcal{E}_\ell\rangle_{\mathbf{x}} \langle \mathrm{Av}\, \mathcal{E}'_\ell\rangle_{\mathbf{x}}$$

and similarly with (6.89), so that (6.85) follows, of which (6.86) is an obvious consequence. As for (6.87), it is proved exactly as in Lemma 6.3.3. □

Proof of Proposition 6.4.4. The strategy is to construct a specific realization of $\mathbf{X}$ for which the quantity $\mathsf{E}\sum_{\ell\le N-k} |v_\ell - w_\ell|$ is small. Consider the set $J = \bigcup_{j\le r} I_j$ (so that card$J \le (p-1)r$). The construction takes place given the set J. By definition of $D(N-k, r(p-1), \gamma_0)$, given J we can construct an i.i.d. sequence $(X_i)_{i\le N-k}$ distributed like μ_- that satisfies

$$\mathsf{E}_- \sum_{i\in J} |X_i - \langle\sigma_i\rangle_-| \le 2D(N-k, r(p-1), \gamma_0) \,. \tag{6.90}$$

We can moreover assume that the sequence $(\theta_j)_{j\ge1}$ is independent of the randomness generated by J and the variables X_i. The sequence $(X_i)_{i\le N-k}$ is our specific realization. It is i.i.d. distributed like μ_-.

It follows from Lemma 6.4.5 that if Ω does not occur,

$$\begin{aligned} |w_\ell - v_\ell| &= \left| \frac{\langle \mathrm{Av}\, \sigma_{N-k+\ell}\mathcal{E}\rangle_{\mathbf{X}}}{\langle \mathrm{Av}\, \mathcal{E}\rangle_{\mathbf{X}}} - \frac{\langle \mathrm{Av}\, \sigma_{N-k+\ell}\, \mathcal{E}\rangle_{\mathbf{Y}}}{\langle \mathrm{Av}\, \mathcal{E}\rangle_{\mathbf{Y}}} \right| \\ &\le \sum \Big(\sum_{i\in I_j} |X_i - \langle\sigma_i\rangle_-| \Big) 2S_j \exp 2S_j \,, \end{aligned}$$

where the first sum is over those $j \le r$ for which $i(j,p) = N-k+\ell$. By summation over $\ell \le k$, we get that when Ω does not occur,

$$\sum_{\ell\le k} |w_\ell - v_\ell| \le 2\sum_{i\in J} |X_i - \langle\sigma_i\rangle_-| S_{j(i)} \exp 2S_{j(i)} \,,$$

where $j(i)$ is the unique $j \le r$ with $i \in I_j$. Denoting by E_θ expectation in the r.v.s $(\theta_j)_{j\ge 1}$ and using independence we get

$$\mathsf{E}_\theta \sum_{\ell\le k} |w_\ell - v_\ell| \le 2\sum_{i\in J} |X_i - \langle\sigma_i\rangle_-| \mathsf{E}S\exp 2S \;.$$

Taking expectation E_- and using (6.90) implies that when Ω does not occur,

$$\mathsf{E}_\theta \mathsf{E}_- \sum_{\ell\le k} |w_\ell - v_\ell| \le 4(\mathsf{E}S\exp 2S) D(N-k, r(p-1), \gamma_0) \;,$$

i.e.

$$\mathbf{1}_{\Omega^c} \mathsf{E}_\theta \mathsf{E}_- \sum_{\ell\le k} |w_\ell - v_\ell| \le 4(\mathsf{E}S\exp 2S) D(N-k, r(p-1), \gamma_0) \;. \tag{6.91}$$

On the other hand, on Ω we have trivially

$$\mathsf{E}_\theta \mathsf{E}_- \sum_{\ell\le k} |w_\ell - v_\ell| \le 2k \;,$$

and combining with (6.91) we see that

$$\mathsf{E}_\theta \mathsf{E}_- \sum_{\ell\le k} |w_\ell - v_\ell| \le 4(\mathsf{E}S\exp 2S) D(N-k, r(p-1), \gamma_0) + 2k\mathbf{1}_\Omega \;.$$

Taking expectation and using (6.79) again yields

$$\mathsf{E} \sum_{\ell\le k} |w_\ell - v_\ell| \le \frac{k^3 K(p,\gamma_0)}{N} + 4(\mathsf{E}S\exp 2S)\mathsf{E}D(N-k, r(p-1), \gamma_0) \;,$$

and this implies (6.83). □

Now comes the key step: by definition of the operator T of (6.51) the r.v.s w_ℓ of (6.82) are nearly independent with law $T(\mu_-)$.

Proposition 6.4.6. *We have*

$$d(\mathcal{L}(w_1,\ldots,w_k), T(\mu_-)\otimes\cdots\otimes T(\mu_-)) \le \frac{k^2}{N} K(p,\gamma_0) \;. \tag{6.92}$$

Proof. Let us define, for $\ell \le k$

$$r(\ell) = \mathrm{card}\big\{j \le r;\ i(j,p-1) \le N-k,\ i(j,p) = N-k+\ell\big\} \;, \tag{6.93}$$

so that when Ω does not occur, $r(\ell)$ is the number of terms in the summation of (6.84), and moreover for different values of ℓ, the sets of indices occurring in (6.84) are disjoint. The sequence $(r(\ell))_{\ell\le k}$ is an i.i.d. sequence of Poisson r.v.s. (and their common mean will soon be calculated).

For $\ell \geq 1$ and $j \geq 1$ let us consider independent copies $\theta_{\ell,j}$ of θ and for $m \geq 1$ let us define, for $\boldsymbol{\sigma} \in \mathbb{R}^{\mathbb{N}}$,

$$\mathcal{E}_{\ell,m} = \mathcal{E}_{\ell,m}(\boldsymbol{\sigma}, \varepsilon) = \exp \sum_{1 \leq j \leq m} \theta_{\ell,j}(\sigma_{(j-1)(p-1)+1}, \ldots, \sigma_{j(p-1)}, \varepsilon) ,$$

a formula that should be compared to (6.49).

For $\ell \leq k$, let us consider sequences $\mathbf{X}_\ell = (X_{i,\ell})_{i \geq 1}$, where the r.v.s $X_{i,\ell}$ are all independent of law μ_-. Let us define $w'_\ell = w_\ell$ when Ω occurs, and otherwise

$$w'_\ell = \frac{\left\langle \mathrm{Av}\, \varepsilon \mathcal{E}_{\ell,r(\ell)} \right\rangle_{\mathbf{X}_\ell}}{\left\langle \mathrm{Av}\, \mathcal{E}_{\ell,r(\ell)} \right\rangle_{\mathbf{X}_\ell}} . \tag{6.94}$$

The basic fact is that the sequences $(w_\ell)_{\ell \leq k}$ and $(w'_\ell)_{\ell \leq k}$ have the same law. This is because they have the same law given the r.v. r and the numbers $i(j,1), \ldots, i(j,p)$ for $j \leq r$. This is obvious when Ω occurs, since then $w'_\ell = w_\ell$. When Ω does not occur we simply observe from (6.85) and the definition of w_ℓ that

$$w_\ell = \frac{\langle \mathrm{Av}\, \sigma_{N-k+\ell} \mathcal{E}_\ell \rangle_{\mathbf{X}}}{\langle \mathrm{Av}\, \mathcal{E}_\ell \rangle_{\mathbf{X}}} .$$

We then compare with (6.94), keeping in mind that there are $r(\ell)$ terms in the summation (6.84), and then using symmetry.

Therefore we have shown that

$$\mathcal{L}(w_1, \ldots, w_k) = \mathcal{L}(w'_1, \ldots, w'_k) . \tag{6.95}$$

Since the sequence $(r(\ell))_{\ell \leq k}$ is an i.i.d. sequence of Poisson r.v.s, the sequence $(w'_\ell)_{\ell \leq k}$ is i.i.d. It has almost law $T(\mu_-)$, but not exactly because the Poisson r.v.s $r(\ell)$ do not have the correct mean. This mean $\gamma' = \mathsf{E} r(\ell)$ is given by

$$\gamma' = \frac{N\gamma}{p} \frac{\binom{N-k}{p-1}}{\binom{N}{p}} = \gamma \frac{(N-k)\cdots(N-k-p+2)}{(N-1)\cdots(N-p+1)} \leq \gamma .$$

To bound the small error created by the difference between γ and γ' we proceed as in the proof of Lemma 6.3.4. We consider independent Poisson r.v.s $(r''(\ell))_{\ell \leq k}$ of mean $\gamma - \gamma'$, so that $s(\ell) = r(\ell) + r''(\ell)$ is an independent sequence of Poisson r.v.s of mean γ. Let

$$w''_\ell = \frac{\left\langle \mathrm{Av}\, \varepsilon \mathcal{E}_{\ell,s(\ell)} \right\rangle_{\mathbf{X}_\ell}}{\left\langle \mathrm{Av}\, \mathcal{E}_{\ell,s(\ell)} \right\rangle_{\mathbf{X}_\ell}} .$$

The sequence $(w''_\ell)_{\ell \leq k}$ is i.i.d. and the law of w''_ℓ is $T(\mu_-)$. Thus (6.95) implies:

$$\begin{aligned} d(\mathcal{L}(w_1, \ldots, w_k), T(\mu_-) \otimes \cdots \otimes T(\mu_-)) &= d(\mathcal{L}(w'_1, \ldots, w'_k), \mathcal{L}(w''_1, \ldots, w''_k)) \\ &\leq \sum_{\ell \leq k} \mathsf{E}|w'_\ell - w''_\ell| . \end{aligned}$$

Now, since $w''_\ell = w'_\ell$ unless Ω occurs or $s(\ell) \neq r(\ell)$, we have

$$\mathsf{E}|w'_\ell - w''_\ell| \leq 2\big(\mathsf{P}(s(\ell) \neq r(\ell)) + \mathsf{P}(\Omega)\big)$$

and

$$\mathsf{P}(s(\ell) \neq r(\ell)) = \mathsf{P}(r''(\ell) \neq 0) \leq \gamma - \gamma' \, .$$

Moreover from (6.75) we see that $\gamma - \gamma' \leq 2\gamma kp/N$. The result follows. □

The next lemma is the last step. It quantifies the fact that $T(\mu_-)$ is nearly μ.

Lemma 6.4.7. *We have*

$$d(T(\mu_-)^{\otimes k}, \mu^{\otimes k}) \leq \frac{4\gamma k^2 p}{N} \, . \tag{6.96}$$

Proof. The left-hand side is bounded by

$$kd(T(\mu_-), \mu) = kd(T(\mu_-), T(\mu)) \leq \frac{k}{2} d(\mu, \mu_-) \leq 2k(\gamma - \gamma_-) \, ,$$

using Lemma 6.3.4. The result follows since by (6.75) we have $\gamma - \gamma_- \leq 2kp\gamma/N$. □

Proof of Theorem 6.4.1. We set $B = 4\mathsf{E}S \exp 2S$. Using the triangle inequality for the transportation-cost distance and the previous estimates, we have shown that for a suitable value of $K_2(\gamma_0, p)$ we have (recalling the definition (6.63) of $A(0)$),

$$d\big(\mathcal{L}(\langle\sigma_{N-k+1}\rangle, \dots, \langle\sigma_N\rangle), \mu^{\otimes k}\big) \;\leq \frac{k^3 A(0)}{N} + B\mathsf{E}D(N-k, r(p-1), \gamma_0) \, . \tag{6.97}$$

Given an integer n we say that property $C^*(N, \gamma_0, n)$ holds if

$$\forall p \leq N' \leq N \, , \; \forall k \leq N' \, , \; D(N', k, \gamma_0) \leq 2^{1-n} k + \frac{k^3 A(n)}{N'} \, . \tag{6.98}$$

Since $D(N', k, \gamma_0) \leq 2k$, $C^*(N, \gamma_0, 0)$ holds for each N. And since $A(n) \geq A(0)$, $C^*(p, \gamma_0, n)$ holds as soon as $K_2(\gamma_0, p) \geq 2p$, since then $D(p, k, \gamma_0) \leq 2k \leq k^3 A(0)/p \leq k^3 A(n)/p$. We will prove that

$$C^*(N-1, \gamma_0, n) \Rightarrow C^*(N, \gamma_0, n+1) \, , \tag{6.99}$$

thereby proving that $C^*(N, \gamma_0, n)$ holds for each N and n, which is the content of the theorem.

To prove (6.99), we assume that $C^*(N-1, \gamma_0, n)$ holds and we consider $k \leq N/2$. It follows from (6.98) used for $N' = N - k \leq N - 1$ and $r(p-1)$ instead of k that since $k \leq N/2$ we have

$$D(N-k, r(p-1), \gamma_0) \leq 2^{1-n} rp + \frac{p^3 r^3 A(n)}{N-k} \leq 2^{1-n} rp + \frac{2p^3 r^3 A(n)}{N} \, , \tag{6.100}$$

and going back to (6.97),

$$d\big(\mathcal{L}(\langle\sigma_{N-k+1}\rangle,\dots,\langle\sigma_N\rangle),\mu^{\otimes k}\big)\le 2^{1-n}pB\mathsf{E}r+\frac{k^3A(0)}{N}+\frac{2p^3A(n)}{N}B\mathsf{E}(r^3)\;. \tag{6.101}$$

Since r is a Poisson r.v., (A.64) shows that $\mathsf{E}r^3 = (\mathsf{E}r)^3 + 3(\mathsf{E}r)^2 + \mathsf{E}r$, so that since $\mathsf{E}r \le 2k\gamma$ we have crudely

$$\mathsf{E}r^3 \le 20(\gamma+\gamma^3)k^3\;, \tag{6.102}$$

using that $\gamma^2 \le \gamma + \gamma^3$. Since $pB\mathsf{E}r = 2pBk\gamma \le k/2$ by (6.62), using (6.102) to bound the last term of (6.101) we get

$$d\big(\mathcal{L}(\langle\sigma_{N-k+1}\rangle,\dots,\langle\sigma_N\rangle),\mu^{\otimes k}\big)\le 2^{-n}k+\frac{k^3}{N}(A(0)+40p^3(\gamma+\gamma^3)BA(n))\;,$$

and since this holds for each $\gamma \le \gamma_0$, the definition of $D(N,k,\gamma_0)$ shows that

$$D(N,k,\gamma_0)\le 2^{-n}k+\frac{k^3}{N}(A(0)+40p^3(\gamma_0+\gamma_0^3)BA(n)) = 2^{-n}k+\frac{k^3A(n+1)}{N}\;. \tag{6.103}$$

We have assumed $k \le N/2$, but since $D(N,k,\gamma_0) \le 2k$ and $A(n+1) \ge A(0)$, (6.103) holds for $k \ge N/2$ provided $K_2(\gamma_0,p) \ge 8$. This proves $C^*(N,\gamma_0,n+1)$ and concludes the proof. □

We now turn to the computation of

$$p_N(\gamma)=\frac{1}{N}\mathsf{E}\log\sum_{\boldsymbol{\sigma}}\exp(-H_N(\boldsymbol{\sigma}))\;. \tag{6.104}$$

We will only consider the situation where (6.66) holds, leaving it to the reader to investigate what kind of rates of convergence she can obtain when assuming only (6.62). We consider i.i.d. copies $(\theta_j)_{j\ge1}$ of the r.v. θ, that are independent of θ, and we recall the notation (6.49). Consider an i.i.d. sequence $\mathbf{X}=(X_i)_{i\ge1}$, where X_i is of law μ_γ (given by Theorem 6.3.1). Recalling the definition (6.49) of $\mathcal{E}_r$ we define

$$p(\gamma)=\log 2-\frac{\gamma(p-1)}{p}\mathsf{E}\log\langle\exp\theta(\sigma_1,\dots,\sigma_p)\rangle_{\mathbf{X}}+\mathsf{E}\log\langle\mathrm{Av}\,\mathcal{E}_r\rangle_{\mathbf{X}}\;. \tag{6.105}$$

Here as usual Av means average over $\varepsilon=\pm1$, the notation $\langle\cdot\rangle_{\mathbf{X}}$ is as in e.g. (6.51), and r is a Poisson r.v. with $\mathsf{E}r=\gamma$.

Theorem 6.4.8. *Under (6.62) and (6.66), for $N \ge 2$, and if $\gamma \le \gamma_0$ we have*

$$|p_N(\gamma)-p(\gamma)|\le\frac{K\log N}{N}\;, \tag{6.106}$$

where K does not depend on N or γ.

As we shall see later, the factor $\log N$ above is parasitic and can be removed.

Let $\gamma_- = \gamma(N-p)/(N-1)$ as in (6.18). Theorem 6.4.8 will be a consequence of the following two lemmas, that use the notation (6.104), and where K does not depend on N or γ.

Lemma 6.4.9. *We have*

$$|Np_N(\gamma) - (N-1)p_{N-1}(\gamma_-) - \log 2 - \mathsf{E}\log\langle \mathrm{Av}\,\mathcal{E}_r\rangle_{\mathbf{X}}| \le \frac{K}{N} \,. \tag{6.107}$$

Lemma 6.4.10. *We have*

$$\Big|(N-1)p_{N-1}(\gamma) - (N-1)p_{N-1}(\gamma_-) \\ - \gamma\frac{p-1}{p}\mathsf{E}\log\langle\exp\theta(\sigma_1,\ldots,\sigma_p)\rangle_{\mathbf{X}}\Big| \le \frac{K}{N} \,. \tag{6.108}$$

Proof of Theorem 6.4.8. Combining the two previous relations we get

$$|Np_N(\gamma) - (N-1)p_{N-1}(\gamma) - p(\gamma)| \le \frac{K}{N} \,,$$

and by summation over N that

$$N|p_N(\gamma) - p(\gamma)| \le K\log N \,. \qquad \square$$

The following prepares for the proof of Lemma 6.4.10.

Lemma 6.4.11. *We have*

$$p'_N(\gamma) = \frac{1}{p}\mathsf{E}\log\langle\exp\theta(\sigma_1,\ldots,\sigma_p)\rangle \,. \tag{6.109}$$

Proof. As N is fixed, it is obvious that $p'_N(\gamma)$ exists. A pretty proof of (6.109) is as follows. Consider $\delta > 0$, i.i.d. copies $(\theta_j)_{j\ge1}$ of θ, sets $\{i(j,1),\ldots,i(j,p)\}$ that are independent uniformly distributed over the subsets of $\{1,\ldots,N\}$ of cardinality p, and define

$$-H_N^\delta(\boldsymbol{\sigma}) = \sum_{j\le u}\theta_j(\sigma_{i(j,1)},\ldots,\sigma_{i(j,p)}) \,, \tag{6.110}$$

where u is a Poisson r.v. of mean $N\delta/p$. All the sources of randomness in this formula are independent of each other and of the randomness in H_N. In distribution, $H_N(\boldsymbol{\sigma}) + H_N^\delta(\boldsymbol{\sigma})$ is the Hamiltonian of an N-spin system with parameter $\gamma+\delta$, so that

$$\frac{p_N(\gamma+\delta) - p_N(\delta)}{\delta} = \frac{1}{N\delta}\mathsf{E}\log\big\langle\exp(-H_N^\delta(\boldsymbol{\sigma}))\big\rangle \,. \tag{6.111}$$

When $u = 0$, we have $H_N^\delta \equiv 0$ so that $\log \left\langle \exp(-H_N^\delta(\boldsymbol{\sigma})) \right\rangle = 0$. For very small δ, the probability that $u = 1$ is at the first order in δ equal to $N\delta/p$. The contribution of this case to the right-hand side of (6.111) is, by symmetry among sites,

$$\frac{1}{p} \mathsf{E} \log \left\langle \exp \theta_1(\sigma_{i(1,1)}, \ldots, \sigma_{i(1,p)}) \right\rangle = \frac{1}{p} \mathsf{E} \log \left\langle \exp \theta(\sigma_1, \ldots, \sigma_p) \right\rangle \ .$$

The contribution of the case $u > 1$ is of second order in δ, so that taking the limit in (6.111) as $\delta \to 0$ yields (6.109). □

Lemma 6.4.12. *Recalling that* $\mathbf{X} = (X_i)_{i \geq 1}$ *where* X_i *are i.i.d. of law* μ_γ *we have*

$$\left| p_N'(\gamma) - \frac{1}{p} \mathsf{E} \log \left\langle \exp \theta(\sigma_1, \ldots, \sigma_p) \right\rangle_{\mathbf{X}} \right| \leq \frac{K}{N} \ . \tag{6.112}$$

Proof. From Lemma 6.4.11 we see that it suffices to prove that

$$\left| \mathsf{E} \log \left\langle \exp \theta(\sigma_1, \ldots, \sigma_p) \right\rangle - \mathsf{E} \log \left\langle \exp \theta(\sigma_1, \ldots, \sigma_p) \right\rangle_{\mathbf{X}} \right| \leq \frac{K}{N} \ . \tag{6.113}$$

Let us denote by E_0 expectation in the randomness of $\langle \cdot \rangle$ (but not in θ), and let $S = \sup |\theta|$. It follows from Theorem 6.2.2 (used as in Proposition 6.2.7) that

$$\mathsf{E}_0 \left| \langle \exp \theta(\sigma_1, \ldots, \sigma_p) \rangle - \left\langle \exp \theta(\sigma_1^1, \ldots, \sigma_p^p) \right\rangle \right| \leq \frac{K}{N} \exp S \ .$$

Here and below, the number K depends only on p and γ_0, but not on S or N. Now

$$\left\langle \exp \theta(\sigma_1^1, \ldots, \sigma_p^p) \right\rangle = \langle \exp \theta(\sigma_1, \ldots, \sigma_p) \rangle_{\mathbf{Y}} \ ,$$

where $\mathbf{Y} = (\langle \sigma_1 \rangle, \ldots, \langle \sigma_p \rangle)$. Next, since

$$\left| \frac{\partial}{\partial x_i} \langle \exp \theta(\sigma_1, \ldots, \sigma_p) \rangle_{\mathbf{x}} \right| \leq \exp S \ ,$$

considering (as provided by Theorem 6.4.1) a joint realization of the sequences $(\mathbf{X}, \mathbf{Y})$ with $\mathsf{E}_0 |X_\ell - \langle \sigma_\ell \rangle| \leq K/N$ for $\ell \leq p$, we obtain as in Section 6.3 that

$$\mathsf{E}_0 \left| \langle \exp \theta(\sigma_1, \ldots, \sigma_p) \rangle_{\mathbf{X}} - \langle \exp \theta(\sigma_1, \ldots, \sigma_p) \rangle_{\mathbf{Y}} \right| \leq \frac{K}{N} \exp S \ .$$

Combining the previous estimates yields

$$\mathsf{E}_0 \left| \langle \exp \theta(\sigma_1, \ldots, \sigma_p) \rangle - \langle \exp \theta(\sigma_1, \ldots, \sigma_p) \rangle_{\mathbf{X}} \right| \leq \frac{K}{N} \exp S \ .$$

Finally for $x, y > 0$ we have

$$|\log x - \log y| \leq \frac{|x - y|}{\min(x, y)}$$

so that

$$\mathsf{E}_0 \left|\log \langle \exp\theta(\sigma_1,\ldots,\sigma_p)\rangle - \log \langle \exp\theta(\sigma_1,\ldots,\sigma_p)\rangle_{\mathbf{X}}\right| \le \frac{K}{N}\exp 2S ,$$

and (6.113) by taking expectation in the randomness of θ. □

Proof of Lemma 6.4.10. We observe that

$$p_{N-1}(\gamma) - p_{N-1}(\gamma_-) = \int_{\gamma-}^{\gamma} p'_{N-1}(t)dt .$$

Combining with Lemma 6.4.12 and Lemma 6.3.4 implies

$$\gamma_- \le t \le \gamma \quad \Rightarrow \quad \left|p'_{N-1}(t) - \frac{1}{p}\log\langle \exp\theta(\sigma_1,\ldots,\sigma_p)\rangle_{\mathbf{X}}\right| \le \frac{K}{N}$$

and we conclude using that

$$\gamma - \gamma_- = \gamma\left(1 - \frac{N-p}{N-1}\right) = \gamma\left(\frac{p-1}{N-1}\right) . \qquad \square$$

Proof of Lemma 6.4.9. Let us denote by $\langle\cdot\rangle_-$ an average for the Gibbs measure of an $(N-1)$-spin system with Hamiltonian (6.13). We recall that we can write in distribution

$$-H_N(\boldsymbol{\sigma}) \stackrel{\mathcal{D}}{=} -H_{N-1}(\boldsymbol{\rho}) + \sum_{j\le r}\theta_j(\sigma_{i(j,1)},\ldots,\sigma_{i(j,p-1)},\sigma_N) ,$$

where $(\theta_j)_{j\ge1}$ are independent distributed like θ, where r is a Poisson r.v. of expectation γ and where the sets $\{i(j,1),\ldots,i(j,p-1)\}$ are uniformly distributed over the subsets of $\{1,\ldots,N-1\}$ of cardinality $p-1$. All these randomnesses, as well as the randomness of H_{N-1} are globally independent. Thus the identity

$$\begin{aligned}\mathsf{E}\log\sum_{\boldsymbol{\sigma}}\exp(-H_N(\boldsymbol{\sigma})) &= \mathsf{E}\log\sum_{\boldsymbol{\rho}}\exp(-H_{N-1}(\boldsymbol{\rho}))\\ &\quad + \log 2 + \mathsf{E}\log\langle \mathrm{Av}\,\mathcal{E}\rangle_- \end{aligned} \tag{6.114}$$

holds, where

$$\mathcal{E} = \mathcal{E}(\boldsymbol{\rho},\varepsilon) = \exp\sum_{j\le r}\theta_j(\sigma_{i(j,1)},\ldots,\sigma_{i(j,p-1)},\varepsilon) .$$

The term $\log 2$ occurs from the identity $a(1)+a(-1) = 2\mathrm{Av}\,a(\varepsilon)$. Moreover (6.114) implies the equality

$$Np_N(\gamma) - (N-1)p_{N-1}(\gamma_-) = \log 2 + \mathsf{E}\log\langle \mathrm{Av}\,\mathcal{E}\rangle_- .$$

Thus (6.107) boils down to the fact that

$$\left|\mathsf{E}\log\langle \mathrm{Av}\,\mathcal{E}\rangle_{-} - \mathsf{E}\log\langle \mathrm{Av}\,\mathcal{E}_r\rangle_{\mathbf{X}}\right| \le \frac{K}{N}\,. \tag{6.115}$$

The reason why the left-hand side is small should be obvious, and the arguments have already been used in the proof of Lemma 6.4.12. Indeed, it follows from Theorems 6.2.2 and 6.4.1 that if F is a function on the $(N-1)$-spin system that depends only on k spins, the law of the r.v. $\langle F\rangle_{-}$ is nearly that of $\langle F\rangle_{\mathbf{Y}}$ where Y_i are i.i.d. r.v.s of law $\mu_{-} = \mu_{\gamma_{-}}$ (which is nearly μ_γ). The work consists in showing that the bound in (6.115) is actually in K/N. Writing the full details is a bit tedious, but completely straightforward. We do not give these details, since the exact rate in (6.107) will never be used. As we shall soon see, all we need in (6.106) is a bound that goes to 0 as $N\to\infty$. □

Theorem 6.4.13. *Under (6.62) and (6.66)we have in fact*

$$|p_N(\gamma) - p(\gamma)| \le \frac{K}{N}\,. \tag{6.116}$$

Proof. It follows from (6.112) that the functions $p'_N(\gamma)$ converge uniformly over the interval $[0,\gamma_0]$. On the other hand, Theorem 6.4.8 shows that $p(\gamma) = \lim p_N(\gamma)$. Thus $p(\gamma)$ has a derivative $p'(\gamma) = \lim_{N\to\infty} p'_N(\gamma)$, so that (6.112) means that $|p'_N(\gamma) - p'(\gamma)| \le K/N$, from which (6.116) follows by integration. □

Comment. In this argument we have used (6.106) only to prove that

$$p'(\gamma) = \frac{1}{p}\mathsf{E}\log\langle \exp\theta(\sigma_1,\ldots,\sigma_p)\rangle_{\mathbf{X}}\ .$$

One would certainly wish to find a simple direct proof of this fact from the definition of (6.105). A complicated proof can be found in [56], Proposition 7.4.9.

6.5 The Franz-Leone Bound

In the previous section we showed that, under (6.62), the value of $p_N(\gamma)$ is nearly given by the value (6.105). In the present section we prove a remarkable fact. If the function θ is nice enough, one can bound $p_N(\gamma)$ by a quantity similar to (6.105) for *all* values of γ. Hopefully this bound can be considered as a first step towards the very difficult problem of understanding the present model without a high-temperature condition. It is in essence a version of Guerra's replica-symmetric bound of Theorem 1.3.7 adapted to the present setting.

We make the following assumptions on the random function θ. We assume that there exists a random function $f:\{-1,1\}\to\mathbb{R}$ such that

$$\exp\theta(\sigma_1,\ldots,\sigma_p) = a(1 + bf_1(\sigma_1)\cdots f_p(\sigma_p))\,, \tag{6.117}$$

where $f_1, \ldots, f_p$ are independent copies of f, b is a r.v. independent of $f_1, \ldots, f_p$ that satisfies the condition

$$\forall n \geq 1, \quad \mathsf{E}(-b)^n \geq 0 , \tag{6.118}$$

and a is any r.v. Of course (6.118) is equivalent to saying that $\mathsf{E}b^{2k+1} \leq 0$ for $k \geq 0$. We also assume two further conditions:

$$|bf_1(\sigma_1) \cdots f_p(\sigma_p)| \leq 1 \quad \text{a.e.}, \tag{6.119}$$

and

$$\text{either} \quad f \geq 0 \quad \text{or} \quad p \quad \text{is even.} \tag{6.120}$$

Let us consider two examples where these conditions are satisfied. First, let

$$\theta(\sigma_1, \ldots, \sigma_p) = \beta J \sigma_1 \cdots \sigma_p ,$$

where J is a symmetric r.v. Then (6.117) holds for $a = \mathrm{ch}(\beta J)$, $b = \mathrm{th}(\beta J)$, $f(\sigma) = \sigma$, (6.118) holds by symmetry and (6.120) holds when p is even.

Second, let

$$\theta(\sigma_1, \ldots, \sigma_p) = -\beta \prod_{j \leq p} \frac{(1 + \eta_j \sigma_j)}{2} ,$$

where η_i are independent random signs. This is exactly the Hamiltonian relevant to the K-sat problem (6.2). We observe that for $x \in \{0, 1\}$ we have the identity $\exp(-\beta x) = 1 + (e^{-\beta} - 1)x$. Let us set $f_j(\sigma) = (1 + \eta_j \sigma)/2 \in \{0, 1\}$. Since $\theta(\sigma_1, \ldots, \sigma_p) = -\beta x$ for $x = f_1(\sigma_1) \cdots f_p(\sigma_p) \in \{0, 1\}$ we see that (6.117) holds for $a = 1$, $b = e^{-\beta} - 1$ and $f_j(\sigma) = (1 + \eta_j \sigma)/2$; (6.118) holds since $b < 0$, and (6.120) holds since $f \geq 0$.

Given a probability measure μ on $[-1, 1]$, consider an i.i.d. sequence $\mathbf{X}$ distributed like μ, and let us denote by $p(\gamma, \mu)$ the right-hand side of (6.105). (Thus, under (6.62), μ_γ is well defined and $p(\gamma) = p(\gamma, \mu_\gamma)$).

Theorem 6.5.1. *Conditions (6.117) to (6.119) imply*

$$\forall \gamma, \forall \mu, \; p_N(\gamma) = \frac{1}{N} \mathsf{E} \log \sum_{\boldsymbol{\sigma}} \exp(-H_N(\boldsymbol{\sigma})) \leq p(\gamma, \mu) + \frac{K\gamma}{N} , \tag{6.121}$$

where K does not depend on N or γ.

Let us introduce for $\varepsilon = \pm 1$ the r.v.

$$U(\varepsilon) = \log \langle \exp \theta(\sigma_1, \ldots, \sigma_{p-1}, \varepsilon) \rangle_{\mathbf{X}} ,$$

and let us consider independent copies $(U_{i,s}(1), U_{i,s}(-1))_{i,s \geq 1}$ of the pair $(U(1), U(-1))$.

Exercise 6.5.2. As a motivation for the introduction of the quantity U prove that if we consider the 1-spin system with Hamiltonian $-\sum_{s\le r} U_{i,s}(\varepsilon)$, the average of ε for this Hamiltonian is equal, in distribution, to the quantity (6.51). (Hence, it is distributed like $T(\mu)$.)

For $0 \le t \le 1$ we consider a Poisson r.v. M_t of mean $\alpha t N = \gamma t N/p$, and independent Poisson r.v.s $r_{i,t}$ of mean $\gamma(1-t)$, independent of M_t. We consider the Hamiltonian

$$-H_{N,t}(\boldsymbol{\sigma}) = \sum_{k\le M_t} \theta_k(\sigma_{i(k,1)}, \ldots, \sigma_{i(k,p)}) + \sum_{i\le N} \sum_{s\le r_{i,t}} U_{i,s}(\sigma_i) \;, \tag{6.122}$$

where as usual the different sources of randomness are independent of each other, and we set

$$\varphi(t) = \frac{1}{N} \mathsf{E} \log \sum_{\boldsymbol{\sigma}} \exp(-H_{N,t}(\boldsymbol{\sigma})) \;.$$

Proposition 6.5.3. *We have*

$$\varphi'(t) \le -\frac{\gamma(p-1)}{p} \,\mathsf{E} \log \langle \exp \theta(\sigma_1, \ldots, \sigma_p) \rangle_{\mathbf{X}} + \frac{K\gamma}{N} \;. \tag{6.123}$$

This is of course the key fact.

Proof of Theorem 6.5.1. We deduce from (6.5.3) that

$$p_N(\gamma) = \varphi(1) \le \varphi(0) - \frac{\gamma(p-1)}{p} \,\mathsf{E} \log \langle \exp \theta(\sigma_1, \ldots, \sigma_p) \rangle_{\mathbf{X}} + \frac{K\gamma}{N} \;.$$

Therefore to prove Theorem 6.5.1 it suffices to show that $\varphi(0) = \log 2 + \mathsf{E} \log \langle \mathrm{Av}\, \mathcal{E}_r \rangle_{\mathbf{X}}$. For $t = 0$ the spins are decoupled, so this reduces to the case $N = 1$. Since $r_{1,0}$ has the same distribution as r, we simply observe that if $(\mathbf{X}_s)_{s\le r}$ are independent copies of $\mathbf{X}$, the quantity

$$\prod_{s\le r} \langle \exp \theta_s(\sigma_1, \ldots, \sigma_{p-1}, \varepsilon) \rangle_{\mathbf{X}_s}$$

has the same distribution as the quantity $\langle \mathrm{Av}\, \mathcal{E}_r \rangle_{\mathbf{X}}$. Therefore,

$$\begin{aligned} \mathsf{E} \log \sum_{\varepsilon=\pm 1} \exp \sum_{s\le r} U_{1,s}(\varepsilon) &= \mathsf{E} \log \sum_{\varepsilon=\pm 1} \prod_{s\le r} \langle \exp \theta_s(\sigma_1, \ldots, \sigma_{p-1}, \varepsilon) \rangle_{\mathbf{X}_s} \\ &= \log 2 + \mathsf{E} \log \langle \mathrm{Av}\, \mathcal{E}_r \rangle_{\mathbf{X}} \;, \end{aligned}$$

and this completes the proof of Theorem 6.5.1. □

We now prepare for the proof of (6.5.3).

Lemma 6.5.4. *We have*

$$\varphi'(t) \le \frac{\gamma}{p}\bigg(\frac{1}{N^p}\sum_{i_1,\dots,i_p=1}^{N} \mathsf{E}\log\big\langle \exp\theta(\sigma_{i_1},\dots,\sigma_{i_p})\big\rangle - \frac{p}{N}\sum_{i\le N}\mathsf{E}\log\langle\exp U(\sigma_i)\rangle\bigg) + \frac{K\gamma}{N} \,. \tag{6.124}$$

Here, as in the rest of the section, we denote by $\langle\cdot\rangle$ an average for the Gibbs measure with Hamiltonian (6.122), keeping the dependence on t implicit. On the other hand, the number K in (6.124) is of course independent of t.

Proof. In $\varphi'(t)$ there are terms coming from the dependence on t of M_t and terms coming from the dependence on t of $r_{i,t}$.

As shown by Lemma 6.4.11, the term created by the dependence of M_t on t is

$$\frac{\gamma}{p}\mathsf{E}\log\langle\exp\theta(\sigma_1,\dots,\sigma_p)\rangle \le \frac{\gamma}{pN^p}\sum_{i_1,\dots,i_p=1}^{N}\mathsf{E}\log\big\langle\exp\theta(\sigma_{i_1},\dots,\sigma_{i_p})\big\rangle + \frac{\gamma K}{N} \,,$$

because all the terms where the indices $i_1,\dots,i_p$ are distinct are equal. The same argument as in Lemma 6.4.11 shows that the term created by the dependence of $r_{i,t}$ on t is $-(\gamma/N)\mathsf{E}\log\langle\exp U(\sigma_i)\rangle$. □

Thus, we have reduced the proof of Proposition 6.5.3 (hence, of Theorem 6.5.1) to the following:

Lemma 6.5.5. *We have*

$$\sum_{i_1,\dots,i_p=1}^{N}\frac{1}{N^p}\mathsf{E}\log\big\langle\exp\theta(\sigma_{i_1},\dots,\sigma_{i_p})\big\rangle - \frac{p}{N}\sum_{i\le N}\mathsf{E}\log\langle\exp U(\sigma_i)\rangle + (p-1)\mathsf{E}\log\langle\exp\theta(\sigma_1,\dots,\sigma_p)\rangle_{\mathbf{X}} \le 0 \,. \tag{6.125}$$

The proof is not really difficult, but it must have been quite another matter when Franz and Leone discovered it.

Proof. We will get rid of the annoying logarithms by power expansion,

$$\log(1+x) = -\sum_{n\ge1}(-1)^n\frac{x^n}{n}$$

for $|x|<1$. Let us denote by E_0 the expectation in the randomness of $\mathbf{X}$ and of the functions f_j of (6.117) only. Let us define

$$C_n = \mathsf{E}_0\,\langle f(\sigma_1)\rangle_{\mathbf{X}}^n \tag{6.126}$$

$$A_{j,n} = A_{j,n}(\boldsymbol{\sigma}^1,\dots,\boldsymbol{\sigma}^n) = \frac{1}{N}\sum_{i\le N}\prod_{\ell\le n} f_j(\sigma_i^\ell) \tag{6.127}$$

$$B_n = B_n(\boldsymbol{\sigma}^1, \ldots, \boldsymbol{\sigma}^n) = \mathsf{E}_0 A_{j,n} \,. \tag{6.128}$$

We will prove that the left-hand side quantity (6.125) is equal to

$$-\sum_{n=1}^{\infty} \frac{\mathsf{E}(-b)^n}{n} \mathsf{E}\left\langle B_n^p - pB_n C_n^{p-1} + (p-1)C_n^p \right\rangle \,. \tag{6.129}$$

The function $x \mapsto x^p$ is convex on $\mathbb{R}^+$, and when p is even it is convex on $\mathbb{R}$. Therefore $x^p - pxy^{p-1} + (p-1)y^p \geq 0$ for all $x, y \in \mathbb{R}^+$, and when p is even this is true for all $x, y \in \mathbb{R}$. Now (6.120) shows that either $B_n \geq 0$ and $C_n \geq 0$ or p is even, and thus it holds that $B_n^p - pB_n C_n^{p-1} + (p-1)C_n^p \geq 0$. Consequently the right-hand side of (6.129) is ≤ 0 because $\mathsf{E}(-b)^n \geq 0$ by (6.118).

By (6.117) we have

$$\exp\theta(\sigma_1, \ldots, \sigma_p) = a(1 + b\prod_{j\leq p} f_j(\sigma_j)) \,, \tag{6.130}$$

so that, taking the average $\langle\cdot\rangle_{\mathbf{X}}$ and logarithm, and using (6.119) to allow the power expansion in the second line,

$$\begin{aligned} \log\left\langle \exp\theta(\sigma_1, \ldots, \sigma_p)\right\rangle_{\mathbf{X}} &= \log a + \log\left(1 + b\left\langle \prod_{j\leq p} f_j(\sigma_j)\right\rangle_{\mathbf{X}}\right) \\ &= \log a - \sum_{n=1}^{\infty} \frac{(-b)^n}{n} \left\langle \prod_{j\leq p} f_j(\sigma_j)\right\rangle_{\mathbf{X}}^n \,. \end{aligned} \tag{6.131}$$

Now, by independence

$$\mathsf{E}_0 \left\langle \prod_{j\leq p} f_j(\sigma_j)\right\rangle_{\mathbf{X}}^n = \mathsf{E}_0 \prod_{j\leq p} \langle f_j(\sigma_j)\rangle_{\mathbf{X}}^n = C_n^p$$

so that

$$\mathsf{E}_0 \log\left\langle \exp\theta(\sigma_1, \ldots, \sigma_p)\right\rangle_{\mathbf{X}} = \mathsf{E}_0 \log a - \sum_{n=1}^{\infty} \frac{(-b)^n}{n} C_n^p \,.$$

As in (6.131),

$$\begin{aligned} &\frac{1}{N^p} \sum_{i_1, \ldots, i_p}^{N} \log\left\langle \exp\theta(\sigma_{i_1}, \ldots, \sigma_{i_p})\right\rangle \\ &= \log a - \sum_{n=1}^{\infty} \frac{(-b)^n}{n} \left(\frac{1}{N^p} \sum_{i_1, \ldots, i_p = 1}^{N} \left\langle \prod_{j\leq p} f_j(\sigma_{i_j})\right\rangle^n\right) . \end{aligned}$$

Using replicas, we get

$$\left\langle\prod_{j\le p} f_j(\sigma_{i_j})\right\rangle^n = \left\langle\prod_{\ell\le n}\prod_{j\le p} f_j(\sigma^\ell_{i_j})\right\rangle ,$$

so that, using (6.127) in the second line yields

$$\begin{aligned}\frac{1}{N^p}\sum_{i_1,\dots,i_p=1}^N \left\langle\prod_{j\le p} f_j(\sigma_{i_j})\right\rangle^n &= \left\langle\frac{1}{N^p}\sum_{i_1,\dots,i_p=1}^N \prod_{\ell\le n}\prod_{j\le p} f_j(\sigma^\ell_{i_j})\right\rangle^n \\ &= \left\langle\prod_{j\le p} A_{j,n}\right\rangle .\end{aligned}$$

Now from (6.128) and independence we get $E_0\prod_{j\le p} A_{j,n} = B^p_n$, so that

$$\mathsf{E}_0\frac{1}{N^p}\sum_{i_1,\dots,i_p=1}^N \log\left\langle\exp\theta(\sigma_{i_1},\dots,\sigma_{i_p})\right\rangle = \mathsf{E}_0\log a - \sum_{n=1}^\infty\frac{(-b)^n}{n}\left\langle B^p_n\right\rangle .$$

In a similar manner, recalling the definition of U, one shows that

$$\mathsf{E}_0\frac{1}{N}\sum_{i\le n}\log\left\langle\exp U(\sigma_i)\right\rangle = \mathsf{E}_0\log a - \sum_{n=1}^\infty\frac{(-b)^n}{n}\left\langle B_n C^{p-1}_n\right\rangle ,$$

and this concludes the proof of Lemma 6.5.5. □

6.6 Continuous Spins

In this section we consider the situation of the Hamiltonian (6.4) when the spins are real numbers. There are two motivations for this. First, the "main parameter" of the system is no longer "a function" but rather "a random function". This is both a completely natural and fun situation. Second, this will let us demonstrate in the next section the power of the convexity tools we developed in Chapters 3 and 4. We consider a (Borel) function θ on $\mathbb{R}^p$, i.i.d. copies $(\theta_k)_{k\ge1}$ of θ, and for $\boldsymbol{\sigma}\in\mathbb{R}^N$ the quantity $H_N(\boldsymbol{\sigma})$ given by (6.4). We consider a given probability measure η on $\mathbb{R}$, and we lighten notation by writing η_N for $\eta^{\otimes N}$, the corresponding product measure on $\mathbb{R}^N$. The Gibbs measure is now defined as the random probability measure on $\mathbb{R}^N$ which has a density with respect to η_N that is proportional to $\exp(-H_N(\boldsymbol{\sigma}))$. Let us fix an integer k and, for large N, let us try to guess the law of $(\sigma_1,\dots,\sigma_k)$ under Gibbs' measure. This is a random probability measure on $\mathbb{R}^k$. We expect that it has a density $Y_{k,N}$ with respect to $\eta_k = \eta^{\otimes k}$. What is the simplest possible structure? It would be nice if we had

$$Y_{k,N}(\sigma_1,\dots,\sigma_k)\simeq X_1(\sigma_1)\cdots X_k(\sigma_k) ,$$

where $X_1, \ldots, X_k \in L^1(\eta)$ are random elements of $L^1(\eta)$, which are probability densities, i.e. $X_i \in \mathcal{D}$, where

$$\mathcal{D} = \left\{ X \in L^1(\eta) \; ; \; X \geq 0 \; ; \int X \mathrm{d}\eta = 1 \right\} . \tag{6.132}$$

The nicest possible probabilistic structure would be that these random elements $X_1, \ldots, X_k$ be i.i.d, with a common law μ, a probability measure on the metric space $\mathcal{D}$. This law μ is the central object, the "main parameter". (If we wish, we can equivalently think of μ as the law of a random element of $\mathcal{D}$.) The case of Ising spins is simply the situation where $\eta(\{1\}) = \eta(\{-1\}) = 1/2$, in which case

$$\mathcal{D} = \{(x(-1), x(1)) \; ; \; x(1), x(-1) \geq 0 \; , \; x(1) + x(-1) = 2\}$$

and

$$\begin{gathered} \mathcal{D} \text{ can be identified with the interval } [-1, 1] \\ \text{by the map } (x(-1), x(1)) \mapsto (x(1) - x(-1))/2 \; . \end{gathered} \tag{6.133}$$

Thus, in that case, as we have seen, the main parameter is a probability measure on the interval $[-1, 1]$.

We will assume in this section that θ is uniformly bounded, i.e.

$$S = \sup_{\sigma_1, \ldots, \sigma_p \in \mathbb{R}} |\theta(\sigma_1, \ldots, \sigma_p)| < \infty \tag{6.134}$$

for a certain r.v. S. Of course $(S_k)_{k \geq 1}$ denote i.i.d. copies of S with $S_k = \sup |\theta_k(\sigma_1, \ldots, \sigma_p)|$. Whether or how this boundedness condition can be weakened remains to be investigated. Overall, once one gets used to the higher level of abstraction necessary compared with the case of Ising spins, the proofs are really not more difficult in the continuous case. In the present section we will control the model under a high-temperature condition and the extension of the methods of the previous sections to this setting is really an exercise. The real point of this exercise is that in the next section, we will succeed to partly control the model *without* assuming a high-temperature condition but assuming instead the concavity of θ, a result very much in the spirit of Section 3.1.

Our first task is to construct the "order parameter" $\mu = \mu_\gamma$. We keep the notation (6.49), that is we write

$$\mathcal{E}_r = \mathcal{E}_r(\boldsymbol{\sigma}, \varepsilon) = \exp \sum_{1 \leq j \leq r} \theta_j(\sigma_{(j-1)(p-1)+1}, \ldots, \sigma_{j(p-1)}, \varepsilon) \; ,$$

where now σ_i and ε are real numbers.

Given a sequence $\mathbf{X} = (X_i)_{i \geq 1}$ of elements of $\mathcal{D}$, for a function f of $\sigma_1, \ldots, \sigma_N$, we define (and this will be fundamental)

$$\langle f\rangle_{\mathbf{X}} = \int f(\sigma_1,\ldots,\sigma_N)X_1(\sigma_1)\cdots X_N(\sigma_N)\mathrm{d}\eta(\sigma_1)\cdots\mathrm{d}\eta(\sigma_N)\ , \tag{6.135}$$

that is, we integrate the generic k-th coordinate with respect to η after making the change of density X_k.

For consistency with the notation of the previous section, for a function $h(\varepsilon)$ we write

$$\mathrm{Av}h = \int h(\varepsilon)\mathrm{d}\eta(\varepsilon)\ . \tag{6.136}$$

Thus

$$\mathrm{Av}\mathcal{E}_r = \int \mathcal{E}_r(\boldsymbol{\sigma},\varepsilon)\mathrm{d}\eta(\varepsilon)$$

is a function of $\boldsymbol{\sigma}$ only, and $\langle \mathrm{Av}\mathcal{E}_r\rangle_{\mathbf{X}}$ means that we integrate in $\sigma_1,\ldots,\sigma_N$, as in (6.135). We will also need the quantity $\langle\mathcal{E}_r\rangle_{\mathbf{X}}$, where we integrate in $\sigma_1,\ldots,\sigma_N$ as in (6.135), but we do *not* integrate this factor in ε. Thus $\langle\mathcal{E}_r\rangle_{\mathbf{X}}$ is a function of ε only, and by Fubini's theorem we have $\mathrm{Av}\langle\mathcal{E}_r\rangle_{\mathbf{X}} = \langle\mathrm{Av}\mathcal{E}_r\rangle_{\mathbf{X}}$. In particular, the function f of ε given by

$$\frac{\langle\mathcal{E}_r\rangle_{\mathbf{X}}}{\langle\mathrm{Av}\mathcal{E}_r\rangle_{\mathbf{X}}} \tag{6.137}$$

is such that $f \geq 0$ and $\mathrm{Av}f = 1$, i.e. $f \in \mathcal{D}$.

Consider a probability measure μ on $\mathcal{D}$, and $(X_i)_{i\geq 1}$ a sequence of elements of $\mathcal{D}$ that is i.i.d. of law μ. We denote by $T(\mu)$ the law (in $\mathcal{D}$) of the random element (6.137) when $\mathbf{X} = (X_i)_{i\geq 1}$. When the spins take only the values ± 1, and provided we then perform the identification (6.133), this coincides with the definition (6.51).

Theorem 6.6.1. *Assuming (6.52), i.e.* $4\gamma p(\mathsf{E}S\exp 2S) \leq 1$, *there exists a unique probability measure* μ *on* $\mathcal{D}$ *such that* $\mu = T(\mu)$.

On $\mathcal{D}$, the natural distance is induced by the L_1 norm relative to η, i.e. for $x, y \in \mathcal{D}$

$$d(x,y) = \|x-y\|_1 = \int |x(\varepsilon)-y(\varepsilon)|\mathrm{d}\eta(\varepsilon)\ . \tag{6.138}$$

The key to prove Theorem 6.6.1 is the following estimate, where we consider a pair (X,Y) of random elements of $\mathcal{D}$, and independent copies $(X_i,Y_i)_{i\geq 1}$ of this pair. Let $\mathbf{X} = (X_i)_{i\geq 1}$ and $\mathbf{Y} = (Y_i)_{i\geq 1}$.

Lemma 6.6.2. *We have*

$$\left\|\frac{\langle\mathcal{E}_r\rangle_{\mathbf{X}}}{\langle\mathrm{Av}\mathcal{E}_r\rangle_{\mathbf{X}}} - \frac{\langle\mathcal{E}_r\rangle_{\mathbf{Y}}}{\langle\mathrm{Av}\mathcal{E}_r\rangle_{\mathbf{Y}}}\right\|_1 \leq 2\sum_{j\leq r} S_j\exp 2S_j \sum_{(j-1)(p-1)<i\leq j(p-1)} \|X_i - Y_i\|_1\ . \tag{6.139}$$

Once this estimate has been obtained we proceed exactly as in the proof of Theorem 6.3.1. Namely, if μ and μ' are the laws of X and Y respectively, and since the law of the quantity (6.137) is $T(\mu)$, the expected value of the left-hand side of (6.139) is an upper bound for the transportation-cost distance $d(T(\mu), T(\mu'))$ associated to the distance d of (6.138) (by the very definition of the transportation-cost distance). Thus taking expectation in (6.139) implies that

$$d(T(\mu), T(\mu')) \le 2\gamma p(\mathsf{E} S \exp 2S)\mathsf{E}\|X - Y\|_1 \ .$$

Since this is true for any choice of X and Y with laws μ and μ' respectively, we obtained that

$$d(T(\mu), T(\mu')) \le 2\gamma p(\mathsf{E} S \exp 2S) d(\mu, \mu') \ ,$$

so that under (6.52) the map T is a contraction for the transportation-cost distance. This completes the proof of Theorem 6.6.1, modulo the fact that the set of probability measures on a complete metric space is itself a complete metric space when provided with the transportation-cost distance.

Proof of Lemma 6.6.2. It is essentially identical to the proof of (6.57), although we find it convenient to write it a bit differently "replacing Y_j by X_j one at a time". Let

$$\mathbf{X}(i) = (X_1, \ldots, X_i, Y_{i+1}, Y_{i+2} \ldots) \ .$$

To ease notation we write

$$\langle \cdot \rangle_i = \langle \cdot \rangle_{\mathbf{X}(i)} \ ,$$

so that

$$\frac{\langle \mathcal{E}_r \rangle_{\mathbf{X}}}{\langle \mathrm{Av}\mathcal{E}_r \rangle_{\mathbf{X}}} = \frac{\langle \mathcal{E}_r \rangle_{r(p-1)}}{\langle \mathrm{Av}\mathcal{E}_r \rangle_{r(p-1)}} \ ; \qquad \frac{\langle \mathcal{E}_r \rangle_{\mathbf{Y}}}{\langle \mathrm{Av}\mathcal{E}_r \rangle_{\mathbf{Y}}} = \frac{\langle \mathcal{E}_r \rangle_0}{\langle \mathrm{Av}\mathcal{E}_r \rangle_0} \ ,$$

and to prove (6.136) it suffices to show that if $(j-1)(p-1) < i \le j(p-1)$ we have

$$\left\| \frac{\langle \mathcal{E}_r \rangle_i}{\langle \mathrm{Av}\mathcal{E}_r \rangle_i} - \frac{\langle \mathcal{E}_r \rangle_{i-1}}{\langle \mathrm{Av}\mathcal{E}_r \rangle_{i-1}} \right\|_1 \le (2S_j \exp 2S_j)\|X_i - Y_i\|_1 \ . \tag{6.140}$$

We bound the left-hand side by $\mathrm{I} + \mathrm{II}$, where

$$\mathrm{I} = \left\| \frac{\langle \mathcal{E}_r \rangle_i - \langle \mathcal{E}_r \rangle_{i-1}}{\langle \mathrm{Av}\mathcal{E}_r \rangle_i} \right\|_1 \tag{6.141}$$

$$\mathrm{II} = \left\| \frac{\langle \mathcal{E}_r \rangle_{i-1}(\langle \mathrm{Av}\mathcal{E}_r \rangle_i - \langle \mathrm{Av}\mathcal{E}_r \rangle_{i-1})}{\langle \mathrm{Av}\mathcal{E}_r \rangle_i \langle \mathrm{Av}\mathcal{E}_r \rangle_{i-1}} \right\|_1 \ . \tag{6.142}$$

Now we observe that to bound both terms by $S_j \exp 2S_j \|X_i - Y_i\|_1$ it suffices to prove that

$$|\langle\mathcal{E}_r\rangle_i - \langle\mathcal{E}_r\rangle_{i-1}| \le S_j \exp 2S_j \|X_i - Y_i\|_1 \langle\mathcal{E}_r\rangle_i \ , \tag{6.143}$$

(where both sides are functions of ε). Indeed to bound the term I using (6.143) we observe that

$$\left\| \frac{\langle\mathcal{E}_r\rangle_i}{\langle \mathrm{Av}\mathcal{E}_r\rangle_i} \right\|_1 = \mathrm{Av}\frac{\langle\mathcal{E}_r\rangle_i}{\langle \mathrm{Av}\mathcal{E}_r\rangle_i} = 1 \tag{6.144}$$

and to bound the term II using (6.143) we observe that

$$\begin{aligned}|\langle \mathrm{Av}\mathcal{E}_r\rangle_i - \langle \mathrm{Av}\mathcal{E}_r\rangle_{i-1}| &\le \mathrm{Av}|\langle\mathcal{E}_r\rangle_i - \langle\mathcal{E}_r\rangle_{i-1}| \\ &\le S_j \exp 2S_j \|X_i - Y_i\|_1 \mathrm{Av}\langle\mathcal{E}_r\rangle_i\end{aligned}$$

and we use (6.144) again (for $i-1$ rather than i).

Thus it suffices to prove (6.143). For this we write $\mathcal{E}_r = \mathcal{E}'\mathcal{E}''$, where

$$\mathcal{E}' = \exp \theta_j(\sigma_{(j-1)(p-1)+1}, \ldots, \sigma_{j(p-1)}, \varepsilon) \ ,$$

and where $\mathcal{E}''$ does not depend on σ_i. Therefore

$$A := \langle\mathcal{E}''\rangle_i = \langle\mathcal{E}''\rangle_{i-1} \ .$$

Since $\mathcal{E}'$ and $\mathcal{E}''$ depend on different sets of coordinates, we have

$$\langle\mathcal{E}_r\rangle_i = \langle\mathcal{E}'\rangle_i\langle\mathcal{E}''\rangle_i = A\langle\mathcal{E}'\rangle_i \ ; \ \langle\mathcal{E}_r\rangle_{i-1} = \langle\mathcal{E}'\rangle_{i-1}\langle\mathcal{E}''\rangle_{i-1} = A\langle\mathcal{E}'\rangle_{i-1} \ .$$

Let us define $B = B(\sigma_i, \varepsilon)$ the quantity obtained by integrating $\mathcal{E}'$ in each spin σ_k, $k < i$, with respect to η, and change of density X_k and each spin σ_k, $k > i$ with respect to η with change of density Y_k. Integrating first in the σ_k for $k \ne i$ we obtain

$$\langle\mathcal{E}'\rangle_i = \int BX_i(\sigma_i)\mathrm{d}\eta(\sigma_i) \quad ; \quad \langle\mathcal{E}'\rangle_{i-1} = \int BY_i(\sigma_i)\mathrm{d}\eta(\sigma_i) \ ,$$

and therefore

$$\langle\mathcal{E}_r\rangle_i = A\int BX_i(\sigma_i)\mathrm{d}\eta(\sigma_i) \quad ; \quad \langle\mathcal{E}_r\rangle_{i-1} = A\int BY_i(\sigma_i)\mathrm{d}\eta(\sigma_i) \ . \tag{6.145}$$

Consequently,

$$\begin{aligned}\langle\mathcal{E}_r\rangle_i - \langle\mathcal{E}_r\rangle_{i-1} &= A\int B(X_i(\sigma_i) - Y_i(\sigma_i))\mathrm{d}\eta(\sigma_i) \\ &= A\int (B-1)(X_i(\sigma_i) - Y_i(\sigma_i))\mathrm{d}\eta(\sigma_i)\end{aligned} \tag{6.146}$$

because $\int X_i \mathrm{d}\eta = \int Y_i \mathrm{d}\eta = 1$. Now, since $|\theta_j| \le S_j$, Jensen's inequality shows that $|\log B| \le S_j$. Using that $|\exp x - 1| \le |x| \exp |x|$ for $x = \log B$ we obtain that $|B-1| \le S_j \exp S_j$. Therefore (6.146) implies

$$|\langle\mathcal{E}_r\rangle_i - \langle\mathcal{E}_r\rangle_{i-1}| \le AS_j \exp S_j \|X_i - Y_i\|_1 \ . \tag{6.147}$$

Finally since $\exp(-S_j) \le \mathcal{E}'$ we have $\exp(-S_j) \le B$, so that $\exp(-S_j) \le \int BX_i(\sigma_i)\mathrm{d}\eta(\sigma_i)$. The first part of (6.145) then implies that $A\exp(-S_j) \le \langle\mathcal{E}_r\rangle_i$ and combining with (6.147) this finishes the proof of (6.143) and Lemma 6.6.2. □

A suitable extension of Theorem 6.2.2 will be crucial to the study of the present model. As in the case of Theorem 6.6.1, once we have found the proper setting, the proof is not any harder than in the case of Ising spins.

Let us consider a probability space $(\mathcal{X}, \lambda)$, an integer n, and a family $(f'_\omega)_{\omega\in\mathcal{X}}$ of functions on $(\mathbb{R}^N)^n$. We assume that there exists $i \le N$ such that for each ω we have

$$f'_\omega \circ T_i = -f'_\omega \ , \tag{6.148}$$

where T_i is defined as in Section 6.2 i.e. T_i exchanges the i^{th} components σ_i^1 and σ_i^2 of the first two replicas and leaves all the other components unchanged. Consider another function $f \ge 0$ on $(\mathbb{R}^N)^n$. We assume that f and the functions f'_ω depend on k coordinates (of course what we mean here is that they depend on the same k coordinates whatever the choice of ω). We assume that for a certain number Q, we have

$$\int |f'_\omega|\mathrm{d}\lambda(\omega) \le Qf \ . \tag{6.149}$$

Theorem 6.6.3. *Under (6.10), and provided that $\gamma \le \gamma_0$, with the previous notation we have*

$$\mathsf{E}\frac{\int |\langle f'_\omega\rangle|\mathrm{d}\lambda(\omega)}{\langle f\rangle} \le \frac{K_0 kQ}{N} \ . \tag{6.150}$$

Proof. The fundamental identity (6.20):

$$\langle f\rangle = \frac{\langle \mathrm{Av}\, f\mathcal{E}\rangle_-}{\langle \mathrm{Av}\, \mathcal{E}\rangle_-}$$

remains true if we define Av as in (6.136). We then copy the proof of Theorem 6.2.2 "by replacing everywhere f' by the average of f'_ω in ω" as follows. First, we define the property $C(N, \gamma_0, B, B^*)$ by requiring that under the conditions of Theorem 6.6.3, rather than (6.9):

$$\mathsf{E}\left|\frac{\langle f'\rangle}{\langle f\rangle}\right| \le \frac{Q(kB + B^*)}{N} \ ,$$

we get instead

$$\mathsf{E}\frac{\int |\langle f'_\omega\rangle|\mathrm{d}\lambda(\omega)}{\langle f\rangle} \le \frac{Q(kB + B^*)}{N} \ .$$

Rather than (6.32) we now define

$$f'_{\omega,s} = (\mathrm{Av}\, f'_\omega \mathcal{E}) \circ \prod_{u \le s-1} U_{i_u} - (\mathrm{Av}\, f'_\omega \mathcal{E}) \circ \prod_{u \le s} U_{i_u} \, .$$

We replace (6.34) by

$$\int |f'_{\omega,s}| \mathrm{d}\lambda(\omega) \le 4QS_v \exp\Bigl(4 \sum_{u \le r} S_u\Bigr) \mathrm{Av}\, f\mathcal{E} \, ,$$

and again the left-hand side of (6.39) by

$$\mathsf{E} \frac{\int |\langle f'_\omega \rangle| \mathrm{d}\lambda(\omega)}{\langle f \rangle} \, . \qquad \square$$

We now describe the structure of the Gibbs measure, under a "high-temperature" condition.

Theorem 6.6.4. *There exists a number $K_1(p)$ such that whenever*

$$K_1(p)(\gamma_0 + \gamma_0^3) \mathsf{E} S \exp 4S \le 1 \, , \tag{6.151}$$

if $\gamma \le \gamma_0$, given any integer k, we can find i.i.d. random elements $X_1, \ldots, X_k$ in $\mathcal{D}$ of law μ such that

$$\begin{aligned} &\mathsf{E} \int |Y_{N,k}(\sigma_1, \ldots, \sigma_k) - X_1(\sigma_1) \cdots X_k(\sigma_k)| \mathrm{d}\eta(\sigma_1) \cdots \mathrm{d}\eta(\sigma_k) \\ &\le \frac{k^3 K(p, \gamma_0)}{N} \mathsf{E} \exp 2S \, , \end{aligned} \tag{6.152}$$

where $Y_{N,k}$ denotes the density with respect to $\eta_k = \eta^{\otimes k}$ of the law of $\sigma_1, \ldots, \sigma_k$ under Gibbs' measure, and μ is as in Theorem 6.6.1 and where $K(p, \gamma_0)$ depends only on p and γ_0.

It is convenient to denote by $\bigotimes_{\ell \le k} X_\ell$ the function

$$(\sigma_1, \ldots, \sigma_k) \mapsto \prod_{\ell \le k} X_\ell(\sigma_\ell) \, ,$$

so that the left-hand side of (6.152) is simply $\mathsf{E} \| Y_{N,k} - \bigotimes_{\ell \le k} X_\ell \|_1$.

Overall the principle of the proof is very similar to that of the proof of Theorem 6.4.1, but the induction hypothesis will not be based on (6.152). The starting point of the proof is the fundamental cavity formula (6.72), where Av now means that $\sigma_{N-k+1}, \ldots, \sigma_N$ are averaged independently with respect to η. When f is a function of k variables, this formula implies that

$$\begin{aligned} \langle f(\sigma_{N-k+1}, \ldots, \sigma_N) \rangle &= \frac{\langle \mathrm{Av} f(\sigma_{N-k+1}, \ldots, \sigma_N) \mathcal{E} \rangle_-}{\langle \mathrm{Av} \mathcal{E} \rangle_-} \\ &= \mathrm{Av} \Bigl(f(\sigma_{N-k+1}, \ldots, \sigma_N) \frac{\langle \mathcal{E} \rangle_-}{\langle \mathrm{Av} \mathcal{E} \rangle_-} \Bigr) \, . \end{aligned} \tag{6.153}$$

The quantity $\langle \mathcal{E} \rangle_-$ is a function of $\sigma_{N-k+1}, \ldots, \sigma_N$ only since $(\sigma_1, \ldots, \sigma_{N-k})$ is averaged for $\langle \cdot \rangle_-$, and (6.153) means that the density with respect to η_k of the law of $\sigma_{N-k+1}, \ldots, \sigma_N$ under Gibbs' measure is precisely the function

$$\frac{\langle \mathcal{E} \rangle_-}{\langle \mathrm{Av}\mathcal{E} \rangle_-} \ . \tag{6.154}$$

Before deciding how to start the proof of Theorem 6.6.4, we will first take full advantage of Theorem 6.6.3. For a function f on $\mathbb{R}^{N-k}$ we denote

$$\langle f \rangle_\bullet = \langle f(\sigma_1^1, \sigma_2^2, \ldots, \sigma_{N-k}^{N-k}) \rangle_- \ ,$$

that is, we average every coordinate in a different replica. We recall the set Ω of (6.79).

Proposition 6.6.5. *We have*

$$\mathsf{E}\mathbf{1}_{\Omega^c} \mathrm{Av} \left| \frac{\langle \mathcal{E} \rangle_-}{\langle \mathrm{Av}\mathcal{E} \rangle_-} - \frac{\langle \mathcal{E} \rangle_\bullet}{\langle \mathrm{Av}\mathcal{E} \rangle_\bullet} \right| \le \frac{k^2 K}{N} \mathsf{E} \exp 2S \ . \tag{6.155}$$

Here and in the sequel, K denotes a constant that depends on p and γ_0 only. This statement approximates the true density (6.154) by a quantity which will be much simpler to work with, since it is defined via integration for the product measure $\langle \cdot \rangle_\bullet$.

The proof of Proposition 6.6.5 greatly resembles the proof of Proposition 6.2.7. Let us state the basic principle behind this proof. It will reveal the purpose of condition (6.149), that might have remained a little bit mysterious.

Lemma 6.6.6. *For $j \le r$ consider sets $I_j \subset \{1, \ldots, N\}$ with* $\mathrm{card} I_j = p$, $\mathrm{card} I_j \cap \{N-k+1, \ldots, N\} = 1$, *and assume that*

$$j \ne j' \quad \Rightarrow \quad I_j \cap I_{j'} \subset \{N-k+1, \ldots, N\} \ ,$$

or, equivalently, that the sets $I_j \backslash \{1, \ldots, N-k\}$ for $j \le r$ are all disjoint. Consider functions $W_j(\boldsymbol{\sigma})$, depending only on the coordinates in I_j, and assume that $\sup_{\boldsymbol{\sigma}} |W_j(\boldsymbol{\sigma})| \le S_j$. *Consider*

$$\mathcal{E} = \exp \sum_{j \le r} W_j(\boldsymbol{\sigma}) \ .$$

Then we have

$$\mathsf{E}_- \mathrm{Av} \left| \frac{\langle \mathcal{E} \rangle_-}{\langle \mathrm{Av}\mathcal{E} \rangle_-} - \frac{\langle \mathcal{E} \rangle_\bullet}{\langle \mathrm{Av}\mathcal{E} \rangle_\bullet} \right| \le \frac{4K_0 k(p-1)}{N-k} \sum_{j \le r} \exp 2S_j \ . \tag{6.156}$$

Here E_- means expectation in the randomness of $\langle\cdot\rangle_-$ only.

Proof. We "decouple the spins one at a time" for $i \le N-k$, that is, we write

$$\mathcal{E}_i = \mathcal{E}(\sigma_1^1, \sigma_2^2, \ldots, \sigma_i^i, \sigma_{i+1}^1, \ldots, \sigma_N^1) ,$$

so that

$$\frac{\langle \mathcal{E}\rangle_-}{\langle \mathrm{Av}\mathcal{E}\rangle_-} = \frac{\langle \mathcal{E}_1\rangle_-}{\langle \mathrm{Av}\mathcal{E}_1\rangle_-} \; ; \qquad \frac{\langle \mathcal{E}\rangle_\bullet}{\langle \mathrm{Av}\mathcal{E}\rangle_\bullet} = \frac{\langle \mathcal{E}_{N-k+1}\rangle_-}{\langle \mathrm{Av}\mathcal{E}_{N-k+1}\rangle_-} \; .$$

We bound

$$\mathsf{E}_- \mathrm{Av} \left| \frac{\langle \mathcal{E}_i\rangle_-}{\langle \mathrm{Av}\mathcal{E}_i\rangle_-} - \frac{\langle \mathcal{E}_{i-1}\rangle_-}{\langle \mathrm{Av}\mathcal{E}_{i-1}\rangle_-} \right| \; . \tag{6.157}$$

When i belongs to no set I_j this is zero because then $\mathcal{E}_i = \mathcal{E}_{i-1}$. Suppose otherwise that $i \in I_j$ for a certain $j \le r$. The term (6.157) is bounded by $\mathrm{I} + \mathrm{II}$, where

$$\mathrm{I} = \mathsf{E}_- \mathrm{Av} \left| \frac{\langle \mathcal{E}_i - \mathcal{E}_{i-1}\rangle_-}{\langle \mathrm{Av}\mathcal{E}_i\rangle_-} \right| \; ; \qquad \mathrm{II} = \mathsf{E}_- \mathrm{Av} \left| \frac{\langle \mathcal{E}_i\rangle_- \langle \mathrm{Av}(\mathcal{E}_i - \mathcal{E}_{i-1})\rangle_-}{\langle \mathrm{Av}\mathcal{E}_i\rangle_- \langle \mathrm{Av}\mathcal{E}_{i-1}\rangle_-} \right| \; .$$

We first bound the term II. We introduce a "replicated copy" $\mathcal{E}_i'$ of $\mathcal{E}_i$ defined by

$$\mathcal{E}_i' = \mathcal{E}(\sigma_1^{N+1}, \sigma_2^{N+2}, \ldots, \sigma_i^{N+i}, \sigma_{i+1}^{N+1}, \ldots, \sigma_N^{N+1})$$

and we write

$$\langle \mathcal{E}_i\rangle_- \langle \mathrm{Av}(\mathcal{E}_i - \mathcal{E}_{i-1})\rangle_- = \langle \mathcal{E}_i' \mathrm{Av}(\mathcal{E}_i - \mathcal{E}_{i-1})\rangle_- \; .$$

Exchanging the variables σ_i^i and σ_i^1 exchanges $\mathcal{E}_i$ and $\mathcal{E}_{i-1}$ and changes the sign of the function $f' = \mathcal{E}_i' \mathrm{Av}(\mathcal{E}_i - \mathcal{E}_{i-1})$. Next we prove the inequality

$$|\mathcal{E}_i - \mathcal{E}_{i-1}| \le (2\exp 2S_j)\mathcal{E}_{i-1} \; .$$

To prove this we observe that $\mathcal{E}$ is of the form AB where A does not depend on the i^{th} coordinate and $\exp(-S_j) \le B \le \exp S_j$. Thus with obvious notation $|B_i - B_{i-1}| \le 2\exp S_j \le 2\exp 2S_j B_{i-1}$ and since A does not depend on the i^{th} coordinate we have $A_i = A_{i-1}$ and thus

$$\begin{aligned} |\mathcal{E}_i - \mathcal{E}_{i-1}| &= |A_i B_i - A_{i-1}B_{i-1}| = A_{i-1}|B_i - B_{i-1}| \\ &\le 2\exp 2S_j A_{i-1}B_{i-1} = (2\exp 2S_j)\mathcal{E}_{i-1} \; . \end{aligned}$$

Therefore

$$|\mathrm{Av}(\mathcal{E}_i - \mathcal{E}_{i-1})| \le (2\exp 2S_j)\mathrm{Av}\mathcal{E}_{i-1}$$

and

$$\mathrm{Av}|f'| \le (2\exp 2S_j)\mathrm{Av}\mathcal{E}_i' \mathrm{Av}\mathcal{E}_{i-1} \; . \tag{6.158}$$

Thinking of Av in the left-hand side as averaging over the parameter $\omega = (\sigma_i^\ell)_{N-k<i\le N, \ell \le N+1}$, we see that (6.158) is (6.149) when $A = 2\exp 2S_j$ and $f = \mathrm{Av}\mathcal{E}_i' \mathrm{Av}\mathcal{E}_{i-1}$. Applying (6.150) to the $(N-k)$-spin system we then obtain

$$\mathrm{II} \le (2\exp 2S_j)\frac{K_0 k}{N-k} .$$

Proceeding similarly we get the same bound for the term I (in a somewhat simpler manner) and this completes the proof of (6.156). □

Proof of Proposition 6.6.5. We take expected values in (6.156), and we remember as in the Ising case (i.e. when $\sigma_i = \pm 1$) that it suffices to consider the case $N \ge 2k$. □

It will be useful to introduce the following random elements $V_1, \dots, V_k$ of $\mathcal{D}$. (These depend also on N, but the dependence is kept implicit.) The function V_ℓ is the density with respect to η of the law of $\sigma_{N-k+\ell}$ under Gibbs' measure. Let us denote by Y_k^* the function (6.154) of $\sigma_{N-k+1}, \dots, \sigma_N$, which, as already noted, is the density with respect to η_k of the law of $\sigma_{N-k+1}, \dots, \sigma_N$ under Gibbs' measure. Thus V_ℓ is the ℓ^{th} -marginal of Y_k^*, that is, it is obtained by averaging Y_k^* over all σ_{N-k+j} for $j \ne \ell$ with respect to η.

Proposition 6.6.7. *We have*

$$\mathsf{E}\left\| Y_k^* - \bigotimes_{\ell\le k} V_\ell \right\|_1 \le \frac{Kk^3}{N}\mathsf{E}\exp 2S . \tag{6.159}$$

Moreover, if $\mathcal{E}_\ell$ is defined as in (6.84), then

$$\forall \ell \le k , \quad \mathsf{E}\left\| V_\ell - \frac{\langle \mathcal{E}_\ell\rangle_\bullet}{\langle \mathrm{Av}\mathcal{E}_\ell\rangle_\bullet}\right\|_1 \le \frac{Kk^2}{N}\mathsf{E}\exp 2S . \tag{6.160}$$

The L_1 norm is computed in $L^1(\eta_k)$ in (6.159) and in $L^1(\eta)$ in (6.160). The function $\bigotimes_{\ell\le k} V_\ell$ in (6.159) is of course given by

$$\left(\bigotimes_{\ell\le k} V_\ell\right)(\sigma_{N-k+1}, \dots, \sigma_N) = \prod_{1\le\ell\le k} V_\ell(\sigma_{N-k+\ell}) .$$

Proof. Consider the event Ω as in (6.79). Using the L_1-norm notation as in (6.159), (6.155) means that

$$\mathsf{E}\mathbf{1}_{\Omega^c}\left\| \frac{\langle \mathcal{E}\rangle_-}{\langle \mathrm{Av}\mathcal{E}\rangle_-} - \frac{\langle \mathcal{E}\rangle_\bullet}{\langle \mathrm{Av}\mathcal{E}\rangle_\bullet}\right\|_1 \le \frac{Kk^2}{N}\mathsf{E}\exp 2S . \tag{6.161}$$

When Ω does not occur, we have $\mathcal{E} = \prod_{\ell\le k}\mathcal{E}_\ell$, and the quantities $\mathcal{E}_\ell$ depend on different coordinates, so that

$$\langle \mathcal{E}\rangle_\bullet = \prod_{\ell\le k}\langle \mathcal{E}_\ell\rangle_\bullet .$$

Also, $\langle \mathcal{E}_\ell\rangle_\bullet$ depends on $\sigma_{N-k+\ell}$ but not on $\sigma_{N-k+\ell'}$ for $\ell \ne \ell'$ and thus

$$\mathrm{Av} \prod_{\ell \le k} \langle \mathcal{E}_\ell \rangle_\bullet = \prod_{\ell \le k} \mathrm{Av} \langle \mathcal{E}_\ell \rangle_\bullet \ .$$

Therefore

$$\frac{\langle \mathcal{E} \rangle_\bullet}{\langle \mathrm{Av}\mathcal{E} \rangle_\bullet} = \prod_{\ell \le k} U_\ell \ , \tag{6.162}$$

where

$$U_\ell = \frac{\langle \mathcal{E}_\ell \rangle_\bullet}{\langle \mathrm{Av}\mathcal{E}_\ell \rangle_\bullet} \ .$$

Let us think of U_ℓ as a function of $\sigma_{N-k+\ell}$ only, so we can write for consistency of notation $\prod_{\ell \le k} U_\ell = \bigotimes_{\ell \le k} U_\ell$. Thus (6.161) means

$$\mathsf{E} \mathbf{1}_{\Omega^c} \left\| Y_k^* - \bigotimes_{\ell \le k} U_\ell \right\|_1 \le \frac{Kk^2}{N} \mathsf{E} \exp 2S \ .$$

Now $\|Y_k^* - \bigotimes_{\ell \le k} U_\ell\|_1 \le 2$, and combining with (6.79) we get

$$\mathsf{E} \left\| Y_k^* - \bigotimes_{\ell \le k} U_\ell \right\|_1 \le \frac{Kk^2}{N} \mathsf{E} \exp 2S \ . \tag{6.163}$$

Now, we have

$$\|V_\ell - U_\ell\|_1 \le \left\| Y_k^* - \bigotimes_{\ell \le k} U_\ell \right\|_1 \ ,$$

because the right-hand side is the average over $\sigma_{N-k+1}, \ldots, \sigma_N$ of the quantity $|Y_k^* - \bigotimes_{\ell \le k} U_\ell|$, and if one averages over $\sigma_{N-k+\ell'}$ for $\ell' \ne \ell$ inside the absolute value rather than outside one gets the left-hand side. Thus (6.160) follows from (6.163). To deduce (6.159) from (6.163) it suffices to prove that

$$\left\| \bigotimes_{\ell \le k} V_\ell - \bigotimes_{\ell \le k} U_\ell \right\|_1 \le \sum_{\ell \le k} \|V_\ell - U_\ell\|_1 \ . \tag{6.164}$$

This inequality holds whenever $V_\ell, U_\ell \in \mathcal{D}$, and is obvious if "one replaces V_ℓ by U_ℓ one at a time" because

$$\begin{aligned} &\|V_1 \otimes \cdots \otimes V_\ell \otimes U_{\ell+1} \otimes \cdots \otimes U_k - V_1 \otimes \cdots \otimes V_{\ell-1} \otimes U_\ell \otimes \cdots \otimes U_k\|_1 \\ &= \|V_1 \otimes \cdots \otimes V_{\ell-1} \otimes (V_\ell - U_\ell) \otimes U_{\ell+1} \otimes \cdots \otimes U_k\|_1 = \|V_\ell - U_\ell\|_1 \end{aligned}$$

since $V_{\ell'}, U_{\ell'} \in \mathcal{D}$ for $\ell' \le k$. □

We recall that $Y_{N,k}$ denotes the density with respect to η_k of the law of $\sigma_1, \ldots, \sigma_k$ under Gibbs' measure. Let us denote by Y_ℓ the density with respect to η of the law of σ_ℓ under Gibbs' measure. We observe that $Y_{N,k}$ corresponds to Y_k^* if we use the coordinates $\sigma_1, \ldots, \sigma_k$ rather than $\sigma_{N-k+1}, \ldots, \sigma_N$, and similarly $Y_1, \ldots, Y_k$ correspond to $V_1, \ldots, V_k$. Thus (6.159) implies

$$\mathsf{E}\|Y_{N,k} - \bigotimes Y_\ell\|_1 \le \frac{Kk^3}{N}\mathsf{E}\exp 2S\ .$$

Using as in (6.164) that

$$\left\|\bigotimes_{\ell\le k} Y_\ell - \bigotimes_{\ell\le k} X_\ell\right\|_1 \le \sum_{\ell\le k}\|Y_\ell - X_\ell\|_1\ ,$$

then (6.159) shows that to prove Theorem 6.6.4, the following estimates suffices.

Theorem 6.6.8. *Assuming (6.151), if $\gamma \le \gamma_0$, given any integer k, we can find i.i.d. random elements $X_1, \ldots, X_k$ in $\mathcal{D}$ with law μ such that*

$$\mathsf{E}\sum_{\ell\le k}\|Y_\ell - X_\ell\|_1 \le \frac{k^3 K(p,\gamma_0)}{N}\mathsf{E}\exp 2S\ . \tag{6.165}$$

We will prove that statement by induction on N. Denoting by $D(N,\gamma_0,k)$ the quantity

$$\sup_{\gamma\le\gamma_0}\ \inf_{X_1,\ldots,X_k}\ \mathsf{E}\sum_{\ell\le k}\|Y_\ell - X_\ell\|_1\ ,$$

one wishes to prove that

$$D(N,\gamma_0,k) \le \frac{k^3 K}{N}\mathsf{E}\exp 2S\ .$$

For this we relate the N-spin system with the $(N-k)$-spin system. For this purpose, the crucial equation is (6.162). The sequence $V_1, \ldots, V_k$ is distributed as $(Y_1, \ldots, Y_k)$. Moreover, if for $i \le N-k$ we denote by Y_i^- the density with respect to η of the law of σ_i under the Gibbs measure of the $(N-k)$-spin system, we have, recalling the notation (6.135)

$$\langle\mathcal{E}_\ell\rangle_\bullet = \langle\mathcal{E}_\ell\rangle_{\mathbf{Y}}\ ,$$

where $\mathbf{Y} = (Y_1^-, \ldots, Y_{N-k}^-)$, so that (6.160) implies

$$\sum_{\ell\le k}\mathsf{E}\mathbf{1}_{\Omega^c}\left\|V_\ell - \frac{\langle\mathcal{E}_\ell\rangle_{\mathbf{Y}}}{\langle \mathrm{Av}\mathcal{E}_\ell\rangle_{\mathbf{Y}}}\right\|_1 \le \frac{Kk^3}{N}\mathsf{E}\exp 2S\ . \tag{6.166}$$

We can then complete the proof of Theorem 6.6.8 along the same lines as in Theorem 6.4.1. The functions $(\mathcal{E}_\ell)_{\ell\le k}$ do not depend on too many spins. We can use the induction hypothesis and Lemma 6.6.2 to show that we can find a sequence $\mathbf{X} = (X_1, \ldots, X_{N-k+1})$ of identically distributed random elements of $\mathcal{D}$, of law μ_- ($= \mu_{\gamma_-}$, where γ_- is given by (6.74)), so that

$$\mathsf{E}\sum_{\ell\le k}\mathbf{1}_{\Omega^c}\left\|V_\ell-\frac{\langle\mathcal{E}_\ell\rangle_{\mathbf{X}}}{\langle\mathrm{Av}\mathcal{E}_\ell\rangle_{\mathbf{X}}}\right\|_1$$

is not too large. Then the sequence $(\langle\mathcal{E}_\ell\rangle_{\mathbf{X}}/\langle\mathrm{Av}\mathcal{E}_\ell\rangle_{\mathbf{X}})_{\ell\le k}$ is nearly i.i.d. with law $T(\mu_-)$, and hence nearly i.i.d. with law μ. Completing the argument really amounts to copy the proof of Theorem 6.4.1, so this is best left as an easy exercise for the motivated reader. There is nothing else to change either to the proof of Theorem 6.4.13.

We end this section by a challenging technical question. The relevance of this question might not yet be obvious to the reader, but it will become clearer in Chapter 8, after we learn how to approach the "spherical model" through the "Gaussian model". Let us consider the sphere

$$\mathbb{S}_N=\{\boldsymbol{\sigma}\in\mathbb{R}^N\;;\;\|\boldsymbol{\sigma}\|=\sqrt{N}\}\tag{6.167}$$

and the uniform probability λ_N on $\mathbb{S}_N$.

Research Problem 6.6.9. Assume that the random function θ is Borel measurable, but not necessarily continuous. Investigate the regularity properties of the function

$$t\mapsto\psi(t)=\frac{1}{N}\mathsf{E}\log\int\exp\sum_{k\le M}\theta_k(t\sigma_{i(k,1)},\ldots,t\sigma_{i(k,p)})\mathrm{d}\lambda_N(\boldsymbol{\sigma})\;.$$

In particular, if M is a proportion on N, $M=\alpha N$, is it true that for large N the difference $\psi(t)-\psi(1)$ becomes small whenever $|t-1|\le 1/\sqrt{N}$?

The situation here is that, even though each of the individual functions $t\mapsto\theta(t\sigma_{i(k,1)},\ldots,t\sigma_{i(k,p)})$ can be wildly discontinuous, these discontinuities should be smoothed out by the integration for λ_N. Even the case θ is not random and $p=1$ does not seem obvious.

6.7 The Power of Convexity

Consider a random convex set V of $\mathbb{R}^p$, and $(V_k)_{k\ge1}$ an i.i.d. sequence of random convex sets distributed like V. Consider random integers $i(k,1)<\ldots<i(k,p)$ such that the sets $I_k=\{i(k,1),\ldots,i(k,p)\}$ are i.i.d. uniformly distributed over the subsets of $\{1,\ldots,N\}$ of cardinality p. Consider the i.i.d. sequence of random convex subsets U_k of $\mathbb{R}^N$ given by

$$\boldsymbol{\sigma}\in U_k\quad\Leftrightarrow\quad(\sigma_{i(k,1)},\ldots,\sigma_{i(k,p)})\in V_k\;.$$

We recall that λ_N is the uniform probability measure on the sphere $\mathbb{S}_N$, and that M is a Poisson r.v. of expectation αN.

Research Problem 6.7.1. (Level 3) Prove that, given p, V and α, there is a number a^* such that for N large

$$\frac{1}{N} \log \lambda_N \left(\mathbb{S}_N \cap \bigcap_{k \le M} U_k \right) \simeq a^* \tag{6.168}$$

with overwhelming probability, and compute a^*.

The value $a^* = -\infty$ is permitted; in that case we expect that given any number $a > 0$, for N large we have $\lambda_N(\mathbb{S}_N \cap \bigcap_{k \le M} U_k) \le \exp(-aN)$ with overwhelming probability. Problem 6.7.1 makes sense even if the random set V is not convex, but we fear that this case could be considerably more difficult.

Consider a number $\kappa > 0$, and the probability measure η $(= \eta_\kappa)$ on $\mathbb{R}$ of density $\sqrt{\kappa/\pi} \exp(-\kappa x^2)$ with respect to Lebesgue measure. After reading Chapter 8, the reader will be convinced that a good idea to approach Problem 6.7.1 is to first study the following, which in any case is every bit as natural and appealing as Problem 6.7.1.

Research Problem 6.7.2. (Level 3) Prove that, given p, V, α and κ there is a number a^* such for large N we have

$$\frac{1}{N} \log \eta^{\otimes N} \left(\bigcap_{k \le M} U_k \right) \simeq a^* \tag{6.169}$$

with overwhelming probability, and compute a^*.

Here again, the value $a^* = -\infty$ is permitted.

Consider a random concave function $\theta \le 0$ on $\mathbb{R}^p$ and assume that

$$V = \{\theta = 0\} \ .$$

Then, denoting by $\theta_1, \ldots, \theta_M$ i.i.d. copies of θ, we have

$$\eta^{\otimes N} \left(\bigcap_{k \le M} U_k \right) = \lim_{\beta \to \infty} \int \exp\left(\beta \sum_{k \le M} \theta_k(\sigma_{i(k,1)}, \ldots, \sigma_{i(k,p)}) \right) \mathrm{d}\eta^{\otimes N}(\boldsymbol{\sigma}) \ . \tag{6.170}$$

Therefore, to prove (6.169) it should be relevant to consider Hamiltonians of the type

$$- H_N(\boldsymbol{\sigma}) = \sum_{k \le M} \theta_k(\sigma_{i(k,1)}, \ldots, \sigma_{i(k,p)}) \ , \tag{6.171}$$

where $\theta_1, \ldots, \theta_k$ are i.i.d. copies of a random concave function $\theta \le 0$. These Hamiltonians never satisfy a condition $\sup_{\sigma_1, \ldots, \sigma_p \in \mathbb{R}} |\theta(\sigma_1, \ldots, \sigma_p)| < \infty$ such as (6.134) unless $\theta \equiv 0$, and we cannot use the results of the previous sections. The purpose of the present section is to show that certain methods we have already used in Chapter 4 allow a significant step in the study of the Hamiltonians (6.171). In particular we will "prove in the limit the fundamental self-consistency equation $\mu = T(\mu)$". We remind the reader that we assume

$$\theta \text{ is concave, } \theta \le 0 \,. \tag{6.172}$$

We will also assume that there exists a non random number A (possibly very large) such that θ satisfies the following Lipschitz condition:

$$\forall \sigma_1, \ldots, \sigma_p, \sigma'_1, \ldots, \sigma'_p \,, \qquad |\theta(\sigma_1, \ldots, \sigma_p) - \theta(\sigma'_1, \ldots, \sigma'_p)| \le A \sum_{j \le p} |\sigma_j - \sigma'_j| \,. \tag{6.173}$$

The Gibbs measure is defined as usual as the probability measure on $\mathbb{R}^N$ with density with respect to $\eta^{\otimes N}$ that is proportional to $\exp(-H_N(\boldsymbol{\sigma}))$, and $\langle \cdot \rangle$ denotes an average for this Gibbs measure.

Lemma 6.7.3. *There exists a number K (depending on p, A, α and κ) such that we have*

$$\mathsf{E} \left\langle \exp \frac{|\sigma_1|}{K} \right\rangle \le K \,. \tag{6.174}$$

Of course it would be nice if we could improve (6.174) into $\mathsf{E}\langle \exp(\sigma_1^2/K) \rangle \le K$.

Lemma 6.7.4. *The density Y with respect to η of the law of σ_1 under Gibbs' measure satisfies*

$$\forall x, y \in \mathbb{R} \,, \qquad Y(y) \le Y(x) \exp rA|y - x| \tag{6.175}$$

where $r = \mathrm{card}\{k \le M \,;\, i(k,1) = 1\}$.

This lemma is purely deterministic, and is true for any realization of the disorder. It is good however to observe right away that r is a Poisson r.v. with $\mathsf{E}r = \gamma$, where as usual $\gamma = \alpha p$ and $\mathsf{E}M = \alpha N$.

Proof. Since the density of Gibbs' measure with respect to $\eta^{\otimes N}$ is proportional to $\exp(-H_N(\boldsymbol{\sigma}))$, the function $Y(\sigma_1)$ is proportional to

$$f(\sigma_1) = \int \exp(-H_N(\boldsymbol{\sigma})) \mathrm{d}\eta(\sigma_2) \cdots \mathrm{d}\eta(\sigma_N) \,.$$

We observe now that the Hamiltonian H_N depends on σ_1 only through the terms $\theta_k(\sigma_{i(k,1)}, \ldots, \sigma_{i(k,p)})$ for which $i(k,1) = 1$ so (6.173) implies that $f(\sigma'_1) \le f(\sigma_1) \exp rA|\sigma'_1 - \sigma_1|$ and this in turn implies (6.175). □

Proof of Lemma 6.7.3. We use (6.175) to obtain

$$Y(0) \exp(-rA|x|) \le Y(x) \le Y(0) \exp rA|x| \,. \tag{6.176}$$

Thus, using Jensen's inequality:

$$
\begin{aligned}
1 = \int Y \mathrm{d}\eta \geq Y(0) \int \exp(-rA|x|)\mathrm{d}\eta(x) &\geq Y(0) \exp\left(-rA \int |x| \mathrm{d}\eta(x)\right) \\
&\geq Y(0) \exp\left(-\frac{LrA}{\sqrt{\kappa}}\right) \\
&\geq Y(0) \exp(-rK) \ ,
\end{aligned}
$$

where, throughout the proof K denotes a number depending on A, κ and p only, that may vary from time to time. Also,

$$
\begin{aligned}
\left\langle \exp \frac{\kappa}{2} \sigma_1^2 \right\rangle &= \int \exp \frac{\kappa}{2} x^2 Y(x) \mathrm{d}\eta(x) \\
&\leq Y(0) \int \exp \frac{\kappa x^2}{2} \exp rA|x| \mathrm{d}\eta(x) \\
&= Y(0)\sqrt{\kappa\pi} \int \exp\left(-\frac{\kappa x^2}{2}\right) \exp rA|x| \mathrm{d}x \\
&\leq KY(0) \exp Kr^2
\end{aligned}
$$

by a standard computation, or simply using that $-\kappa x^2/2 + rA|x| \leq -\kappa x^2/4 + Kx^2$. Combining with (6.176) yields

$$
\left\langle \exp \frac{\kappa}{2} \sigma_1^2 \right\rangle \leq K \exp Kr^2 \tag{6.177}
$$

so that Markov's inequality implies

$$
\langle \mathbf{1}_{\{|\sigma_1| \geq y\}} \rangle \leq K \exp\left(Kr^2 - \frac{\kappa y^2}{2}\right) \ .
$$

Using this for $y = K'x$, we obtain

$$
r \leq x \quad \Rightarrow \quad \langle \mathbf{1}_{\{|\sigma_1| \geq Kx\}} \rangle \leq K \exp(-x^2) \ .
$$

Now, since r is a Poisson r.v. with $\mathsf{E}r = \alpha p$ we have $\mathsf{E} \exp r \leq K$, and thus

$$
\mathsf{E}\langle \mathbf{1}_{\{|\sigma_1| \geq Kx\}} \rangle \leq K \exp(-x^2) + \mathsf{P}(r > x) \leq K \exp(-x) \ ,
$$

from which (6.174) follows. □

The essential fact, to which we turn now, is a considerable generalization of the statement of Theorem 3.1.11 that "the overlap is essentially constant". Throughout the rest of the section, we also assume the following mild condition:

$$
\mathsf{E}\theta^2(0, \ldots, 0) < \infty \ . \tag{6.178}
$$

Proposition 6.7.5. *Consider functions* $f_1, \ldots, f_n$ *on* $\mathbb{R}$*, and assume that for a certain number* D *we have*

$$|f_k^{(\ell)}(x)| \leq D \tag{6.179}$$

for $\ell = 0, 1, 2$ *and* $k \leq n$*. Then the function*

$$R = R(\boldsymbol{\sigma}^1, \ldots, \boldsymbol{\sigma}^n) = \frac{1}{N} \sum_{i \leq N} f_1(\sigma_i^1) \cdots f_n(\sigma_i^n) \tag{6.180}$$

satisfies

$$\mathsf{E}\langle (R - \mathsf{E}\langle R \rangle)^2 \rangle \leq \frac{K}{\sqrt{N}} \,, \tag{6.181}$$

where K *depends only on* κ, n*,* D *and on the quantity (6.178).*

The power of this statement might not be intuitive, but soon we will show that it has remarkable consequences. Throughout the proof, K denotes a number depending only on κ, n, A, D and on the quantity (6.178).

Lemma 6.7.6. *The conditions of Proposition 6.7.5 imply:*

$$\langle (R - \langle R \rangle)^2 \rangle \leq \frac{K}{\sqrt{N}} \,. \tag{6.182}$$

Proof. The Gibbs' measure on $\mathbb{R}^{Nn}$ has a density proportional to

$$\exp\Big(- \sum_{\ell \leq n} H_N(\boldsymbol{\sigma}^\ell) - \kappa \sum_{\ell \leq n} \|\boldsymbol{\sigma}^\ell\|^2 \Big)$$

with respect to Lebesgue's measure. It is straightforward that the gradient of R at every point has a norm $\leq K/\sqrt{N}$, so that

$$R \text{ has a Lipschitz constant } \leq \frac{K}{N} \,. \tag{6.183}$$

Consequently (6.182) follows from (3.17) used for $k = 1$. □

To complete the proof of Proposition 6.7.5 it suffices to show the following.

Lemma 6.7.7. *We have*

$$\mathsf{E}(\langle R \rangle - \mathsf{E}\langle R \rangle)^2 \leq \frac{K}{\sqrt{N}} \,. \tag{6.184}$$

Proof. This proof mimics the Bovier-Gayrard argument of Section 4.5. Writing $\eta_N = \eta^{\otimes N}$, we consider the random convex function

$$\varphi(\lambda) = \frac{1}{N} \log \int \exp\Big(- \sum_{\ell \leq n} H_N(\boldsymbol{\sigma}^\ell) - \kappa \sum_{\ell \leq n} \|\boldsymbol{\sigma}^\ell\|^2 + \lambda N R \Big) \mathrm{d}\boldsymbol{\sigma}^1 \cdots \mathrm{d}\boldsymbol{\sigma}^n \,,$$

so that

$$\varphi'(0) = \langle R \rangle \; .$$

We will deduce (6.184) from Lemma 4.5.2 used for $k = 1$ and $\delta = 0$, $\lambda_0 = 1/K$, $C_0 = K$, $C_1 = K$, $C_2 = K/N$, and much of the work consists in checking conditions (4.135) to (4.138) of this lemma. Denoting by $\langle \cdot \rangle_\lambda$ an average for the Gibbs' measure with density with respect to Lebesgue's measure proportional to

$$\exp\biggl(-\sum_{\ell \le n} H_N(\boldsymbol{\sigma}^\ell) - \kappa \sum_{\ell \le n} \|\boldsymbol{\sigma}^\ell\|^2 + \lambda N R\biggr) \; , \tag{6.185}$$

we have $\varphi'(\lambda) = \langle R \rangle_\lambda$, so $|\varphi'(\lambda)| \le K$ and (4.135) holds for $C_0 = K$. We now prove the key fact that for $\lambda \le \lambda_0 = 1/K$, the function

$$-\sum_{\ell \le n} H_N(\boldsymbol{\sigma}^\ell) - \frac{\kappa}{2} \sum_{\ell \le n} \|\boldsymbol{\sigma}^\ell\|^2 + \lambda N R \tag{6.186}$$

is concave. We observe that (6.179) implies

$$\left| \frac{\partial^2 R}{\partial \sigma_i^\ell \partial \sigma_j^{\ell'}} \right| \le \frac{K}{N} \; ,$$

and that the left-hand side is zero unless $i = j$. This implies in turn that at every point the second differential D of R satisfies $|D(\mathbf{x}, \mathbf{y})| \le K \|\mathbf{x}\| \|\mathbf{y}\| / N$ for every $\mathbf{x}, \mathbf{y}$ in $\mathbb{R}^{Nn}$. On the other hand, the second differential $D^\sim$ of the function $-\kappa \sum_{\ell \le n} \|\boldsymbol{\sigma}^\ell\|^2/2$ satisfies at every point $D^\sim(\mathbf{x}, \mathbf{x}) = -\kappa \|\mathbf{x}\|^2$ for every $\mathbf{x}$ in $\mathbb{R}^{Nn}$. Therefore if $K\lambda \le \kappa$, at every point the second differential D^* of the function (6.186) satisfies $D^*(\mathbf{x}, \mathbf{x}) \le 0$ for every $\mathbf{x}$ in $\mathbb{R}^{Nn}$, and consequently this function is concave. Then the quantity (6.185) is of the type

$$\exp\biggl(U - \frac{\kappa}{2} \sum_{\ell \le n} \|\boldsymbol{\sigma}^\ell\|^2\biggr)$$

where U is concave; we can then use (6.183) and (3.17) to conclude that

$$\varphi''(\lambda) = N \langle (R - \langle R \rangle_\lambda)^2 \rangle_\lambda \le K \; ,$$

and this proves (4.137) with $\delta = 0$ and hence also (4.136). It remains to prove (4.138). For $j \le N$ let us define

$$-H'_j = \sum_{k \le M, i(k,p) = j} \theta_k(\sigma_{i(k,1)}, \ldots, \sigma_{i(k,p)}) \; .$$

The r.v.s H'_j are independent, as is made obvious by the representation of H_N given in Exercise 6.2.3. For $m \le N$ we denote by Ξ_m the σ-algebra generated by the r.v.s H'_j for $j \le m$, and we denote by E_m the conditional expectation given Ξ_m, so that we have the identity

$$\mathsf{E}(\varphi(\lambda) - \mathsf{E}\varphi(\lambda))^2 = \sum_{0 \le m < N} \mathsf{E}(\mathsf{E}_{m+1}\varphi(\lambda) - \mathsf{E}_m\varphi(\lambda))^2 .$$

To prove (4.138), it suffices to prove that for any given value of m we have

$$\mathsf{E}(\mathsf{E}_{m+1}\varphi(\lambda) - \mathsf{E}_m\varphi(\lambda))^2 \le \frac{K}{N^2} .$$

Consider the Hamiltonian

$$-H^{\sim}(\boldsymbol{\sigma}) = -\sum_{j \ne m+1} H'_j \tag{6.187}$$

and

$$\varphi^{\sim}(\lambda) = \frac{1}{N} \log \int \exp\Big(\sum_{\ell \le n} H^{\sim}(\boldsymbol{\sigma}^\ell) - \kappa \sum_{\ell \le n} \|\boldsymbol{\sigma}^\ell\|^2 + \lambda N R\Big) \mathrm{d}\boldsymbol{\sigma}^1 \cdots \mathrm{d}\boldsymbol{\sigma}^n .$$

It should be obvious that (since we have omitted the term H'_{m+1} in (6.187))

$$\mathsf{E}_m\varphi^{\sim}(\lambda) = \mathsf{E}_{m+1}\varphi^{\sim}(\lambda) ,$$

so that

$$\begin{aligned}
\mathsf{E}\big(\mathsf{E}_{m+1}\varphi(\lambda) - \mathsf{E}_m\varphi(\lambda)\big)^2 &= \mathsf{E}\big(\mathsf{E}_{m+1}(\varphi(\lambda) - \varphi^{\sim}(\lambda)) - \mathsf{E}_m(\varphi(\lambda) - \varphi^{\sim}(\lambda))\big)^2 \\
&\le 2\mathsf{E}\big(\mathsf{E}_{m+1}(\varphi(\lambda) - \varphi^{\sim}(\lambda))\big)^2 \\
&\quad + 2\mathsf{E}\big(\mathsf{E}_m(\varphi(\lambda) - \varphi^{\sim}(\lambda))\big)^2 \\
&\le 4\mathsf{E}(\varphi(\lambda) - \varphi^{\sim}(\lambda))^2 .
\end{aligned}$$

Therefore, it suffices to prove that

$$\mathsf{E}(\varphi(\lambda) - \varphi^{\sim}(\lambda))^2 \le \frac{K}{N^2} . \tag{6.188}$$

Thinking of λ as fixed, let us denote by $\langle \cdot \rangle_\sim$ an average on $\mathbb{R}^{Nn}$ with respect to the probability measure on $\mathbb{R}^{Nn}$ of density proportional to

$$\exp\Big(-\sum_{\ell \le n} H^{\sim}(\boldsymbol{\sigma}^\ell) - \kappa \sum_{\ell \le n} \|\boldsymbol{\sigma}^\ell\|^2 + \lambda N R\Big) .$$

We observe the identity

$$\varphi(\lambda) - \varphi^{\sim}(\lambda) = \frac{1}{N} \log \Big\langle \exp\Big(-\sum_{\ell \le n} (H_N(\boldsymbol{\sigma}^\ell) - H^{\sim}(\boldsymbol{\sigma}^\ell))\Big)\Big\rangle_\sim .$$

Now $H_N = H^\sim + H'_{m+1}$ and therefore

$$\varphi(\lambda) - \varphi^\sim(\lambda) = \frac{1}{N} \log \left\langle \exp\Big(-\sum_{\ell \le n} H'_{m+1}(\boldsymbol{\sigma}^\ell)\Big) \right\rangle_\sim . \tag{6.189}$$

Since $-H'_{m+1} \le 0$ we have $\varphi(\lambda) - \varphi^\sim(\lambda) \le 0$. Let us define

$$r = \operatorname{card}\{k \le M \; ; \; i(k,p) = m+1\} ,$$

the number of terms in H'_{m+1}, so that r is a Poisson r.v. with

$$\mathsf{E} r = \alpha N \frac{\binom{m}{p-1}}{\binom{N}{p}} \le \alpha p .$$

From (6.189) and Jensen's inequality it follows that

$$0 \ge \varphi(\lambda) - \varphi^\sim(\lambda) \ge \frac{1}{N} \left\langle -\sum_{\ell \le n} H'_{m+1}(\boldsymbol{\sigma}^\ell) \right\rangle_\sim , \tag{6.190}$$

and thus

$$(\varphi(\lambda) - \varphi^\sim(\lambda))^2 \le \frac{1}{N^2} \left\langle -\sum_{\ell \le n} H'_{m+1}(\boldsymbol{\sigma}^\ell) \right\rangle_\sim^2 \le \frac{1}{N^2} \left\langle \Big(\sum_{\ell \le n} H'_{m+1}(\boldsymbol{\sigma}^\ell)\Big)^2 \right\rangle_\sim .$$

Therefore it suffices to prove that for $\ell \le n$ we have

$$\mathsf{E}\langle H'_{m+1}(\boldsymbol{\sigma}^\ell)^2 \rangle_\sim \le K . \tag{6.191}$$

Writing $a_k = |\theta_k(0,\ldots,0)|$ and using (6.173) we obtain

$$|\theta_k(\sigma^\ell_{i_1}, \ldots, \sigma^\ell_{i_k})| \le a_k + A \sum_{s \le p} |\sigma^\ell_{i_s}| , \tag{6.192}$$

and therefore

$$|H'_{m+1}(\boldsymbol{\sigma}^\ell)| \le \sum_{k \in I} a_k + A \sum_{i \le N} n_i |\sigma^\ell_i| ,$$

where $n_i \in \mathbb{N}$ and $\sum n_i = rp$, because each of the r terms in H'_{m+1} creates at most p terms in the right-hand side. The randomness of H'_{m+1} is independent of the randomness of $\langle \cdot \rangle_\sim$, and since $\mathsf{E} r^2 \le K$ and $\mathsf{E} a_k^2 < \infty$, by (6.178) it suffices to prove that if $i \le N$ then $\mathsf{E}\langle (\sigma^\ell_i)^2 \rangle_\sim \le K$. This is done by basically copying the proof of Lemma 6.7.3. Using (6.183) the density Y with respect to η of the law of σ^ℓ_i under Gibbs' measure satisfies

$$\forall x, y \in \mathbb{R} \; , \; Y(x) \le Y(y) \exp\big((r_i A + K_0/N)|x-y| \big) ,$$

where $r_i = \operatorname{card}\{k \le M \; ; \; \exists s \le p, i(k,s) = i\}$. The rest is as in Lemma 6.7.3.

□

The remarkable consequence of Proposition 6.7.5 we promised can be roughly stated as follows: to make any computation for the Gibbs measure involving only a finite number of spins, we can assume that different spins are independent, both for the Gibbs measure and probabilistically. To make this idea precise, let us recall the notation $\mathcal{D}$ of (6.132) (where now η has density proportional to $\exp(-\kappa x^2)$). Keeping the dependence on N implicit, let us denote by μ $(= \mu_N)$ the law in $\mathcal{D}$ of the density X with respect to η of the law of σ_1 under Gibbs' measure. Let us denote by $\mathbf{X} = (X_1, \ldots, X_N)$ an i.i.d. sequence of random elements of law μ and recall the notation $\langle\cdot\rangle_{\mathbf{X}}$ of (6.135).

Theorem 6.7.8. *Consider two integers n, k. Consider continuous bounded functions $U_1, \ldots, U_k$ from $\mathbb{R}^n$ to $\mathbb{R}$, and a continuous function $V : \mathbb{R}^k \to \mathbb{R}$. Then*

$$\lim_{N\to\infty} |\mathsf{E}V(\langle U_1(\sigma_1,\ldots,\sigma_n)\rangle, \langle U_2(\sigma_1,\ldots,\sigma_n)\rangle, \ldots, \langle U_k(\sigma_1,\ldots,\sigma_n)\rangle) \\ -\mathsf{E}V(\langle U_1(\sigma_1,\ldots,\sigma_n)\rangle_{\mathbf{X}}, \ldots, \langle U_k(\sigma_1,\ldots,\sigma_n)\rangle_{\mathbf{X}})| = 0 \,. \tag{6.193}$$

We leave to the reader to formulate and prove an even more general statement involving functions on several replicas.

Proof. Since $U_1, \ldots, U_k$ are bounded, on their range we can uniformly approximate V by a polynomial, so that it suffices to consider the case where V is a monomial,

$$V(x_1, \ldots, x_k) = x_1^{m_1} \cdots x_k^{m_k} \,. \tag{6.194}$$

The next step is to show that we can assume that for each $j \le k$ we have

$$\lim_{(\sigma_1,\ldots,\sigma_n)\to\infty} U_j(\sigma_1, \ldots, \sigma_n) = 0 \,. \tag{6.195}$$

To see this, we first note that without loss of generality we can assume that $|U_j| \le 1$ for each j. Consider for each $j \le k$ a function $U_j^\sim$ with $|U_j^\sim| \le 1$ and assume that for some number S we have

$$\forall i \le n \;,\; |\sigma_i| \le S \Rightarrow U_j^\sim(\sigma_1, \ldots, \sigma_n) = U_j(\sigma_1, \ldots, \sigma_n) \,. \tag{6.196}$$

Then

$$|U_j(\sigma_1, \ldots, \sigma_n) - U_j^\sim(\sigma_1, \ldots, \sigma_n)| \le \sum_{s\le n} \mathbf{1}_{\{\sigma_s \ge S\}} \;,$$

and therefore

$$|\langle U_j(\sigma_1, \ldots, \sigma_n)\rangle - \langle U_j^\sim(\sigma_1, \ldots, \sigma_n)\rangle| \le \sum_{s\le n} \langle \mathbf{1}_{\{\sigma_s \ge S\}}\rangle \,.$$

We note that for numbers $x_1, \ldots, x_k$ and $y_1, \ldots, y_k$, all bounded by 1, we have the elementary inequality

$$|x_1^{m_1}\cdots x_k^{m_k} - y_1^{m_1}\cdots y_k^{m_k}| \le \sum_{j\le k} m_j |x_j - y_j| \;. \tag{6.197}$$

It then follows that if we set

$$\begin{aligned} C &= \langle U_1(\sigma_1,\ldots,\sigma_n)\rangle^{m_1}\cdots\langle U_k(\sigma_1,\ldots,\sigma_n)\rangle^{m_k} \\ C^\sim &= \langle U_1^\sim(\sigma_1,\ldots,\sigma_n)\rangle^{m_1}\cdots\langle U_k^\sim(\sigma_1,\ldots,\sigma_n)\rangle^{m_k} \end{aligned}$$

then

$$|C - C^\sim| \le \sum_{j\le k} m_j \sum_{s\le n}\langle \mathbf{1}_{\{\sigma_s\ge S\}}\rangle \;,$$

and therefore

$$|\mathsf{E}C - \mathsf{E}C^\sim| \le \sum_{j\le k} m_j \sum_{s\le n} \mathsf{E}\langle \mathbf{1}_{\{\sigma_s\ge S\}}\rangle = n\sum_{j\le k} m_j \mathsf{E}\langle \mathbf{1}_{\{\sigma_1\ge S\}}\rangle \;.$$

By Lemma 6.7.3, the right-hand side can be made small for S large, and since we can choose the functions U_j that satisfy (6.196) and $U_j(\sigma_1,\ldots,\sigma_n) = 0$ if one of the numbers $|\sigma_s|$ is $\ge 2S$, this indeed shows that we can assume (6.195).

A function U_j that satisfies (6.195) can be uniformly approximated by a finite sum of functions of the type

$$f_1(\sigma_1)\cdots f_n(\sigma_n) \;,$$

where $|f_s^{(\ell)}|$ is bounded for $s \le n$ and $\ell = 0,1,2$. By expansion we then reduce to the case where

$$U_j(\sigma_1,\ldots,\sigma_n) = f_{1,j}(\sigma_1)\cdots f_{n,j}(\sigma_n) \tag{6.198}$$

and we can furthermore assume that $|f_{s,j}^{(\ell)}|$ is bounded for $\ell = 0,1,2$ and $s \le n$. Assuming (6.194) and (6.198) we have

$$\begin{aligned} B &:= \mathsf{E}V(\langle U_1(\sigma_1,\ldots,\sigma_n)\rangle,\ldots,\langle U_k(\sigma_1,\ldots,\sigma_n)\rangle) \\ &= \mathsf{E}\langle f_{1,1}(\sigma_1)\cdots f_{n,1}(\sigma_n)\rangle^{m_1}\cdots\langle f_{1,k}(\sigma_1)\cdots f_{n,k}(\sigma_n)\rangle^{m_k} \;. \end{aligned}$$

We will write this expression using replicas. Let $m = m_1 + \cdots + m_k$. Let us write $\{1,\ldots,m\}$ as the disjoint union of sets $I_1,\ldots,I_k$ with $\mathrm{card} I_j = m_j$; and for $\ell \in I_j$ and $s \le n$ let us set

$$g_{s,\ell} = f_{s,j} \;,$$

so that in particular for $\ell \in I_j$ we have $\prod_{s\le n} g_{s,\ell}(\sigma_s) = \prod_{s\le n} f_{s,j}(\sigma_s)$. Then, using independence of replicas in the first equality, we get

$$\begin{aligned} \Big\langle \prod_{\ell\le m}\prod_{s\le n} g_{s,\ell}(\sigma_s^\ell)\Big\rangle &= \prod_{\ell\le m}\Big\langle \prod_{s\le n} g_{s,\ell}(\sigma_s)\Big\rangle \\ &= \Big\langle \prod_{s\le n} f_{s,1}(\sigma_s)\Big\rangle^{m_1}\cdots\Big\langle \prod_{s\le n} f_{s,k}(\sigma_s)\Big\rangle^{m_k} \;, \end{aligned}$$

and therefore

$$B = \mathsf{E}\Big\langle \prod_{\ell \le m} \prod_{s \le n} g_{s,\ell}(\sigma_s^\ell) \Big\rangle = \mathsf{E}\Big\langle \prod_{s \le n} \prod_{\ell \le m} g_{s,\ell}(\sigma_s^\ell) \Big\rangle \,.$$

By symmetry among sites, for any indexes $i_1, \ldots, i_n \le N$, all different, we have

$$B = \mathsf{E}\Big\langle \prod_{s \le n} \prod_{\ell \le m} g_{s,\ell}(\sigma_{i_s}^\ell) \Big\rangle \,. \tag{6.199}$$

Therefore, for a number K that does not depend on N, we have

$$\left| B - \mathsf{E}\frac{1}{N^n} \sum_{i_1,\ldots,i_n} \Big\langle \prod_{s \le n} \prod_{\ell \le m} g_{s,\ell}(\sigma_{i_s}^\ell) \Big\rangle \right| \le \frac{K}{N} \,, \tag{6.200}$$

where the summation is over all values of $i_1, \ldots, i_n$. This is seen by using (6.199) for the terms of the summation where all the indices are different and by observing that there are at most KN^{n-1} other terms. Now

$$\frac{1}{N^n} \sum_{i_1,\ldots,i_n} \prod_{s \le n} \prod_{\ell \le m} g_{s,\ell}(\sigma_{i_s}^\ell) = \prod_{s \le n} \Big(\frac{1}{N} \sum_{i \le N} \prod_{\ell \le m} g_{s,\ell}(\sigma_i^\ell) \Big) \,.$$

Defining

$$R_s = \frac{1}{N} \sum_{i \le N} \prod_{\ell \le m} g_{s,\ell}(\sigma_i^\ell) \,,$$

we obtain from (6.200) that

$$\left| B - \mathsf{E}\Big\langle \prod_{s \le n} R_s \Big\rangle \right| \le \frac{K}{N} \,.$$

Proposition 6.7.5 shows that for each s we have $\mathsf{E}\langle |R_s - \mathsf{E}R_s| \rangle \le KN^{-1/4}$, so that, replacing in turn each R_s by $\mathsf{E}\langle R_s \rangle$ one at a time,

$$\left| \mathsf{E}\Big\langle \prod_{s \le n} R_s \Big\rangle - \prod_{s \le n} \mathsf{E}\langle R_s \rangle \right| \le \frac{K}{N^{1/4}} \,,$$

and therefore

$$\left| B - \prod_{s \le n} \mathsf{E}\langle R_s \rangle \right| \le \frac{K}{N^{1/4}} \,. \tag{6.201}$$

Now, using symmetry among sites in the first equality,

$$\mathsf{E}\langle R_s \rangle = \mathsf{E}\Big\langle \prod_{\ell \le m} g_{s,\ell}(\sigma_s^\ell) \Big\rangle = \mathsf{E} \prod_{\ell \le m} \langle g_{s,\ell}(\sigma_s) \rangle = \mathsf{E} \prod_{j \le k} \langle f_{s,j}(\sigma_s) \rangle^{m_j} \,,$$

and we have shown that

$$\lim_{N\to\infty}\left|B - \prod_{s\le n} \mathsf{E}\prod_{j\le k}\langle f_{s,j}(\sigma_s)\rangle^{m_j}\right| = 0\ . \tag{6.202}$$

In the special case where V is given by (6.194) and U_j is given by (6.198), we have

$$\mathsf{E}V(\langle U_1(\sigma_1,\ldots,\sigma_n)\rangle_{\mathbf{X}},\ldots,\langle U_k(\sigma_1,\ldots,\sigma_n)\rangle_{\mathbf{X}}) = \prod_{s\le n}\mathsf{E}\prod_{j\le k}\langle f_{s,j}(\sigma_s)\rangle^{m_j}\ ,$$

so that (6.202) is exactly (6.193) in this special case. As we have shown, this special case implies the general one. □

Given n, k, and a number C, inspection of the previous argument shows that the convergence is uniform over the families of functions $U_1,\ldots,U_k$ that satisfy $|U_1|,\ldots,|U_k| \le C$.

We turn to the main result of this section, the proof that "in the limit $\mu = T(\mu)$". We recall the definition of $\mathcal{E}_r$ as in (6.49), and that r is a Poisson r.v. of expectation αp. Let us denote by $\mathbf{X} = (X_i)_{i\ge 1}$ an i.i.d. sequence, where $X_i \in \mathcal{D}$ is a random element of law $\mu = \mu_N$ (the law of the density with respect to η of the law of σ_1 under Gibbs' measure), and let us define $T(\mu)$ as follows: if

$$Y = \frac{\langle \mathcal{E}_r\rangle_{\mathbf{X}}}{\langle \mathrm{Av}\mathcal{E}_r\rangle_{\mathbf{X}}} \in \mathcal{D}\ ,$$

then $T(\mu)$ is the law of Y in $\mathcal{D}$. The following asserts in a weak sense that in the limit $T(\mu_N) = \mu_N$.

Theorem 6.7.9. *Consider an integer n, and continuous bounded functions $f_1,\ldots,f_n$ on $\mathbb{R}$. Then*

$$\lim_{N\to\infty}\left|\mathsf{E}\langle f_1(\sigma_1)\rangle\cdots\langle f_n(\sigma_1)\rangle - \mathsf{E}\frac{\langle \mathrm{Av} f_1(\varepsilon)\mathcal{E}_r\rangle_{\mathbf{X}}}{\langle \mathrm{Av}\mathcal{E}_r\rangle_{\mathbf{X}}}\cdots\frac{\langle \mathrm{Av} f_n(\varepsilon)\mathcal{E}_r\rangle_{\mathbf{X}}}{\langle \mathrm{Av}\mathcal{E}_r\rangle_{\mathbf{X}}}\right| = 0\ . \tag{6.203}$$

To relate (6.203) with the statement that "$T(\mu) = \mu$", we note that

$$\frac{\langle \mathrm{Av} f_s(\varepsilon)\mathcal{E}_r\rangle_{\mathbf{X}}}{\langle \mathrm{Av}\mathcal{E}_r\rangle_{\mathbf{X}}} = \int Y f_s \mathrm{d}\eta\ ,$$

so that writing $X = X_1$, (6.203) means that

$$\lim_{N\to\infty}\left|\mathsf{E}\int f_1 X\mathrm{d}\eta\cdots\int f_n X\mathrm{d}\eta - \mathsf{E}\int f_1 Y\mathrm{d}\eta\cdots\int f_n Y\mathrm{d}\eta\right| = 0\ . \tag{6.204}$$

In a weak sense this asserts that in the limit the laws of X (i.e μ) and Y (i.e. $T(\mu)$) coincide.

While we do not know how to prove this directly, in a second stage we will deduce from Theorem 6.7.9 that, as expected,

$$\lim_{N\to\infty} d(\mu_N, T(\mu_N)) = 0 \,, \tag{6.205}$$

where d is the transportation-cost distance.

Let us now explain the strategy to prove (6.203). The basic idea is to combine Theorem 6.7.8 with the cavity method. We find convenient to use the cavity method between an N-spin and an $(N+1)$-spin system. Let us define α' by

$$\alpha'(N+1)\frac{\binom{N}{p}}{\binom{N+1}{p}} = \alpha N \,, \tag{6.206}$$

and let us consider a Poisson r.v. r' with $\mathsf{E}r' = \alpha' p$. The letter r' keeps this meaning until the end of this chapter. For $j \geq 1$, let us consider independent copies θ_j of θ, and sets $\{i(j,1),\ldots,i(j,p-1)\}$ that are uniformly distributed among the subsets of $\{1,\ldots,N\}$ of cardinality $p-1$. Of course we assume that all the randomness there is independent of the randomness of $\langle\cdot\rangle$. Let us define

$$-H(\boldsymbol{\sigma},\varepsilon) = \sum_{j\leq r'} \theta_j(\sigma_{i(j,1)},\ldots,\sigma_{i(j,p-1)},\varepsilon)$$

and $\mathcal{E} = \mathcal{E}(\boldsymbol{\sigma},\varepsilon) = \exp(-H(\boldsymbol{\sigma},\varepsilon))$. Recalling the Hamiltonian (6.171), the Hamiltonian $-H' = -H_N - H$ is the Hamiltonian of an $(N+1)$-spin system, where the value of α has been replaced by α' given by (6.206). Let us denote by $\langle\cdot\rangle'$ an average for the Gibbs measure relative to H'. Writing $\varepsilon = \sigma_{N+1}$, symmetry between sites implies

$$\mathsf{E}\langle f_1(\sigma_1)\rangle'\cdots\langle f_n(\sigma_1)\rangle' = \mathsf{E}\langle f_1(\varepsilon)\rangle'\cdots\langle f_n(\varepsilon)\rangle' \,. \tag{6.207}$$

Now, for a function $f = f(\boldsymbol{\sigma},\varepsilon)$, the cavity formula

$$\langle f\rangle' = \frac{\langle \mathrm{Av} f\mathcal{E}\rangle}{\langle \mathrm{Av}\mathcal{E}\rangle}$$

holds, where Av means integration in ε with respect to η, and where $\mathcal{E} = \mathcal{E}(\boldsymbol{\sigma},\varepsilon) = \exp(-H(\boldsymbol{\sigma},\varepsilon))$. We rewrite (6.207) as

$$\mathsf{E}\frac{\langle f_1(\sigma_1)\mathrm{Av}\mathcal{E}\rangle}{\langle \mathrm{Av}\mathcal{E}\rangle}\cdots\frac{\langle f_n(\sigma_1)\mathrm{Av}\mathcal{E}\rangle}{\langle \mathrm{Av}\mathcal{E}\rangle} = \mathsf{E}\frac{\langle \mathrm{Av} f_1(\varepsilon)\mathcal{E}\rangle}{\langle \mathrm{Av}\mathcal{E}\rangle}\cdots\frac{\langle \mathrm{Av} f_n(\varepsilon)\mathcal{E}\rangle}{\langle \mathrm{Av}\mathcal{E}\rangle} \,. \tag{6.208}$$

We will then use Theorem 6.7.8 to approximately compute both sides of (6.208) to obtain (6.203). However an obstacle is that the denominators can be very small, or, in other words, that the function x/y is not continuous at $y = 0$. To solve this problem we consider $\delta > 0$ and we will replace these denominators by $\delta + \langle \mathrm{Av}\mathcal{E}\rangle$.

We will need to take limits as $\delta \to 0$, and in order to be able to exchange these limits with the limits as $N \to \infty$ we need the following.

Lemma 6.7.10. *Assume that $f = f(\boldsymbol{\sigma}, \varepsilon)$ is bounded. Then*

$$\lim_{\delta\to 0} \sup_N \mathsf{E}\left|\frac{\langle \mathrm{Av} f\mathcal{E}\rangle}{\langle \mathrm{Av}\mathcal{E}\rangle} - \frac{\langle \mathrm{Av} f\mathcal{E}\rangle}{\delta + \langle \mathrm{Av}\mathcal{E}\rangle}\right| = 0 \ .$$

Proof. First, if $|f| \le C$, we have

$$\left|\frac{\langle \mathrm{Av} f\mathcal{E}\rangle}{\langle \mathrm{Av}\mathcal{E}\rangle} - \frac{\langle \mathrm{Av} f\mathcal{E}\rangle}{\delta + \langle \mathrm{Av}\mathcal{E}\rangle}\right| = \frac{\delta|\langle \mathrm{Av} f\mathcal{E}\rangle|}{\langle \mathrm{Av}\mathcal{E}\rangle(\delta + \langle \mathrm{Av}\mathcal{E}\rangle)} \le \frac{C\delta}{\delta + \langle \mathrm{Av}\mathcal{E}\rangle} \ .$$

Next, we have

$$\mathsf{E}\frac{\delta}{\delta + \langle \mathrm{Av}\mathcal{E}\rangle} \le \sqrt{\delta} + \mathsf{P}(\langle \mathrm{Av}\mathcal{E}\rangle \le \sqrt{\delta}) \ ,$$

and, writing $H = H(\boldsymbol{\sigma}, \varepsilon)$,

$$\langle \mathrm{Av}\mathcal{E}\rangle = \langle \mathrm{Av} \exp(-H)\rangle \ge \exp\langle -\mathrm{Av} H\rangle \ ,$$

so that

$$\mathsf{P}\big(\langle \mathrm{Av}\mathcal{E}\rangle \le \sqrt{\delta}\big) \le \mathsf{P}\left(\langle -\mathrm{Av} H\rangle \ge \log\frac{1}{\sqrt{\delta}}\right) \le \frac{\mathsf{E}\langle \mathrm{Av}|H|\rangle}{\log(1/\sqrt{\delta})} \ .$$

It follows from (6.173) that

$$|H(\boldsymbol{\sigma}, \varepsilon)| \le \sum_{j\le r'}\left(|\theta_j(0,\dots,0)| + A\Big(\sum_{s\le p-1} |\sigma_{i(j,s)}| + |\varepsilon|\Big)\right) ,$$

so that (6.178) and Lemma 6.7.3 imply that $\sup_N \mathsf{E}\langle \mathrm{Av}|H|\rangle < \infty$ and the lemma is proved. □

Lemma 6.7.11. *We have*

$$\lim_{N\to\infty}\left|\mathsf{E}\frac{\langle f_1(\sigma_1)\mathrm{Av}\mathcal{E}\rangle}{\langle \mathrm{Av}\mathcal{E}\rangle}\cdots\frac{\langle f_n(\sigma_1)\mathrm{Av}\mathcal{E}\rangle}{\langle \mathrm{Av}\mathcal{E}\rangle} - \mathsf{E}\langle f_1(\sigma_1)\rangle\cdots\langle f_n(\sigma_1)\rangle\right| = 0 \ . \quad (6.209)$$

Proof. Consider the event $\Omega = \Omega_1 \cup \Omega_2 \cap \Omega_3$, where

$$\Omega_1 = \{\exists j \le r' \ , \ i(j,1) = 1\}$$
$$\Omega_2 = \{\exists j, j' \le r' \ , \ j \ne j' \ , \ \exists \ell, \ell' \le p-1 \ , \ i(j,\ell) = i(j',\ell')\} \quad (6.210)$$
$$\Omega_3 = \{(p-1)(r'+1) \le N\} \ , \quad (6.211)$$

so that as we have used many times we have

$$\mathsf{P}(\Omega) \le \frac{K}{N} \ . \quad (6.212)$$

Let us now define

$$U = \mathrm{Av} \exp \sum_{1 \le j \le r'} \theta_j(\sigma_{j(p-1)+1}, \ldots, \sigma_{(j+1)(p-1)}, \varepsilon) \tag{6.213}$$

when $(r'+1)(p-1) \le N$ and $U = 1$ otherwise. The reader observes that U depends only on the spins σ_i for $i \ge p$. On Ω^c we have $i(j,1) > 1$ for all $j < r$, and the indexes $i(j,\ell)$ are all different. Thus symmetry between sites implies that for any $\delta > 0$,

$$\begin{aligned} &\mathsf{E}\left(\mathbf{1}_{\Omega^c} \frac{\langle f_1(\sigma_1)\mathrm{Av}\mathcal{E}\rangle}{\delta + \langle \mathrm{Av}\mathcal{E}\rangle} \cdots \frac{\langle f_n(\sigma_1)\mathrm{Av}\mathcal{E}\rangle}{\delta + \langle \mathrm{Av}\mathcal{E}\rangle}\right) \\ &= \mathsf{E}\left(\mathbf{1}_{\Omega^c} \frac{\langle f_1(\sigma_1)U\rangle}{\delta + \langle U\rangle} \cdots \frac{\langle f_n(\sigma_1)U\rangle}{\delta + \langle U\rangle}\right) . \end{aligned} \tag{6.214}$$

We claim that

$$\begin{aligned} \lim_{N\to\infty} \Bigg| &\mathsf{E}\frac{\langle f_1(\sigma_1)U\rangle}{\delta + \langle U\rangle} \cdots \frac{\langle f_n(\sigma_1)U\rangle}{\delta + \langle U\rangle} \\ &- \mathsf{E}\frac{\langle f_1(\sigma_1)\rangle_{\mathbf{x}}\langle U\rangle_{\mathbf{x}}}{\delta + \langle U\rangle_{\mathbf{x}}} \cdots \frac{\langle f_n(\sigma_1)\rangle_{\mathbf{x}}\langle U\rangle_{\mathbf{x}}}{\delta + \langle U\rangle_{\mathbf{x}}}\Bigg| = 0 . \end{aligned} \tag{6.215}$$

To see this we simply use Theorem 6.7.8 given r' and the functions θ_j, $j \le r'$. Since by (6.212) the influence of Ω vanishes in the limit, we get from (6.214) that

$$\begin{aligned} \lim_{N\to\infty} \Bigg| &\mathsf{E}\frac{\langle f_1(\sigma_1)\mathrm{Av}\mathcal{E}\rangle}{\delta + \langle \mathrm{Av}\mathcal{E}\rangle} \cdots \frac{\langle f_n(\sigma_1)\mathrm{Av}\mathcal{E}\rangle}{\delta + \langle \mathrm{Av}\mathcal{E}\rangle} \\ &- \mathsf{E}\frac{\langle f_1(\sigma_1)\rangle_{\mathbf{x}}\langle U\rangle_{\mathbf{x}}}{\delta + \langle U\rangle_{\mathbf{x}}} \cdots \frac{\langle f_n(\sigma_1)\rangle_{\mathbf{x}}\langle U\rangle_{\mathbf{x}}}{\delta + \langle U\rangle_{\mathbf{x}}}\Bigg| = 0 . \end{aligned} \tag{6.216}$$

Without loss of generality we can assume that $|f_s| \le 1$ for each s. The inequality (6.197) and Lemma 6.7.10 yield

$$\begin{aligned} &\lim_{\delta\to 0} \sup_N \left| \mathsf{E}\frac{\langle f_1(\sigma_1)\mathrm{Av}\mathcal{E}\rangle}{\delta + \langle \mathrm{Av}\mathcal{E}\rangle} \cdots \frac{\langle f_n(\sigma_1)\mathrm{Av}\mathcal{E}\rangle}{\delta + \langle \mathrm{Av}\mathcal{E}\rangle} - \mathsf{E}\frac{\langle f_1(\sigma_1)\mathrm{Av}\mathcal{E}\rangle}{\langle \mathrm{Av}\mathcal{E}\rangle} \cdots \frac{\langle f_n(\sigma_1)\mathrm{Av}\mathcal{E}\rangle}{\langle \mathrm{Av}\mathcal{E}\rangle} \right| \\ &= 0 . \end{aligned} \tag{6.217}$$

Proceeding as in Lemma 6.7.10, we get

$$\lim_{\delta\to 0} \sup_N \mathsf{E}\left| \frac{\langle U\rangle_{\mathbf{x}}}{\delta + \langle U\rangle_{\mathbf{x}}} - 1 \right| = 0 , \tag{6.218}$$

and proceeding as in (6.217) we obtain

$$\lim_{\delta\to 0} \sup_N \left| \mathsf{E}\langle f_1(\sigma_1)\rangle_{\mathbf{x}} \cdots \langle f_n(\sigma_1)\rangle_{\mathbf{x}} - \mathsf{E}\frac{\langle f_1(\sigma_1)\rangle_{\mathbf{x}}\langle U\rangle_{\mathbf{x}}}{\delta + \langle U\rangle_{\mathbf{x}}} \cdots \frac{\langle f_n(\sigma_1)\rangle_{\mathbf{x}}\langle U\rangle_{\mathbf{x}}}{\delta + \langle U\rangle_{\mathbf{x}}} \right| = 0.$$

Combining with (6.217) and (6.216) proves (6.209) since $\langle f_s(\sigma_1)\rangle = \langle f_s(\sigma_1)\rangle_{\mathbf{X}}$. □

To complete the proof of Theorem 6.7.9, we show the following, where we lighten notation by writing $f_s = f_s(\varepsilon)$.

Lemma 6.7.12. *We have*

$$\lim_{N\to\infty}\left|\mathsf{E}\frac{\langle \mathrm{Av} f_1\mathcal{E}\rangle}{\langle \mathrm{Av}\mathcal{E}\rangle}\cdots\frac{\langle \mathrm{Av} f_n\mathcal{E}\rangle}{\langle \mathrm{Av}\mathcal{E}\rangle} - \mathsf{E}\frac{\langle \mathrm{Av} f_1\mathcal{E}_r\rangle_{\mathbf{X}}}{\langle \mathrm{Av}\mathcal{E}_r\rangle_{\mathbf{X}}}\cdots\frac{\langle \mathrm{Av} f_n\mathcal{E}_r\rangle_{\mathbf{X}}}{\langle \mathrm{Av}\mathcal{E}_r\rangle_{\mathbf{X}}}\right| = 0\,.$$

Proof. We follow the method of Lemma 6.7.11, keeping its notation. For $s \le n$ we define

$$U_s = \mathrm{Av} f_s(\varepsilon)\exp\sum_{1\le j\le r'}\theta(\sigma_{j(p-1)+1},\ldots,\sigma_{(j+1)(p-1)},\varepsilon)$$

when $(r'+1)(p-1)\le N$ and $U_s = 1$ otherwise. Consider $\delta > 0$. Recalling (6.211) and (6.213), symmetry between sites yields

$$\begin{aligned}&\mathsf{E}\left(\mathbf{1}_{\Omega^c}\frac{\langle \mathrm{Av} f_1\mathcal{E}\rangle}{\delta+\langle \mathrm{Av}\mathcal{E}\rangle}\cdots\frac{\langle \mathrm{Av} f_n\mathcal{E}\rangle}{\delta+\langle \mathrm{Av}\mathcal{E}\rangle}\right)\\ &= \mathsf{E}\left(\mathbf{1}_{\Omega^c}\frac{\langle U_1\rangle}{\delta+\langle U\rangle}\cdots\frac{\langle U_n\rangle}{\delta+\langle U\rangle}\right)\,. \end{aligned}\tag{6.219}$$

Moreover Theorem 6.7.8 implies

$$\lim_{N\to\infty}\left|\mathsf{E}\frac{\langle U_1\rangle}{\delta+\langle U\rangle}\cdots\frac{\langle U_n\rangle}{\delta+\langle U\rangle} - \mathsf{E}\frac{\langle U_1\rangle_{\mathbf{X}}}{\delta+\langle U\rangle_{\mathbf{X}}}\cdots\frac{\langle U_n\rangle_{\mathbf{X}}}{\delta+\langle U\rangle_{\mathbf{X}}}\right| = 0\,.$$

Since the influence of Ω vanishes in the limit, and exchanging again the limits $N\to\infty$ and $\delta\to 0$ as permitted by Lemma 6.7.10 (and a similar argument for the terms $\mathsf{E}\langle U_s\rangle_{\mathbf{X}}/(\delta+\langle U\rangle_{\mathbf{X}})$), we obtain

$$\lim_{N\to\infty}\left|\mathsf{E}\frac{\langle \mathrm{Av} f_1\mathcal{E}\rangle}{\langle \mathrm{Av}\mathcal{E}\rangle}\cdots\frac{\langle \mathrm{Av} f_n\mathcal{E}\rangle}{\langle \mathrm{Av}\mathcal{E}\rangle} - \mathsf{E}\frac{\langle U_1\rangle_{\mathbf{X}}}{\langle U\rangle_{\mathbf{X}}}\cdots\frac{\langle U_n\rangle_{\mathbf{X}}}{\langle U\rangle_{\mathbf{X}}}\right| = 0\,.$$

It then remains only to show that

$$\lim_{N\to\infty}\left|\mathsf{E}\frac{\langle U_1\rangle_{\mathbf{X}}}{\langle U\rangle_{\mathbf{X}}}\cdots\frac{\langle U_n\rangle_{\mathbf{X}}}{\langle U\rangle_{\mathbf{X}}} - \mathsf{E}\frac{\langle \mathrm{Av} f_1\mathcal{E}_r\rangle_{\mathbf{X}}}{\langle \mathrm{Av}\mathcal{E}_r\rangle_{\mathbf{X}}}\cdots\frac{\langle \mathrm{Av} f_n\mathcal{E}_r\rangle_{\mathbf{X}}}{\langle \mathrm{Av}\mathcal{E}_r\rangle_{\mathbf{X}}}\right| = 0\,,$$

which should be obvious by the definitions of U, $\mathcal{E}_r$ and U_s and since r' is a Poisson r.v. and, as $N\to\infty$, $\mathsf{E}r' = \alpha' p \to \alpha p = \mathsf{E}r$. □

We now state the desired strengthening of Theorem 6.7.9.

Theorem 6.7.13. *If d denotes the transportation-cost distance associated to the L^1 norm in $\mathcal{D}$, we have*

$$\lim_{N\to\infty} d(\mu_N, T(\mu_N)) = 0 . \tag{6.220}$$

As we shall see, the sequence $\mu = \mu_N$ is tight, and (6.220) implies that any cluster point of this sequence is a solution of the equation $\mu = T(\mu)$. If we knew that this equation has a unique solution, we would conclude that the sequence (μ_N) converges to this solution, and we could pursue the study of the model and in particular we could compute

$$\lim_{N\to\infty} \frac{1}{N} \mathsf{E} \log \int \exp(-H_N(\boldsymbol{\sigma}) - \kappa \|\boldsymbol{\sigma}\|^2) \mathrm{d}\boldsymbol{\sigma} .$$

Thus, further results seem to depend on the following.

Research Problem 6.7.14. (Level 2) Prove that the equation $\mu = T(\mu)$ has a unique solution.

One really wonders what kind of methods could be used to approach this question. Even if this can be solved, the challenge remains to find situations where in the relation (see (6.170))

$$\begin{aligned} &\mathsf{E} \frac{1}{N} \log \eta^{\otimes N}\Big(\bigcap_{k\le M} U_k\Big) \\ &= \lim_{\beta\to\infty} \frac{1}{N} \mathsf{E} \log \int \exp \beta \sum_{k\le M} \theta_k(\sigma_{i(k,1)}, \dots, \sigma_{i(k,p)}) \mathrm{d}\eta^{\otimes N}(\boldsymbol{\sigma}) \end{aligned}$$

one can exchange the limits $N \to \infty$ and $\beta \to \infty$. A similar problem in a different context will be solved in Chapter 8.

We turn to the technicalities required to prove Theorem 6.7.13. They are not difficult, although it is hard to believe that these measure-theoretic considerations are really relevant to spin glasses. For this reason it seems that the only potential readers for these arguments will be well versed in measure theory. Consequently the proofs (that use a few basic facts of analysis, which can be found in any textbook) will be a bit sketchy.

Lemma 6.7.15. *Consider a number B and*

$$\mathcal{D}(B) = \{f \in \mathcal{D} \; ; \; \forall x, y \; , \; f(y) \le f(x) \exp B|y - x|\} .$$

Then $\mathcal{D}(B)$ is norm-compact in $L^1(\eta)$.

Proof. A function f in $\mathcal{D}(B)$ satisfies

$$f(0) \exp(-B|x|) \le f(x) \le f(0) \exp B|x| ,$$

so that since $\int f(x) \mathrm{d}\eta(x) = 1$, we have $K^{-1} \le f(0) \le K$ where K depends on B and κ only. Moreover $\mathcal{D}(B)$ is equi-continuous on every interval, so a sequence (f_n) in $\mathcal{D}(B)$ has a subsequence that converges uniformly in any interval; since, given any $\varepsilon > 0$, there exists a number x_0 for which

$$f \in \mathcal{D}(B) \quad \Rightarrow \quad \int_{|x| \ge x_0} |f(x)| \mathrm{d}\eta(x) \le \varepsilon ,$$

it follows that this subsequence converges in $L^1(\eta)$. □

We recall the number A of (6.173).

Lemma 6.7.16. *For each N and each k we have*

$$\mu(\mathcal{D}(kA)) \ge \mathsf{P}(r \le k) , \tag{6.221}$$

where r is a Poisson r.v. of mean αp.

Proof. This is a reformulation of Lemma 6.7.4 since (6.175) means that $Y \in \mathcal{D}(rA)$. □

Proof of Theorem 6.7.13. The set of probability measures μ on $\mathcal{D}$ that satisfy (6.221) for each k is tight (and consequently is compact for the transportation-cost distance). Assuming if possible that (6.220) fails, we can find $\varepsilon > 0$ and a converging subsequence $(\mu_{N(k)})_{k \ge 1}$ of the sequence (μ_N) such that

$$\forall k , \quad d(\mu_{N(k)}, T(\mu_{N(k)})) \ge \varepsilon .$$

We defined $T(\nu)$ for $\nu = \mu_N$. We leave it to the reader to define (in the same manner) $T(\nu)$ for any probability measure ν on $\mathcal{D}$ and to show that the operator T is continuous for d. So that if we define $\nu = \lim_k \mu_{N(k)}$, then $T(\nu) = \lim_k T(\mu_{N(k)})$ and therefore $d(\nu, T(\nu)) \ge \varepsilon$. In particular we have $\nu \ne T(\nu)$. On the other hand, given continuous bounded functions $f_1, \ldots, f_n$ on $\mathbb{R}$, since μ_N is the law of Y (the density with respect to η of the law of σ_1 under Gibbs's measure) in $\mathcal{D}$ we have

$$\begin{aligned} \mathsf{E}\langle f_1(\sigma_1)\rangle \cdots \langle f_n(\sigma_1)\rangle &= \mathsf{E}\left(\int f_1 Y \mathrm{d}\eta \cdots \int f_n Y \mathrm{d}\eta\right) \\ &= \int \left(\int f_1 Y \mathrm{d}\eta \cdots \int f_n Y \mathrm{d}\eta\right) \mathrm{d}\mu_N(Y) . \end{aligned} \tag{6.222}$$

The map

$$\nu \mapsto \psi(\nu) := \int \left(\int f_1 Y \mathrm{d}\eta \cdots \int f_n Y \mathrm{d}\eta\right) \mathrm{d}\nu(Y)$$

is continuous for the transportation-cost distance; in fact if $|f_s| \le 1$ for each s, one can easily show that $|\psi(\nu) - \psi(\nu')| \le n d(\nu, \nu')$. Therefore the limit of the right-hand side of (6.222) along the sequence $(N(k))$ is

$$\int \left(\int f_1 Y \mathrm{d}\eta \cdots \int f_n Y \mathrm{d}\eta \right) \mathrm{d}\nu(Y) .$$

Also, the definition of $T(\mu_N)$ implies

$$\begin{aligned} &\mathsf{E} \frac{\langle \mathrm{Av} f_1 \mathcal{E}_r \rangle_{\mathbf{X}}}{\langle \mathrm{Av} \mathcal{E}_r \rangle_{\mathbf{X}}} \cdots \frac{\langle \mathrm{Av} f_n \mathcal{E}_r \rangle_{\mathbf{X}}}{\langle \mathrm{Av} \mathcal{E}_r \rangle_{\mathbf{X}}} \\ &= \int \left(\int f_1 Y \mathrm{d}\eta \cdots \int f_n Y \mathrm{d}\eta \right) \mathrm{d}T(\nu_N)(Y) \end{aligned} \tag{6.223}$$

and the limit of the previous quantity along the sequence $(N(k))$ is

$$\int \left(\int f_1 Y \mathrm{d}\eta \cdots \int f_n Y \mathrm{d}\eta \right) \mathrm{d}T(\nu)(Y) .$$

Using (6.203) we get

$$\begin{aligned} &\int \left(\int f_1 Y \mathrm{d}\eta \cdots \int f_n Y \mathrm{d}\eta \right) \mathrm{d}\nu(Y) \\ &= \int \left(\int f_1 Y \mathrm{d}\eta \cdots \int f_n Y \mathrm{d}\eta \right) \mathrm{d}T(\nu)(Y) . \end{aligned} \tag{6.224}$$

We will now show that this identity implies $\nu = T(\nu)$, a contradiction which completes the proof of the theorem. Approximating a function on a bounded set by a polynomial yields that if F is a continuous function of n variables, then

$$\begin{aligned} &\int F \left(\int f_1 Y \mathrm{d}\eta, \ldots, \int f_n Y \mathrm{d}\eta \right) \mathrm{d}\nu(Y) \\ &= \int F \left(\int f_1 Y \mathrm{d}\eta, \ldots, \int f_n Y \mathrm{d}\eta \right) \mathrm{d}T(\nu)(Y) . \end{aligned}$$

Consequently,

$$\int \varphi(Y) \mathrm{d}\nu(Y) = \int \varphi(Y) \mathrm{d}T(\nu)(Y) , \tag{6.225}$$

whenever $\varphi(Y)$ is a pointwise limit of a sequence of uniformly bounded functions of the type

$$Y \mapsto F \left(\int f_1 Y \mathrm{d}\eta, \ldots, \int f_n Y \mathrm{d}\eta \right) .$$

These include the functions of the type

$$\varphi(Y) = \min \left(1, \min_{k \le k_1} (a_k + \|Y - Y_k\|_1) \right) , \tag{6.226}$$

where a_k are ≥ 0 numbers. This is because

$$\varphi(Y) = \min\left(1, \min_{k \le k_1}\left(a_k + \max\left|\int fY \mathrm{d}\eta - \int fY_k \mathrm{d}\eta\right|\right)\right),$$

where the maximum is over $|f| \le 1$, f continuous. Any $[0,1]$-valued, 1-Lipschitz function φ on $\mathcal{D}$ is the pointwise limit of a sequence of functions of the type (6.226). It then follows that (6.225) implies that $\nu = T(\nu)$. □

6.8 Notes and Comments

The first paper "solving" a comparable model at high temperature is [153].

A version of Theorem 6.5.1 "with replica symmetry breaking" is presented in [115], where the proof of Theorem 6.5.1 given here can be found. This proof is arguably identical to the original proof of [60], but the computations are much simpler. This is permitted by the identification of which property of θ is really used (i.e. (6.117)). Another relevant paper is [78], but it deals only with a very special model.

An interesting feature of the present chapter is that we gain control of the model "in two steps", the first of which is Theorem 6.2.2. It would be esthetically pleasing to find a proof "in one step" of a statement including both Theorems 6.2.2 and 6.4.1.

There is currently intense interest in specific models of the type considered in this chapter, see e.g. [51] and [102].